MIXED FINITE ELEMENT METHODS FOR LINEAR COSSERAT EQUATIONS [*]

W. M. BOON[†], O. DURAN[‡], AND J. M. NORDBOTTEN[‡ §]

Abstract. We consider the equilibrium equations for a linearized Cosserat material. We identify their structure in terms of a differential complex, which is isomorphic to six copies of the de Rham complex through an algebraic isomorphism. Moreover, we show how the Cosserat materials can be analyzed by inheriting results from linearized elasticity. Both perspectives give rise to mixed finite element methods, which we refer to as strongly and weakly coupled, respectively. We prove convergence of both classes of methods, with particular attention to improved convergence rate estimates, and stability in the limit of vanishing Cosserat material parameters. The theoretical results are fully reflected in the actual performance of the methods, as shown by the numerical verifications.

Key words. Micro-polar materials, Hellinger-Riessner formulation, differential complexes, conforming finite elements

MSC codes. 68Q25, 68R10, 68U05

1. Introduction. In this paper, we will study the equations for Cosserat materials in three spatial dimensions [27, 13], also known as (micro-)polar materials among engineers [19]. Briefly speaking, these are an extension of the equations of linearized elasticity, by considering local rotations as an independent variable compared to the displacement. This additional generality afforded by Cosserat materials is attractive when modeling complex (and potentially multiscale) materials, such as granular media, cellular and crystalline solids, and composite porous media [17].

Previous work on the numerical approximation of the Cosserat equations have been primarily considered with finite-element methods associated with the two-field formulation in terms of displacement and rotation (for an early application, see [14], with a more recent contribution being [24]). Our interest is in the formulation as a four-field system, wherein in addition to the aforementioned variables, also the mechanical stress and the so-called couple stress (associated with spatial variations in rotation) are explicitly represented in the discretization. In terms of regular elastic materials, it is well-known that such formulations with explicit stress variables avoid locking phenomena associated with nearly incompressible materials (see e.g. [6, 1, 2]). Herein, we prove that in the context of Cosserat materials, the four-field formulation has the advantage that it is robust also in the limit of vanishing Cosserat material parameters.

Our contributions in this regards as follows:
1. We identify that the operators appearing in the Cosserat equations can be associated with a differential complex, which we denote as the *Cosserat complex*. This complex is isomorphic to six copies of the de Rham complex, and as such inherits all the properties of the de Rham complex. In particular, the L^2 version of the Cosserat complex provides the proper setting to prove

[†]MOX Laboratory, Department of Mathematics, Politecnico di Milano, Italy.

[‡]Centre for Modeling of Coupled Subsurface Dynamics, Department of Mathematics, University of Bergen, Norway

[§]Norwegian Research Center, Postboks 22 Nygårdgaten, 5838, Bergen, Norway

well-posedness.
2. We also provide an alternative proof of well-posedness of the four-field formu-
   lation of the Cosserat equations using Ladyzhenskaya–Babuška–Brezzi (LBB)
   theory, through a relationship to the Hellinger-Riessner formulation of elastic-
   ity with so-called weak symmetry. This relationship shows that the Cosserat
   equations can be treated as a continuous generalization of the elasticity equa-
   tions.
3. We show that the two perspectives on well-posedness stated in 1) and 2) nat-
   urally lead to different families of stable and convergent mixed finite element
   discretizations, which we refer to as *strongly coupled* and *weakly coupled*, re-
   spectively. Importantly, the weakly coupled discretization is stable in the
   continuous transition between Cosserat and elastic materials.

The rest of the paper is structured as follows. In the next section, we introduce
the relevant notation and spaces, and state the linear Cosserat equations and their
relationship to linearized elasticity. Thereafter, in Section 3, we elaborate the con-
struction of the Cosserat complex, and the resulting discretization. In Section 4, we
elaborate the well-posedness proof of the Cosserat equation by use of the LBB the-
ory, and the resulting discretization. Sections 3 and 4 can be read independently of
each other. In Section 5, the expected optimal-order convergence rates and stability
properties are demonstrated in numerical examples.

2. Notation and problem definition. We first summarize the notation needed
for the later developments, whereafter we state the equations of linear Cosserat ma-
terials.

In the interest of space and clarity, we will exclusively detail the case of three spa-
tial dimensions, wherein both displacement and rotation are characterized by three-
dimensional vectors. The reduction to the simpler case of 2D problems is obtained
by considering line symmetry along the third dimension, whereafter the problem re-
duces to being characterized by a two-dimensional displacement vector, together with
a single scalar rotation variable around the out-of-plane dimension. This reduction
is straight-forward for all sections of the paper except for Section 3.2, where some
modifications would be needed in the statement of the theoretical results.

2.1. General notation. As a preliminary remark, we note that we will as far as
possible adhere to the standard conventions of "vector-matrix calculus", as opposed to
either tensor or exterior calculus. This choice it justified by avoiding the proliferation
of indices in tensor calculus, and by being accessible to a broader audience than
exterior calculus.

Throughout the paper, we will consider the domain $\Omega \subset \mathbb{R}^3$ to be open, bounded,
simply connected and contractible, and we will consistently denote the spatial coordi-
nate as $x \in \Omega$. As the domain will not change, we will often suppress the dependence
on Ω in the statement of function spaces. That said, we denote spaces of vector fields
by $\mathbb{V} := \{v : \Omega \to \mathbb{R}^3\}$, and spaces of matrix fields by $\mathbb{M} := \{\sigma : \Omega \to \mathbb{R}^{3\times3}\}$. To
codify material constants, will also need four-tensors fields (i.e. mappings between
matrices), the space of which we denote as $\mathbb{A} = \{A : \mathbb{M} \to \mathbb{M}\}$. As indicated, we will
use lower-case Latin letters for vectors, Greek letters for matrices, and capital Latin
letters for fourth-order tensors.

