

Mathematical modeling of a Cosserat method in finite-strain holonomic plasticity

Thomas Blesgen · Ada Amendola* (*Corresponding Author, Orcid ID:
0000-0002-2562-881X)

Submitted to Springer

Abstract This article deals with the mathematical derivation and the validation over benchmark examples of a numerical method for the solution of a finite-strain holonomic (rate-independent) Cosserat plasticity problem for materials, possibly with microstructure. Two improvements are made in contrast to earlier approaches: First, the micro-rotations are parameterized with the help of an Euler-Rodrigues formula related to quaternions. Secondly, as main result, a novel two-pass preconditioning scheme for searching the energy-minimizing solutions based on the limited memory Broyden-Fletcher-Goldstein-Shanno quasi-Newton method is proposed that consists of a predictor step and a corrector-iteration. After outlining the necessary adaptations to the model, numerical simulations compare the performance and efficiency of the new and the old algorithm. The proposed numerical model can be effectively employed for studying the mechanical response of complicated materials featuring large size effects.

Keywords Micropolar materials · Crystal plasticity · Quaternions · Cosserat theory · Numerical simulations · Preconditioning

Thomas Blesgen
Department of Applied Mathematics, Bingen University of Applied Sciences, Bingen, Germany

Ada Amendola
Department of Civil Engineering, University of Salerno, Fisciano, Italy
E-mail: adaamendola1@unisa.it

1 Introduction

In recent years, the scientific interest towards sophisticated and heterogeneous materials featuring multiple internal length scales has grown significantly, mainly due to the possibility of playing with the internal microstructure of these materials to model and engineer structures that exhibit properties not found in conventional materials (refer, e.g., to [1,2] and references therein). Such materials include cellular solids, fibrous and particle composites, biological materials, robots, and also building-scale systems made of masonry structures [3,4,5,6,7,8]. The mechanical modeling of these materials and structures calls for the introduction of degrees of freedom that are not accounted for in classical continuum mechanics, typically rotation of points (or micro-rotations) and couple stresses [9,10,11]. A viable continuum description of such phenomena is provided by the micropolar theories of Cosserat continua [12], which have been intensively applied since their introduction in 1909 to a variety of different problems in solid and structural mechanics, fluid dynamics, liquid crystals, granular materials, powders, etc. (cf. [13,14,15] for an overview). Particularly interesting is the Cosserat modeling of chiral honeycomb lattices with bending-dominated behavior whose mechanical response cannot be accurately described by classical continuum theories due to large size effects, [3]. So far, physical models of these exciting materials have been fabricated through additive manufacturing (AM) technologies in polymeric materials and have been described through Cosserat elasticity, [3]. The numerical model

presented in this work allows for simulating the response of ductile versions of such metamaterials, assuming radial loading and holonomic plasticity, [16,17,18], which are, e.g., fabricated via AM techniques manual assembling methods employing metallic materials, [19,20,21].

Since the Cosserat model of a micropolar material induces sensitivity to the microrotation strain gradient, such generalized continua are endowed with an internal length scale such that localization zones have a finite width. The Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm is a well-known quasi-Newton method where instead of storing the full Hessian matrix H (a big matrix for large dimensions) an approximation is computed by the sum of two rank-one matrices. In the limited-memory (L-BFGS) variant, [22,23], the approximation to H is constructed from a small number of vectors by a rank-one update formula, see Eqn. (32) below. The resulting algorithm is still considered the state-of-the-art method when huge systems of equations with a very large number of unknowns need to get solved.

In [24], a L-BFGS algorithm is developed for the solution of a finite-strain rate-independent Cosserat model of finite plasticity. Therein, the *elastic Cosserat micro-rotations* R_e are parameterized by a vector $\alpha = (\alpha_1, \alpha_2, \alpha_3) \in \mathbb{R}^3$ of Euler angles,

$$\begin{aligned} \tilde{R}_e(\alpha) &:= R_3(\alpha_3)R_2(\alpha_2)R_1(\alpha_1) \\ &:= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha_3 & \sin \alpha_3 \\ 0 & -\sin \alpha_3 & \cos \alpha_3 \end{pmatrix} \begin{pmatrix} \cos \alpha_2 & 0 & -\sin \alpha_2 \\ 0 & 1 & 0 \\ \sin \alpha_2 & 0 & \cos \alpha_2 \end{pmatrix} \\ &\quad \begin{pmatrix} \cos \alpha_1 & \sin \alpha_1 & 0 \\ -\sin \alpha_1 & \cos \alpha_1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned} \quad (1)$$

Two main criticisms of the approach in [24] are eminent. The first is that Euler angles are not well-suited to parameterize the rotation group $\mathcal{SO}(3)$ and have several shortcomings. Especially the parameterization may degenerate and become non-unique.

In other areas of mechanics such as unmanned aerial vehicle (UAV) control, quaternion-based descriptions have demonstrated their superior performance, see [25,26]. There-

fore, in this article, the alternative parameterization

$$R_e(q) := \begin{pmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1q_2 - q_0q_3) & 2(q_1q_3 + q_0q_2) \\ 2(q_1q_2 + q_0q_3) & q_0^2 - q_1^2 + q_2^2 - q_3^2 & 2(q_2q_3 - q_0q_1) \\ 2(q_1q_3 - q_0q_2) & 2(q_2q_3 + q_0q_1) & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{pmatrix} \quad (2)$$

is studied which is based on an Euler-Rodrigues vector $q = (q_1, q_2, q_3, q_4)$ defined on the unit sphere

$$S^3 := \left\{ q = (q_1, q_2, q_3, q_4) \in \mathbb{R}^4 \mid |q|^2 = 1 \right\}.$$

Formula (2) goes back to historical work by L. Euler in 1775, [27]. The approach was independently reinvented by Rodrigues in 1840, [28]. As was already discovered early, it can also be derived from quaternion theory, [29].

The second major criticism to [24] is that the quasi-Newton iteration may get stuck in a local minimum of the mechanical energy without finding the global minimizer. Preconditioning of the numerical scheme may help to speed up the code and correctly compute the global minimizer. While there is vast literature on preconditioning in general, only a few articles deal with preconditioning of the L-BFGS-method, [30,31,32,33], especially when directly related to energy minimization, [34].

The first goal of this paper is to study the implications of (1), (2) on the finite-strain Cosserat algorithm, assuming radial loading and holonomic-type plasticity [16,17,18]. Secondly, as main result, a two-step preconditioning strategy of the L-BFGS algorithm is proposed that consists of a predictor step followed by a corrector iteration for solving the time-discrete problem. This two-pass approach effectively defines a non-linear preconditioning strategy.

This article is organized in the following way. In Section 2, the finite-strain Cosserat model is reviewed. Section 3 derives background theory on a quaternion-based Cosserat theory. Section 4 revisits the L-BFGS update scheme and derives the aforementioned preconditioning method. Section 5 performs various numerical tests, followed by a discussion of the results and an outlook. At the end of the paper, a complete list of symbols with explanations can be found. The generalization of the present approach to more general cases of gradient-type plasticity [44,45,46,47,48] is addressed to future work.

2 The finite-strain Cosserat model of holonomic plastic materials with microstructure

The deformation mapping of the current material from the reference configuration $\Omega \subset \mathbb{R}^3$ to the deformed state Ω_t is described by a diffeomorphism $\varphi \in \mathcal{G}l^+(3)$, for times $t \geq 0$. Throughout, Ω is assumed a smooth Lipschitz domain.

Assuming radial loading and holonomic-type plasticity [16, 17], the fundamental relationship of the Cosserat approach is the multiplicative decomposition

$$F = F_e F_p = R_e U_e F_p \quad (3)$$

of the deformation tensor $F := D\varphi$, where F_e, F_p are the elastic and the plastic deformation tensors, $U_e = R_e^t D\varphi F_p^{-1} \in \mathcal{G}l(3)$ is the stretching component, and

$$R_e \in \mathcal{S}\mathcal{O}(3) := \{R \in \mathcal{G}l(3) \mid \det(R) = 1, R^t R = \mathbb{I}\}$$

are the micro-rotations. In (3), U_e need not be symmetric and positive definite, i.e. the decomposition $F_e = R_e U_e$ is in general *not* the polar decomposition.

