

A computational framework for polyconvex large strain elasticity for geometrically exact beam theory

Rogelio Ortigosa¹ · Antonio J. Gil¹ · Javier Bonet¹ · Christian Hesch²

Received: 5 March 2015 / Accepted: 21 August 2015

© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract In this paper, a new computational framework is presented for the analysis of nonlinear beam finite elements subjected to large strains. Specifically, the methodology recently introduced in Bonet et al. (Comput Methods Appl Mech Eng 283:1061–1094, 2015) in the context of three dimensional polyconvex elasticity is extended to the geometrically exact beam model of Simo (Comput Methods Appl Mech Eng 49:55–70, 1985), the starting point of so many other finite element beam type formulations. This new variational framework can be viewed as a continuum degenerate formulation which, moreover, is enhanced by three key novelties. First, in order to facilitate the implementation of the sophisticated polyconvex constitutive laws particularly associated with beams undergoing large strains, a novel tensor cross product algebra by Bonet et al. (Comput Methods Appl Mech Eng 283:1061–1094, 2015) is adopted, leading to an elegant and physically meaningful representation of an otherwise complex computational framework. Second, the paper shows how the novel algebra facilitates the re-expression of any invariant of the deformation gradient, its cofactor and its determinant in terms of the classical beam strain measures. The latter being very useful whenever a classical beam implementation is preferred. This is particularised for the case of a Mooney–Rivlin model although the technique can be straightforwardly generalised to other more complex isotropic and anisotropic polyconvex models. Third, the connection between the two most accepted

restrictions for the definition of constitutive models in three dimensional elasticity and beams is shown, bridging the gap between the continuum and its degenerate beam description. This is carried out via a novel insightful representation of the tangent operator.

Keywords Polyconvexity · Geometrically exact beam theory · Continuum degenerate beam formulation · Finite elements

1 Introduction

Most classical beam theories [1–8] are based on the definition of so called beam strain measures, namely the axial-shear and torsional–bending strain vectors. Work conjugates to those variables (typically denoted as resultant contact forces and resultant contact couples [9]) can be derived from a hyperelastic energy functional which is defined in terms of the beam strain measures [1–8]. Other authors prefer an alternative continuum degenerate approach [10,11] where, typically, the Lagrangian strain tensor is retained as the main strain measure and particularised for the kinematic description of the beam. In these latter formulations, the second Piola–Kirchhoff stress tensor emerges as the work conjugate variable and can be derived from a hyperelastic energy functional expressed as a function of the Lagrangian strain tensor.

The present manuscript aims to extend the variational and computational framework recently introduced in Bonet et al. [12] in the context of three dimensional elasticity to the geometrically exact Simo [13] beam model, the starting point of so many other finite element beam type formulations [1–7]. The variational approach proposed herein can therefore be viewed as a continuum degenerate formulation emerging from that presented in [12].

✉ Antonio J. Gil
a.j.gil@swansea.ac.uk

¹ Zienkiewicz Centre for Computational Engineering,
College of Engineering, Swansea University,
Bay Campus SA1 8EN, UK

² Institute of Mechanics, Karlsruhe Institute of Technology,
Karlsruhe, Germany

The formulation by Bonet et al. [12] has been particularly formulated for use in large strain scenarios [14, 15], where appropriate convexity criteria [16–18] are necessary to ensure the well posedness of the problem. The most well-established of these criteria is the concept of polyconvexity [19–25] whereby the strain energy is expressed as a convex multi-value function of an extended kinematic set defined by the deformation gradient (fibre map), its determinant (volume map) and its cofactor (area map).

Numerous authors have previously incorporated the concept of polyconvexity into computational models for both isotropic and non-isotropic materials for a variety of applications [26–31]. However, the classical approach consists of ensuring that the strain energy satisfies the polyconvexity condition first but then proceeds towards a computational solution by re-expressing the energy in terms of the deformation gradient alone. Pioneered in [29] and [12, 32, 33], an alternative framework is proposed based on maintaining as independent variables the extended kinematic set on which the strain energy is expressed as a convex function, namely, the deformation gradient, its cofactor and its determinant.

In this paper, the latter approach is particularised for a degenerate beam description. The paper also presents the work conjugates of the extended kinematic set in the context of beam theory as well as novel expressions for the tangent operators. The present formulation utilises a novel algebra based on a tensor cross product operation pioneered in [34] and reintroduced and exploited for the first time in [12, 32, 33] in the context of solid mechanics. The new tensor cross product operation is particularly helpful when dealing with polyconvex constitutive laws, where invariants of the cofactor and the determinant of the deformation feature heavily in the representation of the strain energy functional.

In addition, the paper shows how the novel algebra facilitates the re-expression of any invariant of the deformation gradient, its cofactor and its determinant in terms of the classical beam strain measures. The latter being very useful whenever a classical beam implementation is preferred. As an example, and to the best of the authors knowledge, this is the first time that the strain energy function for a polyconvex Mooney–Rivlin (or Neo-Hookean) material is expressed entirely in terms of the classical beam strain measures.

Convexity of a beam constitutive model with respect to the classical beam strain measures is a well accepted condition [35] that must be satisfied by admissible strain energy functionals in the large strain regime. The present manuscript shows the relationship between a polyconvex constitutive model defined in terms of the deformation gradient, the cofactor and the determinant of the mapping in the continuum and its degenerate beam counterpart defined in terms of the beam strain measures. Hence, the connection between the two most accepted restrictions for the definition of constitutive models in three dimensional elasticity and beams is shown, bridg-

ing the gap between the continuum and its degenerate beam description.

The paper is organised as follows. Section 2 briefly revises the computational framework developed in [12] for large strain scenarios and especially tailored for polyconvex constitutive models. Section 3 presents the new degenerate beam formulation which extends the variational framework presented in [12] to the case of geometrically exact beam theory where the deformation gradient, its cofactor and its determinant are retained as the main strain variables. Section 4 presents the classical framework for the geometrically exact beam theory where the physically meaningful axial-shear and torsional–bending strain vectors are used as the main strain measures. Additionally, a link is established between the convexity criteria required for the constitutive model in both continuum and beam descriptions. Section 5 briefly presents the variational principle associated with the proposed continuum degenerate formulation. Section 6 presents the Finite Element discretisation, where the use of the novel tensor cross product algebra leads to alternative tangent operator representations. Section 7 includes representative numerical examples in large strain scenarios, including a comparison against a recently published mixed continuum based formulation [12]. Finally, Sect. 8 provides some concluding remarks and a summary of the key contributions of this paper.

2 Continuum mechanics

2.1 Continuum kinematics

Consider the three dimensional deformation of an elastic body from its initial configuration occupying a volume V , of boundary ∂V , into a final configuration at volume v , of boundary ∂v , where \mathbf{x} represents the current position of a particle originally at \mathbf{X} (see Fig. 1). Virtual and incremental variations of \mathbf{x} will be denoted by $\delta \mathbf{u}$ and \mathbf{u} , respectively. It will be assumed that \mathbf{x} , $\delta \mathbf{u}$ and \mathbf{u} satisfy appropriate essential (displacement) boundary conditions on $\partial_u V$. Additionally, the body is under the action of certain body forces per unit undeformed volume \mathbf{f}_0 and traction per unit undeformed area \mathbf{t}_0 on $\partial_t V$, where $\partial_t V \cup \partial_u V = \partial V$ and $\partial_t V \cap \partial_u V = \emptyset$. The deformation gradient tensor \mathbf{F} (or fibre map) and its determinant J (or volume map) are defined as [16]

$$\mathbf{F} = \frac{\partial \mathbf{x}}{\partial \mathbf{X}} = \nabla_0 \mathbf{x}; \quad J = \det \mathbf{F} = \frac{dv}{dV}, \quad (1)$$

where ∇_0 denotes the gradient with respect to material coordinates and dv and dV represent elemental volumes in the initial and final configurations, respectively. The elemental area vector is mapped from initial $d\mathbf{A}$ to final $d\mathbf{a}$ configurations by means of the cofactor tensor \mathbf{H} (or area map), which

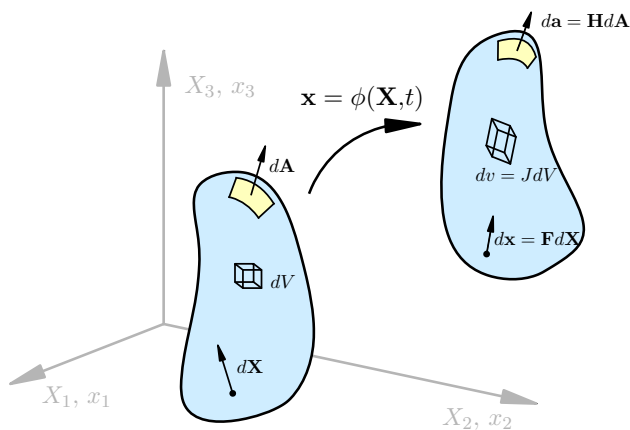


Fig. 1 Deformation mapping of a continuum and associated kinematic variables: F, H, J

is related to the deformation gradient F and its determinant J via the Nanson’s rule [16] as

$$da = HdA; \quad H = JF^{-T}. \tag{2}$$

In references [12, 32], the authors employ an alternative definition of the cofactor H , namely $H = \frac{1}{2}F \times F$, which simplifies considerably the algebra [33]. The new definition of the area map H is based on a tensor cross product \times introduced in [34] and applied within the context of nonlinear elasticity for the first time in [12]. The properties of this tensor cross product developed in [12] have been included in Appendix 1 for completeness.

Crucially, the first and second directional derivatives of H with respect to geometry changes are now easily evaluated as¹

$$\begin{aligned} DH[\delta u] &= F \times DF[\delta u] = F \times \nabla_0 \delta u; \\ D^2H[\delta u; u] &= DF[u] \times DF[\delta u] = \nabla_0 \delta u \times \nabla_0 u. \end{aligned} \tag{3}$$

Similarly, the directional derivatives of the volume map J are easily expressed with the help of (139) and (142) as

$$\begin{aligned} DJ[\delta u] &= H : \nabla_0 \delta u; \\ D^2J[\delta u; u] &= F : (\nabla_0 \delta u \times \nabla_0 u). \end{aligned} \tag{4}$$

2.2 Polyconvex elasticity

Polyconvexity is now well accepted as a fundamental mathematical requirement that must be satisfied by admissible strain energy functionals used to describe elastic materials in the large strain regime [18–22, 24, 25, 36, 37]. The strain

¹ The first and second directional derivatives of the deformation gradient tensor F are well known to be $DF[\delta u] = \nabla_0 \delta u$ and $D^2F[\delta u; u] = \mathbf{0}$ [16].

energy Ψ per unit undeformed volume must be a function of the deformation gradient $\nabla_0 x$ via a convex multi-valued function W as

$$\Psi(\nabla_0 x) = W(F, H, J). \tag{5}$$

Crucially, frame invariance (objectivity) implies that W must be independent of the rotational components of F and H , which is typically achieved by ensuring that W depends on F and H via the symmetric tensors $F^T F$ and $H^T H$, respectively. The three strain measures F, H , and J have work conjugate stresses Σ_F, Σ_H , and Σ_J defined by [12, 32]

$$\begin{aligned} \Sigma_F(F, H, J) &= \frac{\partial W}{\partial F}; \\ \Sigma_H(F, H, J) &= \frac{\partial W}{\partial H}; \\ \Sigma_J(F, H, J) &= \frac{\partial W}{\partial J}. \end{aligned} \tag{6}$$

It is then possible [12, 32, 33] to express the first Piola–Kirchhoff stress tensor P in terms of the extended strain measures $\{F, H, J\}$ and conjugate stresses $\{\Sigma_F, \Sigma_H, \Sigma_J\}$ as

$$P = \Sigma_F + \Sigma_H \times F + \Sigma_J H. \tag{7}$$

A tangent elasticity operator $D^2\Psi[\delta u; u]$ is usually needed [16] in order to ensure quadratic convergence of a Newton–Raphson type of solution process, derived in [12, 32, 33] as follows

$$\begin{aligned} D^2\Psi[\delta u; u] &= [(\nabla_0 \delta u) : (\nabla_0 \delta u \times F) : (\nabla_0 \delta u : H)] \\ &+ [\mathbb{H}_W] \begin{bmatrix} : (\nabla_0 u) \\ : (\nabla_0 u \times F) \\ (\nabla_0 u : H) \end{bmatrix} \\ &+ (\Sigma_H + \Sigma_J F) : (\nabla_0 \delta u \times \nabla_0 u), \end{aligned} \tag{8}$$

where the Hessian operator $[\mathbb{H}_W]$ denotes the symmetric positive semi-definite operator containing the second derivatives of $W(F, H, J)$ as

$$[\mathbb{H}_W] = \begin{bmatrix} W_{FF} & W_{FH} & W_{FJ} \\ W_{HF} & W_{HH} & W_{HJ} \\ W_{JF} & W_{JH} & W_{JJ} \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 W}{\partial F \partial F} & \frac{\partial^2 W}{\partial F \partial H} & \frac{\partial^2 W}{\partial F \partial J} \\ \frac{\partial^2 W}{\partial H \partial F} & \frac{\partial^2 W}{\partial H \partial H} & \frac{\partial^2 W}{\partial H \partial J} \\ \frac{\partial^2 W}{\partial J \partial F} & \frac{\partial^2 W}{\partial J \partial H} & \frac{\partial^2 W}{\partial J \partial J} \end{bmatrix}. \tag{9}$$

Note that the first term on the right hand side of (8) is necessarily positive for $\delta u = u$ and, therefore, buckling can only be induced by the second “initial stress” term

$(\Sigma_H + \Sigma_J F) : (\nabla_0 \delta u \times \nabla_0 u)$. Finally, it is also possible to re-write the tangent elasticity operator (8) as

$$D^2\Psi[\delta u; u] = \nabla_0 \delta u : [\mathbb{H}_W^{eq}] : \nabla_0 u + (\Sigma_H + \Sigma_J F) : (\nabla_0 \delta u \times \nabla_0 u), \tag{10}$$

where use of the tensor cross product operations (133) and (134) yields

$$\begin{aligned} [\mathbb{H}_W^{eq}] &= W_{FF} + F \times W_{HH} \times F \\ &+ W_{JJ} H \otimes H + W_{FH} \times F + F \times W_{HF} \\ &+ W_{FJ} \otimes H + H \otimes W_{JF} + (F \times W_{HJ}) \otimes H \\ &+ H \otimes (W_{JH} \times F). \end{aligned} \tag{11}$$

2.2.1 A polyconvex constitutive model: Mooney–Rivlin material

The well known Mooney–Rivlin constitutive model is polyconvex. An elegant representation of its strain energy functional in terms of the extended kinematic set $\{F, H, J\}$ [12,32,33] is

$$W_{MR}(F, H, J) = \alpha F : F + \beta H : H + f(J), \tag{12}$$

where²

$$f(J) = -2\alpha \ln J - 4\beta J + \frac{\lambda}{2} (J - 1)^2 - 3(\alpha + \beta), \tag{13}$$

with α and β suitable non-negative material parameters and $f(J)$ a convex function of J (refer to remark 1). The conjugate stresses $\{\Sigma_F, \Sigma_H, \Sigma_J\}$ (6) for this model are obtained as

$$\Sigma_F = 2\alpha F; \quad \Sigma_H = 2\beta H; \quad \Sigma_J = -\frac{2\alpha}{J} - 4\beta + \lambda (J - 1) \tag{14}$$

and the first Piola–Kirchhoff (7) results in

$$P = 2\alpha F + 2\beta H \times F - \left(\frac{2\alpha}{J} + 4\beta - \lambda (J - 1)\right) H. \tag{15}$$

The Hessian operator $[\mathbb{H}_W]$ (9) adopts the following simple expression

$$[\mathbb{H}_W] = \begin{bmatrix} 2\alpha \mathcal{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 2\beta \mathcal{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \frac{2\alpha}{J^2} + \lambda \end{bmatrix}, \tag{16}$$

² As a particular degenerate case, the Neo-Hookean constitutive model is obtained for $\beta = 0$.

where $\mathcal{I}_{i1j1} = \delta_{ij} \delta_{11}$ and $[\mathbb{H}_W^{eq}]$ (11) is expressed as

$$[\mathbb{H}_W^{eq}] = 2\alpha \mathcal{I} + 2\beta F \times \mathcal{I} \times F + \left(\frac{2\alpha}{J^2} + \lambda\right) H \otimes H. \tag{17}$$

The clear benefits of employing the new tensor cross product algebra in the context of polyconvex elasticity have been presented in [12,32] and thoroughly detailed in [33].

Remark 1 An alternative function $f(J)$ to that used in (13) is

$$f(J) = -(2\alpha + 4\beta) \ln J + \frac{\lambda}{2} (J - 1)^2. \tag{18}$$

Following [12], the material parameters α , β and λ used in (12) and (18) can be related to the classical Lamé parameters μ_{lin} and λ_{lin} in linearised elasticity (in the reference configuration) through

$$2\alpha + 2\beta = \mu_{lin}; \quad \lambda + 4\beta = \lambda_{lin}. \tag{19}$$

The Poisson ratio ν in the reference configuration is related to material parameters λ_{lin} and μ_{lin} as

$$\nu = \frac{\lambda_{lin}}{2(\lambda_{lin} + \mu_{lin})}. \tag{20}$$

Substitution of Eq. (19) into (20) leads to the following expression for the Poisson ratio in terms of the material parameters α , β and λ

$$\nu = \frac{\lambda + 4\beta}{2(\lambda + 6\beta + 2\alpha)}. \tag{21}$$

Classical beam theories do not consider dilatations or contractions of the beam section. Hence, a Poisson ratio of value $\nu = 0$ is consistent with such kinematical description of the beam. However, from Eq. (21), it can be inferred that it is not possible to find any combination of non-negative (and hence, satisfying polyconvexity) values for the material parameters α , β and λ compatible with $\nu = 0$. On the contrary, if the volumetric function $f(J)$ in (13) is used instead, the relation between the material parameters α , β and λ relate to μ_{lin} and λ_{lin} as

$$2\alpha + 2\beta = \mu_{lin}; \quad \lambda = \lambda_{lin}, \tag{22a}$$

which leads to an expression for the Poisson ratio ν_{lin}

$$\nu_{lin} = \frac{\lambda}{2(\lambda + 2\alpha + 2\beta)}. \tag{23}$$

The definition of $f(J)$ in (13) allows to model the particular case $v_{\text{lin}} = 0$ without violating polyconvexity (non-negative values for α and β with $\lambda = 0$).

2.2.2 A non-polyconvex constitutive model: Saint Venant–Kirchhoff material

A popular constitutive model for beams is that of a Saint Venant–Kirchhoff material [1,3,8,10,38], where the strain energy functional per unit undeformed volume Ψ is defined in terms of the Green–Lagrange strain tensor \mathbf{E} as

$$\begin{aligned} \Psi_{SVK}(\nabla_0 \mathbf{x}) &= \frac{\lambda}{2} (\text{tr} \mathbf{E})^2 + \mu \text{tr}(\mathbf{E} \mathbf{E}); \\ \mathbf{E} &= \frac{1}{2} (\mathbf{C} - \mathbf{I}); \quad \mathbf{C} = \mathbf{F}^T \mathbf{F}, \end{aligned} \tag{24}$$

where \mathbf{C} is the right Cauchy–Green deformation tensor, λ and μ the Lamé coefficients and \mathbf{I} the identity tensor. Alternatively, Ψ_{SVK} can be expressed in terms of \mathbf{C} as

$$\Psi_{SVK}(\mathbf{C}) = \frac{\mu}{4} (\text{tr}(\mathbf{C}\mathbf{C}) - 2\text{tr}(\mathbf{C}) + 3) + \frac{\lambda}{8} (\text{tr}(\mathbf{C}) - 3)^2. \tag{25}$$

As shown in [30], it is possible to re-express the term $\text{tr}(\mathbf{C}\mathbf{C})$ in (25) as

$$\text{tr}(\mathbf{C}\mathbf{C}) = (\mathbf{F} : \mathbf{F})^2 - 2(\mathbf{H} : \mathbf{H}). \tag{26}$$

Introduction of identity (26) into (25) enables to rewrite the strain energy functional in terms of \mathbf{F} and \mathbf{H} as

$$\begin{aligned} W_{SVK}(\mathbf{F}, \mathbf{H}) &= \frac{\mu}{4} \left((\mathbf{F} : \mathbf{F})^2 - 2(\mathbf{F} : \mathbf{F}) - 2(\mathbf{H} : \mathbf{H}) + 3 \right) \\ &\quad + \frac{\lambda}{8} ((\mathbf{F} : \mathbf{F}) - 3)^2. \end{aligned} \tag{27}$$

The conjugate stresses $\{\Sigma_F, \Sigma_H, \Sigma_J\}$ (6) for this model are obtained as

$$\begin{aligned} \Sigma_F &= \left(\mu + \frac{\lambda}{2} \right) (\mathbf{F} : \mathbf{F}) \mathbf{F} - \left(\frac{3\lambda}{2} + \mu \right) \mathbf{F}; \\ \Sigma_H &= -\mu \mathbf{H}; \quad \Sigma_J = 0 \end{aligned} \tag{28}$$

and the first Piola–Kirchhoff (7) results in

$$\mathbf{P} = \left(\mu + \frac{\lambda}{2} \right) (\mathbf{F} : \mathbf{F}) \mathbf{F} - \left(\frac{3\lambda}{2} + \mu \right) \mathbf{F} - \mu \mathbf{H} \times \mathbf{F}. \tag{29}$$

The Hessian operator $[\mathbb{H}_W]$ (9) adopts the following simple expression

$$\begin{aligned} [\mathbb{H}_W] &= \begin{bmatrix} 2\left(\mu + \frac{\lambda}{2}\right) \mathbf{F} \otimes \mathbf{F} - \left(\frac{\lambda}{2}(3 - \mathbf{F} : \mathbf{F}) + \mu(1 - \mathbf{F} : \mathbf{F})\right) \mathcal{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mu \mathcal{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}. \end{aligned} \tag{30}$$

Unfortunately this model is not polyconvex as can be observed from the lack of positive (semi)-definiteness of the Hessian operator in Eq. (30). Finally, the tensor $[\mathbb{H}_W^{eq}]$ (11) is expressed as

$$\begin{aligned} [\mathbb{H}_W^{eq}] &= 2 \left(\mu + \frac{\lambda}{2} \right) \mathbf{F} \otimes \mathbf{F} - \left(\frac{\lambda}{2} (3 - \mathbf{F} : \mathbf{F}) + \mu (1 - \mathbf{F} : \mathbf{F}) \right) \mathcal{I} \\ &\quad - \mu \mathbf{F} \times \mathcal{I} \times \mathbf{F}. \end{aligned} \tag{31}$$

3 Continuum degenerate polyconvex beam formulation

In this section, a continuum degenerate beam model is presented. Contrary to the work in [10,11], where the strain energy functional is presented in terms of the Green–Lagrange strain tensor, the formulation hereby presented exploits the extended set of kinematic strain measures $\{\mathbf{F}, \mathbf{H}, \mathbf{J}\}$ along with the computational framework outlined in [12,32,33].