When no preface is given, we assume that the spaces have "sufficient" regularity
for the relevant expressions to make sense. On the other hand, when the regularity
of the fields needs to be treated explicitly, this is indicated by prefaces, e.g. $L^2\mathbb{V} :=
\{v \in \mathbb{V} : \int_\Omega v \cdot v \, dV < \infty\}$. The subscript "0" on some space V, i.e. V_0, denotes the

subspace of that space with compact support in Ω (i.e. that the function takes zero values on the boundary $\partial\Omega$).

For a vector field $v \in \mathbb{V}$, with components v_i and a matrix field $\sigma \in \mathbb{M}$ with components $\sigma_{i,j}$, we define the vector generalization of gradient and divergence, together with the *asymmetry operator* $S : \mathbb{M} \to \mathbb{V}$, as:

$$(2.1) \qquad (\nabla v)_{i,j} = \frac{\partial v_i}{\partial x_j}, \qquad (\nabla \cdot \sigma)_i = \sum_j \frac{\partial \sigma_{i,j}}{\partial x_j}, \qquad (S\sigma)_i = \sigma_{i-1,i+1} - \sigma_{i+1,i-1}.$$

In the last definition, indices are understood modulo 3.

By standard arguments [23], we generalize the definition of the differential operators to their weak counterparts, which we still denote as ∇ for the gradient and $\nabla\cdot$ for the divergence. These are densely defined in $L^2\mathbb{V}$ and $L^2\mathbb{M}$, respectively and their domains of definition are the vector extensions of $H^1(\Omega)$ and $H(\nabla\cdot;\Omega)$:

$$(2.2) \quad H(\nabla;\mathbb{V}) := \left\{ v \in L^2\mathbb{V} : \nabla v \in L^2\mathbb{M} \right\}, \quad H(\nabla\cdot;\mathbb{M}) := \left\{ \sigma \in L^2\mathbb{M} : \nabla \cdot \sigma \in L^2\mathbb{V} \right\}$$

For all variables, we denote the standard L^2 inner product with parentheses, and the norm with double vertical strokes. As an example, for $\sigma, \tau \in L^2\mathbb{M}$:

$$(2.3) \qquad (\sigma, \tau) := \int_\Omega \sigma : \tau \, dV, \qquad\qquad \|\sigma\| := (\sigma, \sigma)^{1/2},$$

where the matrix double dot product : indicates the sum of element-wise products.

As a convention, we will omit subscript on the norm if the L^2 norm is used, and include subscripts whenever a different norm is considered. A second convention is that we will denote by the letter $0 < C < \infty$ a generic computable constant. At various points, it will be emphasized within a concrete context what C depends on.

For conciseness, we will primarily consider homogeneous boundary conditions on the displacement and rotation fields, and we thus summarize the following adjoint, or integration by parts, relationships for $(v, \sigma) \in \mathbb{V}_0 \times \mathbb{M}$:

$$(2.4) \qquad\qquad (\nabla \cdot \sigma, v) = (\sigma, -\nabla v)$$

Similarly, we denote the adjoint (with respect to the standard L^2 inner product) of S from (2.1) as S^*, which has the explicit form

$$(2.5) \qquad\qquad S^* v = \begin{pmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{pmatrix}$$

We note that in this case the adjoint is also (up to a factor 2) the right inverse of S, so that $SS^* v = 2v$.

Finally, we will often consider a pair of vector or matrix spaces on Ω. We denote these by script font, so that

$$(2.6) \qquad \mathscr{V} := \mathbb{V} \times \mathbb{V} \qquad\qquad \text{and} \qquad\qquad \mathscr{M} := \mathbb{M} \times \mathbb{M}.$$

All the previous definitions and conventions carry over to these compound spaces.

For the exposition of mixed finite element spaces, we will need the definition of the relevant families of discrete subspaces for $H(\nabla\cdot;\mathbb{M})$ and $L^2(\Omega)$. Thus given a simplicial tesselation $\mathfrak{F}_h$ of Ω, we denote the standard mixed finite element spaces according to the following conventions [9]:

- The space $P_{-k}(\mathfrak{F}_h) \subset L^2(\Omega)$ consists of k-order polynomials on each 3-simplex $\Delta \in \mathfrak{F}_h$.
- The space $P_k(\mathfrak{F}_h) := P_{-k}\mathfrak{F}_h \cap C^0(\mathbb{R})$ consists of k-order polynomials on each 3-simplex $\Delta \in \mathfrak{F}_h$ that are continuous across all element boundaries.
- The Brezzi-Douglas-Marini space of order k is defined as

$$BDM_k(\mathfrak{F}_h) := P_{-k}(\mathfrak{F}_h)^3 \cap H(\nabla\cdot; \mathbb{M}).$$

- The Raviart-Thomas space of order k is defined as

$$RT_k(\mathfrak{F}_h) := \left(P_{-k}(\mathfrak{F}_h)^3 \oplus \hat{P}_{-k}(\mathfrak{F}_h)\, \boldsymbol{x} \right) \cap H(\nabla\cdot; \mathbb{M})$$

where $\hat{P}_{-k} = P_{-k} \setminus P_{-(k-1)}$ are the homogeneous polynomials of exactly order k, and $\boldsymbol{x}$ is the position x interpreted as a vector.

We note that the convection for enumerating the RT-spaces varies in literature, here we adopt the most common convention where the lowest-order space is indexed as RT_0. The notation introduced so far suffices for the majority of the paper. Additional notation is needed for the introduction of the Cosserat complex in Section 3, which will be elaborated there.