We fundamentally assume that the mechanical energy depends on the elastic part F_e of the deformation, only. With κ denoting the density of the (geometrically necessary) dislocations, it follows by frame indifference that the stored mechanical energy is of the form, [35],

$$W(F_e, \kappa) = W_{\text{st}}(U_e) + W_c(K_e) + V(\kappa),$$

where $K_e = (R_e^t \partial_{x_l} R_e)_{1 \leq l \leq 3}$ is the (right) curvature tensor, W_{st} denotes the stretching energy, W_c the curvature energy due to bending and torsion of the material, and V the energy of stored dislocations. For these functionals we make the ansatz, cf. [14, 36],

$$W_{\text{st}}(U_e) := \mu \|\text{sym} U_e - \mathbb{I}\|^2 + \mu_c \|\text{skw}(U_e - \mathbb{I})\|^2 + \frac{\lambda}{2} |\text{tr}(U_e - \mathbb{I})|^2, \quad (4)$$

$$W_c(q) := \mu_2 \|K_e(q)\|^2 = \mu_2 \|\nabla R_e(q)\|^2 = \mu_2 \sum_{l=1}^3 \|\partial_l R_e(q)\|^2, \quad (5)$$

$$V(\kappa) := \rho \kappa^2. \quad (6)$$

In (4), (5), $\mu_2 := \frac{\mu}{2} L_c^2$ with the internal length scale $L_c > 0$, the Cosserat couple modulus $\mu_c > 0$, and $\lambda > 0, \mu > 0$ are the Lamé parameters; $\partial_l := \frac{\partial}{\partial x_l}, 1 \leq l \leq 3$ for short; $\rho > 0$ is a constant. In (4), $\text{sym}(A) := \frac{1}{2}(A + A^t), \text{skw}(A) := \frac{1}{2}(A - A^t)$ denote the symmetric and skew-symmetric part of a tensor A , respectively; $\text{tr}(A) := \sum_i A_{ii}$ is the trace operator, $\|A\| := \sqrt{\text{tr}(A^t A)}$ the Frobenius matrix norm; $\mathbf{u} \cdot \mathbf{v} := \sum_{i=1}^3 u_i v_i$ is the inner product in \mathbb{R}^3, \mathbb{I} the real 3×3 identity matrix. For $A, B \in \mathbb{R}^{3 \times 3}, A : B := \text{tr}(A^t B) = \sum_{i,j=1}^3 A_{ij} B_{ij}$ denotes the inner product in $\mathbb{R}^{3 \times 3}$. For a general introduction to tensor calculus in plasticity, we recommend [37, 38].

Applying ideas from [39], see also [40], the time evolution of the deformed material can be computed by a sequence of minimization problems for the mechanical energy. If $h > 0$ is a fixed time step, for given (F_p^0, κ^0) of the previous time step, the values of $(\varphi, R_e, F_p, \kappa)$ need to be calculated at time $t + h$. Let $P := F_p^{-1}$ be the plastic backstress, and $P^0 := (F_p^0)^{-1}$. Then the approximations

$$d_t^h(F_p) := \frac{\mathbb{I} - P^{-1} P^0}{h}, \quad \partial_t^h \kappa := \frac{\kappa - \kappa^0}{h}$$

of the time derivatives are used. Other forms of time integrators are discussed in [41]. We obtain the minimization problem

$$\begin{aligned} \mathcal{E}(\varphi, q, F_p, \kappa) := & \int_{\Omega} W_{\text{st}}(U_e(\varphi, q, F_p)) + W_c(K_e(q)) + V(\kappa) \\ & + \Lambda (1 - |q|^2)^2 - f_{\text{ext}} \cdot \varphi - M_{\text{ext}} : R_e(q) \\ & + h Q^*(d_t^h(F_p), \partial_t^h \kappa) dx - \int_{\Gamma_t} \mathbf{t} \cdot \varphi dS \\ & - \int_{\Gamma_c} M_t : R_e(q) dS \rightarrow \min \end{aligned} \quad (7)$$

subject to the initial and Dirichlet boundary conditions

$$\begin{aligned} \varphi(x, 0) = x, \quad \kappa(\cdot, 0) = \kappa^0 & \quad \text{in } \Omega, \\ \varphi = g_D, \quad q = q_D & \quad \text{on } \Gamma_D \end{aligned} \quad (8)$$

with fixed Dirichlet boundary data q_D and g_D . As is typical of a variational theory, the functional \mathcal{E} represents the total mechanical energy of the system minus the ground state energy. In (7), (8), $\Gamma_D \subset \partial\Omega$ is that part of the boundary where Dirichlet conditions are applied; Γ_t is the part of the boundary where traction boundary conditions apply; $\Gamma_c \subset \partial\Omega$ the boundary where surface couples are applied. It must hold

$\Gamma_D \cap \Gamma_t = \emptyset$, $\Gamma_D \cap \Gamma_C = \emptyset$. For simplicity, we assume from now on $\Gamma_D = \partial\Omega$ and $\Gamma_t = \Gamma_C = \emptyset$.

In (7), the term $\Lambda(|q|^2 - 1)^2$ ensures the validity of the constraint $|q| = 1$ in Ω , where $\Lambda > 0$ is a constant. By $f_{\text{ext}} = f_{\text{ext}}(t)$, $M_{\text{ext}} = M_{\text{ext}}(t)$, the external volume force density and external volume couples are specified, respectively. The term $hQ^*(d_t^h(F_p), \partial_t^h \kappa)$ is the dissipated mechanical energy in the time interval from t to $t+h$. It is the Legendre-Fenchel dual

$$Q^*(F_p, \kappa) := \sup_{(X, \xi)} \{X : F_p + \xi \kappa - Q(X, \xi)\} \quad (9)$$

of the plastic potential

$$Q(X, \xi) := \begin{cases} 0, & \text{for } Y(X, \xi) \leq 0, \\ \infty, & \text{else,} \end{cases}$$

where $Y \leq 0$ is the yield function with $Y = 0$ indicating plastic flow. In case of the van Mises condition,

$$Y(\sigma, \xi) := \|\text{dev sym } \sigma\| - \sigma_Y - \xi$$

with $\text{dev } \sigma := \sigma - \frac{1}{3}\mathbb{I}$ the deviatoric part of σ . The above formulas establish a rate-independent theory where the material responds immediately (infinitely fast) to applied forces.

As a result of plastic deformation due to structural changes within the material like the increase of immobilized dislocations inside the lattice structure, hardening occurs, [42, 43]. It is assumed throughout the text that plastic deformations only occur along one a-priori given material-dependent single-slip system, specified by a normal vector \mathbf{n} and a slip vector \mathbf{m} with $|\mathbf{m}| = |\mathbf{n}| = 1$ and $\mathbf{m} \cdot \mathbf{n} = 0$, see [44].

The real parameter γ determines the plastic slip and the plastic deformation tensor by

$$F_p = F_p(\gamma) := \mathbb{I} + \gamma \mathbf{m} \otimes \mathbf{n}. \quad (10)$$

Formula (10) is obtained from $\dot{F}_p = \dot{\gamma} \mathbf{m} \otimes \mathbf{n}$ by integration from the initial state $F_p(t=0) = \mathbb{I}$ to time t .

In contrast to [36], we restrict here to the case of one slip system, by leaving the multislip case for future work.

As can be checked, [45], the dissipated energy satisfies the relationship

$$Q^*(\dot{A}, \dot{\kappa}) = \begin{cases} \sigma_Y |\dot{\gamma}|, & \text{if } \dot{A} = \dot{\gamma} \mathbf{m} \otimes \mathbf{n} \text{ and } |\dot{\gamma}| + \dot{\kappa} \leq 0, \\ \infty, & \text{else.} \end{cases} \quad (11)$$

As is well known, plastic deformations always occur on the boundary of the set of feasible deformations. Consequently, see [36], the constraint $|\gamma - \gamma^0| + \kappa - \kappa^0 \leq 0$ appearing in the definition of Q^* has to be satisfied with equality, leading to

$$\kappa = -|\gamma - \gamma^0| + \kappa^0, \quad (12)$$

which allows us to define κ as a function of γ , γ^0 , and κ^0 . Plugging in (12) in $V(\kappa)$ and dropping an inconsequential constant $\rho(\kappa^0)^2$, we end up with the optimization problem

$$\begin{aligned} \mathcal{E}(\varphi, q, \gamma) := & \int_{\Omega} \left[W_{\text{st}}(R_c^t(q) D\varphi F_p(\gamma)^{-1}) + W_c(q) \right. \\ & + \Lambda(|q|^2 - 1)^2 - f_{\text{ext}} \cdot \varphi - M_{\text{ext}} : R_c(q) \\ & \left. + \rho(\gamma - \gamma^0)^2 + |\gamma - \gamma^0|(\sigma_Y - 2\rho\kappa^0) \right] dx \rightarrow \min \end{aligned} \quad (13)$$

subject to the initial and boundary conditions (8) for $\Gamma_D = \partial\Omega$.

The functional \mathcal{E} in (13) coincides with the one in [49] except for the new term $\Lambda(|q|^2 - 1)^2$ and the parameterization (2) instead of (1) for the micro-rotations.

For a fixed discrete time step $h > 0$ and known (γ^0, κ^0) at time t , the new (φ, q, γ) representing values at time $t+h$ are calculated from (13). Finally, the new κ is computed from (12) and (γ, κ) become the initial values of the next time step.