In addition, a co-rotational formulation in terms of an extended alternative set of co-rotational kinematic entities is also presented. This co-rotational approach will prove to be a natural link between the continuum degenerate approach and a more classical beam description where well-known engineering strain measures feature as the main kinematic entities.

3.1 Beam kinematics

Let us now consider that the elastic body introduced in the previous section is a beam structural element. In particular, it is assumed that the beam in the reference configuration has a straight axis of length L which is completely characterised by the position of a centre line $\mathbf{X}_0(s)$, parametrised in terms of $s \in [0, L]$, and an orthonormal triad $\{\mathbf{D}_1, \mathbf{D}_2, \mathbf{D}_3\}$ where the vectors \mathbf{D}_1 and \mathbf{D}_2 lay parallel to the cross sectional area $A(s)$ of the beam (with boundary $\partial A(s)$) and \mathbf{D}_3 is aligned along the undeformed centre line, see Fig. 2. In the following, summation over Greek indexes α ranges from 1 to 2 whereas summation over Latin indexes i ranges from 1 to 3 and Einstein’s convention is assumed for repeated indices unless otherwise stated. The beam reference configuration is

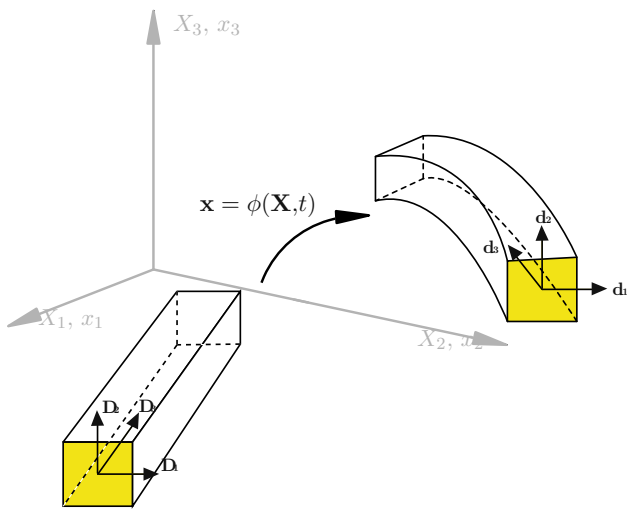


Fig. 2 Reference and current configuration of a beam

defined in terms of the convective coordinates $\{\eta^\alpha, \alpha = 1, 2\}$ [8] by

$$X(\eta^\alpha, s) = X_0(s) + \bar{X}; \quad \bar{X} = \eta^\alpha D_\alpha. \tag{32}$$

For the current configuration, an orthonormal triad $\{d_1(s), d_2(s), d_3(s)\}$ can be introduced with $d_1(s)$ and $d_2(s)$ laying on the cross sectional area $a(s)$ (with boundary $\partial a(s)$) and with $d_3(s) = d_1(s) \times d_2(s)$, see Fig. 2. Hence, the motion of the beam can be defined as [8,39]

$$x(\eta^\alpha, s) = x_0(s) + \bar{x}(s); \quad \bar{x}(s) = \eta^\alpha d_\alpha(s), \tag{33}$$

where the orthonormal triads defined above are related via a rotation tensor $R \in SO(3)$ defined as $R = d_i \otimes D_i$ [1].

It is known that the rotation tensor R can be defined in terms of the Rodrigues parametrisation of a rotation vector θ [40], with associated skew symmetric second order tensor $\hat{\theta} = I \times \theta$,³ as

$$R(\theta) = I + \frac{\sin \|\theta\|}{\|\theta\|} (I \times \theta) + \frac{1 - \cos \|\theta\|}{\|\theta\|^2} (I \times \theta)(I \times \theta). \tag{34}$$

The deformation gradient tensor F (1) derived from the mapping (33) is obtained as [8]

$$F = x'_0 \otimes D_3 + d_\alpha \otimes D_\alpha + \eta^\alpha d'_\alpha \otimes D_3, \tag{35}$$

where $(\bullet)' := \frac{d(\bullet)(s)}{ds}$. Above Eq. (35) in conjunction with (34) leads to a description of the deformation gradient tensor

F in terms of the centre line x_0 and the rotation vector θ , typically used in continuum degenerate descriptions [10,11] of beam structural models. Notice that a computational framework built upon the representation (35) of the deformation gradient tensor is not restricted to beams with a straight axis in the reference undeformed configuration.

It can be interesting to re-write F in a more meaningful manner from the physical standpoint. With that in mind, following the algebraic manipulations in [8], the deformation gradient tensor F (35) can be re-expressed as

$$F = R \left(\Gamma \otimes D_3 + \hat{K} \bar{X} \otimes D_3 + I \right), \tag{36}$$

where the beam strain measures Γ and \hat{K} are defined as

$$\Gamma = R^T x'_0 - D_3; \quad \hat{K} = R^T R', \tag{37}$$

with \hat{K} a skew-symmetric tensor that can be re-written in terms of a vector K as $\hat{K} = I \times K$.⁴ The beam strain measure Γ is known as the axial-shear strain vector whilst K is known as the torsional–bending strain vector. Thus, an alternative representation of (36) expressed in terms of the beam strain measures $\{\Gamma, K\}$ is [8]

$$F = R \bar{U}; \quad \bar{U} = B \otimes D_3 + I; \quad B = \Gamma + (K \times \bar{X}). \tag{38}$$

Above Eq. (38) is known as the right extended polar decomposition of the deformation gradient F in terms of the beam strain measures $\{\Gamma, K\}$ as presented in [8]. Notice that whilst R is a rotation tensor, the non-symmetric strain tensor \bar{U} is not a pure stretch tensor (as in the classical polar decomposition theorem [16]). Note that \bar{U} is only symmetric in the case that B is colinear with D_3 , which only happens in the absence of shear strain and torsion.

Formulae (38) for the representation of the deformation gradient tensor F are very useful in the context of geometrically exact beam models when the strain energy functional is defined in terms of the classical beam strain measures $\{\Gamma, K\}$ [1,3,41]. Notice that a computational framework built upon the representation (36) of the deformation gradient tensor is restricted to beams with a straight axis in the reference undeformed configuration. A more general expression for this representation of the deformation gradient tensor can be found in [1].

3.2 Linearisation of the beam kinematics

As stated above (33), the mapping of the beam is defined in terms of the centre line x_0 and the triad $\{d_i\}$ (or the rota-

³ Use of the tensor cross product \times between a second order tensor and a vector has been made.

⁴ Use of the tensor cross product \times between a second order tensor and a vector has been made.

tion vector θ). Virtual or incremental variations of the centre line x_0 will be denoted by δu_0 or u_0 , respectively. In addition, virtual or incremental variations of the rotation vector θ will be denoted by $\delta\theta$ or $\Delta\theta$, respectively. All of these fields must satisfy appropriate essential and natural boundary conditions at $s = 0$ and/or $s = L$ [42]. Thus, the virtual or incremental variations of the mapping (33) are then defined as

$$\begin{aligned} \delta u &= \delta u_0 + \eta^\alpha Dd_\alpha[\delta\theta]; \\ u &= u_0 + \eta^\alpha Dd_\alpha[\Delta\theta]. \end{aligned} \tag{39}$$

The first and second directional derivatives of the triad with respect to geometry changes are defined by⁵

$$\begin{aligned} Dd_i[\delta\theta] &= \delta\theta \times d_i; \\ D^2d_i[\delta\theta; \Delta\theta] &= \delta\theta \times (\Delta\theta \times d_i). \end{aligned} \tag{40}$$

Analogously, for the derivative of the triad with respect to the beam axis, the first and second directional derivatives are

$$\begin{aligned} Dd'_i[\delta\theta] &= \delta\theta' \times d_i + \delta\theta \times d'_i; \\ D^2d'_i[\delta\theta; \Delta\theta] &= \delta\theta' \times (\Delta\theta \times d_i) + \delta\theta \times (\Delta\theta' \times d_i) \\ &\quad + \delta\theta \times (\Delta\theta \times d'_i). \end{aligned} \tag{41}$$

The first and second directional derivatives of the deformation gradient F can be obtained as

$$\begin{aligned} DF[\delta u_0, \delta\theta] &= \delta u'_0 \otimes D_3 + Dd_\alpha[\delta\theta] \otimes D_\alpha \\ &\quad + \eta^\alpha Dd'_\alpha[\delta\theta] \otimes D_3; \\ D^2F[\delta u_0, \delta\theta; u_0, \Delta\theta] &= D^2d_\alpha[\delta\theta; \Delta\theta] \otimes D_\alpha \\ &\quad + \eta^\alpha D^2d'_\alpha[\delta\theta; \Delta\theta] \otimes D_3. \end{aligned} \tag{42}$$

where the directional derivatives of the triad were previously computed in (40)–(41). Notice how the kinematics of the beam introduces, in contrast to the continuum formulation, additional non-linearities into the problem through the non-vanishing second directional derivative of F (recall that this term vanishes in the more general continuum description). The first and second directional derivatives of the cofactor H are computed as

⁵ These directional derivatives need to comply with the orthogonality condition $d_i \cdot d_j - \delta_{ij} = 0$ [1].

$$\begin{aligned} DH[\delta u_0, \delta\theta] &= F \times DF[\delta u_0, \delta\theta]; \\ D^2H[\delta u_0, \delta\theta; u_0, \Delta\theta] &= DF[\delta u_0, \delta\theta] \times DF[u_0, \Delta\theta] \\ &\quad + F \times D^2F[\delta u_0, \delta\theta; u_0, \Delta\theta]. \end{aligned} \tag{43}$$

Finally, the first and second directional derivatives of the determinant J are computed as

$$\begin{aligned} DJ[\delta u_0, \delta\theta] &= H : DF[\delta u_0, \delta\theta]; \\ D^2J[\delta u_0, \delta\theta; u_0, \Delta\theta] &= F : (DF[\delta u_0, \delta\theta] \times DF[u_0, \Delta\theta]) \\ &\quad + H : D^2F[\delta u_0, \delta\theta; u_0, \Delta\theta]. \end{aligned} \tag{44}$$

Due to the kinematics introduced by the beam problem, further nonlinearities are observed in (43) and (44) with respect to (3) and (4), respectively. Alternatively, a formulation exclusively in terms of the centre line x_0 and the triad $\{d_i\}$ has been presented in references [1, 3]. This formulation avoids relating directional derivatives of the triad $\{d_i\}$ with respect to the rotation vector θ as in Eq. (40).

3.3 Transition from polyconvex continuum model to polyconvex beam theory

For the case of a polyconvex constitutive model satisfying Eq. (5), it is possible to express the internal virtual work as follows

$$\begin{aligned} P : DF[\delta u_0, \delta\theta] &= D\Psi[\delta u_0, \delta\theta] \\ &= DW[DF[\delta u_0, \delta\theta], DH[\delta u_0, \delta\theta], DJ[\delta u_0, \delta\theta]] \\ &= \Sigma_F : DF[\delta u_0, \delta\theta] + \Sigma_H : DH[\delta u_0, \delta\theta] \\ &\quad + \Sigma_J DJ[\delta u_0, \delta\theta] \\ &= \Sigma_F : DF[\delta u_0, \delta\theta] + \Sigma_H : (F \times DF[\delta u_0, \delta\theta]) \\ &\quad + \Sigma_J H : DF[\delta u_0, \delta\theta] \\ &= (\Sigma_F + \Sigma_H \times F + \Sigma_J H) : DF[\delta u_0, \delta\theta], \end{aligned} \tag{45}$$

which leads to an identical expression for the first Piola–Kirchhoff stress tensor P as in (7). The tangent elasticity operator is obtained as

$$\begin{aligned} D^2\Psi[\delta u_0, \delta\theta; u_0, \Delta\theta] &= DF[\delta u_0, \delta\theta] : C : DF[u_0, \Delta\theta] \\ &\quad + P : D^2F[\delta u_0, \delta\theta; u_0, \Delta\theta], \end{aligned} \tag{46}$$

where $C = \frac{\partial P}{\partial F} = \frac{\partial^2 \Psi}{\partial F \partial F}$. Use of Eq. (45) yields

$$\begin{aligned}
 & D^2\Psi[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}; \mathbf{u}_0, \Delta\boldsymbol{\theta}] \\
 &= [\mathbb{S}_\delta]_W^T [\mathbb{H}_W] [\mathbb{S}_\Delta]_W \\
 &+ (\boldsymbol{\Sigma}_H + \boldsymbol{\Sigma}_J \mathbf{F}) : (D\mathbf{F}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}] \times D\mathbf{F}[\mathbf{u}_0, \Delta\boldsymbol{\theta}]) \\
 &+ (\boldsymbol{\Sigma}_F + \boldsymbol{\Sigma}_H \times \mathbf{F} + \boldsymbol{\Sigma}_J \mathbf{H}) : D^2\mathbf{F}[\delta\boldsymbol{\theta}; \Delta\boldsymbol{\theta}],
 \end{aligned} \tag{47}$$

where

$$\begin{aligned}
 & [\mathbb{S}_\delta]_W^T \\
 &= [D\mathbf{F}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}] : (D\mathbf{F}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}] \times \mathbf{F}) : D\mathbf{F}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}] : \mathbf{H}]; \\
 & [\mathbb{S}_\Delta]_W = \begin{bmatrix} : D\mathbf{F}[\mathbf{u}_0, \Delta\boldsymbol{\theta}] \\ : (D\mathbf{F}[\mathbf{u}_0, \Delta\boldsymbol{\theta}] \times \mathbf{F}) \\ D\mathbf{F}[\mathbf{u}_0, \Delta\boldsymbol{\theta}] : \mathbf{H} \end{bmatrix}.
 \end{aligned} \tag{48}$$

Note that the first term on the right hand side of (47) is necessarily positive for $\delta\mathbf{u}_0 = \mathbf{u}_0$ and $\delta\boldsymbol{\theta} = \Delta\boldsymbol{\theta}$ and therefore buckling can only be induced by the “initial stress” term associated with the last two terms on the right hand side of (47). In comparison with Eq. (8), an extra nonlinearity present in above Eq. (47) can be observed, namely the last term on the right hand side.

Finally, proceeding as in Sect. 2.2, it is possible to re-express Eq. (47) in terms of the alternative Hessian operator $[\mathbb{H}_W^{eq}]$ defined in Eq. (11) as

$$\begin{aligned}
 & D^2\Psi[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}; \mathbf{u}_0, \Delta\boldsymbol{\theta}] \\
 &= D\mathbf{F}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}] : [\mathbb{H}_W^{eq}] : D\mathbf{F}[\mathbf{u}_0, \Delta\boldsymbol{\theta}] \\
 &+ (\boldsymbol{\Sigma}_H + \boldsymbol{\Sigma}_J \mathbf{F}) : (D\mathbf{F}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}] \times D\mathbf{F}[\mathbf{u}_0, \Delta\boldsymbol{\theta}]) \\
 &+ (\boldsymbol{\Sigma}_F + \boldsymbol{\Sigma}_H \times \mathbf{F} + \boldsymbol{\Sigma}_J \mathbf{H}) : D^2\mathbf{F}[\delta\boldsymbol{\theta}; \Delta\boldsymbol{\theta}].
 \end{aligned} \tag{49}$$

3.4 Polyconvex co-rotational beam formulation

In this Section, an alternative polyconvex co-rotational formulation is presented. This new approach will prove to be very useful establishing a link between the continuum degenerate description and a classical beam description in terms of the beam strain measures $\boldsymbol{\Gamma}$ and \mathbf{K} defined above. Taking into consideration the right extended decomposition theorem (38) and the necessary objectivity requirement, an equivalent energy functional to $W(\mathbf{F}, \mathbf{H}, J)$ can be defined in terms of the strain tensor $\bar{\mathbf{U}} = \mathbf{R}^T \mathbf{F}$ (38), its cofactor $\bar{\mathbf{W}} = \mathbf{R}^T \mathbf{H}$ and its determinant J ,⁶ namely $\hat{W}(\bar{\mathbf{U}}, \bar{\mathbf{W}}, J)$. Notice that both representations $W(\mathbf{F}, \mathbf{H}, J)$ and $\hat{W}(\bar{\mathbf{U}}, \bar{\mathbf{W}}, J)$ have identical expressions in their own arguments. Analogously to (6), a new set of stress variables $\{\boldsymbol{\Sigma}_{\bar{\mathbf{U}}}, \boldsymbol{\Sigma}_{\bar{\mathbf{W}}}, \Sigma_J\}$ conjugate to $\{\bar{\mathbf{U}}, \bar{\mathbf{W}}, J\}$ can be defined as

⁶ Notice that the determinants of \mathbf{F} and $\bar{\mathbf{U}}$ are equal, namely $J = \det \mathbf{F} = \det \bar{\mathbf{U}}$.

$$\begin{aligned}
 \boldsymbol{\Sigma}_{\bar{\mathbf{U}}}(\bar{\mathbf{U}}, \bar{\mathbf{W}}, J) &= \frac{\partial \hat{W}}{\partial \bar{\mathbf{U}}} = \mathbf{R}^T \frac{\partial W}{\partial \mathbf{F}}; \\
 \boldsymbol{\Sigma}_{\bar{\mathbf{W}}}(\bar{\mathbf{U}}, \bar{\mathbf{W}}, J) &= \frac{\partial \hat{W}}{\partial \bar{\mathbf{W}}} = \mathbf{R}^T \frac{\partial W}{\partial \mathbf{H}}; \\
 \Sigma_J(\bar{\mathbf{U}}, \bar{\mathbf{W}}, J) &= \frac{\partial \hat{W}}{\partial J} = \frac{\partial W}{\partial J}.
 \end{aligned} \tag{50}$$

Considering the right extended decomposition theorem (38), the internal virtual work can be written as

$$\begin{aligned}
 \mathbf{P} : D\mathbf{F}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}] &= \mathbf{P} : \mathbf{R} D\bar{\mathbf{U}}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}] \\
 &+ \mathbf{P} : D\mathbf{R}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}] \bar{\mathbf{U}}.
 \end{aligned} \tag{51}$$