2.2. The linearized Cosserat equations. The Cosserat model of an elastic material concerns both (macroscopic) displacements u, as well as local (microscopic) rotations r as primary state variables. These are associated with a linearized elastic stress field σ and, respectively, a "couple stress" field ω. In their linearized form, the model can be written as a set of four first-order equations in terms of the variables $(\sigma, \omega, u, r) \in \mathbb{M} \times \mathbb{M} \times \mathbb{V}_0 \times \mathbb{V}_0$. The strong form of these equations is as follows:

(2.7a)   Constitutive law for linearized elastic stress: $A_\sigma \sigma = \nabla u + S^* r$

(2.7b)   Constitutive law for linearized couple stress: $A_\omega \omega = \nabla r$

(2.7c)   Balance of linear momentum: $-\nabla \cdot \sigma = f_u$

(2.7d)   Balance of angular momentum: $-\nabla \cdot \omega + S\sigma = f_r$

We see that the model equations depend on material parameters A_σ and A_ω. Based on physical arguments [22, 20, 17, 19], we expect these to have the properties summarized in Assumptions 2.1-2.3 below.

ASSUMPTION 2.1 (Elastic stress parameter). *The elastic material parameter $A_\sigma \in \mathbb{A}$ is pointwise symmetric positive definite, and satisfies uniform upper and lower bounds, i.e. for any $\sigma, \tau \in L^2\mathbb{M}$ the following three statements hold with $A_\sigma^-, A_\sigma^+ \in \mathbb{R}$:*

$$(A_\sigma \sigma, \tau) = (A_\sigma \tau, \sigma), \qquad \frac{(A_\sigma \sigma, \sigma)}{\|\sigma^2\|} \geq A_\sigma^- > 0, \qquad \frac{(A_\sigma \sigma, \tau)}{\|\sigma\|\|\tau\|} \leq A_\sigma^+ < \infty.$$

ASSUMPTION 2.2 (Non-degenerate couple stress parameter). *The couple stress material parameter $A_\omega \in \mathbb{A}$ is pointwise symmetric positive definite, and satisfies uniform upper and lower bounds, i.e. for any $\omega, \rho \in L^2\mathbb{M}$ the following three statements hold with $A_\omega^-, A_\omega^+ \in \mathbb{R}$:*

$$(A_\omega \omega, \rho) = (A_\omega \rho, \omega), \qquad \frac{(A_\omega \omega, \omega)}{\|\omega^2\|} \geq A_\omega^- > 0, \qquad \frac{(A_\omega \omega, \rho)}{\|\omega\|\|\rho\|} \leq A_\omega^+ < \infty.$$

An important degeneracy is the limiting case when a Cosserat material reduces to a standard linearly elastic material, as we will detail in Sections 2.3 and 4. This requires a weaker assumption on the material parameters for the couple stress:

ASSUMPTION 2.3 (Degenerate couple stress parameter). *The inverse of the couple stress material parameter A_ω can be decomposed as*

$$A_\omega^{-1} = \epsilon^2 \tilde{A}_\omega^{-1},$$

in which

- $\epsilon \in P_1(\mathfrak{E})$ *is a scalar function with $\mathfrak{E}$ being a non-overlapping partitioning of the domain, satisfying (almost everywhere)*

$$0 \leq \epsilon < 1, \qquad\qquad |\nabla \epsilon| \leq C_\epsilon < \infty.$$

- $\tilde{A}_\omega \in \mathbb{A}$ *is pointwise symmetric positive definite and satisfies uniform upper and lower bounds, i.e. for any $\omega, \rho \in L^2\mathbb{M}$ the following three statements hold with $A_\omega^-, A_\omega^+ \in \mathbb{R}$:*

$$(\tilde{A}_\omega \omega, \rho) = (\tilde{A}_\omega \rho, \omega), \qquad \frac{(\tilde{A}_\omega \omega, \omega)}{\|\omega^2\|} \geq A_\omega^- > 0, \qquad \frac{(\tilde{A}_\omega \omega, \rho)}{\|\omega\|\|\rho\|} \leq A_\omega^+ < \infty.$$

COROLLARY 2.4 (Invertibility of material parameters). *Whenever Assumption 2.1 holds, A_σ^{-1} is well defined, symmetric positive definite, and has upper and lower bounds (given by $(A_\sigma^-)^{-1}$ and $(A_\sigma^+)^{-1}$, respectively). Similarly, if Assumption 2.2 (or 2.3) holds, A_ω^{-1} (or $\tilde{A}_\omega^{-1}$), is symmetric positive semi-definite, with upper and lower bounds.*

We note that while the assumptions 2.1 and 2.2 are sufficient for invertibility, they are not strictly necessary and may be relaxed to appropriate inf-sup conditions on A_ω or $\tilde{A}_\omega$. We make the slightly stronger assumptions since these are satisfied by the examples considered herein.

EXAMPLE 2.5 (Isotropic Cosserat material). *An isotropic material satisfies the condition that the material tensors A_σ and A_ω are invariant under coordinate rotations. As is well-known (see e.g. [26]), isotropic fourth-order tensors have only two free parameters, known as the Lamé parameters in the context of elasticity. Denoting the elastic Lamé parameters as μ_σ and λ_σ, we thus obtain for the elastic material parameter the action on any $\sigma \in L^2\mathbb{M}$:*

$$(2.8) \qquad A_\sigma \sigma = \frac{1}{2\mu_\sigma} \left(\sigma - \frac{\lambda_\sigma}{2\mu_\sigma + 3\lambda_\sigma} (\sigma : I) I \right)$$

Similarly for the couple stress parameter we introduce μ_ω and λ_ω, such that for any $\omega \in L^2\mathbb{M}$:

$$(2.9) \qquad A_\omega \sigma = \frac{1}{2\epsilon^2 \mu_\omega} \left(\omega - \frac{\lambda_\omega}{2\mu_\omega + 3\lambda_\omega} (\omega : I) I \right)$$

We remark that Assumption 2.1 holds if $\mu_\sigma > 0$ and $2\mu_\sigma + 3\lambda_\sigma > 0$, and similarly Assumption 2.2 (with $\epsilon = 1$) holds if $\mu_\omega > 0$ and $2\mu_\omega + 3\lambda_\omega > 0$. Importantly, $\lambda_\omega = 0$ is allowed in either case. Furthering the example, subject to Assumption 2.1 the elastic material law can be inverted to obtain the more familiar:

$$(2.10) \qquad A_\sigma^{-1}\tau = 2\mu_\sigma \tau + \lambda_\sigma (\tau : I) I$$

Remark 2.6. To model nearly incompressible materials, it is important to consider robustness with respect to the limit $\lambda_\sigma \to \infty$. This case is handled naturally in this

work since we consider the mixed formulation of the elasticity equations. In particular, for an isotropic material as characterized in Example 2.5, the continuity and coercivity constants appearing in the proofs can be shown to be independent of λ due to [9, Prop. 9.1.1]. We will therefore not consider the incompressible limit explicitly in the analyses of Sections 3 and 4, but verify this property numerically in Section 5.2.