If the material is initially free of dislocations, $\kappa(\cdot, 0) = 0$, the hardening law (12) implies $\kappa(t+h) \leq \kappa(t) \leq 0$ for all times t . Hence, $-2\rho\kappa^0 \geq 0$ in (13) represents the increase of the yield stress σ_Y due to stored dislocations.

3 An application of the Euler-Rodrigues formula

Following the classical notation in [50,51], let

$$\begin{aligned} \mathbb{H} &:= \text{span}_{\mathbb{R}}\{1, \mathbf{i}, \mathbf{j}, \mathbf{k}\} \\ &= \{q = q_0 + q_1 \mathbf{i} + q_2 \mathbf{j} + q_3 \mathbf{k} \mid q_0, q_1, q_2, q_3 \in \mathbb{R}\} \end{aligned}$$

denote the *space of quaternions*, where the quaternion imaginary units satisfy $\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1$. Let

$$\mathbb{H}_p := \{q = q_0 + q_1 \mathbf{i} + q_2 \mathbf{j} + q_3 \mathbf{k} \in \mathbb{H} \mid q_0 = 0\}$$

be the space of *pure quaternions* and

$$q = q_0 + \hat{q} := q_0 + q_1 \mathbf{i} + q_2 \mathbf{j} + q_3 \mathbf{k}. \quad (14)$$

The set \mathbb{H} is equipped with the multiplication (for $p, q \in \mathbb{H}$)

$$pq := p_0q_0 - \hat{p} \cdot \hat{q} + p_0\hat{q} + q_0\hat{p} + \hat{p} \times \hat{q}, \quad (15)$$

where $\hat{p} \cdot \hat{q} := p_1q_1 + p_2q_2 + p_3q_3$ specifies as above the inner product and $\hat{p} \times \hat{q}$ the vector product of \mathbb{R}^3 , respectively. In general, $pq \neq qp$, so \mathbb{H} is an associative, non-commutative algebra. Let $\bar{q} := q_0 - \hat{q}$ be the *conjugate* of q and

$$|q| := (q\bar{q})^{1/2} = (\bar{q}q)^{1/2} = (q_0^2 + q_1^2 + q_2^2 + q_3^2)^{1/2} \quad (16)$$

be the *modulus* of q . By Formula (15), $q \in \mathbb{H}^* := \mathbb{H} \setminus \{0\}$ possesses the multiplicative inverse $q^{-1} = \frac{\bar{q}}{|q|^2}$. Let

$$\text{so}(3) := \{\omega \in \mathbb{R}^{3 \times 3} \mid \omega^t = -\omega\}$$

be the Lie algebra of $\mathcal{SO}(3)$. The *alternating skew tensor* $\varepsilon : \mathbb{H}_p \rightarrow \text{so}(3)$ is defined by

$$\varepsilon(\hat{q}) := \begin{pmatrix} 0 & -q_3 & q_2 \\ q_3 & 0 & -q_1 \\ -q_2 & q_1 & 0 \end{pmatrix}. \quad (17)$$

Evidently,

$$\varepsilon(\hat{q})v = \hat{q} \times v \quad \text{for } v \in \mathbb{R}^3 \simeq \mathbb{H}_p. \quad (18)$$

By direct inspection, it is straightforward to verify that for every $q \in S^3$

$$R_e(q)v := qv\bar{q} \quad \text{for } v \in \mathbb{R}^3 \simeq \mathbb{H}_p \quad (19)$$

defines a rotation in $\mathcal{SO}(3)$. Using (15), this leads to

$$R_e(q) = (2q_0^2 - |q|^2)\mathbb{I} + 2\hat{q} \otimes \hat{q} + 2q_0\varepsilon(\hat{q}). \quad (20)$$

Plugging in the above definitions, this coincides with Formula (2).

The mapping R_e thus introduced has the properties

$$R_e(1) = \mathbb{I}, \quad R_e(\bar{q}) = R_e(q)^t, \quad R_e(pq) = R_e(p)R_e(q)$$

and is therefore an algebra-homomorphism. It is a double cover of $\mathcal{SO}(3)$, especially it is non-unique, since

$$R_e(q) = R_e(-q) \quad \text{for } q \in S^3. \quad (21)$$

In comparison, the parameterization (1) breaks down for $\alpha_2 = \frac{\pi}{2}$, in which case α_1 and α_3 denote a rotation around the same axis. In summary, both (2) and (1) set up rivaling charts on the manifold $\mathcal{SO}(3)$ which have certain disadvantages when used globally.

Formula (2) can be used to interpolate between rotations and allows to introduce a distance in $\mathcal{SO}(3)$, see, e.g., [52]. This is a prerequisite to studying surface energies between grains or particles of different orientations, [53].

For $x \in \mathbb{R}^3$ and a quaternion field $q = q(x)$, the m -th material *curvature vector* or *Darboux vector* is given by

$$K_e^m(q) := 2\bar{q}\partial_m q \in \mathbb{H}_p, \quad 1 \leq m \leq 3. \quad (22)$$

The following lemma computes the derivatives of $R_e(q)$ and $K_e(q)$ in \mathbb{H} with $|q| = 1$.

Lemma 1 (Lie Derivatives of R_e and K_e^m) *Let $q = q(x) : \mathbb{R}^3 \rightarrow S^3$ and $1 \leq l, m \leq 3$. Then*

$$\partial_l R_e(q) = R_e(q)\varepsilon(K_e^l(q)), \quad (23)$$

$$\partial_l K_e^m(q) = 2\bar{q}[\partial_l \partial_m q - \partial_l q \bar{q} \partial_m q]. \quad (24)$$

Proof An elementary proof of (23) can be found in [54], Chapter 11. The following proof is a modification of an argument in [55]. Let $v \in \mathbb{R}^3 \simeq \mathbb{H}_p$ and let $w \in \mathbb{R}^3$ denote various

changing vectors. Then it holds

$$\begin{aligned}
\varepsilon(K_e^l(q))v &= \varepsilon(2\bar{q}\partial_l q)v && \text{by Eqn. (22)} \\
&= 2\bar{q}\partial_l q \times v && \text{by Eqn. (18)} \\
&= 2\bar{q}\partial_l qv && \text{by Eqn. (15)} \\
&= \widehat{2\bar{q}\partial_l qv} && \text{by Eqn. (14)} \\
&= \bar{q}\partial_l qv - \overline{\bar{q}\partial_l qv} && \text{since } w - \bar{w} = 2\widehat{w} \\
&= \bar{q}\partial_l qv + v\partial_l \bar{q}q && \text{since } \bar{v} = -v \\
&= \bar{q}(\partial_l qv\bar{q} + qv\partial_l \bar{q})q && \text{since } \bar{q}q = |q|^2 = 1 \\
&= \bar{q}(\partial_l(qv\bar{q}))q && \text{since } \partial_l v = 0 \\
&= \bar{q}(\partial_l R_e(q)v)q && \text{by Eqn. (19)} \\
&= R_e(q)^t \partial_l R_e(q)v && \text{since } (R_e(q)w)^t = \bar{q}wq.
\end{aligned}$$

As this is true for every $v \in \mathbb{R}^3 \simeq \mathbb{H}_p$, this shows

$$\varepsilon(K_e^l(q)) = R_e(q)^t \partial_l R_e(q).$$

Multiplication with $R_e(q)$ from the left yields (23).

In order to show (24), multiplying (22) with q from the left yields

$$2\partial_m q = qK_e^m(q).$$

Consequently,

$$2\partial_l \partial_m q = \partial_l q K_e^m(q) + q \partial_l K_e^m(q)$$

or equivalently

$$q \partial_l K_e^m(q) = 2\partial_l \partial_m q - \partial_l q K_e^m(q).$$

Multiplication of this identity with \bar{q} from the left leads to

$$\partial_l K_e^m(q) = 2\bar{q}\partial_l \partial_m q - \bar{q}\partial_l q K_e^m(q).$$

With (22), this shows (24). \square

Applying the results of Lemma 1 to W_c , it holds by Eqns. (23) and (17),

$$\begin{aligned}
W_c(q) &= \mu_2 \sum_{l=1}^3 \|\partial_l R_e(q)\|^2 = \mu_2 \sum_{l=1}^3 \|R_e(q)\varepsilon(K_e^l(q))\|^2 \\
&= \mu_2 \sum_{l=1}^3 \|\varepsilon(K_e^l(q))\|^2 \\
&= 2\mu_2 \sum_{l=1}^3 \left[(K_{e1}^l(q))^2 + (K_{e2}^l(q))^2 + (K_{e3}^l(q))^2 \right] \\
&= 2\mu_2 \sum_{l=1}^3 |K_e^l(q)|^2. \tag{25}
\end{aligned}$$

For the first derivative, using (22) and (24), this results in

$$\begin{aligned}
\partial_m W_c(q) &= 4\mu_2 \sum_{l=1}^3 \widehat{\partial_m K_e^l(q)} \cdot \widehat{K_e^l(q)} \\
&= 16\mu_2 \sum_{l=1}^3 [\bar{q}(\partial_m \partial_l q - \partial_m q \bar{q} \partial_l q)] \cdot [\bar{q} \partial_l q]. \tag{26}
\end{aligned}$$

4 Preconditioning

When implementing the L-BFGS method for the Cosserat problem (13), frequently situations are encountered where the algorithm requires many iterations to converge. Also it may happen that the iteration is stopped before a correct minimizer has been reached. Therefore, in this section, certain modifications of the L-BFGS algorithm are discussed. It is noteworthy that this does not only increase the speed of the code, but may be an essential step to correctly compute the minimizers.