Conservation of angular momentum in the continuum implies

$$\begin{aligned}
 \mathbf{P} : D\mathbf{R}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}] \bar{\mathbf{U}} &= \mathbf{F} \mathbf{S} : D\mathbf{R}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}] \bar{\mathbf{U}} \\
 &= \underbrace{\bar{\mathbf{U}}^T \mathbf{S} \bar{\mathbf{U}}}_{\text{Symmetric tensor}} : \underbrace{\mathbf{R}^T D\mathbf{R}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}]}_{\text{Skew symmetric tensor}} \\
 &= 0
 \end{aligned} \tag{52}$$

where \mathbf{S} is the second Piola–Kirchhoff stress tensor. Hence, an alternative co-rotational representation for the internal virtual work (51) is as follows

$$\mathbf{P} : D\mathbf{F}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}] = \bar{\mathbf{P}} : D\bar{\mathbf{U}}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}]; \quad \bar{\mathbf{P}} = \mathbf{R}^T \mathbf{P}, \tag{53}$$

which results in

$$\begin{aligned}
 & \bar{\mathbf{P}} : D\bar{\mathbf{U}}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}] \\
 &= D\Psi[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}] \\
 &= D\hat{W}[D\bar{\mathbf{U}}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}], D\bar{\mathbf{W}}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}], DJ[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}]] \\
 &= \boldsymbol{\Sigma}_{\bar{\mathbf{U}}} : D\bar{\mathbf{U}}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}] + \boldsymbol{\Sigma}_{\bar{\mathbf{W}}} : D\bar{\mathbf{W}}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}] \\
 &+ \Sigma_J DJ[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}] \\
 &= \boldsymbol{\Sigma}_{\bar{\mathbf{U}}} : D\bar{\mathbf{U}}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}] \\
 &+ \boldsymbol{\Sigma}_{\bar{\mathbf{W}}} : (\bar{\mathbf{U}} \times D\bar{\mathbf{U}}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}]) + \Sigma_J \bar{\mathbf{W}} : D\bar{\mathbf{U}}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}] \\
 &= (\boldsymbol{\Sigma}_{\bar{\mathbf{U}}} + \boldsymbol{\Sigma}_{\bar{\mathbf{W}}} \times \bar{\mathbf{U}} + \Sigma_J \bar{\mathbf{W}}) : D\bar{\mathbf{U}}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}],
 \end{aligned} \tag{54}$$

which leads to the following expression for the co-rotational stress tensor $\bar{\mathbf{P}}$

$$\bar{\mathbf{P}} = \boldsymbol{\Sigma}_{\bar{\mathbf{U}}} + \boldsymbol{\Sigma}_{\bar{\mathbf{W}}} \times \bar{\mathbf{U}} + \Sigma_J \bar{\mathbf{W}}. \tag{55}$$

Following a similar procedure to that of Eq. (47), an alternative expression for the tangent operator of the energy Ψ can be obtained in terms of $\{\bar{\mathbf{U}}, \bar{\mathbf{W}}, J\}$ as follows

$$\begin{aligned}
 & D^2\Psi[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}; \mathbf{u}_0, \Delta\boldsymbol{\theta}] \\
 &= [\mathbb{S}_\delta]_{\hat{W}}^T [\mathbb{H}_{\hat{W}}] [\mathbb{S}_\Delta]_{\hat{W}} \\
 &+ (\boldsymbol{\Sigma}_{\bar{W}} + \boldsymbol{\Sigma}_J \bar{U}) : (D\bar{U}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}] \times D\bar{U}[\mathbf{u}_0, \Delta\boldsymbol{\theta}]) \\
 &+ (\boldsymbol{\Sigma}_{\bar{U}} + \boldsymbol{\Sigma}_{\bar{W}} \times \bar{U} + \boldsymbol{\Sigma}_J \bar{W}) : D^2\bar{U}[\delta\boldsymbol{\theta}; \Delta\boldsymbol{\theta}],
 \end{aligned} \tag{56}$$

where

$$\begin{aligned}
 & [\mathbb{S}_\delta]_{\hat{W}}^T \\
 &= [D\bar{U}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}] : (D\bar{U}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}] \times \bar{U}) : D\bar{U}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}] : \bar{W}]; \\
 & [\mathbb{S}_\Delta]_{\hat{W}} = \begin{bmatrix} : D\bar{U}[\mathbf{u}_0, \Delta\boldsymbol{\theta}] \\ : (D\bar{U}[\mathbf{u}_0, \Delta\boldsymbol{\theta}] \times \bar{U}) \\ D\bar{U}[\mathbf{u}_0, \Delta\boldsymbol{\theta}] : \bar{W} \end{bmatrix}
 \end{aligned} \tag{57}$$

and with the Hessian operator $[\mathbb{H}_{\hat{W}}]$ denoting the symmetric positive semi-definite operator containing the second derivatives of $\hat{W}(\bar{U}, \bar{W}, J)$ as

$$[\mathbb{H}_{\hat{W}}] = \begin{bmatrix} \hat{W}_{\bar{U}\bar{U}} & \hat{W}_{\bar{U}\bar{W}} & \hat{W}_{\bar{U}J} \\ \hat{W}_{\bar{W}\bar{U}} & \hat{W}_{\bar{W}\bar{W}} & \hat{W}_{\bar{W}J} \\ \hat{W}_{J\bar{U}} & \hat{W}_{J\bar{W}} & \hat{W}_{JJ} \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 \hat{W}}{\partial \bar{U} \partial \bar{U}} & \frac{\partial^2 \hat{W}}{\partial \bar{U} \partial \bar{W}} & \frac{\partial^2 \hat{W}}{\partial \bar{U} \partial J} \\ \frac{\partial^2 \hat{W}}{\partial \bar{W} \partial \bar{U}} & \frac{\partial^2 \hat{W}}{\partial \bar{W} \partial \bar{W}} & \frac{\partial^2 \hat{W}}{\partial \bar{W} \partial J} \\ \frac{\partial^2 \hat{W}}{\partial J \partial \bar{U}} & \frac{\partial^2 \hat{W}}{\partial J \partial \bar{W}} & \frac{\partial^2 \hat{W}}{\partial J \partial J} \end{bmatrix}. \tag{58}$$

As mentioned above, the identical representation of the energy functionals $W(\mathbf{F}, \mathbf{H}, J)$ and $\hat{W}(\bar{U}, \bar{W}, J)$ with respect to their respective arguments results in identical expressions of the Hessian operators $[\mathbb{H}_W]$ and $[\mathbb{H}_{\hat{W}}]$ in their respective arguments. Therefore, both Hessian operators $[\mathbb{H}_W]$ and $[\mathbb{H}_{\hat{W}}]$ will be positive semi-definite, provided that $W(\mathbf{F}, \mathbf{H}, J)$ and hence, $\hat{W}(\bar{U}, \bar{W}, J)$ satisfy polyconvexity. The latter representation of the strain energy functional will be further pursued in the next Section.

4 Relationship with the classical beam formulation

Alternatively to continuum degenerate based formulations, as that presented in Sect. 3, classical beam formulations employ a kinematical description based on the strain measures $\{\boldsymbol{\Gamma}, \mathbf{K}\}$ by means of the definition of the deformation gradient tensor presented in Eq. (38).

The objective of this Section is twofold. On the one hand, the strain energy per unit undeformed volume (i.e. $W(\mathbf{F}, \mathbf{H}, J) = \hat{W}(\bar{U}, \bar{W}, J)$) will be re-written in terms of the alternative beam strain vector \mathbf{B} defined in (38). Further manipulations will lead to the introduction of the strain energy per unit undeformed length in terms of the beam strain measures $\boldsymbol{\Gamma}$ and \mathbf{K} . On the second hand, relationships will be established between the tangent operators of the different

energy representations to relate the concept of polyconvexity at a continuum and beam level.

4.1 Beam kinematics and linearisation

4.1.1 The classical beam strain measures

Manipulation of equations in (37) enable the strain measures $\{\boldsymbol{\Gamma}, \mathbf{K}\}$ to be expressed in terms of the centre line and the triads [1] as

$$\boldsymbol{\Gamma} = (\mathbf{d}_i \cdot \mathbf{x}'_0) \mathbf{D}_i - \mathbf{D}_3; \quad \mathbf{K} = \frac{1}{2}(\mathbf{d}_i \cdot \mathbf{d}'_j)(\mathbf{D}_j \times \mathbf{D}_i). \tag{59}$$

From equations in (59), the first and second directional derivatives of the beam strain measures $\{\boldsymbol{\Gamma}, \mathbf{K}\}$ follow as

$$\begin{aligned}
 D\boldsymbol{\Gamma}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}] &= (D\mathbf{d}_i[\delta\boldsymbol{\theta}] \cdot \mathbf{x}'_0 + \mathbf{d}_i \cdot \delta\mathbf{u}'_0) \mathbf{D}_i; \\
 D^2\boldsymbol{\Gamma}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}; \mathbf{u}_0, \boldsymbol{\theta}] &= (D\mathbf{d}_i[\delta\boldsymbol{\theta}] \cdot \mathbf{u}'_0 + D\mathbf{d}_i[\Delta\boldsymbol{\theta}] \cdot \delta\mathbf{u}'_0 \\
 &+ D^2\mathbf{d}_i[\delta\boldsymbol{\theta}, \Delta\boldsymbol{\theta}] \cdot \mathbf{x}'_0) \mathbf{D}_i,
 \end{aligned} \tag{60}$$

and

$$\begin{aligned}
 D\mathbf{K}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}] &= \frac{1}{2} \left(D\mathbf{d}_i[\delta\boldsymbol{\theta}] \cdot \mathbf{d}'_j + \mathbf{d}_i \cdot D\mathbf{d}'_j[\delta\boldsymbol{\theta}] \right) (\mathbf{D}_j \times \mathbf{D}_i); \\
 D^2\mathbf{K}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}; \mathbf{u}_0, \Delta\boldsymbol{\theta}] &= \frac{1}{2} \left(D\mathbf{d}_i[\delta\boldsymbol{\theta}] \cdot D\mathbf{d}'_j[\Delta\boldsymbol{\theta}] + D\mathbf{d}_i[\Delta\boldsymbol{\theta}] \cdot D\mathbf{d}'_j[\delta\boldsymbol{\theta}] \right. \\
 &+ D^2\mathbf{d}_i[\delta\boldsymbol{\theta}; \Delta\boldsymbol{\theta}] \cdot \mathbf{d}'_j + \mathbf{d}_i \cdot D^2\mathbf{d}'_j[\delta\boldsymbol{\theta}; \Delta\boldsymbol{\theta}]) \\
 &(\mathbf{D}_j \times \mathbf{D}_i).
 \end{aligned} \tag{61}$$

4.1.2 The strain vector \mathbf{B}

The directional derivative of the strain vector \mathbf{B} defined in (38) can be written as

$$\begin{aligned}
 D\mathbf{B}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}] &= D\boldsymbol{\Gamma}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}] + D\mathbf{K}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}] \times \bar{\mathbf{X}} \\
 &= [\mathbf{I} - \mathbf{I} \times \bar{\mathbf{X}}] \begin{bmatrix} D\boldsymbol{\Gamma}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}] \\ D\mathbf{K}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}] \end{bmatrix}; \\
 D^2\mathbf{B}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}; \mathbf{u}_0, \Delta\boldsymbol{\theta}] &= D^2\boldsymbol{\Gamma}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}; \mathbf{u}_0, \Delta\boldsymbol{\theta}] \\
 &+ D^2\mathbf{K}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}; \mathbf{u}_0, \Delta\boldsymbol{\theta}] \times \bar{\mathbf{X}} \\
 &= [\mathbf{I} - \mathbf{I} \times \bar{\mathbf{X}}] \begin{bmatrix} D^2\boldsymbol{\Gamma}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}; \mathbf{u}_0, \Delta\boldsymbol{\theta}] \\ D^2\mathbf{K}[\delta\mathbf{u}_0, \delta\boldsymbol{\theta}; \mathbf{u}_0, \Delta\boldsymbol{\theta}] \end{bmatrix}
 \end{aligned} \tag{62}$$

where use of the tensor cross product formula (130) has been made above.

4.1.3 Co-rotational strain measures

It is possible to obtain explicit expressions of the strain tensor \bar{U} , its cofactor \bar{W} and its determinant J in terms of the strain vector B (linearly related to Γ and K). Recalling equation (38) which gives an explicit representation of the strain tensor $\bar{U} = B \otimes D_3 + I$ in terms of the strain vector B , the cofactor \bar{W} can be evaluated making use of property (143) as

$$\begin{aligned} \bar{W} &= \frac{1}{2} \bar{U} \times \bar{U} \\ &= \frac{1}{2} (B \otimes D_3 + I) \times (B \otimes D_3 + I) \\ &= \frac{1}{2} (B \otimes D_3) \times (B \otimes D_3) + (B \otimes D_3) \times I \\ &\quad + \frac{1}{2} I \times I \\ &= \frac{1}{2} \underbrace{(B \times B)}_0 \otimes \underbrace{(D_3 \times D_3)}_0 + (B \otimes D_3) \times I + I \\ &= (B \otimes D_3) \times I + I, \end{aligned} \tag{63}$$

where use of the properties (141) and (144) has been made. Similarly, the determinant J can be computed by means of the property (142) and using the above result (63)

$$\begin{aligned} J &= \frac{1}{3} \bar{W} : \bar{U} \\ &= \frac{1}{3} [(B \otimes D_3) \times I + I] : [B \otimes D_3 + I] \\ &= \frac{1}{3} [(B \otimes D_3) \times I] : (B \otimes D_3) \\ &\quad + [(B \otimes D_3) \times I] : I + B \cdot D_3 + 3 \\ &= \frac{1}{3} \underbrace{(B \times B)}_0 \cdot \underbrace{(D_3 \times D_3)}_0 + \underbrace{[(B \otimes D_3) \times I] : I}_{2B \cdot D_3} \\ &\quad + B \cdot D_3 + 3 = B \cdot D_3 + 1, \end{aligned} \tag{64}$$

where use of properties (139), (141) and (144) has been made. The directional derivatives of $\{\bar{U}, \bar{W}, J\}$ are obtained from above equations as

$$D\bar{U}[\delta u_0, \delta \theta] = DB[\delta u_0, \delta \theta] \otimes D_3; \tag{65}$$

$$D^2\bar{U}[\delta u_0, \delta \theta; u_0, \Delta \theta] = D^2B[\delta u_0, \delta \theta; u_0, \Delta \theta] \otimes D_3,$$

$$\begin{aligned} D\bar{W}[\delta u_0, \delta \theta] &= (DB[\delta u_0, \delta \theta] \otimes D_3) \times I; \\ D^2\bar{W}[\delta u_0, \delta \theta; u_0, \Delta \theta] & \\ &= (D^2B[\delta u_0, \delta \theta; u_0, \Delta \theta] \otimes D_3) \times I, \end{aligned} \tag{66}$$

and

$$\begin{aligned} DJ[\delta u_0, \delta \theta] &= DB[\delta u_0, \delta \theta] \cdot D_3; \\ D^2J[\delta u_0, \delta \theta; u_0, \Delta \theta] &= D^2B[\delta u_0, \delta \theta; u_0, \Delta \theta] \cdot D_3. \end{aligned} \tag{67}$$

4.2 Polyconvex beam theory

It is customary in classical beam formulations to define a strain energy functional per unit undeformed length $W_b(\Gamma, K)$ dependent on the classical beam strain measures Γ and K . However, it is insightful to obtain the relationship between the strain energy functional per unit undeformed length $W_b(\Gamma, K)$ and the strain energy functional per unit undeformed volume $W(F, H, J)$ (5) (or its alternative equivalent representation $\hat{W}(\bar{U}, \bar{W}, J)$) defined previously.

Moreover, by use of Eq. (38), where the strain tensor \bar{U} is expressed in terms of the strain vector B , in conjunction with the definitions of the cofactor (143) and the determinant of a tensor (142), it is feasible to obtain an explicit expression of the strain energy functional in terms of the strain vector B (linearly related to Γ and K), namely $\bar{W}(B(\Gamma, K))$. The relationship between the alternative strain energy functionals can be summarised as

$$\begin{aligned} W_b(\Gamma, K) &= \int_{A(s)} \bar{W}(B(\Gamma, K)) dA; \\ \bar{W}(B(\Gamma, K)) &= \hat{W}(\bar{U}, \bar{W}, J) = W(F, H, J). \end{aligned} \tag{68}$$

4.2.1 Convexity in terms of the strain vector B

In the context of beam mechanics, an appropriate restriction for a constitutive model is convexity with respect to the beam strain vector B . Essentially, the strain energy Ψ per unit undeformed volume of the beam must be a function of the deformation gradient via a convex function \bar{W} as

$$\Psi(\nabla_0 x) = \bar{W}(B), \tag{69}$$

where \bar{W} is convex with respect to its three variables, namely the 3×1 components of B . Following a similar approach to that in Sect. 2.2, it is possible to define a new work conjugate stress vector Σ_B to the beam strain vector B as

$$\Sigma_B(B) = \frac{\partial \bar{W}}{\partial B}. \tag{70}$$

This set of conjugate measures enables the directional derivative of the strain energy \bar{W} to be expressed as

$$D\bar{W}[DB[\delta u_0, \delta \theta]] = \Sigma_B \cdot DB[\delta u_0, \delta \theta]. \tag{71}$$

Recalling that the directional derivative of the strain energy Ψ fulfils

$$D\Psi[\delta u_0, \delta \theta] = \bar{P} : D\bar{U}[\delta u_0, \delta \theta], \tag{72}$$

enables, after comparison of Eqs. (72) and (71) (making use of Eq. (65)), to conclude

$$\Sigma_B = \bar{P} D_3. \tag{73}$$

It transpires from Eq. (73) and the definition of \bar{P} in (53) that the stress vector Σ_B represents the traction vector on the cross section of the beam rotated back to the reference undeformed configuration.

Alternatively to Eqs. (47) and (56), the tangent operator for the strain energy Ψ can be re-expressed in terms of the directional derivatives of the strain vector B as

$$\begin{aligned} D^2\Psi[\delta u_0, \delta\theta; u_0, \Delta\theta] &= D(\Sigma_B \cdot DB[\delta u_0, \delta\theta])[u_0, \Delta\theta] \\ &= DB[\delta u_0, \delta\theta] \cdot [\mathbb{H}_{\bar{W}}] DB[u_0, \Delta\theta] \\ &\quad + \Sigma_B \cdot D^2B[\delta u_0, \delta\theta; u_0, \Delta\theta], \end{aligned} \tag{74}$$

where the Hessian operator $[\mathbb{H}_{\bar{W}}]$ denotes the symmetric positive semi-definite operator containing the second derivatives of $\bar{W}(B)$

$$[\mathbb{H}_{\bar{W}}] = \frac{\partial^2 \bar{W}}{\partial B \partial B}. \tag{75}$$

Note that the first term on the last right hand side of Eq. (74) is necessarily positive for $\delta u_0 = u_0$ and $\delta\theta = \Delta\theta$. Therefore, buckling can only be induced by the ‘‘initial stress’’ term in (74) defined by $\Sigma_B \cdot D^2B[\delta u_0, \delta\theta; u_0, \Delta\theta]$, where the second directional derivatives of the beam strain vector B were previously evaluated in (62).

4.2.2 Polyconvexity in terms of the beam strain measures Γ and K

Equivalently, an appropriate restriction for a constitutive model is polyconvexity with respect to the beam strain measures $\{\Gamma, K\}$ [35]. It is well known that this condition ensures ellipticity of the static problem and, equivalently, hyperbolicity of its dynamic counterpart [35]. Essentially, the strain energy Ψ_b ⁷ per unit undeformed length of the beam must be a function of the deformation gradient via a convex multi-valued function W_b as

$$\Psi_b(\nabla_0 x) = W_b(\Gamma, K), \tag{76}$$

where W_b is convex with respect to its 6 variables, namely, the 3×1 components of Γ and K . Moreover, invariance with respect to rotations is automatically satisfied since Γ

and K are defined in the reference configuration. Following a similar approach to that in Sect. 2.2, it is possible to define work conjugates Σ_Γ and Σ_K to the beam strain measures Γ and K , respectively, as

$$\Sigma_\Gamma(\Gamma, K) = \frac{\partial W_b}{\partial \Gamma}; \quad \Sigma_K(\Gamma, K) = \frac{\partial W_b}{\partial K}. \tag{77}$$

It is easy to realise that Σ_Γ represents the axial-shear force vector rotated back to the reference undeformed configuration, whilst Σ_K is the torsional–bending moment rotated back to the reference undeformed configuration, work conjugates of the axial-shear strain vector Γ and the torsional–bending strain vector K , respectively. This set of conjugate measures enables the directional derivative of the strain energy W_b to be expressed as

$$\begin{aligned} DW_b[D\Gamma[\delta u_0, \delta\theta], DK[\delta u_0, \delta\theta]] &= \Sigma_\Gamma \cdot D\Gamma[\delta u_0, \delta\theta] \\ &\quad + \Sigma_K \cdot DK[\delta u_0, \delta\theta]. \end{aligned} \tag{78}$$

Recalling the relation between the strain energies W_b and \bar{W} in (68) and hence, between Ψ_b and Ψ and Eq. (71), it is possible to obtain the following expression for the directional derivative of Ψ_b

$$\begin{aligned} D\Psi_b[\delta u_0, \delta\theta] &= \int_{A(s)} D\Psi[\delta u_0, \delta\theta] dA \\ &= \int_{A(s)} \Sigma_B \cdot DB[\delta u_0, \delta\theta] dA. \end{aligned} \tag{79}$$

The directional derivative of B can be expressed in terms of the directional derivatives of $\{\Gamma, K\}$ through Eq. (62), leading to the following modified version of Eq. (79)

$$\begin{aligned} D\Psi_b[\delta u_0, \delta\theta] &= \int_{A(s)} \Sigma_B dA \cdot D\Gamma[\delta u_0, \delta\theta] \\ &\quad - \int_{A(s)} \Sigma_B dA \cdot (I \times \bar{X}) DK[\delta u_0, \delta\theta] \\ &= \int_{A(s)} \Sigma_B dA \cdot D\Gamma[\delta u_0, \delta\theta] \\ &\quad + \int_{A(s)} (\bar{X} \times \Sigma_B) dA \cdot DK[\delta u_0, \delta\theta] \end{aligned} \tag{80}$$

where use of the tensor cross product formula (130) has been made. Comparison of Eq. (80) with (78) leads to the relation between the conjugate beam measures Σ_Γ and Σ_K (77) with Σ_B (70) as

$$\Sigma_\Gamma = \int_{A(s)} \Sigma_B dA; \quad \Sigma_K = \int_{A(s)} \bar{X} \times \Sigma_B dA. \tag{81}$$

⁷ It is possible to relate the strain energy per unit undeformed length Ψ_b with the strain energy per unit of undeformed volume Ψ as $\Psi_b := \int_{A(s)} \Psi dA$, where $A(s)$ is a generic beam cross sectional area.