2.3. Connection to linear elasticity. An important aspect of the Cosserat equations, as well as the later analysis, is that they can be considered as a generalization of the equations of linearized elasticity. To see this, we first recall the Hellinger-Riessner form of the equations for elasticity. Let therefore as before $(\sigma, u, r) \in \mathbb{M} \times \mathbb{V}_0 \times \mathbb{V}$ represent stress, displacement and rotation, then a static equilibrium of a linearly elastic system satisfies [2, 9]:

(2.11a)   Constitutive law for linearized elastic stress: $A_\sigma \sigma = \nabla u + S^* r$

(2.11b)   Balance of linear momentum: $-\nabla \cdot \sigma = f_u$

(2.11c)   Symmetry of the stress tensor: $S\sigma = 0$

By inspection of equations (2.7) and (2.11), we note that, formally, the equations of linearized elasticity can be obtained as the limit of the linearized Cosserat equations when in a suitable sense the parameter $A_\omega \to \infty$ and thus $\omega \to 0$. This motivates the distinction highlighted by Assumptions 2.2 and 2.3.

This paper will not provide rigorous analysis of the relationship between equations (2.7) and (2.11), but rather use the formal observation to motivate the need for numerical methods that are stable and convergent independent of an upper bound on A_ω. This topic will be discussed in detail in Section 4.

2.4. Problem definition. For future reference, we state the linearized Cosserat equations and highlight their structure. Consider first equations (2.7) written as a 4×4 block system: Find $(\sigma, \omega, u, r) \in \mathbb{M} \times \mathbb{M} \times \mathbb{V}_0 \times \mathbb{V}_0$ such that

$$
(2.12) \qquad
\begin{pmatrix}
A_\sigma & & -\nabla & -S^* \\
& A_\omega & & -\nabla \\
-\nabla \cdot & & & \\
S & -\nabla \cdot & &
\end{pmatrix}
\begin{pmatrix}
\sigma \\ \omega \\ u \\ r
\end{pmatrix}
=
\begin{pmatrix}
0 \\ 0 \\ f_u \\ f_r
\end{pmatrix}
$$

We highlight the structure of this system by introducing the block operators

$$
(2.13) \qquad \mathscr{A} := \begin{pmatrix} A_\sigma & \\ & A_\omega \end{pmatrix}, \qquad\qquad \mathscr{S} := \begin{pmatrix} -\nabla \cdot & \\ S & -\nabla \cdot \end{pmatrix}.
$$

This allows us to rewrite the problem to: For $\mathscr{f} = (f_u, f_r)$, find $(\sigma, \omega) = \mathscr{p} \in \mathscr{M}$ and $(u, r) = \mathscr{u} \in \mathscr{V}_0$ such that

$$
(2.14) \qquad \begin{pmatrix} \mathscr{A} & -\mathscr{S}_0^* \\ \mathscr{S} & \end{pmatrix} \begin{pmatrix} \mathscr{p} \\ \mathscr{u} \end{pmatrix} = \begin{pmatrix} 0 \\ \mathscr{f} \end{pmatrix}
$$

We emphasize with a subscript 0 on the adjoint that this operator is only defined for functions that satisfy zero boundary conditions in the sense of integration by parts, cf. (2.4) [23].

Whenever both Assumptions 2.1 and 2.2 hold, Corollary 2.4 applies and the Schur-complement formulation of (2.14) is given by

$$
(2.15) \qquad\qquad \mathscr{S} \left(\mathscr{A}^{-1} \mathscr{S}_0^* \right) \mathscr{u} = \mathscr{f}
$$

After solving for u, the variable p can be postprocessed as $p = \mathscr{A}^{-1}\mathcal{S}_0^* u$.

Our objective in the next sections is to provide well-posedness analysis and mixed finite element (MFE) methods for the system given in (2.14), both in the context of Assumption 2.2 (Section 4), as well as in the context of Assumption 2.3 (Section 3).

Remark 2.7 (Boundary conditions). We reiterate that the adjoint relationship between vector calculus operators only holds if one of the spaces has zero boundary conditions, due to integration-by-parts formulas such as (2.4). To minimize discussion of boundary conditions, we will consistently consider boundary conditions on the adjoint operator, however the opposite choice could equally well have been made, with minor modifications to the resulting presentation. As such, (2.15) implies so-called Dirichlet conditions on u, even though these are not explicitly stated.

3. Analysis and MFE discretization via the Cosserat complex. In this section, we will identify that the Cosserat equations can be seen as a Hodge-Laplace operator on a differential complex. This has the implication that general results from both analysis and MFE discretization of Hilbert complexes can be invoked. In particular, we obtain from general arguments both well-posedness and stability of both the continuous and discrete Cosserat equations, as well as quasi-optimal approximation properties of the resulting MFE solution. Readers who are not familiar with differential complexes may skip this section on first reading, as Section 4 can be read independently.

Since this section requires a somewhat richer theory than the rest of the manuscript, we start by introducing some extra notation, and present the analysis of the continuous and discrete systems afterward. Throughout this section, both Assumptions 2.1 and 2.2 will be assumed to be valid.