Starting point is the minimization problem (13) written as

$$\mathcal{E}(x) \rightarrow \min, \tag{27}$$

where $x \in \mathbb{R}^D$ corresponds to a spatial discretization of (φ, q, γ) by finite elements or finite differences. The L-BFGS algorithm is a quasi-Newton method and constructs a minimizing sequence $(x_k)_{k \in \mathbb{N}} \subset \mathbb{R}^D$ by setting

$$\begin{aligned}
d_k &:= -H_k \nabla \mathcal{E}(x_k), \\
x_{k+1} &:= x_k + \alpha d_k. \tag{28}
\end{aligned}$$

Here, H_k approximates the inverse Hessian $(D^2 \mathcal{E}(x_k))^{-1}$ and is constructed from rank-one updates, d_k is a descent direc-

tion, and $\alpha \in \mathbb{R}$ is a parameter computed by a linesearch algorithm. The iteration (28) stops if for chosen small $\varepsilon_0 > 0$

$$|\nabla \mathcal{E}(x_k)| < \varepsilon_0 \max\{1, |x_k|\}. \quad (29)$$

Letting

$$\begin{aligned} s_{k-1} &:= x_k - x_{k-1}, \\ y_{k-1} &:= g_k - g_{k-1} := \nabla \mathcal{E}(x_k) - \nabla \mathcal{E}(x_{k-1}), \end{aligned}$$

the *BFGS-update* is given by

$$\begin{aligned} H_k &= H_{k-1} + \left(\frac{y_{k-1}^t H_{k-1} y_{k-1}}{y_{k-1}^t s_{k-1}} + 1 \right) \frac{s_{k-1} s_{k-1}^t}{y_{k-1}^t s_{k-1}} \\ &\quad - \frac{1}{y_{k-1}^t s_{k-1}} \left[s_{k-1} y_{k-1}^t H_{k-1} + H_{k-1} y_{k-1} s_{k-1}^t \right] \quad (30) \\ &= (\mathbb{I} - \rho_{k-1} s_{k-1} y_{k-1}^t) H_{k-1} (\mathbb{I} - \rho_{k-1} y_{k-1} s_{k-1}^t) \\ &\quad + \rho_{k-1} s_{k-1} s_{k-1}^t \\ &=: V_{k-1}^t H_{k-1} V_{k-1} + \rho_{k-1} s_{k-1} s_{k-1}^t \\ &= (V_{k-1}^t \dots V_0^t) H_0 (V_0 \dots V_{k-1}) \quad (31) \\ &\quad + \sum_{l=1}^{k-1} (V_{k-1}^t \dots V_l^t) s_{l-1} s_{l-1}^t (V_l \dots V_{k-1}) + \rho_{k-1} s_{k-1} s_{k-1}^t \end{aligned}$$

with $\rho_l := \frac{1}{y_l^t s_l}$ and $V_l := \mathbb{I} - \rho_l y_l s_l^t$.

In the limited-memory variant of (30), the matrices H_k are not stored explicitly. Instead, given a small number $m \in \mathbb{N}$ and vectors s_0, \dots, s_{m-1} , y_0, \dots, y_{m-1} , the multiplication

$$H_k \nabla \mathcal{E}(x_k)$$

is carried out by the two-loop iteration, see [22],[56],

$$\begin{aligned} g_k &:= \nabla \mathcal{E}(x_k) \\ \text{FOR } i &= m-1, \dots, 0 \\ \alpha_i &:= \rho_i s_i^t g_k \\ g_k &:= g_k - \alpha_i y_i \\ r_k &:= H_k^0 g_k \quad (32) \\ \text{FOR } i &= 0, \dots, m-1 \\ \beta_k &:= \rho_i y_i^t r_k \\ r_k &:= r_k + (\alpha_i - \beta_k) s_i \\ H_k \nabla \mathcal{E}(x_k) &:= r_k. \end{aligned}$$

The first FOR-loop of the above scheme for determining $r_k = H_k g_k$ computes and stores $(V_l \dots V_{m-1}) g_k$ for $0 \leq l \leq m-1$. After carrying out the multiplication (32), the second FOR-loop then computes (31).

The above scheme is considered one of the most effective update formulas of numerical analysis and requires only $\mathcal{O}(mD)$ operations. The parameter m is usually chosen as $3 \leq m \leq 7$, see [57], and increasing m further does not improve the quality of the update.

In (32), for each iteration step k , one is free to pick H_k^0 . In the original implementation of the algorithm, in order to reduce the condition numbers of H_k , the diagonal is scaled with the *Cholesky factor* δ_k , [58],

$$H_k^0 = \delta_k \mathbb{I}, \quad \delta_k := \frac{s_{k-1}^t y_{k-1}}{y_{k-1}^t y_{k-1}}. \quad (33)$$

Instead, another matrix or non-linear scheme such as a fixed point iteration may be used in place of H_k^0 in (32) such that ideally, $H_k^0 \sim D^2 \mathcal{E}(x_k)$.

In order to find an efficient preconditioning method, it is helpful to study the particular features of the Cosserat functional \mathcal{E} . From physical insight and numerical investigations, it is evident that the hardest part in solving (13) is the computation of the optimal rotations, i.e. finding the quaternion field q . Therefore, the following two-step strategy for the solution of one time-step is effective:

Step 1 (Predictor): Fix (φ, γ) .

Solve with the L-BFGS-method the optimization problem

$$\mathcal{E}_{\varphi, \gamma}(q) \rightarrow \min.$$

Step 2 (Corrector): Solve with the L-BFGS-method the full problem (27). Pick the solution q_{opt} of Step 1 as initial values for q .

Typically, the solution of Step 1 is very fast in comparison to Step 2 since far less variables need to be optimized and the complicated dependence of q on (φ, γ) is eliminated. Step 1 provides a reasonable approximation to the solution of the full problem (27). In the conducted tests, the combined numerical costs for solving Step 1 and Step 2 turned out significantly lower than for solving the original minimization problem directly in one pass with the un-precon-

ditioned L-BFGS method. This is discussed below in more detail.

In Step 1, (φ, γ) is fixed with data of the previous time step. At the first time step, γ is loaded with the initial values γ^0 and φ is initialized with an extension of the boundary values g_D in $\overline{\Omega}$ that satisfies the Cauchy-Born rule.

Both Step 1 and Step 2 are preconditioned. In Step 1, a special preconditioning matrix Z replacing H_k^0 is chosen that resembles the common discretization of the Laplace operator on structured grids. Step 2 is preconditioned with the final converged matrix H_k computed in Step 1. As this matrix is obtained from a L-BFGS-procedure, it has a data-sparse representation by vectors $(s_0, y_0), \dots, (s_{m-1}, y_{m-1})$.

In order to derive the preconditioning-matrix Z of Step 1, recall the computation of the total curvature energy by finite differences in 3D

$$\int_{\Omega} W_c(q) dx \approx \frac{w}{8} \sum_{i=0}^{d_1} \sum_{j=0}^{d_2} \sum_{k=0}^{d_3} N_{ijk} W_c(q(y_{ijk})) \quad (34)$$

used in [24], where $N_{ijk} \in \mathbb{N}$ are numerical weights derived from a Newton-Cotes formula, $y_{ijk} \in \overline{\Omega}$ are points of the numerical mesh with equal spacings

$$\eta_1 := \frac{L_1}{d_1}, \quad \eta_2 := \frac{L_2}{d_2}, \quad \eta_3 := \frac{L_3}{d_3}, \quad (35)$$

$\Omega = (0, L_1) \times (0, L_2) \times (0, L_3)$ is assumed, $d_l \in \mathbb{N}$ is the number of discretization points in direction l , $l = 1, 2, 3$, and $w := \eta_1 \eta_2 \eta_3$ is an integration factor.