Note that the tangent operator of the strain energy W_b naturally emerges as

$$\begin{aligned}
 & D^2 W_b[\delta \mathbf{u}_0, \delta \boldsymbol{\theta}; \mathbf{u}_0, \Delta \boldsymbol{\theta}] \\
 &= \int_{A(s)} D^2 \Psi[\delta \mathbf{u}_0, \delta \boldsymbol{\theta}; \mathbf{u}_0, \Delta \boldsymbol{\theta}] dA \\
 &= [\mathbb{S}_\delta]_{W_b}^T [\mathbb{H}_{W_b}] [\mathbb{S}_\Delta]_{W_b} \\
 &\quad + \boldsymbol{\Sigma}_\Gamma \cdot D^2 \boldsymbol{\Gamma}[\delta \mathbf{u}_0, \delta \boldsymbol{\theta}; \mathbf{u}_0, \Delta \boldsymbol{\theta}] \\
 &\quad + \boldsymbol{\Sigma}_K \cdot D^2 \mathbf{K}[\delta \mathbf{u}_0, \delta \boldsymbol{\theta}; \mathbf{u}_0, \Delta \boldsymbol{\theta}],
 \end{aligned} \tag{82}$$

with

$$\begin{aligned}
 [\mathbb{S}_\delta]_{W_b}^T &= [D\boldsymbol{\Gamma}[\delta \mathbf{u}_0, \delta \boldsymbol{\theta}] \cdot D\mathbf{K}[\delta \mathbf{u}_0, \delta \boldsymbol{\theta}]]; \\
 [\mathbb{S}_\Delta]_{W_b} &= \begin{bmatrix} D\boldsymbol{\Gamma}[\mathbf{u}_0, \Delta \boldsymbol{\theta}] \\ D\mathbf{K}[\mathbf{u}_0, \Delta \boldsymbol{\theta}] \end{bmatrix}.
 \end{aligned} \tag{83}$$

and the Hessian operator $[\mathbb{H}_{W_b}]$ is

$$[\mathbb{H}_{W_b}] = \begin{bmatrix} \frac{\partial^2 W_b}{\partial \boldsymbol{\Gamma} \partial \boldsymbol{\Gamma}} & \frac{\partial^2 W_b}{\partial \boldsymbol{\Gamma} \partial \mathbf{K}} \\ \frac{\partial^2 W_b}{\partial \mathbf{K} \partial \boldsymbol{\Gamma}} & \frac{\partial^2 W_b}{\partial \mathbf{K} \partial \mathbf{K}} \end{bmatrix} \tag{84}$$

4.3 Relationship between tangent elasticity operators

Use of Eqs. (62) and (81) in the second term on the right hand side of Eq. (82) enables both geometric terms of the tangent operators for $W_b(\boldsymbol{\Gamma}, \mathbf{K})$ and $\bar{W}(\mathbf{B})$ to be related as

$$\begin{aligned}
 & \boldsymbol{\Sigma}_\Gamma \cdot D^2 \boldsymbol{\Gamma}[\delta \mathbf{u}_0, \delta \boldsymbol{\theta}; \mathbf{u}_0, \Delta \boldsymbol{\theta}] + \boldsymbol{\Sigma}_K \cdot D^2 \mathbf{K}[\delta \mathbf{u}_0, \delta \boldsymbol{\theta}; \mathbf{u}_0, \Delta \boldsymbol{\theta}] \\
 &= \int_{A(s)} \boldsymbol{\Sigma}_B \cdot D^2 \mathbf{B}[\delta \mathbf{u}_0, \delta \boldsymbol{\theta}; \mathbf{u}_0, \Delta \boldsymbol{\theta}] dA
 \end{aligned} \tag{85}$$

Hence, the constitutive term for the tangent operators for both $W_b(\boldsymbol{\Gamma}, \mathbf{K})$ and $\bar{W}(\mathbf{B})$ must coincide, namely

$$[\mathbb{S}_\delta]_{W_b}^T [\mathbb{H}_{W_b}] [\mathbb{S}_\Delta]_{W_b} = D\mathbf{B}[\delta \mathbf{u}_0, \delta \boldsymbol{\theta}] \cdot [\mathbb{H}_{\bar{W}}] D\mathbf{B}[\mathbf{u}_0, \Delta \boldsymbol{\theta}]. \tag{86}$$

Use of Eqs. (83) and (62) into (86) finally yields

$$\begin{aligned}
 [\mathbb{H}_{W_b}] &= \begin{bmatrix} \frac{\partial^2 W_b}{\partial \boldsymbol{\Gamma} \partial \boldsymbol{\Gamma}} & \frac{\partial^2 W_b}{\partial \boldsymbol{\Gamma} \partial \mathbf{K}} \\ \frac{\partial^2 W_b}{\partial \mathbf{K} \partial \boldsymbol{\Gamma}} & \frac{\partial^2 W_b}{\partial \mathbf{K} \partial \mathbf{K}} \end{bmatrix} \\
 &= \int_{A(s)} \begin{bmatrix} \mathbf{I} \\ \mathbf{I} \times \bar{\mathbf{X}} \end{bmatrix} [\mathbb{H}_{\bar{W}}] [\mathbf{I} \quad -\mathbf{I} \times \bar{\mathbf{X}}] dA.
 \end{aligned} \tag{87}$$

Equation (87) establishes a relationship between the tangent operators $[\mathbb{H}_{W_b}]$ and $[\mathbb{H}_{\bar{W}}]$ for the internal energies $W_b(\boldsymbol{\Gamma}, \mathbf{K})$ and $\bar{W}(\mathbf{B})$, respectively. Notice that these two energies are defined in terms of the classical beam strain measures. In addition, it is also possible to relate these two Hessian operators with that of the energy $\hat{W}(\bar{\mathbf{U}}, \bar{\mathbf{W}}, J)$ (56)

defined in terms of the more general continuum strain measures. This can be obtained by re-expressing the tangent operator for the energy $\hat{W}(\bar{\mathbf{U}}, \bar{\mathbf{W}}, J)$ (56) in terms of the strain vector \mathbf{B} .

Notice first that consideration of Eq. (65) enables the simplification of the second term in the right hand side of Eq. (56) as

$$\begin{aligned}
 & (\boldsymbol{\Sigma}_{\bar{W}} + \boldsymbol{\Sigma}_J \bar{\mathbf{U}}) : (D\bar{\mathbf{U}}[\delta \mathbf{u}_0, \delta \boldsymbol{\theta}] \times D\bar{\mathbf{U}}[\mathbf{u}_0, \Delta \boldsymbol{\theta}]) \\
 &= (D\mathbf{B}[\delta \mathbf{u}_0, \delta \boldsymbol{\theta}] \times D\mathbf{B}[\mathbf{u}_0, \Delta \boldsymbol{\theta}]) \\
 &\quad \cdot (\boldsymbol{\Sigma}_{\bar{W}} + \boldsymbol{\Sigma}_J \bar{\mathbf{U}}) (D_3 \times D_3) = 0.
 \end{aligned} \tag{88}$$

Regarding the third term on the right-hand side in Eq. (56), use of Eqs. (55), (73) and (65) leads to the following identity

$$\begin{aligned}
 & (\boldsymbol{\Sigma}_{\bar{U}} + \boldsymbol{\Sigma}_{\bar{W}} \times \bar{\mathbf{U}} + \boldsymbol{\Sigma}_J \bar{\mathbf{W}}) : D^2 \bar{\mathbf{U}}[\delta \boldsymbol{\theta}; \Delta \boldsymbol{\theta}] \\
 &= \boldsymbol{\Sigma}_B \cdot D^2 \bar{\mathbf{B}}[\delta \boldsymbol{\theta}; \Delta \boldsymbol{\theta}].
 \end{aligned} \tag{89}$$

Finally, introduction of the expressions in (65), (66) and (67) for the directional derivatives of $\bar{\mathbf{U}}$, $\bar{\mathbf{W}}$ and J , respectively, together with Eqs. (88) and (89) enables both constitutive terms in Eqs. (56) and (74) to be related as

$$D\mathbf{B}[\delta \mathbf{u}_0, \delta \boldsymbol{\theta}] \cdot [\mathbb{H}_{\bar{W}}] D\mathbf{B}[\mathbf{u}_0, \Delta \boldsymbol{\theta}] = [\mathbb{S}_\delta]_{\bar{W}}^T [\mathbb{H}_{\bar{W}}] [\mathbb{S}_\Delta]_{\bar{W}}, \tag{90}$$

with the Hessian operator $[\mathbb{H}_{\bar{W}}]$ (75) related to the Hessian operator $[\mathbb{H}_{\hat{W}}]$ via

$$\begin{aligned}
 [\mathbb{H}_{\bar{W}}] &= \left[\mathbb{H}_{\hat{W}}^{eq} \right]_{D_3 D_3}; \\
 \left[\mathbb{H}_{\hat{W}}^{eq} \right]_{D_3 D_3} &= D_{3K} D_{3L} \left[\mathbb{H}_{\hat{W}}^{eq} \right]_{IKJL}
 \end{aligned} \tag{91}$$

where

$$\begin{aligned}
 [\mathbb{H}_{\hat{W}}^{eq}] &= \hat{W}_{\bar{U}\bar{U}} + \mathbf{I} \times \hat{W}_{\bar{W}\bar{W}} \times \mathbf{I} + \hat{W}_{JJ} \mathbf{I} \otimes \mathbf{I} \\
 &\quad + \hat{W}_{\bar{U}\bar{W}} \times \mathbf{I} + \mathbf{I} \times \hat{W}_{\bar{W}\bar{U}} \\
 &\quad + \hat{W}_{\bar{U}J} \otimes \mathbf{I} + \mathbf{I} \otimes \hat{W}_{J\bar{U}} + \left(\mathbf{I} \times \hat{W}_{\bar{W}J} \right) \otimes \mathbf{I} \\
 &\quad + \mathbf{I} \otimes \left(\hat{W}_{J\bar{W}} \times \mathbf{I} \right).
 \end{aligned} \tag{92}$$

Notice the similarities between above equation (92) and (11), where the second order tensors \mathbf{F} and \mathbf{H} have been replaced with the identity tensor \mathbf{I} .

The relationship between the different tangent operators $[\mathbb{H}_{\bar{W}}]$, $[\mathbb{H}_{\hat{W}}]$ and $[\mathbb{H}_{W_b}]$ as summarised in above Eqs. (85) and (91) establishes an elegant link between the beam and the more general continuum models. For a given continuum polyconvex constitutive model (characterised by positive semi-definite Hessian operators $[\mathbb{H}_{\hat{W}}]$ (58) and $[\mathbb{H}_W]$ (9)), the associated (beam) Hessian operators $[\mathbb{H}_{\bar{W}}]$ (91)–(92)

and $[\mathbb{H}_{W_b}]$ (87) will be positive semi-definite too. Therefore, as expected, satisfaction of the ellipticity condition [19] in the continuum through a polyconvex constitutive law implies satisfaction of the ellipticity condition [35] in the consistently derived beam problem. The opposite cannot be inferred. Further inclusion of suitable growth conditions into the constitutive model guarantees well posedness of the overall problem [19].

4.4 Polyconvex constitutive models in classical beam theory

In general, given an arbitrary objective energy functional expressed as a function of the extended kinematic set $\{\bar{U}, \bar{W}, J\}$, it is possible to obtain an equivalent expression in terms of the strain vector \mathbf{B} (and hence in terms of $\{\Gamma, \mathbf{K}\}$) after some algebraic manipulations.⁸

4.4.1 A Mooney–Rivlin constitutive model

As presented in Appendix 2, an expression for the strain energy functional $\bar{W}(\mathbf{B})$ can be obtained for a Mooney–Rivlin model (154) as (where use has been made of the function f as defined in (18))

$$\begin{aligned} \bar{W}(\mathbf{B}) = & \alpha (\mathbf{B} \cdot \mathbf{B} + 2\mathbf{B} \cdot \mathbf{D}_3 + 3) \\ & - 2(\alpha + 2\beta) \ln(\mathbf{B} \cdot \mathbf{D}_3 + 1) + \frac{\lambda}{2} (\mathbf{B} \cdot \mathbf{D}_3)^2 \\ & + \beta(2\mathbf{B} \cdot \mathbf{B} - \underbrace{(\mathbf{D}_3 \times \mathbf{B}) \cdot (\mathbf{D}_3 \times \mathbf{B})}_{\xi}) \\ & + 4\mathbf{B} \cdot \mathbf{D}_3 + 3). \end{aligned} \tag{93}$$

As shown in Appendix 2, the stress vector $\Sigma_{\mathbf{B}}$ is obtained as

$$\begin{aligned} \Sigma_{\mathbf{B}} = & 2\alpha (\mathbf{B} + \mathbf{D}_3) + 2\beta [(I + \mathbf{D}_3 \otimes \mathbf{D}_3) \mathbf{B} + 2\mathbf{D}_3] \\ & + \left[\lambda (\mathbf{B} \cdot \mathbf{D}_3) - \frac{2(\alpha + 2\beta)}{(\mathbf{B} \cdot \mathbf{D}_3 + 1)} \right] \mathbf{D}_3, \end{aligned} \tag{94}$$

and the Hessian operator $[\mathbb{H}_{\bar{W}}]$ as

$$\begin{aligned} [\mathbb{H}_{\bar{W}}] = & 2\alpha I + 2\beta (I + \mathbf{D}_3 \otimes \mathbf{D}_3) \\ & + \left[\lambda + \frac{2(\alpha + 2\beta)}{(\mathbf{B} \cdot \mathbf{D}_3 + 1)^2} \right] \mathbf{D}_3 \otimes \mathbf{D}_3. \end{aligned} \tag{95}$$

Hence, convexity is guaranteed provided that the material parameters α, β and λ are non-negative. Therefore, conforming to the conclusions obtained in Sect. 4.3, both Hessian operators $[\mathbb{H}_{\bar{W}}]$ (16) and $[\mathbb{H}_{\bar{W}}]$ (95) are positive semi-definite for the Mooney–Rivlin model.

⁸ Algebraic operations are greatly facilitated via application of the properties of the tensor cross product as presented in 2.

For the degenerate case of a Neo-Hookean model ($\alpha = \mu/2, \beta = 0$), the stress vector $\Sigma_{\mathbf{B}}$ and the Hessian operator $[\mathbb{H}_{\bar{W}}]$ are obtained as

$$\begin{aligned} \Sigma_{\mathbf{B}} = & \mu (\mathbf{B} + \mathbf{D}_3) + \left[\lambda (\mathbf{B} \cdot \mathbf{D}_3) - \frac{\mu}{(\mathbf{B} \cdot \mathbf{D}_3 + 1)} \right] \mathbf{D}_3; \\ [\mathbb{H}_{\bar{W}}] = & \mu I + \left[\lambda + \frac{\mu}{(\mathbf{B} \cdot \mathbf{D}_3 + 1)^2} \right] \mathbf{D}_3 \otimes \mathbf{D}_3. \end{aligned} \tag{96}$$

4.4.2 A Saint-Venant constitutive model

The Saint-Venant–Kirchhoff strain energy in Eq. (24) can be expressed in terms of the strain vector \mathbf{B} using Eq. (38)

$$E = \frac{1}{2} (\mathbf{B} \otimes \mathbf{D}_3 + \mathbf{D}_3 \otimes \mathbf{B}) + \frac{1}{2} (\mathbf{B} \cdot \mathbf{B}) \mathbf{D}_3 \otimes \mathbf{D}_3. \tag{97}$$

Introduction of Eq. (97) into (24) leads to a final expression of the strain energy \bar{W}_{SVK} in terms of \mathbf{B} . Successive derivations of this expression yield the following Hessian operator $[\mathbb{H}_{\bar{W}}]$

$$\begin{aligned} [\mathbb{H}_{\bar{W}}] = & \lambda[(\mathbf{D}_3 + \mathbf{B}) \otimes (\mathbf{D}_3 + \mathbf{B}) + (\mathbf{B} \cdot \mathbf{D}_3 + \frac{1}{2} \mathbf{B} \cdot \mathbf{B}) I] \\ & + \mu[\mathbf{D}_3 \otimes \mathbf{D}_3 + I + 2(\mathbf{B} \cdot \mathbf{D}_3) I] \\ & + 2(\mathbf{B} \otimes \mathbf{D}_3 + \mathbf{D}_3 \otimes \mathbf{B}) \\ & + 2\mathbf{B} \otimes \mathbf{B} + (\mathbf{B} \cdot \mathbf{B}) I]. \end{aligned} \tag{98}$$

Notice that for the particular case of $\mathbf{B} = -\mathbf{D}_3$

$$[\mathbb{H}_{\bar{W}}] \Big|_{\mathbf{B}=-\mathbf{D}_3} = -\frac{\lambda}{2} I - \mu \mathbf{D}_3 \otimes \mathbf{D}_3. \tag{99}$$

which is clearly non positive semi-definite. Since positive semi-definiteness of $\mathbb{H}_{\bar{W}}$ is not guaranteed for all values of the strain vector \mathbf{B} , it can be concluded that the constitutive model is not convex in \mathbf{B} . In order to overcome this shortcoming, it is customary [8] to neglect the higher order terms in (97) (last term on the right hand side of (97)) resulting in a linearised version of the Saint Venant–Kirchhoff model as

$$\bar{W}_{SVK}^{\text{lin}}(\mathbf{B}) = \frac{\lambda}{2} \Psi_1^{\text{lin}} + \mu \Psi_2^{\text{lin}}, \tag{100}$$

where

$$\Psi_1^{\text{lin}} = (\mathbf{B} \cdot \mathbf{D}_3)^2 \quad \Psi_2^{\text{lin}} = \frac{1}{2} [(\mathbf{B} \cdot \mathbf{D}_3)^2 + (\mathbf{B} \cdot \mathbf{B})]. \tag{101}$$

The Hessian operator $\mathbb{H}_{\bar{W}}$ (75) for the linearised Saint-Venant constitutive model is finally computed as:

$$[\mathbb{H}_{\bar{W}}] = (2\lambda + \mu) \mathbf{D}_3 \otimes \mathbf{D}_3 + \mu I. \tag{102}$$

Therefore, this linearised version of the Saint-Venant–Kirchhoff constitutive model is convex with respect to \mathbf{B} and hence, according to Sect. 4.3, $W_b(\Gamma, \mathbf{K})$ will be convex with respect to $\{\Gamma, \mathbf{K}\}$. However, on the contrary to the Mooney–Rivlin and Neo-Hookean models, the energy functional Ψ_{SVK}^{lin} still fails to satisfy adequate growth conditions, namely the coercivity condition⁹ [30,36]

$$\lim_{J(\nabla_0 \mathbf{x}) \rightarrow 0^+} \Psi(\nabla_0 \mathbf{x}) \rightarrow \infty. \tag{103}$$

Hence, under high compression scenarios, the Saint Venant–Kirchhoff constitutive model does not perform well. Given the limitations of this model in large strain scenarios, more suitable constitutive laws need to be considered.