3.1. Notation for differential complexes. In this section, we will need additional operators to those defined in Section 2.2. In particular, for a matrix field $\sigma \in \mathbb{M}$ with components $\sigma_{i,j}$, we define the matrix generalization of the curl operator (acting on the second index), as

$$(3.1) \qquad (\nabla \times \sigma)_{i,j} = \partial_{j+1}\sigma_{i,j-1} - \partial_{j-1}\sigma_{i,j+1}$$

where the indices are again understood modulo 3.

Similarly, we will also need to expand on the definition of the asymmetry operator, and for this purpose introduce a self-adjoint operator common to the analysis of elasticity, denoted $T : \mathbb{M} \to \mathbb{M}$

$$(3.2) \qquad (T\sigma)_{i,j} = (I : \sigma)\, I - \sigma_{j,i}$$

where $I \in \mathbb{M}$ is the identity matrix field.

Next, we will require the cross-product between the vector-interpretation of the spatial coordinate and vectors, thus we define for $v \in \mathbb{V}$ (or similarly $v \in \mathbb{R}^3$) the element $Uv \in \mathbb{V}$ as:

$$(3.3a) \qquad (Uv)_i = \boldsymbol{x} \times v = x_{i+1}v_{i-1} - x_{i-1}v_{i+1}$$

We extend the definition to matrix fields, by letting the cross product act on the first index, i.e.

$$(3.3b) \qquad (U\sigma)_{i,j} = x_{i+1}\sigma_{i-1,j} - x_{i-1}\sigma_{i+1,j}$$

Furthermore, we denote the embedding of constant (vectors) into vector-valued fields by $\iota : \mathbb{R}^3 \to \mathbb{V}$.

Remark 3.1 (Bounded operators). We will use in the following the fact that S, T, S^* and U are bounded operators. In particular, the following continuity constants can be easily verified:

$$(3.4) \qquad \|S\| = \|S^*\| = \sqrt{2}, \qquad \|T\| = 2, \qquad \|U\| = \sqrt{2} \sup_{x \in \Omega} |x|.$$

Finally, we introduce a weighted inner product on $L^2 \mathscr{M}$ using the material weights. For $\varkappa = (\sigma, \omega)$, $\rho = (\tau, \rho) \in L^2 \mathscr{M}$, we define

$$(3.5) \qquad (\varkappa, \rho)_{\mathscr{A}} = (A_\sigma \sigma, \tau) + (A_\omega \omega, \rho)$$

with the associated norm $\|\rho\|_{\mathscr{A}} = (\rho, \rho)_{\mathscr{A}}^{1/2}$.

3.2. The Cosserat complex. In this section we will detail the construction of a differential complex in which the operator $\mathscr{S}$ from (2.13) forms a part. We emphasize the blanket assumption from Section 2.2 that all domains considered herein are for simplicity contractible, the extensions to more general domains requiring a treatment of the cohomology spaces.

While a permutation of the complex we construct has appeared previously as an intermediate calculation when deriving an *elasticity complex* [5, 16], we believe this is the first time its stand-alone importance, as well as the connection to the Cosserat equations, has been identified [1]. We therefore introduce as follows:

DEFINITION 3.2 (The Cosserat Complex). *The following differential complex is denoted the Cosserat complex:*

$$(3.6)$$

$$0 \longrightarrow \mathbb{R}^6 \xrightarrow{\left(\begin{smallmatrix} U & \iota \\ \iota & 0 \end{smallmatrix} \right)} \mathscr{V} \xrightarrow{\left(\begin{smallmatrix} -\nabla & -S^* \\ 0 & -\nabla \end{smallmatrix} \right)} \mathscr{M} \xrightarrow{\left(\begin{smallmatrix} 0 & -\nabla \times \\ -\nabla \times & T \end{smallmatrix} \right)} \mathscr{M} \xrightarrow{\left(\begin{smallmatrix} -\nabla \cdot & 0 \\ S & -\nabla \cdot \end{smallmatrix} \right)} \mathscr{V} \longrightarrow 0$$

The claimed differential complex structure follows from the calculus identities $\nabla \iota \to 0$, $\nabla \cdot \nabla \times \to 0$ and $\nabla \times \nabla \to 0$, together with the commutative properties stated below. For all $y \in \mathbb{R}^3$, $v \in \mathbb{V}$ and $\sigma \in \mathbb{M}$:

$$(3.7) \qquad S \nabla \times \sigma + \nabla \cdot (T\sigma) = 0, \qquad \nabla \times S^* v - T \nabla v = 0, \qquad S^* \iota y + \nabla U y = 0$$

Before we identify the link between the Cosserat complex and the linearized Cosserat equations, we recall the definition of the Hodge-Laplace operator. Let (V^k, d^k) be a differential complex with vector spaces V^k and differential operators $d^k : V^k \to V^{k+1}$. Moreover, let $d_k = (d^k)^*$ be the adjoint with respect to the inner products on the spaces. Then the Hodge-Laplace equation on vector space V^k is defined as (see e.g. [4]):

$$(3.8) \qquad \left(d_k d^k + d^{k-1} d_{k-1} \right) u = f.$$

COROLLARY 3.3 (Cosserat equations). *The linearized Cosserat equations (2.7) arise in two ways from the Cosserat complex (3.6):*

---

[1] After uploading this preprint, we were made aware of very recent work by Čap and Hu also discussing this same complex and its relationship to the Cosserat equations [28]. However, as they do not discuss finite-dimensional subspaces, the main results of the current contribution remain independent of their work. A future update of this preprint will contain a more careful referencing of their work.