Since for the preconditioning matrix only a reasonably good approximation of the second derivative is needed, in the following $N_{ijk} = 8$ is assumed (the value of N_{ijk} in $\overline{\Omega} \setminus \partial\Omega$). First, let

$$W_c(q) := 2\mu_2 |\partial_x q|^2.$$

Then, by a straightforward computation, for fixed subscripts $0 \leq I \leq d_1$, $0 \leq J \leq d_2$, $0 \leq K \leq d_3$ and fixed component $0 \leq b \leq 3$ of q ,

$$\begin{aligned} \frac{\partial}{\partial q_{IJK}^b} \int_{\Omega} W_c(q(x)) dx &\approx \frac{w\mu_2}{2\eta_1^2} \frac{\partial}{\partial q_{IJK}^b} \sum_i \left(q_{i+1,J,K}^b - q_{i-1,J,K}^b \right)^2 \\ &= \frac{w\mu_2}{\eta_1^2} \sum_i \left(q_{i+1,J,K}^b - q_{i-1,J,K}^b \right) \left(\delta_{i+1,J} - \delta_{i-1,J} \right) \\ &= \frac{w\mu_2}{\eta_1^2} \left(-q_{I-2,J,K}^b + 2q_{I,J,K}^b - q_{I+2,J,K}^b \right) \end{aligned} \quad (36)$$

with the short-hand notation $q_{ijk} \equiv q(y_{ijk})$. In the same way the second derivative

$$\frac{\partial^2}{(\partial q_{IJK}^b)^2} \int_{\Omega} W_c(q(x)) dx$$

can be computed. Let $D_1 \hat{=} (I, J, K)$ be the line index and $D_2 \hat{=} (I_2, J_2, K_2)$ be the column index of the 2nd derivative matrix Z . Then, from (36),

$$Z_{D_1, D_2} := \frac{w\mu_2}{\eta_1^2} \begin{cases} +2, & \text{if } D_1 = D_2, \\ -1, & \text{if } |I - I_2| = 2, \\ 0, & \text{otherwise.} \end{cases}$$

Likewise, if W_c is given by (38), then up to a pre-factor, Z is 2 on the diagonal, equals -1 if $|I - I_1| = 2$ or $|J - J_1| = 2$ or $|K - K_1| = 2$, and is 0 otherwise.

In the implementation, Z is not stored explicitly. The multiplication Zg for a vector $g \in \mathbb{R}^D$ is carried out by exploiting the band structure of Z .

5 Numerical tests

Subsequently, different algorithms for the solution of (13) are investigated. First, the following general remarks are in place.

Remark 1 Following [24], for small $\varepsilon > 0$, in (13) the modulus $|\cdot|$ is replaced by

$$r_{\varepsilon}(x) := \begin{cases} x, & x > \varepsilon, \\ x^2/\varepsilon, & -\varepsilon \leq x \leq +\varepsilon, \\ -x, & x < -\varepsilon. \end{cases}$$

This removes the singularity at the origin and allows the application of Newton's method.

Remark 2 Since the quasi-Newton method applied in this article computes variations of q that are not in S^3 , the parameterization (2) is not applicable unmodified in the numerical code. Instead, the mapping

$$\widehat{R}_e(q) := \frac{1}{|q|^2} \begin{pmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1 q_2 - q_0 q_3) & 2(q_1 q_3 + q_0 q_2) \\ 2(q_1 q_2 + q_0 q_3) & q_0^2 - q_1^2 + q_2^2 - q_3^2 & 2(q_2 q_3 - q_0 q_1) \\ 2(q_1 q_3 - q_0 q_2) & 2(q_2 q_3 + q_0 q_1) & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{pmatrix} \quad (37)$$

is used which is defined for all $q \in \mathbb{R}^4 \setminus \{0\}$. When minimizing \mathcal{E}_ε , due to the term $\Lambda(|q|^2 - 1)^2$, the computed optimal q lies (approximately) in S^3 .

Remark 3 All plastic deformations considered in this section satisfy $\det(F_p) = 1$. Hence the plastic deformations preserve the volume.

5.1 Comparison of the parameterizations by Euler angles and Euler-Rodrigues formula

The quaternion-based algorithm, due to its additional component in the representation of R_e , requires about 14% more computer memory. Table 1 has the exact figures for different spatial resolutions. Let *Algorithm 1* denote the algorithm of [24] which is based on finite differences in 3D, the L-BFGS method, Euler angles, and the curvature energy (5), *Algorithm 2a* be the analogous quaternion-based algorithm that solves (13); finally *Algorithm 2b* be identical to Alg. 2a, but with the simplified curvature energy

$$\widetilde{W}_c(q) := 2\mu_2 \sum_{l=1}^3 |\partial_l q|^2. \quad (38)$$

This choice is motivated by the fact that Euler angles permit to write (5) as

$$W_c(\alpha) = 2\mu_2 \sum_{l=1}^2 |\partial_l \alpha|^2, \quad (39)$$

see Eqn. (25) or [49]. As the numerical costs for computing (38) and (39) are very similar, this permits an unbiased comparison of the two parameterizations.

In [49], a class of 3D analytic solutions to (13) is calculated for an ultra-soft material with $\sigma_Y = \rho = 0$ subject to the boundary conditions

$$D\varphi(t) = \mathbb{I} + \beta(t)m \otimes n \quad \text{on } \partial\Omega. \quad (40)$$

This represents a simple shear problem for prescribed values $\beta(t) \in \mathbb{R}$. The Cauchy-Born rule is valid here and (40) is satisfied in $\overline{\Omega}$.

The above test constitutes a benchmark problem. The following simulation compares the performance and speed

of convergence for both Alg. 1 and Alg. 2. The stopping criterion is (29) with $\varepsilon_0 := 10^{-7}$.

Parameters (Benchmark test):

$$\begin{aligned} \Omega &= (0, 1)^3, \quad t \in [0, 1], \quad \mu = 10^4, \quad \mu_c = 2 \cdot 10^4, \\ \mu_2 &:= \mu \frac{L_c^2}{2} = 100, \quad \lambda = 10^3, \quad \rho = \sigma_Y = f_{\text{ext}} = 0, \quad M_{\text{ext}} = \mathbf{0}, \\ \mathbf{m} &= (1, 0, 0)^t, \quad \mathbf{n} = (0, 1, 0)^t, \quad \beta(t) = 0.25 * t, \quad h = 0.1, \\ \varepsilon &= 10^{-4}, \quad \Lambda = 20, \quad q_D = (1, 0, 0, 0). \end{aligned}$$

Initial values: $\varphi_0 \equiv \mathbb{I}$, $\kappa^0 = \gamma^0 = 0$ in Ω .

Results: $\gamma(\cdot, t) = \beta(t)$, $R_e = U_e = \mathbb{I}$, $W_{\text{st}} = W_c = 0$ in $\overline{\Omega}$,

$\varphi(x, t) = (x_1 + \beta(t)x_2, x_2, x_3)$ in Ω , i.e. the validity of the Cauchy-Born rule.

Table 2 summarizes the required number of iterations and computation times for all variants. The stopping criterion is (29) with $\varepsilon_0 := 10^{-7}$. As can be seen, Alg. 2b requires about 20% less iterations, Alg. 2a about 10% less iterations than Alg. 1. This behavior is typical. In our numerical tests, the quaternion-based algorithms reveal superior convergence. Table 3 illustrates the deviation of the numerical solution from the constraint $|q| = 1$.

5.2 The effect of preconditioning

This section conducts numerical tests of the preconditioning strategy presented in Section 4. While for large values of the stop parameter ε_0 the code usually converges after a small number of iterations, preconditioning becomes mandatory when ε_0 is chosen small. Fig 1 demonstrates that reducing ε_0 may go along with an exponential increase of the number of iterations. Simultaneously, fine properties of the physical solution may be missed when ε_0 is set too large, cf. also Table 3. The following bending problem of a 3D rod, see [36, Eqn. (27)], serves as a test problem. For given $\beta(t)$ as in (40), φ at $\partial\Omega$ is prescribed by

$$g_D^{\text{bend}}(x_1, x_2, x_3, t) := \begin{pmatrix} x_1 \\ x_2 + \frac{2L_1}{\pi} \left[\sin\left(\frac{3\pi}{2} + \frac{\pi}{2} \frac{x_1}{L_1}\right) + 1 \right] \beta(t) \\ x_3 \end{pmatrix}. \quad (41)$$

In order to determine the boundary conditions on q , let

$$R_D^{\text{bend}} := \text{polar}(Dg_D^{\text{bend}} F_p^{-1}),$$

where $\text{polar}(\cdot)$ is the polar decomposition, computed with the algorithm in [59]. Then set

$$q = q_D^{\text{bend}} \text{ on } \partial\Omega,$$

where $R_e(q_D^{\text{bend}}) = R_D^{\text{bend}}$ and q_D^{bend} is computed from R_D^{bend} with the algorithm in [60].