5 Variational formulation

The objective of this section is to succinctly present the variational formulation of the geometrically exact beam theory following a continuum degenerate approach. This will be used as a starting point when describing the discretisation strategy employed. As a starting point, let us consider the following standard total energy minimisation variational principle

$$\begin{aligned} &\Pi_M(\mathbf{x}_0^*, \boldsymbol{\theta}^*) \\ &= \min_{\mathbf{x}_0, \boldsymbol{\theta}} \left\{ \int_V \Psi(\nabla_0 \mathbf{x}) \Big|_{\substack{x = x_0 + \bar{x} \\ d_i = R(\boldsymbol{\theta}) D_i}} dV - W^{\text{ext}} \Big|_{\substack{x = x_0 + \bar{x} \\ d_i = R(\boldsymbol{\theta}) D_i}} \right\}. \end{aligned} \tag{104}$$

where \mathbf{x}_0^* and $\boldsymbol{\theta}^*$ denote the exact solution and W^{ext} the work done by external forces. The stationary condition of this functional leads to the principle of virtual work, written as¹⁰

$$\begin{aligned} D\Pi_M[\delta \mathbf{u}_0, \delta \boldsymbol{\theta}] &= \int_V \mathbf{P} : D\mathbf{F}[\delta \mathbf{u}_0, \delta \boldsymbol{\theta}] dV \\ &- DW^{\text{ext}}[\delta \mathbf{u}_0, \delta \boldsymbol{\theta}] = 0. \end{aligned} \tag{105}$$

Following the continuum degenerate approach described in Sect. 3, substitution of $D\mathbf{F}[\delta \mathbf{u}_0, \delta \boldsymbol{\theta}]$ by its expression in Eq. (42) allows to re-write the first time on the right hand side of (105) as

$$\begin{aligned} D\Pi_M^{\text{int}}[\delta \mathbf{u}_0, \delta \boldsymbol{\theta}] &= \int_0^L \delta \mathbf{u}'_0 \cdot \left[\int_{A(s)} \mathbf{P} \mathbf{D}_3 dA \right] ds \\ &+ \int_0^L D\mathbf{d}'_\alpha[\delta \boldsymbol{\theta}] \cdot \left[\int_{A(s)} \mathbf{P} \mathbf{D}_\alpha dA \right] ds \\ &+ \int_0^L D\mathbf{d}'_\alpha[\delta \boldsymbol{\theta}] \cdot \left[\int_{A(s)} \eta^\alpha \mathbf{P} \mathbf{D}_3 dA \right] ds, \end{aligned} \tag{106}$$

where the first Piola–Kirchhoff stress tensor \mathbf{P} is evaluated via Eq. (7) in terms of the conjugate stresses $\boldsymbol{\Sigma}_F, \boldsymbol{\Sigma}_H$ and $\boldsymbol{\Sigma}_J$. An iterative¹¹ Newton–Raphson process is applied to obtain the solution. This is usually achieved by solving a linearised system for the increments $\{\mathbf{u}_0, \Delta \boldsymbol{\theta}\}$ as [1]

$$\begin{aligned} &D^2\Pi_M(\mathbf{x}_0^k, \boldsymbol{\theta}^k)[\delta \mathbf{u}_0, \delta \boldsymbol{\theta}; \mathbf{u}_0, \Delta \boldsymbol{\theta}] \\ &= -D\Pi_M(\mathbf{x}_0^k, \boldsymbol{\theta}^k)[\delta \mathbf{u}_0, \delta \boldsymbol{\theta}], \end{aligned} \tag{107}$$

where the update of the position of the center line \mathbf{x}_0 and of the rotation matrix \mathbf{R} is obtained via

$$\begin{aligned} \mathbf{x}_0^{k+1} &= \mathbf{x}_0^k + \mathbf{u}_0; \quad \mathbf{R}^{k+1} = \mathbf{R}(\Delta \boldsymbol{\theta}) \mathbf{R}^k; \\ \mathbf{d}_i^{k+1} &= \mathbf{R}^{k+1} \mathbf{D}_i. \end{aligned} \tag{108}$$

and with the rotation matrix $\mathbf{R}(\Delta \boldsymbol{\theta})$ obtained particularising Eq. (34) to the incremental rotation vector $\Delta \boldsymbol{\theta}$. Finally, in the absence of follower loads, the second derivative of the total energy functional is given by

$$\begin{aligned} &D^2\Pi_M(\mathbf{x}_0^k, \boldsymbol{\theta}^k)[\delta \mathbf{u}_0, \delta \boldsymbol{\theta}; \mathbf{u}, \Delta \boldsymbol{\theta}] \\ &= \int_V D^2\Psi(\mathbf{x}_0^k, \boldsymbol{\theta}^k)[\delta \mathbf{u}_0, \delta \boldsymbol{\theta}; \mathbf{u}, \Delta \boldsymbol{\theta}] dV, \end{aligned} \tag{109}$$

where the tangent operator is evaluated using Eq. (49).

6 Finite element discretisation

Traditional finite element discretisations based on rotations do not satisfy objectivity [1,2,4,6]. In order to circumvent this shortcoming, a formulation in which the nodal degrees of freedom correspond to displacements of the centre line \mathbf{x}_0 and the director triad $\{\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3\}$ is considered following reference [1,6]. In this approach, \mathbf{x}_0 and the director vectors are interpolated as

$$\mathbf{x}_0 = \mathbf{x}_0^a N_u^a(s); \quad \mathbf{d}_i = \mathbf{d}_i^a N_d^a(s). \tag{110}$$

¹¹ The letter k will indicate iteration number.

⁹ The more general strain energy functional Ψ_{SVK} also fails to satisfy the coercivity condition.

¹⁰ The expression of the external virtual work $DW^{\text{ext}}[\delta \mathbf{u}_0, \delta \boldsymbol{\theta}]$ is well known and, hence, omitted.

where N_u^a and N_d^a are standard Finite Element shape functions. Using a Galerkin approach

$$\delta \mathbf{u}_0 = \delta \mathbf{u}_0^a N_u^a(s); \quad D \mathbf{d}_i[\delta \boldsymbol{\theta}] = D \mathbf{d}_i^a[\delta \boldsymbol{\theta}] N_d^a(s). \quad (111)$$

Particularisation of Eqs. (40) to the triads \mathbf{d}_k collocated at the nodes of the underlying mesh, it results in

$$\begin{aligned} D \mathbf{d}_i^a[\delta \boldsymbol{\theta}^a] &= \delta \boldsymbol{\theta}^a \times \mathbf{d}_i^a; \\ D^2 \mathbf{d}_i^a[\delta \boldsymbol{\theta}^a; \Delta \boldsymbol{\theta}^b] &= \Delta \boldsymbol{\theta}^a (\delta \boldsymbol{\theta}^a \cdot \mathbf{d}_i^a) \delta_{ab} - \mathbf{d}_i^a (\delta \boldsymbol{\theta}^a \cdot \Delta \boldsymbol{\theta}^a) \delta_{ab} \end{aligned} \quad (112)$$

where the cross product property $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \mathbf{b} (\mathbf{a} \cdot \mathbf{c}) - \mathbf{c} (\mathbf{a} \cdot \mathbf{b})$ has been used in Eq. (112) in order to obtain the expression for $D^2 \mathbf{d}_i^a[\delta \boldsymbol{\theta}^a; \Delta \boldsymbol{\theta}^b]$.

6.1 Discretised kinematics

Based on the discretisation presented and making use of Eq. (112), the discretised form of the first and second directional derivatives of \mathbf{F} in Eq. (42) can be obtained as

$$\begin{aligned} D \mathbf{F}[\delta \mathbf{u}_0^a, \delta \boldsymbol{\theta}^a] &= \delta \mathbf{u}_0^a \otimes N_u^{a'} \mathbf{D}_3 + (\delta \boldsymbol{\theta}^a \times \mathbf{d}_\alpha^a) \otimes \mathbf{Q}_\alpha^a; \\ D^2 \mathbf{F}[\delta \mathbf{u}_0^a, \delta \boldsymbol{\theta}^a; \mathbf{u}_0^b, \Delta \boldsymbol{\theta}^b] &= (\Delta \boldsymbol{\theta}^a (\delta \boldsymbol{\theta}^a \cdot \mathbf{d}_\alpha^a) \delta_{ab} - \mathbf{d}_\alpha^a (\delta \boldsymbol{\theta}^a \cdot \Delta \boldsymbol{\theta}^a) \delta_{ab}) \otimes \mathbf{Q}_\alpha^a, \end{aligned} \quad (113)$$

with

$$\mathbf{Q}_\alpha^a = (N_d^a \mathbf{D}_\alpha + \eta^\alpha N_d^{a'} \mathbf{D}_3). \quad (114)$$

6.2 Discretisation of the weak forms

Substitution of the discretisation expressions for \mathbf{x}_0 and \mathbf{d}_k and its virtual variations in (110) and (111) respectively, into (106) leads to the final discretised form of the principle of virtual work as

$$\begin{aligned} D \Pi_M^{int}[\delta \mathbf{u}_0^a, \delta \boldsymbol{\theta}^a] &= \delta \mathbf{u}_0^a \cdot \int_0^L N_u^{a'} \left[\int_A \mathbf{P} \mathbf{D}_3 dA \right] ds \\ &+ \delta \boldsymbol{\theta}^a \cdot \mathbf{d}_\alpha^a \times \int_0^L \left[\int_A \mathbf{P} \mathbf{Q}_\alpha^a dA \right] ds. \end{aligned} \quad (115)$$

where \mathbf{P} is evaluated according to Eq. (7). Notice that the first term in the right hand side of Eq. (115) can be identified as the weak form of conservation of linear momentum whereas the second term can be identified as the weak form of the conservation of angular momentum.

6.3 Discretisation of linearisations

Bearing in mind Eq. (49), the discretisation of the second directional derivative of the potential Π_M in Eq. (109) will

be presented in this section. For instance, second derivatives with respect to the displacement of the centre line are discretised as

$$\begin{aligned} D^2 \Pi_M[\delta \mathbf{u}_0^a; \mathbf{u}_0^b] &= \int_V D \mathbf{F}[\delta \mathbf{u}_0^a] : [\mathbb{H}_W^{eq}] : D \mathbf{F}[\mathbf{u}_0^b] dV \\ &+ \int_V \boldsymbol{\Sigma}_g : (D \mathbf{F}[\delta \mathbf{u}_0^a] \times D \mathbf{F}[\mathbf{u}_0^b]) dV, \end{aligned} \quad (116)$$

where $\boldsymbol{\Sigma}_g = \boldsymbol{\Sigma}_H + \boldsymbol{\Sigma}_J \mathbf{F}$. Equation (116) can be equivalently expressed as

$$D^2 \Pi_M[\delta \mathbf{u}_0^a; \mathbf{u}_0^b] = \delta \mathbf{u}_0^a \cdot \mathbf{K}_{uu}^{ab} \mathbf{u}_0^b, \quad (117)$$

where the stiffness matrix \mathbf{K}_{uu}^{ab} is obtained as

$$\begin{aligned} \mathbf{K}_{uu}^{ab} &= \int_0^L N_u^{a'} N_u^{b'} \left[\int_{A(s)} ([\mathbb{H}_W^{eq}] \mathbf{D}_3 \mathbf{D}_3 + \boldsymbol{\mathcal{E}} \right. \\ &\quad \left. : \boldsymbol{\Sigma}_g (\mathbf{D}_3 \times \mathbf{D}_3)) dA \right] ds, \end{aligned} \quad (118)$$

where $\boldsymbol{\mathcal{E}}$ is the third order permutation tensor. It is easy to observe that the second (geometric) term in (118) vanishes, leading to the final expression for \mathbf{K}_{uu}^{ab} as

$$\mathbf{K}_{uu}^{ab} = \int_0^L N_u^{a'} N_u^{b'} \left[\int_{A(s)} \mathbf{D}_3 : [\mathbb{H}_W^{eq}] : \mathbf{D}_3 dA \right] ds. \quad (119)$$

Similarly, the discretisation of the cross derivatives of the potential Π_M with respect to changes in the displacement of the centre line and rotations would lead to an expression of the form

$$\begin{aligned} D^2 \Pi_M[\delta \mathbf{u}_0^a; \Delta \boldsymbol{\theta}^b] &= \int_V D \mathbf{F}[\delta \mathbf{u}_0^a] : [\mathbb{H}_W^{eq}] : D \mathbf{F}[\Delta \boldsymbol{\theta}^b] dV \\ &+ \int_V \boldsymbol{\Sigma}_g : (D \mathbf{F}[\delta \mathbf{u}_0^a] \times D \mathbf{F}[\Delta \boldsymbol{\theta}^b]) dV, \end{aligned} \quad (120)$$

which can be written equivalently as $D^2 \Pi_M[\delta \mathbf{u}_0^a; \Delta \boldsymbol{\theta}^b] = \delta \mathbf{u}_0^a \cdot \mathbf{K}_{u\theta}^{ab} \Delta \boldsymbol{\theta}^b$, with the stiffness matrix $\mathbf{K}_{u\theta}^{ab}$ being

$$\begin{aligned} \mathbf{K}_{u\theta}^{ab} &= - \int_0^L N_u^{a'} \left(\int_{A(s)} \mathbf{D}_3 : [\mathbb{H}_W^{eq}] : \mathbf{Q}_\alpha^b dA \right) \times \mathbf{d}_\alpha^b ds \\ &+ \int_0^L \left[\int_{A(s)} (\mathbf{d}_\alpha^b \otimes \mathbf{b}_\alpha^{ab} - (\mathbf{d}_\alpha^b \cdot \mathbf{b}_\alpha^{ab}) \mathbf{I}) dA \right] ds, \end{aligned} \quad (121)$$

where

$$\mathbf{b}_\alpha^{ab} = \boldsymbol{\Sigma}_g (N_u^{a'} \mathbf{D}_3 \times \mathbf{Q}_\alpha^b). \quad (122)$$

Finally, the discretisation of the second derivatives with respect to changes in rotations is obtained as

$$\begin{aligned}
 D^2\Pi_M[\delta\theta^a; \Delta\theta^b] &= \int_V DF[\delta\theta^a] : [\mathbb{H}_W^{eq}] : DF[\Delta\theta^b] dV \\
 &+ \int_V \Sigma_g : (DF[\delta\theta^a] \times DF[\Delta\theta^b]) dV \\
 &+ \int_V \mathbf{P} : D^2\mathbf{F}[\delta\theta^a; \Delta\theta^b] dV.
 \end{aligned}
 \tag{123}$$

Equation (123) is equivalent to $D^2\Pi_M[\delta\theta^a; \Delta\theta^b] = \delta\theta^a \cdot \mathbf{K}_{\theta\theta}^{ab} \Delta\theta^b$, with the stiffness matrix $\mathbf{K}_{\theta\theta}^{ab}$ being defined as

$$\begin{aligned}
 \mathbf{K}_{\theta\theta}^{ab} &= - \int_0^L \left(\int_{A(s)} \mathbf{Q}_\alpha^a : [\mathbb{H}_W^{eq}] : \mathbf{Q}_\beta^b dA \right) \times \left(\mathbf{d}_\alpha^a \otimes \mathbf{d}_\beta^b \right) ds \\
 &+ \int_0^L \left[\int_{A(s)} \left(\boldsymbol{\varepsilon} : \Sigma_g \left(\mathbf{Q}_\alpha^a \times \mathbf{Q}_\beta^b \right) \right) \times \left(\mathbf{d}_\alpha^a \otimes \mathbf{d}_\beta^b \right) dA \right] ds \\
 &+ \int_0^L \left[\int_{A(s)} \left(\mathbf{d}_\alpha^a \otimes \mathbf{P} \mathbf{Q}_\alpha^a - \left(\mathbf{d}_\alpha^a \cdot \mathbf{P} \mathbf{Q}_\alpha^a \right) \mathbf{I} \right) \delta_{ab} dA \right] ds.
 \end{aligned}
 \tag{124}$$

7 Numerical examples

Note that it is not the primary aim of this paper to propose a new finite element discretisation. Nevertheless, the objective of this section is to present a series of numerical examples in order to show the applicability of the variational and computational frameworks presented in Sects. 3, 5 and 6. For all the examples included, linear shape functions have been used for the interpolation of the centre line and the director triad as presented in [1]. The consideration of dynamic effects for one of the examples presented has been implemented in a straightforward manner [4] and thus, it is not further discussed. Reduced numerical integration [1] of the internal virtual work contributions has been carried out in order to alleviate locking effects, whilst exact integration of the inertial contributions has been implemented. For the last example presented, a comparison of the beam model against a sophisticated mixed continuum polyconvex computational framework as reported in [12] is carried out.

7.1 Bending test

This example consists of a straight cantilever beam of length L with squared cross sectional area of side $\frac{1}{25}m$. This example has already been reported in [1,3] using a linearised Saint Venant–Kirchhoff constitutive model (100) with $\lambda_{lin} = 1 \text{ N/m}^2$ and $\mu_{lin} = \frac{1}{2} \text{ N/m}^2$ the Lamé coefficients of the material defined in the reference configuration. With these parameters, the Young’s modulus is computed

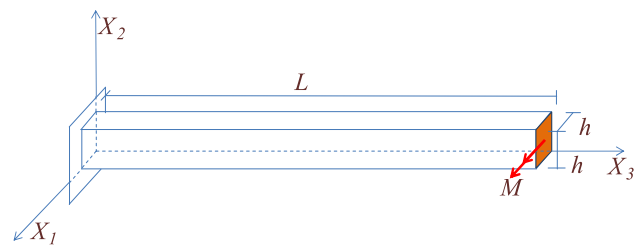


Fig. 3 Bending test example. Geometry and boundary conditions. $L = 1 \text{ m}$ and $h = \frac{1}{25} \text{ m}$

as $E = 1 \text{ N/m}^2$. The geometry and boundary conditions of the problem are shown in Fig. 3. The maximum bending moment applied at the free end of the beam is defined by $M_{max} = \frac{2\pi EI}{L}$, where I denotes the second moment of area of the beam.

For this example, four nonlinear constitutive models are compared by employing the computational framework described above. For consistency with the reference [1,3], the material characterisation of the different constitutive models employed is carried out at the origin utilising the Lamé coefficients λ_{lin} and μ_{lin} defined above. Specifically, the four nonlinear constitutive models are: (i) a Saint Venant–Kirchhoff model defined in (27) with material parameters λ_{SVK} and μ_{SVK} ; (ii) a Mooney–Rivlin model defined in (12)–(13) with material parameters α_{MR} , β_{MR} and λ_{MR} ; (iii) a Neo-Hookean model defined in (12)–(13) with material parameters α_{NH} , $\beta_{NH} = 0$ and λ_{NH} , and iv) a modified Mooney–Rivlin model defined by the expression

$$\begin{aligned}
 W_{MMR}(\mathbf{F}, \mathbf{H}, J) &= \frac{\alpha_{MMR}}{4} \left((\mathbf{F} : \mathbf{F})^2 - 9 \right) \\
 &+ \frac{\beta_{MMR}}{4} \left((\mathbf{H} : \mathbf{H})^2 - 9 \right) + f(J); \\
 f(J) &= -(\alpha_{MMR} + 2\beta_{MMR}) \ln(J) \\
 &+ \frac{\lambda_{MMR}}{2} (J - 1)^2.
 \end{aligned}
 \tag{125}$$

The strain energy defined above in Eq. (125) has been extracted from a modified Mooney–Rivlin model given in [29]. Material characterisation renders the following equivalences between the respective material parameters and those of the linearised model (i.e. λ_{lin} and μ_{lin}), as

$$\begin{aligned}
 \mu_{SVK} &= \mu_{lin}; \quad \lambda_{SVK} = \lambda_{lin}; \\
 \alpha_{NH} &= \mu_{lin}; \quad \lambda_{NH} = \lambda_{lin}; \\
 2\alpha_{MR} + 2\beta_{MR} &= \mu_{lin}; \quad \lambda_{MR} = \lambda_{lin}; \\
 \alpha_{MMR} + \beta_{MMR} &= \mu_{lin}; \\
 \lambda_{MMR} &= \lambda_{lin} - 2/3\alpha_{MMR} - 14/3\beta_{MMR}.
 \end{aligned}
 \tag{126}$$

In order to close the definition of the material parameters for the Mooney–Rivlin and modified Mooney–Rivlin

Fig. 4 Bending test for a discretisation of 24 linear elements (25×6 degrees of freedom defined in the variables x_0 and $\Delta\theta$). Results shown for a Saint Venant–Kirchhoff constitutive model defined in (27) with material properties μ_{SVK} and λ_{SVK} defined in (126). **a** Deformed shape for $M = M_{max}$. **b** Contour plot of the hydrostatic pressure distribution p (N/m²) for $M = 0.8M_{max}$

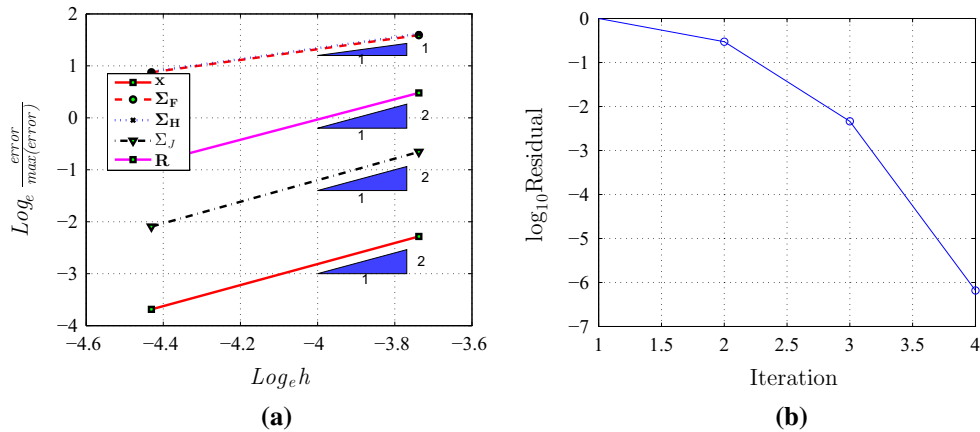
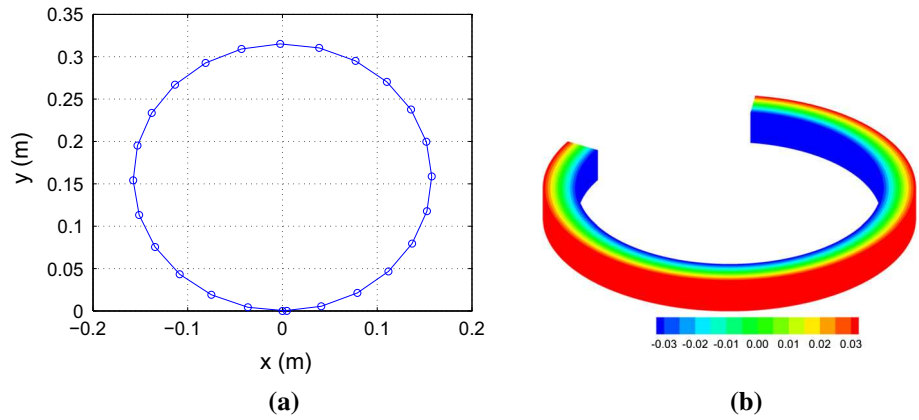