1) *The Hodge-Laplace equation on the right-most space $\mathscr{V}$, subject to $\mathscr{A}$-weighted inner-products on $\mathscr{M}$, is equivalent to the linearized Cosserat equations with Dirichlet boundary conditions stated in* (2.15),

2) *The Hodge-Laplace equation on the left-most space $\mathscr{V}$, subject to $\mathscr{A}$-weighted inner-products on $\mathscr{M}$, is equivalent to the linearized Cosserat equation with Neumann boundary conditions and rigid body motion constraints.*

Proof. 1) We note that the rightmost differential operator is the operator $\mathscr{S}$ defined in Section 2.4. The Hodge-Laplace equation is then defined for $u \in \mathscr{V}$, and a right-hand side $f \in \mathscr{V}$, as

$$(3.9) \qquad \mathscr{S}\mathscr{S}^*_{\mathscr{A},0} u = f$$

Where $\mathscr{S}^*_{\mathscr{A},0}$ is the adjoint of $\mathscr{S}$ with respect to the weighted inner product, i.e. $\mathscr{S}^*_{\mathscr{A},0}$ satisfies for $p \in \mathscr{M}$ and $u \in \mathscr{V}$:

$$(3.10) \qquad (\mathscr{S}^*_0 p, u) = (p, \mathscr{S}u) = \left(\mathscr{S}^*_{\mathscr{A},0} p, u\right)_{\mathscr{A}} = \left(\mathscr{A}\mathscr{S}^*_{\mathscr{A},0} p, u\right)$$

Thus $\mathscr{S}^*_{\mathscr{A},0} = \mathscr{A}^{-1}\mathscr{S}^*_0$, and by substitution we see that (3.9) coincides with (2.15), i.e. the Cosserat equations with Dirichlet boundary conditions.

2) To state the Hodge-Laplace equation for the leftmost space $\mathscr{V}$ concisely, we use the notation

$$(3.11) \qquad \mathcal{S}^* := \begin{pmatrix} \nabla & S^* \\ 0 & \nabla \end{pmatrix}, \qquad\qquad \mathscr{U} := \begin{pmatrix} U & \iota \\ \iota & 0 \end{pmatrix}.$$

For which the Hodge-Laplace equation becomes:

$$(3.12) \qquad \left((\mathcal{S}^*)^*_{\mathscr{A},0}\, \mathcal{S}^* + \mathscr{U}\mathscr{U}^*\right) u = f$$

Expanding the definitions, we obtain

$$(3.13) \qquad \left(\begin{pmatrix} -\nabla & -S^* \\ 0 & -\nabla \end{pmatrix}^*_{\mathscr{A},0} \begin{pmatrix} -\nabla & -S^* \\ 0 & -\nabla \end{pmatrix} + \begin{pmatrix} U & \iota \\ \iota & 0 \end{pmatrix}\begin{pmatrix} U & \iota \\ \iota & 0 \end{pmatrix}^*\right) u = f$$

Here we recognize that the first term in (3.13) has the same structure as (2.15), but with the role of the "operator" and "adjoint" reversed, that is to say, with Neuman boundary conditions imposed on the stress and couple stress.

The second term of (3.13) constrains the six degrees of freedom associated with rigid body motions. Concretely, for $(u, r) \in \mathscr{V}$ and $y, z \in \mathbb{R}^3$, the adjoint satisfies

$$(3.14) \qquad \begin{pmatrix} U & \iota \\ \iota & 0 \end{pmatrix}^* \begin{pmatrix} u \\ r \end{pmatrix} \cdot \begin{pmatrix} y \\ z \end{pmatrix} = \left(\begin{pmatrix} u \\ r \end{pmatrix}, \begin{pmatrix} U & \iota \\ \iota & 0 \end{pmatrix}\begin{pmatrix} y \\ z \end{pmatrix}\right) = \left(\begin{pmatrix} u \\ r \end{pmatrix}, \begin{pmatrix} Uy + \iota z \\ \iota y \end{pmatrix}\right)$$

from which it follows that

$$(3.15) \qquad \begin{pmatrix} U & \iota \\ \iota & 0 \end{pmatrix}^* \begin{pmatrix} u \\ r \end{pmatrix} = \begin{pmatrix} \int_\Omega r - u \times \boldsymbol{x}\, dV \\ \int_\Omega u\, dV \end{pmatrix}$$

where we identify the rows of the adjoint as the angular and linear displacements, respectively. □

Recognizing the connection between the Cosserat complex and the Cosserat equations, we proceed to study the former. The initial starting point is the following.

LEMMA 3.4 (De Rham complex isomorphism).   *The Cosserat complex* (3.6) *is isomorphic to six copies of the de Rham complex.*

Proof. The isomorphism between a permutation of (3.6) and six copies of the de Rham complex was perhaps first presented in [16], and the argument was reproduced in [5]. We restate the proof here for completeness since our complex differs slightly from those in the stated references, and since the structure of the isomorphism will be needed later in the analysis of the mixed-finite element discretization. Moreover, we will use the explicit expressions for the isomorphism in the subsequent analysis of finite element subspaces.

Consider the map U from (3.3). For both $\mathscr{V}$ and $\mathscr{M}$ we can then define the automorphisms

$$(3.16a) \qquad \Phi = \begin{pmatrix} I & 0 \\ U & I \end{pmatrix} \qquad \text{with inverse} \qquad \Phi^{-1} = \begin{pmatrix} I & 0 \\ -U & I \end{pmatrix}$$

$$(3.16b) \qquad \Psi = \begin{pmatrix} U & I \\ I & 0 \end{pmatrix} \qquad \text{with inverse} \qquad \Psi^{-1} = \begin{pmatrix} 0 & I \\ I & -U \end{pmatrix}$$

Moreover, we calculate the bounds:

$$(3.17) \qquad \|\Phi\| = \sup_{v \in \mathscr{V}} \frac{\|\Phi v\|}{\|v\|} \leq \sup_{\substack{v \in \mathscr{V} \\ x \in \Omega}} \frac{(1 + \|x\|)\|v\|}{\|v\|} = \sup_{x \in \Omega}(1 + \|x\|) = C < \infty$$

where C depends only on the domain. By a similar calculation we find that also $\|\Phi\| \leq C$, and identical bounds on Ψ.