Parameters (Bending problem):

$$\Omega = (0, 5) \times (0, 1) \times (0, 2), \quad t \in [0, 1], \quad \lambda = \mu = 0.025,$$

$$\mu_c = 0.4, \quad \mu_2 = 0.02; \quad \rho = \sigma_Y = f_{\text{ext}} = 0, \quad M_{\text{ext}} = \mathbf{0},$$

$$\mathbf{m} = (1, 0, 0)^t, \quad \mathbf{n} = (0, 1, 0)^t, \quad \beta(t) = 0.25 * t, \quad h = 0.1,$$

$$\varepsilon = 10^{-4}, \quad \Lambda = 20, \quad W_c(q) = 2\mu \sum_{l=1}^3 |\partial_l q|^2,$$

$$\text{Initial values: } \varphi_0 \equiv \mathbb{I}, \quad \kappa^0 = \gamma^0 = 0 \text{ in } \Omega.$$

$$\text{Boundary values: } \varphi = g_D^{\text{bend}}, \quad q = q_D^{\text{bend}} \text{ on } \partial\Omega.$$

$$\text{Results: } \gamma(x, t) = \sin\left(\frac{\pi x_1}{2 L_1}\right) \beta(t), \quad q = (1, 0, 0, q_4), \quad U_e = \mathbb{I}, \\ W_{\text{st}} \equiv 0, \quad \varphi = g_D^{\text{bend}} \text{ in } \overline{\Omega}.$$

Table 4 compares the numerical costs for solving the first time step of the bending problem with the original L-BFGS-algorithm (where H_k^0 is defined by (33)) and with the preconditioned two-step L-BFGS-algorithm of Section 4 when $\varepsilon_0 := 10^{-11}$. Again, this behavior is typical. In our numerical tests, the two-step preconditioner leads to a significant speed-up, often accompanied with increased precision.

6 Discussion

In this paper, a parameterization by quaternions is applied to a strongly non-linear finite-strain Cosserat model of plastic materials, possibly with microstructure. Despite increased memory requirements, in the conducted numerical tests the quaternion-based algorithm needed less iterations and converged faster. As main result, a novel two-level preconditioning scheme is proposed that exploits the physical properties of the Cosserat model. The preconditioner solves a simpli-

fied problem for R_e with fixed (φ, γ) which represents the most complicated step in computing a global minimizer of \mathcal{E}_ε . Note that the degrees of freedom in the micro-rotations are responsible for the occurrence of a large number of local minima. With this reasonably good guess for R_e , the preconditioned algorithm is eventually able to succeed to a global minimizer. The preconditioning strategy is compatible with the L-BFGS update scheme and can be regarded as a non-linear preconditioning technique. Numerical tests support that this scheme significantly reduces the algorithmic costs and is essential to computing the physical solution when high precision is required. Similar two-step L-BFGS-algorithms may also be applicable to other classes of problems that depend in an un-symmetrical way on its variables. Fig. 2 documents a further important numerical feature: Since the energy landscape of \mathcal{E}_ε consists of many flat plateaus, the L-BFGS-scheme stagnates for a long time with each step only slightly decreasing \mathcal{E}_ε . It is unknown if and when an iteration significantly decreases the energy. When ε_0 in (29) is taken too large, the algorithm may wrongly interpret this stagnation as convergence. It would be desirable to have analytic results on the choice of ε_0 , or even better an algorithm that is capable to prevent this stagnation period. Finally, it may be desirable to develop a specialized L-BFGS algorithm that restricts the variations of the functional w.r.t. certain variables to the tangent space.

We address the above generalizations and enrichments of the numerical model presented in this study, together with the analysis of the more general case of gradient-type plasticity and hysteretic response under general loading, [44, 45, 46], to future work. Additional future research lines will be devoted to applying the current Cosserat model to bending-dominated lattices with plastic behavior which exhibit arbitrarily large size effects and consist, e.g., of cubical modules/particles connected by deformable links or Sarrus linkages tessellating triangular lattice structures [3]. Physical models of such systems will be fabricated through AM in ductile materials, [19]. These mockups will be laboratory-tested in order to validate the accuracy of numerical simulations and to demonstrate the presence of size effects that cannot be described through classical continuum or homogenization theories. Recent results revealing metamaterial-type

behaviors of the above systems, which are related to auxetic response [3] and/or high strength effects induced by bending and twisting of the material, will be extended to the plastic regime on accounting for a ductile response of the background material.

Figure captions

Parameterizations	Resolution (d_1, d_2, d_3)	No. Unknowns	No. Nodes	Memory (MB)
Euler angles	(32, 32, 32)	214683	35937	1.64
Euler angles	(64, 64, 64)	1774907	274625	13.5
Euler angles	(128, 128, 128)	14436987	2146689	110
Euler angles	(256, 256, 256)	116462843	16974593	889
Quaternions	(32, 32, 32)	244476	35937	1.86
Quaternions	(64, 64, 64)	2024956	274625	15.4
Quaternions	(128, 128, 128)	16485372	2146689	126
Quaternions	(256, 256, 256)	133044220	16974593	1015

Table 1: Comparison between the parameterizations by Euler angles (1) and by quaternions (2). 'No. Unknowns' is the total number of unknowns in the discrete model, 'Memory' the total memory for storing the data in case of 64 bit precision, 'No Nodes' the total number of discretization points in the finite difference mesh.

Algorithm	$10 \times 10 \times 10$	$30 \times 30 \times 30$	$50 \times 50 \times 50$	$70 \times 70 \times 70$
Alg. 1	359 (0.91s)	1022 (119s)	1228 (723s)	1383 (2040s)
Alg. 2a	332 (0.93s)	919 (107s)	1150 (650s)	1239 (1973s)
Alg. 2b	299 (0.89s)	812 (81s)	891 (541s)	1119 (1370s)

Table 2: Averaged number of iterations for Alg. 1 and Alg. 2a/b for different spatial resolutions $d_1 \times d_2 \times d_3$ and the benchmark problem over the time interval $[0, 1]$. Averaged computation times are in brackets.

Acknowledgements Part of this article was written while TB visited the Hausdorff Research Institute for Mathematics (HIM), University of Bonn, in 2019. This visit was supported by the HIM. TB gratefully acknowledges both this support and the hospitality of HIM. AA gratefully acknowledges financial support from the Italian Ministry of Education, University and Research (MIUR) under the 'Departments of Excellence' grant L.232/2016.

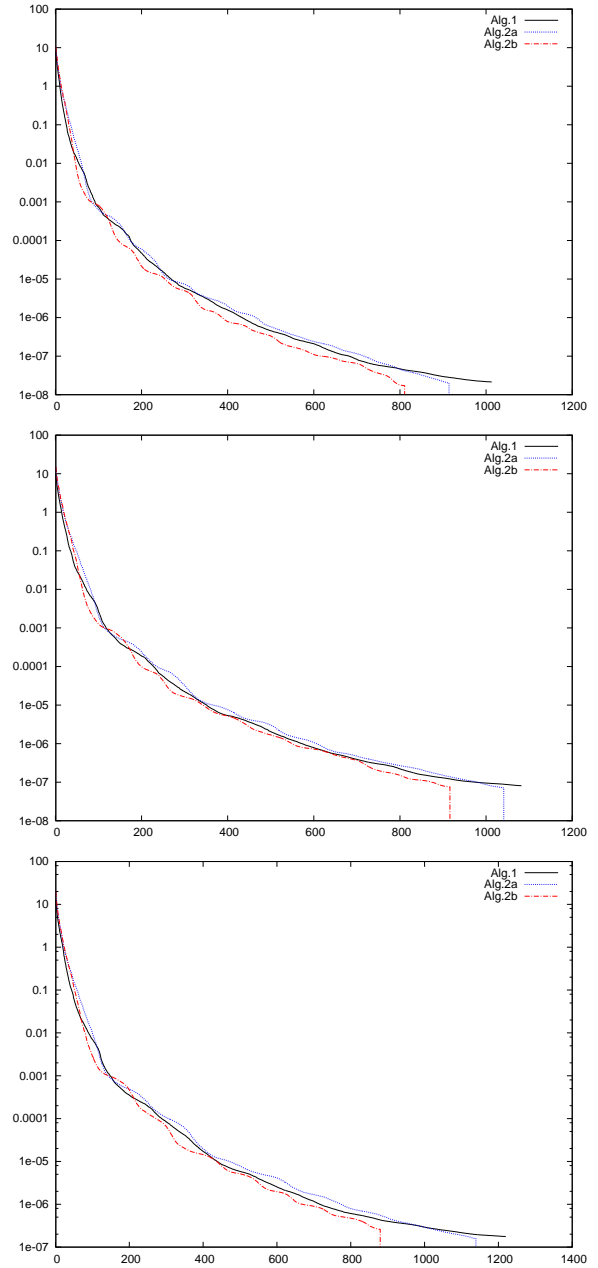


Fig. 1: Convergence of Alg. 1 and the two variants of Alg. 2 for the first time step of the benchmark problem. The values of \mathcal{E}_ε are rendered on the ordinate as a function of the L-BFGS-iterations on the abscissa. Top left: Spatial resolution $d_1 = d_2 = d_3 = 30$. Top right: Resolution $d_1 = d_2 = d_3 = 40$. Bottom: Resolution $d_1 = d_2 = d_3 = 50$. The exact solution in all cases is $\mathcal{E}_\varepsilon = 0$.