Fig. 5 Bending test for a Mooney–Rivlin model defined in (12) and (13) with material properties α_{MR} , β_{MR} and λ_{MR} defined in (126). **a** Order of accuracy for the variables x , x_0 , R , Σ_F , Σ_H and Σ_J . Two

meshes analysed for the maximum applied moment $M = M_{max}$: 80 and 160 linear elements. Reference solution taken from a 640 element mesh. **b** Quadratic convergence of the Newton–Raphson algorithm

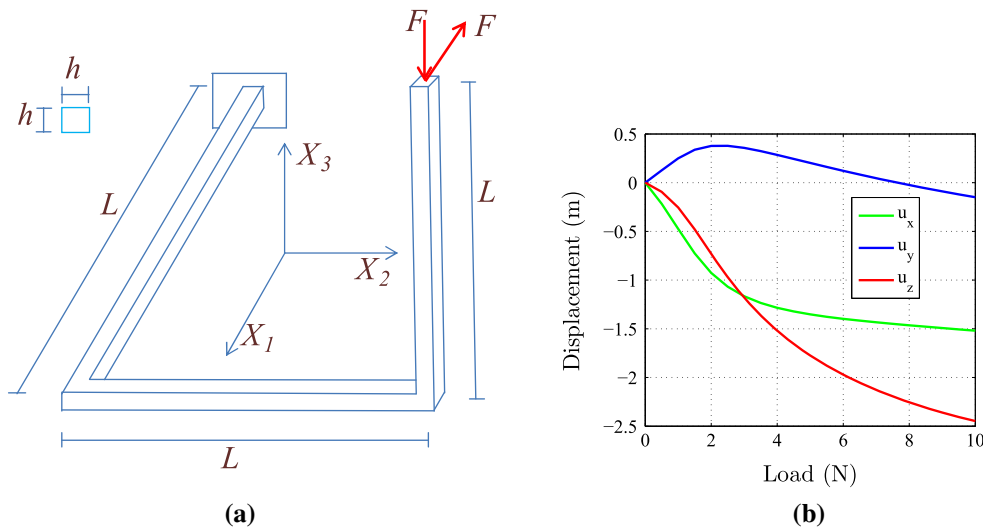
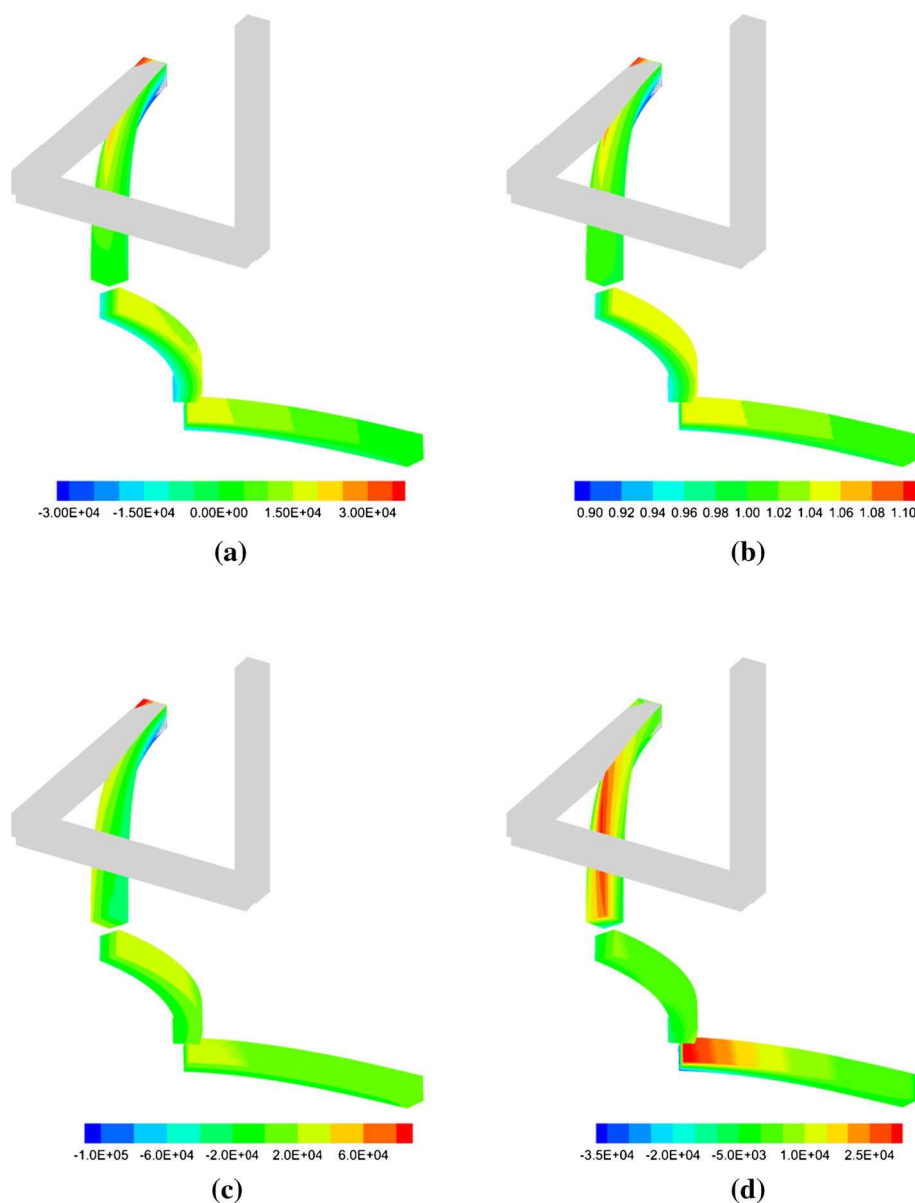


Fig. 6 Beam with slope discontinuity. **a** Geometry and boundary conditions. $L = 1$ m and $h = 0.1$ m. **b** Absolute value of the tip displacement versus absolute value of the applied force F (N). Material parameters of $\mu_{lin} = 0.5 \times 10^6$ N/m² and $\nu = 0$ in the reference

configuration. Results shown for the Mooney–Rivlin model defined in Eqs. (12) and (13) for a discretisation of 18 linear elements (19×6 degrees of freedom)

Fig. 7 Beam with slope discontinuity. Reference configuration (*shadowed*) and deformed configuration. Contour plot of **a** hydrostatic pressure p (N/m²), **b** Jacobian J , **c** stress component σ_{11} (N/m²) and **d** stress component σ_{22} (N/m²). Material parameters of $\mu_{lin} = 0.5 \times 10^6$ N/m² and $\nu = 0$ in the reference configuration. Results shown for the Mooney–Rivlin model defined in Eqs. (12) and (13) for a discretisation of 480 linear elements (481 × 6 degrees of freedom)



models, the following choice of material parameters is preferred $\alpha_{MR} = \beta_{MR} = \mu_{lin}$ and $\alpha_{MMR} = 3/4\mu_{lin}$ and $\beta_{MMR} = \mu_{lin}/4$.

As reported in reference [43], for a linearised Saint Venant–Kirchhoff model (100) with Poisson ratio $\nu = 0$ and an applied moment of value $M = M_{max}$, the beam closes on itself forming a closed loop which can be described with an available analytical solution. In general, for a nonlinear constitutive model (or when the Poisson ratio ν is not equal to zero), the exact closure of the beam configuration cannot be a priori guaranteed. Figure 4a shows the deformed shape of the beam for the general nonlinear Saint Venant–Kirchhoff model (27). As can be observed in this figure, there can be observed a small interpenetration due to the nonlinearity of the considered model.

With respect to the stress distribution, Fig. 4b shows the contour plot of the hydrostatic pressure for the Saint Venant–Kirchhoff constitutive model. For the other three constitutive models, the hydrostatic pressure distribution is almost identical, with hardly any differences, and is thus not displayed. This can be explained due to the moderate strains undergone through the deformation process. For this particular example, the choice of a specific model does not seem to be relevant for the overall solution.

The objective of this example is also to demonstrate the p -order of accuracy of the formulation, as a function of the chosen finite element approximation spaces. For this purpose, and particularising for the Mooney–Rivlin model, the beam is initially discretised with 80 elements and, subsequently, h -refinement is carried out generating a total of 3

discretisations. As a closed form solution is not available for this problem (due to the nonlinearity of the constitutive model adopted), the finest mesh is used to generate numerically the so-called “benchmark” solution for comparison purposes. The two coarsest meshes are compared against the finest mesh. The error between the benchmark solution and the other discretisations is measured in the L^2 norm for all the unknown variables. Let us define for a tensor (e.g. scalar, vector or second order) field, the L^2 norm as

$$\|\zeta\|_{L^2} = \left[\int_V (\zeta : \zeta) dV \right]^{1/2} \tag{127}$$

associated with the magnitude of the tensor field ζ . Although the integrands associated to the internal virtual work have been underintegrated along the centre line, the integral defined in Eq. (127) for the evaluation of the error is computed exactly. Notice that in the evaluation of the error norm, a reconstruction of the continuum is carried out. In our case, ζ can be any of the kinematic or kinetic variables, namely \mathbf{x} ,¹² \mathbf{x}_0 , \mathbf{R} , Σ_F , Σ_H and Σ_J . This enables the definition of the following error norm $\|\zeta_i - \zeta_b\|_{L^2} / \|\zeta_b\|_{L^2}$, where ζ_b stands for the benchmark solution and ζ_i the solution of the i -th mesh, with $i = 1, \dots, (b - 1)$. This can then be used to assess the convergence of the algorithm under h -refinement.

Figure 5a shows the order of accuracy of the variables \mathbf{x} , \mathbf{x}_0 , \mathbf{R} , Σ_F , Σ_H and Σ_J . The convergence observed is $p + 1$ in \mathbf{x} , \mathbf{x}_0 , \mathbf{R} and Σ_J , and p for Σ_F , Σ_H . As expected, primary variables lead to $p + 1$ order of convergence whereas, with the exception of Σ_J for this particular example, derived magnitudes (i.e. stresses) lead to a reduced order of convergence p . For completeness, the quadratic convergence of the Newton–Raphson algorithm is shown in Fig. 5b.

7.2 Beam with slope discontinuity with zero Poisson ratio

In this example, a series of interconnected beams with a Young modulus of $E = 10^6 \text{ N/m}^2$ and a Poisson ratio of $\nu = 0$ are considered. The geometry and boundary conditions of the problem are shown in Fig. 6a. In order to model the connection between the beams at any angle, continuity of the incremental rotation angle $\Delta\theta$ is strongly imposed. This example has been presented in [3] using the linearised Saint Venant–Kirchhoff constitutive model (100). In this case, the Mooney–Rivlin constitutive model defined in Eqs. (12) and (13) is considered, with material parameters obtained after material characterisation in the undeformed reference configuration and the consideration of $\alpha_{MR} = \beta_{MR}$.

¹² Note that \mathbf{x} is computed according to Eq. (33) namely $\mathbf{x} = \mathbf{x}_0 + \bar{\mathbf{x}}$.

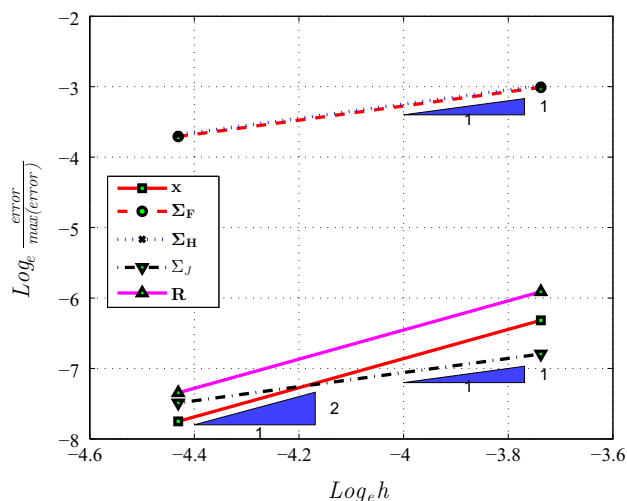


Fig. 8 Beam with slope discontinuity: order of accuracy of the \mathbf{x} , \mathbf{x}_0 , \mathbf{R} , Σ_F , Σ_H and Σ_J . Mooney–Rivlin model defined in (12) and (13). Two meshes analysed: 120 and 240 linear elements. Reference solution taken from a 480 element mesh

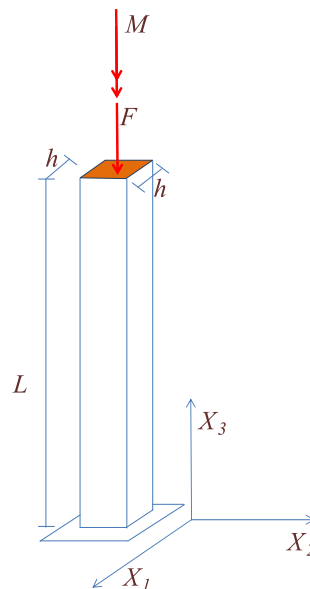
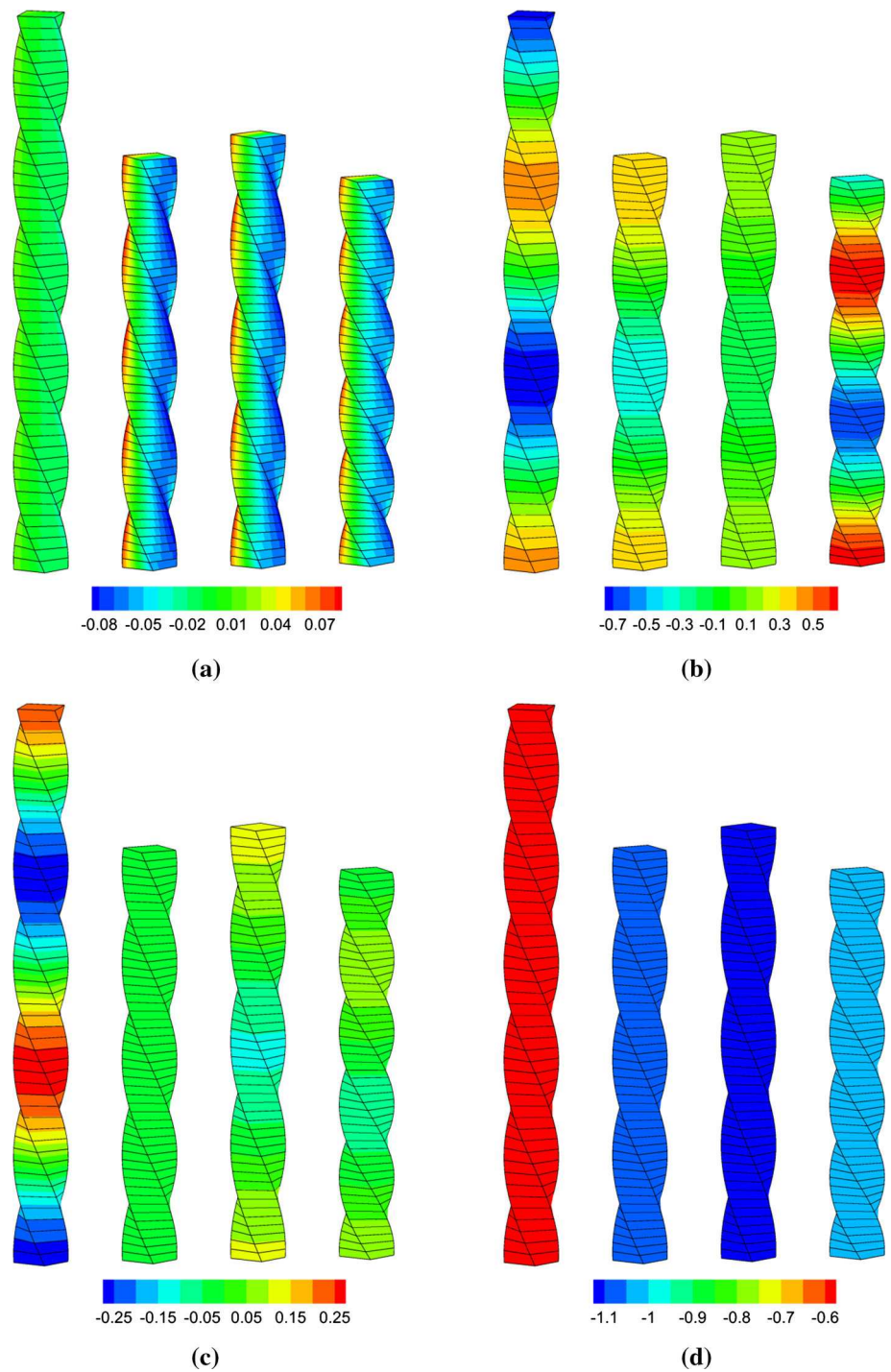


Fig. 9 Constrained torsion–compression example. Geometry and boundary conditions. $L = 20 \text{ m}$ and $h = 1 \text{ m}$

The objective of this example is to demonstrate the use of a polyconvex model with null Poisson ratio in the reference configuration for the case of a non-straight beam configuration. Notice the excellent agreement between the results presented in reference [3] for a linearised Saint Venant–Kirchhoff model and those for the current Mooney–Rivlin model in Fig. 6b.

Figure 7 shows the contour plot for different representative variables, namely the hydrostatic pressure p , the Jacobian J and different stress components (i.e. σ_{11} and σ_{22}) (notice that axis in reference $\{OX_1, OX_2, OX_3\}$ and deformed

Fig. 10 Constrained torsion–compression example. From *left to right*: Saint Venant–Kirchhoff with 30% of total loading, Neo-Hookean (100% of the loading), Mooney–Rivlin (100% of the loading) and modified Mooney–Rivlin (100% of the loading) models. Contour plot of **a** stress component σ_{32} (N/m²), **b** conjugate stress Σ_{F11} (N/m²), **c** conjugate stress Σ_{H22} (N/m²) and **d** conjugate stress Σ_J (N/m²). Material parameters $\mu_{lin} = 1/2.7$ N/m² and $E = 1$ N/m² in the reference configuration. Results for a discretisation of 50 linear elements (51×6 degrees of freedom associated to the spatial coordinates \mathbf{x}_0 and incremental rotations $\Delta\theta$)



$\{ox_1, ox_2, ox_3\}$ are coincident). In order to emphasize the high deformation that the beam is subjected to, a shadowed representation of the beam in its undeformed configuration is included. For visualisation purposes, the different spans of the structural system in the deformed configuration are shown slightly disconnected, enabling a clearer observation of the contour plot of the different variables.

Finally, Fig. 8 shows the expected $p + 1$ order of accuracy for the primary variables \mathbf{x} , \mathbf{x}_0 and \mathbf{R} , and the reduced p order of accuracy for the derived magnitudes Σ_F , Σ_H and Σ_J , where two meshes of 120 and 240 elements are compared with a reference solution defined from a mesh of 480 elements.

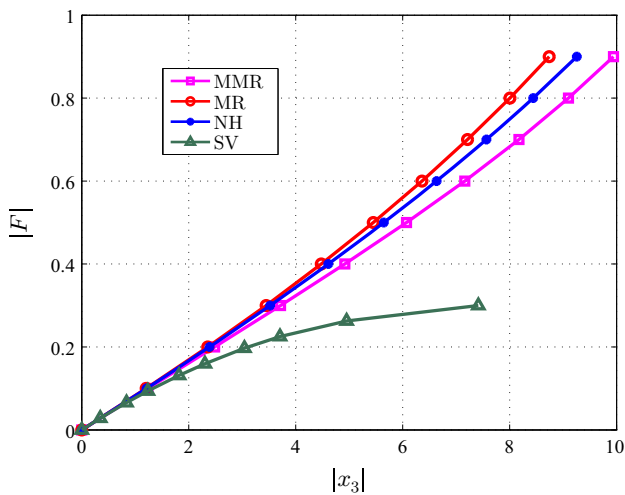


Fig. 11 Constrained torsion–compression example. Curve relating the absolute value of the compressible load at the free end of the beam and its associated displacement in the OX_3 axis. Saint-Venant model (SV), Neo-Hookean model (NH), Mooney–Rivlin model (MR) and Modified Mooney–Rivlin model (MMR). Material parameters $\mu_{lin} = 1/2.7 \text{ N/m}^2$ and $E = 1 \text{ N/m}^2$ in the reference configuration. Results for a discretisation of 50 linear elements (51×6 degrees of freedom associated to the spatial coordinates \mathbf{x}_0 and incremental rotations $\Delta\theta$)

7.3 Constrained torsion–compression example: coercivity of the models

In this example, a straight beam is subjected to high compressive and torsional effects. The objective of this example is to show the un-realistic behaviour of the Saint Venant–Kirchhoff (24) model specially under high compression scenarios in contrast to that of the Neo-Hookean, Mooney–Rivlin (12)–(13) and the modified Mooney–Rivlin (125) models.