We now propose the following commuting diagram, where the upper sequence is six (negated) copies of the de Rham complex, while the lower sequence is the Cosserat complex.

(3.18)

$$\begin{array}{ccccccccccc}
0 & \longrightarrow & \mathbb{R}^6 & \xrightarrow{\left(\begin{smallmatrix} \iota & 0 \\ 0 & \iota \end{smallmatrix}\right)} & \mathscr{V} & \xrightarrow{\left(\begin{smallmatrix} -\nabla & 0 \\ 0 & -\nabla \end{smallmatrix}\right)} & \mathscr{M} & \xrightarrow{\left(\begin{smallmatrix} -\nabla\times & 0 \\ 0 & -\nabla\times \end{smallmatrix}\right)} & \mathscr{M} & \xrightarrow{\left(\begin{smallmatrix} -\nabla\cdot & 0 \\ 0 & -\nabla\cdot \end{smallmatrix}\right)} & \mathscr{V} & \longrightarrow & 0 \\
& & \downarrow{\scriptstyle I} & & \downarrow{\scriptstyle \Psi} & & \downarrow{\scriptstyle \Psi} & & \downarrow{\scriptstyle \Phi} & & \downarrow{\scriptstyle \Phi} & \\
0 & \longrightarrow & \mathbb{R}^6 & \xrightarrow{\left(\begin{smallmatrix} U & \iota \\ \iota & 0 \end{smallmatrix}\right)} & \mathscr{V} & \xrightarrow{\left(\begin{smallmatrix} -\nabla & -S^* \\ 0 & -\nabla \end{smallmatrix}\right)} & \mathscr{M} & \xrightarrow{\left(\begin{smallmatrix} 0 & -\nabla\times \\ -\nabla\times & T \end{smallmatrix}\right)} & \mathscr{M} & \xrightarrow{\left(\begin{smallmatrix} -\nabla\cdot & 0 \\ S & -\nabla\cdot \end{smallmatrix}\right)} & \mathscr{V} & \longrightarrow & 0
\end{array}$$

Note in particular the extension of the de Rham (and Cosserat) complexes to include the space of constants $\mathbb{R}^6$ as the left-most non-trivial space. This ensures that the de Rham (and by isomorphism Cosserat) complex is exact, and in practice provides control over the six degrees of freedom associated with rigid body motions.

The commutativity of diagram (3.18) is verified by direct calculation, using the following relationships for all $v \in \mathbb{V}$ and $\sigma \in \mathbb{M}$:

$$(3.19a) \qquad\qquad -S^*v = \nabla Uv - U\nabla v$$

$$(3.19b) \qquad\qquad T\sigma = \nabla \times U\sigma - U\nabla \times \sigma$$

$$(3.19c) \qquad\qquad S\sigma = \nabla \cdot U\sigma - U\nabla \cdot \sigma$$

Thus it follows that Φ and Ψ form an (invertible) complex map and that the six de Rham complexes and the Cosserat complexes are equivalent.   □

From Lemma 3.4, it follows that on the continuous level, the properties of the de Rham complex have a counterpart for the Cosserat complex. We will not review

the theory of the de Rham complex herein (see e.g. [10, 25] for some perspectives), but merely allude to the fact that this ensures that the complex is closed, that the cohomology of the complex is well understood, a Hodge (Helmholtz) decomposition is available, and that Poincaré inequalities can be established for all parts of the complex [4].

Our main concern is the Cosserat complex with L^2 regularity, which leads to a Hilbert-Cosserat complex (for the extension of continuous complexes to the Hilbert setting, see [12, 4]). Thus, we consider the weak generalization of the calculus operators ∇, $\nabla\times$ and $\nabla\cdot$ to densely defined operators between distributions of L^2 regularity. This leads to the following definition.

DEFINITION 3.5 (Hilbert-Cosserat complex). *The* Hilbert-Cosserat complex *is given by*

(3.20)

$$0 \longrightarrow \mathbb{R}^6 \xrightarrow{\left(\begin{smallmatrix} U & \iota \\ \iota & 0 \end{smallmatrix}\right)} L^2\mathscr{V} \xrightarrow{\left(\begin{smallmatrix} -\nabla & -S^* \\ 0 & -\nabla \end{smallmatrix}\right)} L^2\mathscr{M} \xrightarrow{\left(\begin{smallmatrix} 0 & -\nabla\times \\ -\nabla\times & T \end{smallmatrix}\right)} L^2\mathscr{M} \xrightarrow{\left(\begin{smallmatrix} -\nabla\cdot & 0 \\ S & -\nabla\cdot \end{smallmatrix}\right)} L^2\mathscr{V} \longrightarrow 0$$

Next, we characterize the domains of the differential operators appearing in diagram (3.20).

LEMMA 3.6 (Domain complex). *The domains of the generatlized differential operators coincide with the domains of the differential operators of the 6-fold de Rham complex. The* domain complex *of* (3.20) *is thus:*

(3.21)

$$\mathbb{R}^6 \xrightarrow{\left(\begin{smallmatrix} U & \iota \\ \iota & 0 \end{smallmatrix}\right)} H(\nabla;\mathscr{V}) \xrightarrow{\left(\begin{smallmatrix} -\nabla & -S^* \\ 0 & -\nabla \end{smallmatrix}\right)} H(\nabla\times;\mathscr{M}) \xrightarrow{\left(\begin{smallmatrix} 0 & -\nabla\times \\ -\nabla\times & T \end{smallmatrix}\right)} H(\nabla\cdot;\mathscr{M}) \xrightarrow{\left(\begin{smallmatrix} -\nabla\cdot & 0 \\ S & -\nabla\cdot \end{smallmatrix}\right)} L^2\mathscr{V}$$

Proof. The key point is to show that the algebraic operators S^*, T and S do not alter the domain nor range of the differential operators in the 6-fold de Rham complex. This follows directly from the commuting properties stated in (3.7) and the boundedness of the operators in Remark 3.1. As an example, for any $v \in H(\nabla;\mathbb{V})$, it follows that $S^*v \in H(\nabla\times;\mathbb{M})$ since $\|\nabla \times S^*v\| = \|T\nabla v\| \leq \|T\|\|\nabla v\| < \infty$. □

In view of Lemma 3.6, we will prefer the more standard shorthand notation $H(\nabla\cdot;\mathscr{M}) := \{p = (\sigma, \omega) \in L^2\mathscr{M} : (\nabla\cdot\sigma, \nabla\cdot\omega) \in L^2\mathscr{V}\}$ over the equivalent $H(\mathscr{S};\mathscr{M})$.