Algorithm	$10 \times 10 \times 10$	$30 \times 30 \times 30$	$50 \times 50 \times 50$	$70 \times 70 \times 70$
Alg. 2a ($\varepsilon_0 = 10^{-7}$)	$8.42 \cdot 10^{-9}$	$4.64 \cdot 10^{-8}$	$5.01 \cdot 10^{-7}$	$1.04 \cdot 10^{-6}$
Alg. 2a ($\varepsilon_0 = 10^{-9}$)	$8.21 \cdot 10^{-14}$	$4.81 \cdot 10^{-12}$	$1.22 \cdot 10^{-12}$	$3.05 \cdot 10^{-11}$

Table 3: Value of $\max_{t \in [0,1]} \int_{\Omega} \Lambda(|q(x,t)|^2 - 1)^2 dx$ for different spatial resolutions, two stop values (cf. Eqn. (29)), and Alg. 2a.

Resolution	Iterations L-BFGS	Iterations pc-L-BFGS	Time L-BFGS	Time pc-L-BFGS
$10 \times 10 \times 10$	22166995	41/13554468	383min 28s	234min 37s
$20 \times 20 \times 20$	15773252	133/162642	2656min 31s	24min 56s
$30 \times 30 \times 30$	62300391	269/229012	41131min 17s	128min 57s

Table 4: The first time step of the bending problem for the original ('L-BFGS') and the preconditioned ('pc-L-BFGS') scheme in comparison for $\varepsilon_0 = 10^{-11}$. For the preconditioned scheme, both predictor and corrector iterations are listed. 'Time' is the total computation time for the solution of one time step.

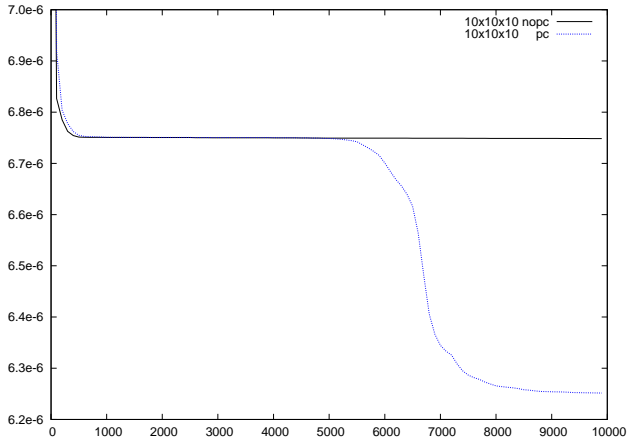


Fig. 2: The progression of \mathcal{E}_ε (ordinate) for the first 10000 L-BFGS-iterations (abscissa) of the original L-BFGS method (black) and the preconditioned L-BFGS method (blue) for $d_1 = d_2 = d_3 = 10$, $\varepsilon_0 = 10^{-11}$, and the first time step of the bending problem. As can be seen, even the preconditioned algorithm requires many iterations to overcome local minima of the energy.

Appendix - List of symbols

$A : B$ tensor product of A, B , below (6)
 $u \cdot v$ inner product of $u, v \in \mathbb{R}^3$

$\text{sym}(\sigma)$ symmetric part of a tensor σ , (4)
 $\text{skw}(\sigma)$ skew-symmetric part of σ , (4)
 $\text{tr}(\sigma)$ trace of tensor σ
 σ^t transpose of σ ; $R^t = R^{-1}$ for $R \in \mathcal{SO}(3)$
 $\|\cdot\|$ Frobenius matrix norm, (4)
 $|\cdot|$ Euclidean vector norm in \mathbb{R}^4 , (16)
 $\Omega \subset \mathbb{R}^3$ reference domain, undeformed solid
 (x, t) space and time coordinates
 φ deformation vector of the solid, (3)
 $F = D\varphi$ deformation tensor, (3)
 F_e elasticity tensor, (3)
 F_p plasticity tensor, (3)
 R_e rotation tensor, (1), (2), (3)
 U_e (right) stretching tensor, (3)
 K_e (right) curvature tensor, (22)
 \mathbb{I} identity tensor, $(\mathbb{I})_{kl} = (\delta_{kl})_{kl}$, (10)
 α Euler angle parameterization of R_e , (1)
 γ single-slip parameterization of F_p , (10)
 q Quaternion parameterization of R_e , (2)
 q_D Dirichlet boundary values of q , (8)
 \mathcal{E} mechanical energy, (13)
 $h > 0$ discrete (fixed) time step, (13)
 γ^0 values of γ at old time t , (12)
 κ^0 values of κ at old time t , (12)
 κ dislocation density, (12)
 $V(\kappa)$ dislocation energy, (6)
 W_{st} stretching energy, (4)
 W_c curvature energy, (5)
 X back stress (dual variable to F_p), (9),
 ξ hardening (dual variable to κ), (9)
 f_{ext} external volume forces, (13)
 M_{ext} external volume couples, (13)
 σ_Y yield stress, (13)
 Q^* dissipated energy, (11)
 m slip vector, (10)
 n slip normal, (10)
 $\rho > 0$ dislocation energy constant, (13)
 g_D Dirichlet boundary values of φ , (8)
 $\varepsilon > 0$ regularization of $|\cdot|$, Remark 1
 $\Lambda > 0$ Lagrange parameter to $|q|^2 = 1$, (13)
 λ, μ Lamé parameters, (4)
 μ_c Cosserat couple modulus, (4)

L_c	internal length scale, (5)
μ_2	parameter μ scaled by L_c^2 , (5)
d_1, d_2, d_3	spatial resolution, (34)
η_1, η_2, η_3	points on the numerical mesh, (35)
N_{IJK}	discrete numerical weights, (34)
$\beta(t)$	deformation parameter, (40), (41).

Compliance with Ethical Standards

The authors declare that they have no conflict of interest.