The beam configuration and geometry and the boundary conditions of the problem are shown in Fig. 9. The free end is subjected to a compressive force $F = -9 \times 10^{-1} \text{ N}$ in the OX_3 axis and to a torsional moment $M = -2.7 \times 10^{-2} \text{ Nm}$ about the OX_3 axis. The displacement of all the nodes along the centre line is constrained to remain aligned with the straight axis of the undeformed beam (i.e. $\mathbf{x}_0 \cdot \mathbf{D}_\alpha = 0$). In this way, physical buckling is prevented and the lack of coercivity of the models can be explored. The material properties in the reference configuration are $E = 1 \text{ N/m}^2$ and $\nu = 0.35$.

Figure 10a shows the contour plot of σ_{12} (notice that axis in reference $\{OX_1, OX_2, OX_3\}$ and deformed $\{ox_1, ox_2, ox_3\}$ are coincident) and the deformed configuration of the beam for 100 % of the total load for the Neo-Hookean, Mooney–Rivlin and modified Mooney–Rivlin models. For the Saint Venant–Kirchhoff model, only 30% of the total loading has been applied. Figure 10b–d shows the contour plot for some representative conjugate stresses for the differ-

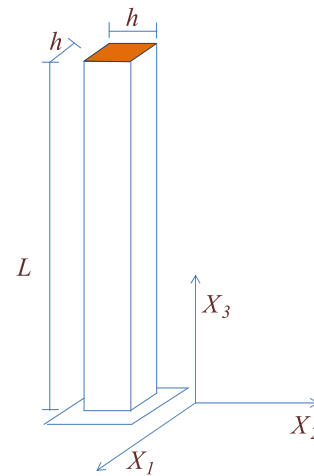


Fig. 12 Twisting column. $L = 6 \text{ m}$, $h = 1 \text{ m}$ and Dirichlet boundary conditions described in Eq. (128) for the static case in Sect. 7.4.1. $L = 15 \text{ m}$, $h = 1 \text{ m}$ and initial angular velocity Ω_0 for the dynamic case in Sect. 7.4.2

ent constitutive models. Note that the conjugate stress Σ_{H22} vanishes for the Neo-Hookean model (see Fig. 10c) whilst the conjugate stress Σ_J vanishes for the Saint-Venant model (see Fig. 10d).

Figure 11 displays the curve relating the absolute value of the displacement in the OX_3 axis of the free end of the beam versus the absolute value of the force F applied. For small strains, the constitutive response of the four constitutive models is, as expected, almost identical. However, the lack of the required coercivity and convexity requirements for the Saint Venant–Kirchhoff constitutive model leads to the unphysical buckling observed in Fig. 11.

7.4 Static and dynamic twisting column

7.4.1 Static twisting column

This example includes the twisting of a cantilever beam of length $L = 6 \text{ m}$ and a squared cross sectional area of side $a = 1 \text{ m}$ as shown in Fig. 12. The beam is clamped at $X_3 = 0$ and subjected to a torsion on its free end, namely $X_3 = 6$. The torsion at the free end is generated through Dirichlet boundary conditions as follows

$$(\mathbf{I} - \mathbf{D}_3 \otimes \mathbf{D}_3) \mathbf{x} = \theta \mathbf{D}_3 \times \mathbf{X}, \tag{128}$$

where θ is the angle of rotation. As can be observed, the section is not restricted to in-plane torsion and zero Neumann boundary conditions are imposed normal to the cross sectional area. The material properties of the beam are compatible with a shear modulus and Poisson ratio in the reference configuration of $E = 1 \text{ Pa}$ and $\nu = 0.35$, respec-

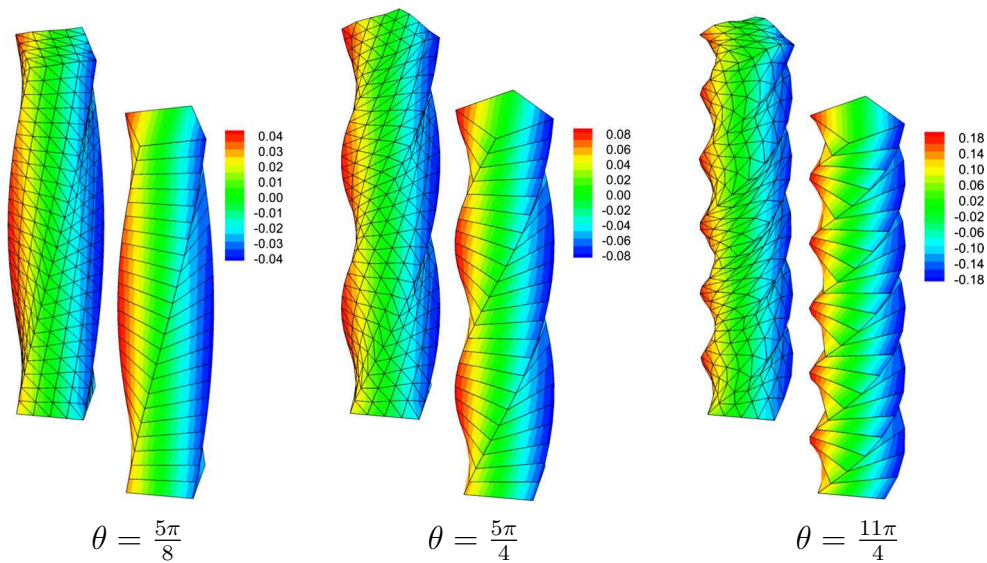


Fig. 13 Static twisting column. Comparison between beam and continuum representations. Contour plot of stress Σ_{F13} (N/m^2) for (left) a Hu–Washizu mixed formulation [12] with 3969×3 degrees of freedom associated to the displacement \mathbf{x} , $2304 \times 4 \times 9$ degrees of freedom for

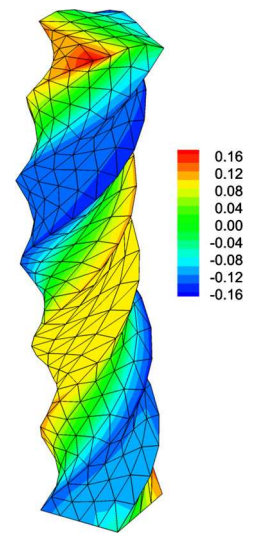
F, H, Σ_F and Σ_H , 2304 degrees of freedom associated to J and Σ_J ; (right) results displayed for the beam model with a discretisation of 24 linear elements (25×6 degrees of freedom associated to the spatial coordinates \mathbf{x}_0 and incremental rotations $\Delta\theta$)

tively. A similar example has been presented by the authors in previous reference [12].

The objective of this example is to benchmark the current beam formulation testing the capabilities of the beam representation against a very robust and precise continuum formulation. In particular, a mixed formulation associated to a special Hu–Washizu type of variational principle as presented in reference [12] is considered for that purpose. In this mixed formulation, the unknowns are displacements \mathbf{x} , the fibre, area and volume maps $\{F, H, J\}$ and their stress conjugates $\{\Sigma_F, \Sigma_H, \Sigma_J\}$. Regarding the selection of functional spaces: continuous quadratic interpolation of the displacement field \mathbf{x} , piecewise linear interpolation of the strain and stress fields F, H, Σ_F and Σ_H and piecewise constant interpolation of the Jacobian J and its associated stress conjugate Σ_J are considered.

The Mooney–Rivlin model defined in (12)–(13) has been considered. Despite the obvious differences between beam and continuum formulations (i.e. different kinematical description, different interpolation spaces), Fig. 13 shows reasonable agreement between both beam and continuum representations in terms of tangential stresses. This agreement is excellent for stages of the deformation in which the kinematical assumptions of the beam model are applicable. Hence, when warping of the cross section of the column is pronounced, as in Fig. 13c, the comparison is still reasonable but not as accurate, as expected. The warping of the cross sectional area shown by the continuum representation leads eventually to the buckled configuration represented in Fig. 14. Appropriate incorporation of warping effects into

Fig. 14 Static twisting column. Comparison between beam and continuum representations. Contour plot of stress Σ_{H32} (N/m^2) for a Hu–Washizu mixed formulation [12] with 3969×3 degrees of freedom associated to the displacement \mathbf{x} , $2304 \times 4 \times 9$ degrees of freedom for F, H, Σ_F and Σ_H , 2304 degrees of freedom associated to J and Σ_J



the kinematics of the beam [9] would enable to capture this nonlinear behaviour.

7.4.2 Dynamic twisting column

Finally, a beam with geometry depicted in Fig. 12 is considered. In this time dependent problem, an initial angular velocity $\Omega_0 = 35 \sin\left(\frac{\pi X_3}{2L}\right) \mathbf{D}_3$ compatible with the boundary conditions is prescribed. The material properties of the beam are compatible with a shear modulus and Poisson ratio in the reference configuration of $E = 0.0179$ GPa and $\nu = 0.3$, respectively. A similar example has been presented by the authors in reference [32].

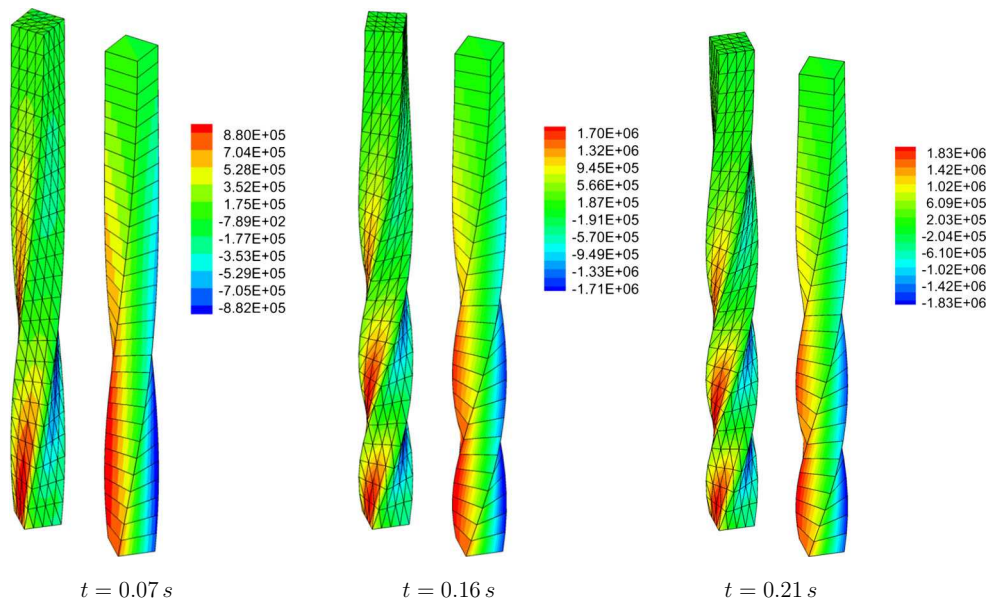


Fig. 15 Dynamic twisting column. Contour plot of stress σ_{31} (N/m^2) for (left) a Hu–Washizu mixed formulation [12] with 3969×3 degrees of freedom associated to the displacement \mathbf{x} , $2304 \times 4 \times 9$ degrees of freedom for \mathbf{F} , \mathbf{H} , $\Sigma_{\mathbf{F}}$ and $\Sigma_{\mathbf{H}}$, 2304 degrees of freedom associated to \mathbf{J} and $\Sigma_{\mathbf{J}}$. Generalised alpha method with $\rho_{\infty} = 1$ and $\Delta t = 10^{-3}$ s;

(right) results displayed for the beam model with a discretisation of 30 linear elements (31×6 degrees of freedom associated to the spatial coordinates \mathbf{x}_0 and incremental rotations $\Delta\theta$). Generalised alpha method with $\rho_{\infty} = 1$ and $\Delta t = 10^{-2}$ s

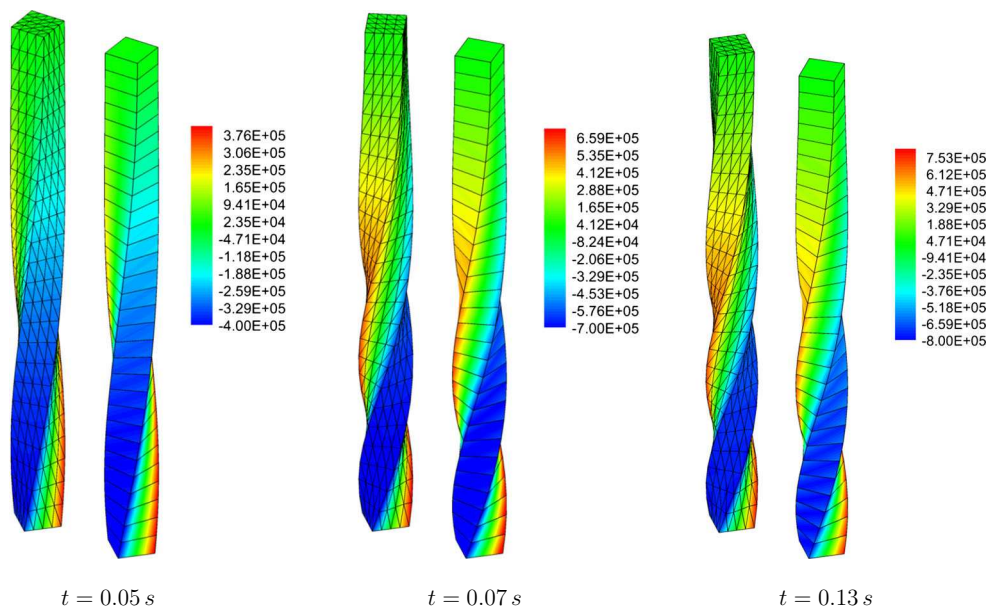


Fig. 16 Dynamic twisting column. Contour plot of stress $\Sigma_{\mathbf{H}31}$ (N/m^2) for (left) a Hu–Washizu mixed formulation [12] with 3969×3 degrees of freedom associated to the displacement \mathbf{x} , $2304 \times 4 \times 9$ degrees of freedom for \mathbf{F} , \mathbf{H} , $\Sigma_{\mathbf{F}}$ and $\Sigma_{\mathbf{H}}$, 2304 degrees of freedom associated to \mathbf{J} and $\Sigma_{\mathbf{J}}$. Generalised alpha method with

$\rho_{\infty} = 1$ and $\Delta t = 10^{-3}$ s; (right) results displayed for the beam model with a discretisation of 30 linear elements (31×6 degrees of freedom associated to the spatial coordinates \mathbf{x}_0 and incremental rotations $\Delta\theta$). Generalised alpha method with $\rho_{\infty} = 1$ and $\Delta t = 10^{-2}$ s

The objective of this section is to benchmark the current beam formulation with the continuum formulation described in Sect. 7.4.1 in a time dependent problem. For both the continuum and beam degenerate problems, a generalised alpha method time integrator is employed [44].

The Mooney–Rivlin model defined in (12)–(13) is considered with material parameters defined in (126). Figures 15 and 16 compare well in terms of displacements between the continuum and beam descriptions. In addition, Figs. 15 and 16 also show a good agreement between the contour

plots of the tangential stress σ_{31} (notice that axis in reference $\{OX_1, OX_2, OX_3\}$ and deformed $\{ox_1, ox_2, ox_3\}$ are coincident) and the conjugate stress $\Sigma_{H_{31}}$, respectively, for the continuum and beam models. However, as expected, the simplifications introduced in the kinematical description of the beam model (lack of contraction/expansion of the cross section of the beam) lead to different results between both continuum and beam models regarding normal stresses.

8 Concluding remarks

This paper has provided a novel variational and computational approach to formulate polyconvex large strain geometrically exact beam theory, extending the original ideas introduced by Bonet et al. [12]. In addition, three key novel contributions are incorporated in the present work. First, the deformation gradient, its cofactor and its determinant, namely $\{F, H, J\}$ are used for the first time as the main strain measures in the context of beam theory. Their respective work conjugates, namely $\{\Sigma_F, \Sigma_H, \Sigma_J\}$ also feature for the first time in a geometrically exact beam formulation. Moreover, their co-rotational strain $\{\bar{U}, \bar{W}, J\}$ and stress $\{\Sigma_{\bar{U}}, \Sigma_{\bar{W}}, \Sigma_J\}$ counterparts are also presented.

For the first time, the strain energy of a Mooney–Rivlin model has been presented in terms of the classical beam strain measures $\{\Gamma, K\}$ by taking advantage of the novel algebra associated to the tensor cross product operation presented in [12] and thoroughly detailed in [33]. Notice that this re-expression procedure can be generalised to more complex constitutive models (i.e. anisotropy, higher nonlinearities).

Finally, the authors have shown that polyconvexity of a continuum constitutive model defined via a continuum strain energy functional $W(F, H, J)$ implies convexity with respect to the classical beam strain measures of the equivalent beam strain energy functional $W_b(\Gamma, K)$, stating explicitly the relationship between alternative tangent operators.

Further extension of our work will include multi-physics electro-magneto-mechanical effects.

Acknowledgments The second author acknowledges the financial support received through “The Leverhulme Prize” awarded by The Leverhulme Trust, UK. The first three authors acknowledge the financial support provided by the Sêr Cymru National Research Network for Advanced Engineering and Materials.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix 1: Tensor cross product

One of the key elements of the framework proposed in [12] and extended to geometrically exact beam theory is the extension of the standard vector cross product to define the cross product between second order tensors and between tensors and vectors [34]. For instance, the left cross product of a vector v and a second order tensor A to give a second order tensor denoted $v \times A$ is defined so that when applied to a general vector w gives:

$$(v \times A) w = v \times (Aw); \quad (v \times A)_{ij} = \mathcal{E}_{ikl} v_k A_{lj} \quad (129)$$

where \mathcal{E}_{ikl} denote the standard third order alternating tensor components, repeated indices indicate summation and \times is the standard vector cross product. The effect of the above operation is to replace the columns of A by the cross products between v and the original columns of A . Similarly, the right cross product of a second order tensor A by a vector v to give a second order tensor denoted $A \times v$ is defined so that for every vector w the following relationship applies:

$$(A \times v) w = A (v \times w); \quad (A \times v)_{ij} = \mathcal{E}_{jkl} A_{ik} v_l. \quad (130)$$

The effect is now to replace the rows of A by the cross products of its original rows by v . Finally, the cross product of two second order tensors A and B to give a new second order tensor denoted $A \times B$ is defined so that for any arbitrary vectors v and w gives:

$$v \cdot (A \times B) w = (v \times A) : (B \times w); \quad (A \times B)_{ij} = \mathcal{E}_{ikl} \mathcal{E}_{jmn} A_{km} B_{ln}. \quad (131)$$

In this paper, the tensor cross product will be mostly applied between two-point tensors. For this purpose, the above definition can be particularised to second order two-point tensors or material tensors as:

$$(A \times B)_{iI} = \mathcal{E}_{ijk} \mathcal{E}_{IJK} A_{jJ} B_{kK}; \quad (A \times B)_{IJ} = \mathcal{E}_{IKL} \mathcal{E}_{JMN} A_{KM} B_{LN}. \quad (132)$$

When applied to a second order tensor A and a fourth order tensors \mathbb{H} , two possible operations are defined as:

$$(\mathbb{H} \times A)_{pPiI} = \mathcal{E}_{ijk} \mathcal{E}_{IJK} \mathbb{H}_{pPjJ} A_{kK}; \quad (A \times \mathbb{H})_{iIpP} = \mathcal{E}_{ijk} \mathcal{E}_{IJK} A_{jJ} \mathbb{H}_{kKpP}. \quad (133)$$

Moreover, the double application of the tensor cross product between a fourth order tensor and two second order tensors is associative, namely:

$$A \times \mathbb{H} \times B = (A \times \mathbb{H}) \times B = A \times (\mathbb{H} \times B). \quad (134)$$

Remark 2 It is easy to show using simply algebraic manipulations based on the permutation properties of \mathcal{E} or the fact that $\mathcal{E}_{ijk}\mathcal{E}_{klm} = \delta_{il}\delta_{jn} - \delta_{in}\delta_{jl}$, that the above tensor cross products satisfy the following properties (note that $\mathbf{v}, \mathbf{v}_1, \mathbf{v}_2, \mathbf{w}, \mathbf{w}_1$ and \mathbf{w}_2 denote arbitrary vectors and $\mathbf{A}, \mathbf{A}_1, \mathbf{A}_2, \mathbf{B}, \mathbf{B}_1, \mathbf{B}_2$ and \mathbf{C} are second order tensors):

$$\mathbf{A} \times \mathbf{B} = \mathbf{B} \times \mathbf{A} \tag{135}$$

$$(\mathbf{A} \times \mathbf{B})^T = \mathbf{A}^T \times \mathbf{B}^T \tag{136}$$

$$\mathbf{A} \times (\mathbf{B}_1 + \mathbf{B}_2) = \mathbf{A} \times \mathbf{B}_1 + \mathbf{A} \times \mathbf{B}_2 \tag{137}$$

$$\alpha (\mathbf{A} \times \mathbf{B}) = (\alpha \mathbf{A}) \times \mathbf{B} = \mathbf{A} \times (\alpha \mathbf{B}) \tag{138}$$

$$(\mathbf{A} \times \mathbf{B}) : \mathbf{C} = (\mathbf{B} \times \mathbf{C}) : \mathbf{A} = (\mathbf{A} \times \mathbf{C}) : \mathbf{B} \tag{139}$$

$$\mathbf{A} \times \mathbf{I} = (\text{tr} \mathbf{A}) \mathbf{I} - \mathbf{A}^T \tag{140}$$

$$\mathbf{I} \times \mathbf{I} = 2\mathbf{I} \tag{141}$$

$$(\mathbf{A} \times \mathbf{A}) : \mathbf{A} = 6 \det \mathbf{A} \tag{142}$$

$$\text{Cof} \mathbf{A} = \frac{1}{2} \mathbf{A} \times \mathbf{A} \tag{143}$$

$$(\mathbf{v}_1 \otimes \mathbf{v}_2) \times (\mathbf{w}_1 \otimes \mathbf{w}_2) = (\mathbf{v}_1 \times \mathbf{w}_1) \otimes (\mathbf{v}_2 \times \mathbf{w}_2) \tag{144}$$

$$\mathbf{v} \times (\mathbf{A} \times \mathbf{w}) = (\mathbf{v} \times \mathbf{A}) \times \mathbf{w} = \mathbf{v} \times \mathbf{A} \times \mathbf{w} \tag{145}$$

$$\mathbf{A} \times (\mathbf{v} \otimes \mathbf{w}) = -\mathbf{v} \times \mathbf{A} \times \mathbf{w} \tag{146}$$

$$(\mathbf{A} \times \mathbf{B}) (\mathbf{v} \times \mathbf{w}) = (\mathbf{A}\mathbf{v}) \times (\mathbf{B}\mathbf{w}) + (\mathbf{B}\mathbf{v}) \times (\mathbf{A}\mathbf{w}) \tag{147}$$

$$\begin{aligned} (\mathbf{A}_1 \times \mathbf{A}_2) (\mathbf{B}_1 \times \mathbf{B}_2) \\ = (\mathbf{A}_1 \mathbf{B}_1) \times (\mathbf{A}_2 \mathbf{B}_2) + (\mathbf{A}_1 \mathbf{B}_2) \times (\mathbf{A}_2 \mathbf{B}_1) \end{aligned} \tag{148}$$

$$(\mathbf{A}_1 \mathbf{B}) \times (\mathbf{A}_2 \mathbf{B}) = (\mathbf{A}_1 \times \mathbf{A}_2) \text{Cof} \mathbf{B} \tag{149}$$

Appendix 2: Transformation of invariants for the Mooney–Rivlin model in terms of the beam strain measures

The objective of this section is to re-express the strain energy functional for a Mooney–Rivlin constitutive model $W_{MR}(\mathbf{F}, \mathbf{H}, J)$, as presented in Eq. (12), in terms of the beam strain measures $\mathbf{\Gamma}$ and \mathbf{K} , namely $\bar{W}_{MR}(\mathbf{\Gamma}, \mathbf{K})$. Notice first, that an intermediate equivalent expression for (12) can be obtained in terms of the strain tensor $\bar{\mathbf{U}}$, its cofactor $\bar{\mathbf{W}}$ and its determinant J , namely $\hat{W}_{MR}(\bar{\mathbf{U}}, \bar{\mathbf{W}}, J)$.¹³ By making use of (38), it can be shown that

$$\bar{\mathbf{U}} : \bar{\mathbf{U}} = \mathbf{B} \cdot \mathbf{B} + 2\mathbf{B} \cdot \mathbf{D}_3 + 3. \tag{150}$$

¹³ It is easy to show from the use of the right extended polar decomposition theorem (38) that $\mathbf{F} : \mathbf{F} = \bar{\mathbf{U}} : \bar{\mathbf{U}}, \mathbf{H} : \mathbf{H} = \bar{\mathbf{W}} : \bar{\mathbf{W}}$.

Similarly, it can be shown that

$$\begin{aligned} \bar{\mathbf{W}} : \bar{\mathbf{W}} &= [(\mathbf{B} \otimes \mathbf{D}_3) \times \mathbf{I} + \mathbf{I}] : [(\mathbf{B} \otimes \mathbf{D}_3) \times \mathbf{I} + \mathbf{I}] \\ &= \underbrace{[(\mathbf{B} \otimes \mathbf{D}_3) \times \mathbf{I}] : [(\mathbf{B} \otimes \mathbf{D}_3) \times \mathbf{I}]}_{\xi} \\ &\quad + 2 \underbrace{[(\mathbf{B} \otimes \mathbf{D}_3) \times \mathbf{I}] : \mathbf{I}}_{2\mathbf{B} \cdot \mathbf{D}_3} + 3. \end{aligned} \tag{151}$$

where use of properties (139), (141) and (144) has been made. The ξ term above (151) can be further manipulated as follows

$$\begin{aligned} \xi &= \text{tr}([(\mathbf{B} \otimes \mathbf{D}_3) \times \mathbf{I}]^T [(\mathbf{B} \otimes \mathbf{D}_3) \times \mathbf{I}]) \\ &= \text{tr}([(\mathbf{D}_3 \otimes \mathbf{B}) \times \mathbf{I}] [(\mathbf{B} \otimes \mathbf{D}_3) \times \mathbf{I}]) \\ &= \text{tr}([(\mathbf{D}_3 \otimes \mathbf{B}) (\mathbf{B} \otimes \mathbf{D}_3)] \times \mathbf{I} + (\mathbf{D}_3 \otimes \mathbf{B}) (\mathbf{B} \otimes \mathbf{D}_3)) \\ &= \underbrace{[(\mathbf{D}_3 \otimes \mathbf{B}) (\mathbf{B} \otimes \mathbf{D}_3)] \times \mathbf{I} : \mathbf{I}}_{2[(\mathbf{D}_3 \otimes \mathbf{B}) (\mathbf{B} \otimes \mathbf{D}_3)] : \mathbf{I}} \\ &\quad + \underbrace{[(\mathbf{D}_3 \otimes \mathbf{B}) (\mathbf{B} \otimes \mathbf{D}_3)] : \mathbf{I}}_{(\mathbf{D}_3 \times \mathbf{B}) \otimes (\mathbf{B} \times \mathbf{D}_3)} \\ &= 2 (\mathbf{B} \cdot \mathbf{B}) \underbrace{(\mathbf{D}_3 \cdot \mathbf{D}_3)}_1 - (\mathbf{D}_3 \times \mathbf{B}) \cdot (\mathbf{D}_3 \times \mathbf{B}) \\ &= 2 (\mathbf{B} \cdot \mathbf{B}) - (\mathbf{D}_3 \times \mathbf{B}) \cdot (\mathbf{D}_3 \times \mathbf{B}) \end{aligned} \tag{152}$$

where use of properties (136), (148), (139), (141) and (144) has been made (in that order). Substitution of (152) into (151) leads to

$$\bar{\mathbf{W}} : \bar{\mathbf{W}} = 2\mathbf{B} \cdot \mathbf{B} - (\mathbf{D}_3 \times \mathbf{B}) \cdot (\mathbf{D}_3 \times \mathbf{B}) + 4\mathbf{B} \cdot \mathbf{D}_3 + 3. \tag{153}$$

Finally, substitution of (150), (153) and (64) into (12) leads to

$$\begin{aligned} \bar{W}(\mathbf{B}(\mathbf{\Gamma}, \mathbf{K})) &= \alpha (\mathbf{B} \cdot \mathbf{B} + 2\mathbf{B} \cdot \mathbf{D}_3 + 3) \\ &\quad + \beta (2\mathbf{B} \cdot \mathbf{B} - (\mathbf{D}_3 \times \mathbf{B}) \cdot (\mathbf{D}_3 \times \mathbf{B}) \\ &\quad + 4\mathbf{B} \cdot \mathbf{D}_3 + 3) + f(\mathbf{B} \cdot \mathbf{D}_3 + 1), \end{aligned} \tag{154}$$

where $\mathbf{B} = \mathbf{\Gamma} + (\mathbf{K} \times \bar{\mathbf{X}})$. The stress vector Σ_B can then be obtained following Eq. (70) as

$$\begin{aligned} \Sigma_B &= 2\alpha (\mathbf{B} + \mathbf{D}_3) + \beta \underbrace{\left[4\mathbf{B} - \frac{\partial \xi}{\partial \mathbf{B}} + 4\mathbf{D}_3 \right]}_x \\ &\quad + \left[\lambda (\mathbf{B} \cdot \mathbf{D}_3) - \frac{2(\alpha + 2\beta)}{(\mathbf{B} \cdot \mathbf{D}_3 + 1)} \right] \mathbf{D}_3. \end{aligned} \tag{155}$$

where the partial derivative term contained within the second term on the right hand side of above Eq. (155) can be further

manipulated to give

$$\begin{aligned} \frac{\partial \xi}{\partial \mathbf{B}} &= 2(\mathbf{D}_3 \times \mathbf{B}) \cdot \frac{\partial}{\partial \mathbf{B}} (\mathbf{D}_3 \times \mathbf{B}) \\ &= 2(\mathbf{D}_3 \times \mathbf{B}) \cdot (\mathbf{D}_3 \times \mathbf{I})\mathbf{B} \\ &= 2(\mathbf{D}_3 \times \mathbf{I})^T (\mathbf{D}_3 \times \mathbf{I})\mathbf{B} \\ &= -2(\mathbf{I} \times \mathbf{D}_3)(\mathbf{I} \times \mathbf{D}_3)\mathbf{B}, \end{aligned} \tag{156}$$

where use of the relationships $(\mathbf{I} \times \mathbf{D}_3)^T = (\mathbf{D}_3 \times \mathbf{I})^T = -(\mathbf{I} \times \mathbf{D}_3) = -(\mathbf{D}_3 \times \mathbf{I})$ has been made for the last step in above Eq. (156).

Furthermore, use of the identity $\mathbf{I} \times \mathbf{D}_3 = -\mathbf{D}_1 \otimes \mathbf{D}_2 + \mathbf{D}_2 \otimes \mathbf{D}_1$ and the subsequent relationship $(\mathbf{I} \times \mathbf{D}_3)(\mathbf{I} \times \mathbf{D}_3) = -\mathbf{D}_1 \otimes \mathbf{D}_1 - \mathbf{D}_2 \otimes \mathbf{D}_2$, enables to write the second term on the right hand side of above Eq. (155) as

$$\begin{aligned} \chi &= 2\mathbf{B} + 2\mathbf{B} - \frac{\partial \xi}{\partial \mathbf{B}} + 4\mathbf{D}_3 \\ &= 2\mathbf{B} + 2\mathbf{B} + 2(\mathbf{I} \times \mathbf{D}_3)(\mathbf{I} \times \mathbf{D}_3)\mathbf{B} + 4\mathbf{D}_3 \\ &= 2\mathbf{B} + 2\mathbf{B} - 2(\mathbf{D}_1 \otimes \mathbf{D}_1 + \mathbf{D}_2 \otimes \mathbf{D}_2)\mathbf{B} + 4\mathbf{D}_3 \\ &= 2\mathbf{B} + (\mathbf{D}_3 \otimes \mathbf{D}_3)\mathbf{B} + 4\mathbf{D}_3 \\ &= 2(\mathbf{I} + \mathbf{D}_3 \otimes \mathbf{D}_3)\mathbf{B} + 4\mathbf{D}_3. \end{aligned} \tag{157}$$

Substitution of Eq. (157) in above Eq. (155) yields

$$\begin{aligned} \Sigma_{\mathbf{B}} &= 2\alpha (\mathbf{B} + \mathbf{D}_3) + 2\beta [(\mathbf{I} + \mathbf{D}_3 \otimes \mathbf{D}_3) \mathbf{B} + 2\mathbf{D}_3] \\ &\quad + \left[\lambda (\mathbf{B} \cdot \mathbf{D}_3) - \frac{2(\alpha + 2\beta)}{(\mathbf{B} \cdot \mathbf{D}_3 + 1)} \right] \mathbf{D}_3. \end{aligned} \tag{158}$$

The Hessian operator $[\mathbb{H}_{\bar{W}}]$ (75) is then obtained as

$$\begin{aligned} [\mathbb{H}_{\bar{W}}] &= 2\alpha \mathbf{I} + 2\beta (\mathbf{I} + \mathbf{D}_3 \otimes \mathbf{D}_3) \\ &\quad + \left[\lambda + \frac{2(\alpha + 2\beta)}{(\mathbf{B} \cdot \mathbf{D}_3 + 1)^2} \right] \mathbf{D}_3 \otimes \mathbf{D}_3. \end{aligned} \tag{159}$$

References

- Betsch P, Steinmann P (2002) Frame-indifferent beam finite elements based upon the geometrically exact beam theory. *Int J Numer Methods Eng* 54:1775–1788
- Crisfield MA, Jelenić G (1999) Objectivity of strain measures in the geometrically exact three-dimensional beam theory and its finite-element implementation. *Proc R Soc Lond Ser A* 455:1125–1147
- Eugster S, Hesch C, Betsch P, Glocker Ch (2014) Director-based beam finite elements relying on the geometrically exact beam theory formulated in skew coordinates. *Int J Numer Methods Eng* 97:111–129
- Jelenić G, Crisfield M (1999) Geometrically exact 3D beam theory: implementation of a strain-invariant finite element for statics and dynamics. *Comput Methods Appl Mech Eng* 171:141–171
- Jelenić G, Saje M (1995) A kinematically exact space finite strain beam model-finite element formulation by generalized virtual work principle. *Comput Methods Appl Mech Eng* 120:131–161
- Romero I, Armero F (2002) An objective finite element approximation of the kinematics of geometrically exact rods and its use in the formulation of an energy-momentum conserving scheme in dynamics. *Int J Numer Methods Eng* 54:1683–1716
- Simo J, Vu-Quoc L (1986) A three-dimensional finite-strain rod model. Part II: computational aspects. *Comput Methods Appl Mech Eng* 58:79–116
- Auricchio F, Carotenuto P, Reali A (2008) On the geometrically exact beam model: a consistent, effective and simple derivation from three-dimensional finite-elasticity. *Int J Solids Struct* 45:4766–4781
- Eugster SR (2014) On the foundations of continuum mechanics and its application to beam theories. Ph.D. thesis, ETH-Zürich
- Gruttmann F, Sauer R, Wagner W (1998) A geometrical nonlinear eccentric 3D-beam element with arbitrary cross-sections. *Comput Methods Appl Mech Eng* 160:383–400
- Balzani D, Gruttmann F, Schröder J (2008) Analysis of thin shells using anisotropic polyconvex energy densities. *Comput Methods Appl Mech Eng* 197:1015–1032
- Bonet J, Gil AJ, Ortigosa R (2015) A computational framework for polyconvex large strain elasticity. *Comput Methods Appl Mech Eng* 283:1061–1094
- Simo J (1985) A finite strain beam formulation. The three-dimensional dynamic problem. Part I. *Comput Methods Appl Mech Eng* 49:55–70
- Bonet J, Gil AJ, Wood RD, Said R, Curtis RV (2006) Simulating superplastic forming. *Comput Methods Appl Mech Eng* 195:6580–6603
- Gil AJ, Curtis RV, Bonet J, Coward T (2010) Finite element superplastic forming (FE-SPF) of patient-specific maxillofacial prostheses. *Int J Numer Methods Biomed Eng* 26:139–155
- Bonet J, Wood RD (2008) *Nonlinear continuum mechanics for finite element analysis*, 2nd edn. Cambridge University Press, Cambridge
- Gonzalez O, Stuart AM (2008) *A first course in continuum mechanics*. Cambridge University Press, Cambridge
- Marsden JE, Hughes TJR (1994) *Mathematical foundations of elasticity*. Dover Civil and Mechanical Engineering
- Ball JM (1977) Convexity conditions and existence theorems in nonlinear elasticity. *Arch Ration Mech Anal* 63:337–403
- Ball JM (1983) Energy-minimising configurations in nonlinear elasticity. In *Proceedings of the international congress of mathematicians, Warsaw*
- Ball JM, Murat F (1984) W^1, p -quasiconvexity and variational problems for multiple integrals. *J Funct Anal* 58:225–253
- Ball JM (2002) Some open problems in elasticity. In: Newton P, Holmes P, Weinstein A (eds) *Geometry, mechanics and dynamics*. Springer, New York, pp 3–59
- Zhang K (1992) A construction of quasiconvex functions with linear growth at infinity. *Annali della Scuola Normale Superiore di Pisa, Classe di Scienze* 97:313–326
- Dacorogna B (2008) *Direct methods in the calculus of variations*. Springer, Berlin
- Ciarlet P (2010) Existence theorems in intrinsic nonlinear elasticity. *Journal des mathématiques pures et appliquées* 94:229–243
- Schröder J, Neff P (2003) Invariant formulation of hyperelastic transverse isotropy based on polyconvex free energy functions. *Int J Solids Struct* 40:401–445
- Schröder J, Neff P, Ebbing V (2008) Anisotropic polyconvex energies on the basis of crystallographic motivated structural tensors. *J Mech Phys Solids* 56:3486–3506
- Schröder J, Neff P, Ebbing V (2010) Polyconvex energies for trigonal, tetragonal and cubic symmetry groups. In: *IUTAM-symposium on variational concepts with applications to the mechanics of materials: proceedings of the IUTAM-symposium*

29. Schröder J, Wriggers P, Balzani D (2011) A new mixed finite element based on different approximations of the minors of deformation tensors. *Comput Methods Appl Mech Eng* 200:3583–3600
30. Schröder J (2010) Anisotropic polyconvex energies, In: Schröder J, Neff P (eds) Poly-, quasi- and rank-one convexity in applied mechanics. Volume 516 of CISM courses and lectures, volume 516 of CISM courses and lectures. Springer, Vienna
31. Kambouchev N, Radovitzky R, Fernández J (2006) Anisotropic materials which can be modeled by polyconvex strain energy functions. In: 47th AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics, and materials conference, volume AIAA 2006-2250
32. Bonet J, Gil AJ, Lee CH, Aguirre M, Ortigosa R (2015) A first order hyperbolic framework for large strain computational solid dynamics. Part I: total Lagrangian isothermal elasticity. *Comput Methods Appl Mech Eng* 283:689–732
33. Bonet J, Gil AJ, Ortigosa R (2015) On a tensor cross product based formulation of large strain solid mechanics. *Int J Solids Struct* (Under review)
34. de Boer R (1982) *Vektor- und Tensorrechnung für Ingenieure*. Springer, Berlin
35. Antmann S (2005) *Nonlinear problems of elasticity*. Springer, New York
36. Ciarlet PG (1988) *Mathematical elasticity. Volume 1: three dimensional elasticity*. Elsevier, Amsterdam
37. Schröder J, Neff P, Balzani D (2005) A variational approach for materially stable anisotropic hyperelasticity. *Int J Solids Struct* 42:4352–4371
38. Betsch P, Sängner N (2009) On the use of geometrically exact shells in a conserving framework for flexible multibody dynamics. *Comput Methods Appl Mech Eng* 198:1609–1630
39. Poya R, Gil AJ, Ledger PD (2015) A computational framework for the analysis of linear piezoelectric beams using hp-FEM. *Comput Struct* 152:155–172
40. Stuelpnagel J (1964) On the parametrization of the three-dimensional rotation group. *SIAM Rev* 6:422–430
41. Auricchio F, de Veiga LB, Lovadina C, Reali A (2005) A stability study of some mixed finite elements for large deformation elasticity problems. *Comput Methods Appl Mech Eng* 194:1075–1092
42. Betsch P, Sängner N (2013) On the consistent formulation of torques in a rotationless framework for multibody dynamics [Special Issue IASS-IACM-2012]. *Comput Struct* 127:29–38
43. Ibrahimbegović A, Frey F, Kőzar I (1995) Computational aspects of vector-like parametrization of three-dimensional finite rotations. *Int J Numer Methods Eng* 38:3653–3673
44. Hughes TJR (2000) *The finite element method: linear static and dynamic finite element analysis*. Dover Publications, New York