References

- Lakes, R.S.: Experimental micro mechanics methods for conventional and negative Poisson's ratio cellular solids as Cosserat continua, *J. Engineering Materials and Technology*, 113, 148-155, (1991).
- Bardella, L., Paggi, M., Vena, P.: Special issue on 'Recent advances on the mechanics of materials'. *Meccanica*, 53(3), 509-510, (2018).
- Rueger Z. and Lakes R.S. (2017) Strong Cosserat elastic effects in a unidirectional composite. *Zeitschrift für Angewandte Mathematik und Physik (ZAMP)* 68(54).
- Trovalusci P., Pau A. (2014) Derivation of microstructured continua from lattice systems via principle of virtual works: The case of masonry-like materials as micropolar, second gradient and classical continua. *Acta Mechanica* 225(1):157–177.
- Leonetti L., Greco F., Trovalusci P., Luciano R., Masiani R. (2018) A multiscale damage analysis of periodic composites using a couple-stress/Cauchy multidomain model: Application to masonry structures. *Composites Part B: Engineering* 141:50-59.
- Trovalusci P., Varano V., Rega G. (2010) A generalized continuum formulation for composite microcracked materials and wave propagation in a bar. *Journal of Applied Mechanics, Transactions ASME*, 77(6):061002, .
- Trovalusci P., Augusti G. (1998) A continuum model with microstructure for materials with flaws and inclusions. *Journal De Physique. IV: JP*, 8(8):383–390.
- Minga E., Macorini L., Izzuddin B. A. (2018). A 3D mesoscale damage-plasticity approach for masonry structures under cyclic loading. *Meccanica*, 53(7):1591-1611, .
- Lakes R.S. (1995) Experimental methods for study of Cosserat elastic solids and other generalized continua, in *Continuum models for materials with micro-structure*. Ed. H. Mhlhaus, J. Wiley, N. Y. Ch. 1, 1-22.
- Mindlin R.D. (1964) Micro-structure in linear elasticity. *Archive of Rational Mechanics and Analysis* 16:51–78.
- Eringen A.C. (1999) *Microcontinuum Field Theories*. Springer, New York.
- Cosserat E., Cosserat F. (1909) *Théorie des corps déformables*, Librairie Scientifique A. Hermann et Fils, Paris (English version: *Theory of deformable bodies*, NASA TT F-11 561 (1968)).
- Maugin G.A. (2010) *Mechanics of Generalized Continua – One Hundred years after the Cosserats*, Springer publishing.
- Neff P. (2006) A finite-strain elastic-plastic Cosserat theory for polycrystals with grain rotations. *Int. Journal of Eng. Science* 44:574–594.
- Ask A., Forest S., Appolaire B., Ammar K., Salmand O.U. (2018) A Cosserat crystal plasticity and phase field theory for grain boundary migration, *Journal of the Mechanics and Physics of Solids* 115:167-194.
- Corradi L., Genna F. (1990) Kinematic extremum theorems for holonomic plasticity. *International Journal of Plasticity*, 6(1):63-82.
- Corradi L., Genna F., Annovazzi L. (1991) Holonomic Versus Rate Relations in Plasticity. In: Del Piero G., Maceri F. (eds) *Unilateral Problems in Structural Analysis IV. International Series of Numerical Mathematics / Internationale Schriftenreihe zur Numerischen Mathematik – Série Internationale d'Analyse Numérique*, Vol 101. Birkhäuser Basel.
- Gidoni P., DeSimone A. (2017) Stasis domains and slip surfaces in the locomotion of a bio-inspired two-segment crawler. *Meccanica*, 52(3):587-601.
- Amendola A., Smith C.J., Goodall R., Auricchio F., Feo L., Benzoni G., Fraternali, F. (2016) Experimental response of additively manufactured metallic pentamode materials confined between stiffening plates. *Composite Structures* 142:254-262, .
- Modano M., Mascolo I., Fraternali F., Bieniek Z. (2018) Numerical and Analytical Approaches to the Self-Equilibrium Problem of Class $\theta = 1$ Tensegrity Metamaterials. *Front. Mater.* 5:5.
- Mascolo I., Amendola A., Zuccaro G., Feo L., Fraternali F. (2018) On the Geometrically Nonlinear Elastic Response of Class $\theta = 1$ Tensegrity Prisms. *Front. Mater.* 5:16.
- Nocedal J. (1980) Updating quasi-Newton matrices with limited storage. *Mathematics of Computation* 35:773–782.
- Nocedal J. (1989) On the Limited Memory Method for Large Scale Optimization. *Math. Programming B* 45:503–528.
- Blesgen T. (2015) On rotation deformation zones for finite-strain Cosserat plasticity. *Acta Mechanica* 226:2421–2434.
- Alaimo A., Artale V., Milazzo C., Ricciardello A. (2013) Comparison between Euler and Quaternion parametrization in UAV dynamics. *AIP Conference Proceedings* 1558, 1228.
- Dargham R., Sayouti A., Medromi H. (2015) Euler and Quaternion Parameterization in VTOL UAV Dynamics with Test Model Efficiency, *IJAIS* 9.
- Euler L. (1775) Nova methodus motum corporum rigidorum determinandi, novi commentari, *Acad. Imp. Petrop.* 20, 208–238, .
- Rodrigues O. (1840) Des lois géométriques qui régissent les déplacements d'un système solide dans l'espace, et de la variation des coordonnées provenant de ces déplacements considérés indépendamment des causes qui peuvent les produire, *J. Math.* 5:380–440.
- Hamilton W.R. (1844) On quaternions; or on a new system of imaginaries in algebra, *Phil. Mag.* 3:489–495.
- Andrei, N. (2007) A scaled BFGS preconditioned conjugate gradient algorithm for unconstrained optimization. *Appl. Math. Letters* 20:645–650.

31. Erway J., Marcia R.F.: Solving Limited-Memory BFGS Systems with Generalized Diagonal Updates, Proceedings World Congress on Engineering 2012 (1), (2012).
32. De Sterck H., Howse A.J.M. (2018) Nonlinearly Preconditioned L-BFGS as an Acceleration Mechanism for Alternating Least Squares, with Application to Tensor Decomposition. *Numer. Linear Algebra Appl.* 25:1–31.
33. Marjugi S.M., Leong W.J. (2013) Diagonal Hessian Approximation for Limited Memory Quasi-Newton via Variational Principle. *J. Appl. Math.* Vol. 2013, <http://dx.doi.org/10.1155/2013/523476>.
34. Jiang L., Byrd R.H., Eskow E., Schnabel R.B. (2004) A Preconditioned L-BFGS Algorithm with Application to Molecular Energy Minimization, Technical Report CU-CS-982-04, University of Colorado.
35. Kessel S. (1964) Lineare Elastizitätstheorie des anisotropen Cosserat-Kontinuums, *Abhandlung Braunschweigische Wiss. Gesell.* 16:1-22.
36. Blesgen T. (2013) Deformation patterning in Cosserat plasticity. *Modelling Simulation Mater. Sci. Eng.* 21:035001.
37. Han W., Reddy D. (1999) *Plasticity. Mathematical theory and numerical analysis*, Springer publishing.
38. Lubliner J. (2008) *Plasticity Theory*, Dover publications.
39. Fonseca, I., Francfort, G.A. (1995) Relaxation in BV versus quasiconvexification in $W^{1,p}$; a model for the interaction between fracture and damage. *Calculus of Variations* 3(4):407–446.
40. Ortiz M., Repetto E. (1999) Nonconvex energy minimization and dislocation structures in ductile single crystals. *Journal. Mech. Phys. Solids* 47: 397–462.
41. Weber G., Anand L. (1990) Finite deformation constitutive equations and a time integration procedure for isotropic, hyperelastic-viscoplastic solids, *Comput. Methods Appl. Mech. Eng.* 79:173–202.
42. Crumbach M., Goerdeler, M., Gottstein G. (2006) Modelling of recrystallisation textures in aluminium alloys: I. Model set-up and integration. *Acta Materialia* 54:3275–3289.
43. Blesgen T., Luckhaus S. (2006) The Dynamics of Transition Layers in Solids with Discontinuous Chemical Potentials. *Math. Meth. Appl. Sciences* 29:525–536.
44. Gurtin, M.E. (2002) A gradient theory of single-crystal viscoplasticity that accounts for geometrically necessary dislocations. *Journal of the Mechanics and Physics of Solids* 50:5–32.
45. Carstensen C., Hackl K., Mielke A. (2002) Non-convex potentials and microstructures in finite-strain plasticity. *Proceedings Roy. Soc. London A* 458:3275–3289.
46. Panteghini A., Bardella L. (2018) On the role of higher-order conditions in distortion gradient plasticity. *Journal of the Mechanics and Physics of Solids* 118:293-321.
47. Schatz F., Silveira J.L. (2018) Variational principles for shakedown analysis of materials with internal rotation. *Meccanica* 53(4-5):1093-1104.
48. Borokinni A. S., Fadodun O., Akinola A. P. (2018) Distortion-gradient plasticity theory for an isotropic body in finite deformation. *Meccanica* 53(11-12):3145-3155.
49. Blesgen T. (2014) Deformation patterning in three-dimensional large-strain Cosserat plasticity. *Mechanics Research Communications* 62(C):37–43.
50. Hamilton W.R. (1844) On quaternions, *The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science* 25:10–13.
51. Ebbinghaus H.-D., Hermes H., Hirzebruch F., Koecher M., Mainzer K., Neukirch J., Prestel A., Remmert R.: *Numbers*. Springer, (1991).
52. Dam E.B., Koch M., Lillholm M. (1998) Quaternions, Interpolation and Animation, Technical Report DIKU-TR-98/5, University of Copenhagen.
53. Blesgen T. (2017) A variational model for dynamic recrystallization based on Cosserat plasticity. *Composites Part B* 115:236–243.
54. Kuipers J.B. (1999) *Quaternions and Rotation Sequences: a primer with Applications to Orbits, Aerospace, and Virtual Reality*, Princeton university press.
55. Lang H., Linn J. (2009) Lagrangian field theory in space and time for geometrically exact Cosserat rods. *Reports of Fraunhofer ITWM* 150.
56. Liu D.C., Nocedal J. (1989) On the limited memory BFGS method for large scale optimization methods. *Math. Prog.* 45:503–528.
57. Byrd R.H., Nocedal J., Schnabel R.B. (1994) Representations of quasi-Newton matrices and their use in limited-memory methods. *Math. Program.* 63:129–156.
58. Oren S.S., Luenberger D.G. (1974) Self-scaling variable metric (SSVM) algorithms I: criteria and sufficient conditions for scaling a class of algorithms. *Management Science* 20:845–862.
59. Dongarra J.J., Bunch J.R., Moler C.B., Stewart G.W. (1979) *LINPACK Users' Guide*, SIAM, Philadelphia.
60. Shoemake K. (1985) Animating Rotations with Quaternion Curves. In B.A. Barsky: *Computer Graphics (SIGGRAPH 85 Proceedings)* 19:245–254.