COROLLARY 3.7 (Well-posedness with Dirichlet conditions). *Let the mixed variational form of the Cosserat equations with Dirichlet conditions* (2.15), *be stated as follows: For* $f \in L^2\mathscr{V}$, *find* $(p,u) \in H(\nabla\cdot;\mathscr{V}) \times L^2\mathscr{V}$ *such that:*

(3.22a) $$(p, p')_{\mathscr{A}} - (u, \mathscr{S}p') = 0 \qquad \forall p' \in H(\nabla\cdot;\mathscr{V})$$
(3.22b) $$(\mathscr{S}p, u') = (f, u') \qquad \forall u' \in L^2\mathscr{V}$$

Then (3.22) *admits a unique solution that satisfies*

(3.23) $$\|p\|_{\mathscr{A}} + \|\mathscr{S}p\| + \|u\| \leq C\|f\|$$

Where the constant C depends on the Poincaré constant of the Cosserat complex, and thus scales with $\sqrt{\max\left(A_\sigma^+, A_\omega^+\right)}$.

Proof. Due to Corollary 3.3, point 1), system (3.22) is a Hodge-Laplace equation on the Hilbert-Cosserat complex (3.20), and due to Lemma 3.4 this complex is closed and exact. The existence of unique solutions is therefore guaranteed by standard theory, see e.g. [4, Thm 4.8]. □

COROLLARY 3.8 (Well-posedness with Neumann conditions). *Let the mixed variational form of the Cosserat equations with Neumann boundary conditions* (3.12), *be stated as follows: For $f \in L^2\mathcal{V}$, find $(y, u) \in \mathbb{R}^6 \times H(\nabla; \mathcal{V})$ such that:*

(3.24a)
$$y - \mathcal{U}^* u = 0$$

(3.24b)
$$(\mathcal{S}^* u, \mathcal{S}^* v)_{\mathcal{A}} + (\mathcal{U} y, v) = (f, v) \qquad \forall v \in H(\nabla; \mathcal{V})$$

Then (3.24) *admits a unique solution that satisfies*

(3.25)
$$\|z\| + \|\mathcal{U} z\| + \|u\| + \|\mathcal{S}^* u\|_{\mathcal{A}} \leq C\|f\|$$

Where the constant $C < \infty$ depends on the Poincaré constant of the Cosserat complex.

Proof. The argument is identical to Corollary 3.7, using point 2) of Corollary 3.3. □

Remark 3.9 (Boundary conditions). We note that equations (3.22) are a first-order system, while equations (3.24) are a second-order system (due to the presence of the term $(\mathcal{S}^* u, \mathcal{S}^* v)_{\mathcal{A}}$). This is not a consequence of the boundary conditions – indeed by reversing the convention for placing boundary conditions on the operator or the adjoint, one can equally well obtain a first order system for Neumann conditions and a second order system incorporating Dirichlet conditions, or even combinations of the two. In the interest of space, these variations are not elaborated further.

3.3. Strongly coupled mixed finite element approximation. Our next objective is to obtain a discrete approximation to equations (3.22). That is to say, for a family of finite-dimensional spaces $X_h^p \subset H(\nabla\cdot; \mathcal{V})$ and $X_h^u \subset L^2\mathcal{V}$, both indexed by h, we approximate equations (3.22) by the following finite-dimensional system: For $f \in L^2\mathcal{V}$, find $(p_h, u_h) \in X_h^p \times X_h^u$ such that:

(3.26a)
$$(p_h, p_h')_{\mathcal{A}} - (u_h, \mathcal{S} p_h') = 0 \qquad \forall p_h' \in X_h^p$$

(3.26b)
$$(\mathcal{S} p_h, u_h') = (f, u_h') \qquad \forall u_h' \in X_h^u$$

We will in this section consider finite element pairs that satisfy the following definition:

DEFINITION 3.10 (Strongly coupled spaces). *The finite element spaces $X_h^p \times X_h^u \subset H(\nabla\cdot; \mathcal{V}) \times L^2\mathcal{V}$ form a strongly coupled pair if $\mathcal{S} X_h^p \subseteq X_h^u$ and there exist linear mappings $\pi^\star$ with $\star \in \{p, u\}$ such that the following diagram arises:*

(3.27)
$$
\begin{array}{ccccc}
H(\nabla\cdot; \mathcal{M}) & \xrightarrow{\ \mathcal{S}\ } & L^2\mathcal{V} & \longrightarrow & 0 \\
\downarrow{\scriptstyle \pi^p} & & \downarrow{\scriptstyle \pi^u} & & \\
X_h^p & \xrightarrow{\ \mathcal{S}\ } & X_h^u & \longrightarrow & 0
\end{array}
$$

Remark 3.11. For a strongly coupled pair of spaces, the condition $\mathcal{S} X_h^p \subseteq X_h^u$ implies that linear (2.7c) and angular (2.7d) momentum are balanced pointwise if $f \in X_h^u$. The nomenclature adopted here therefore relates to the strong or weak satisfaction of these balance equations.

If the diagram (3.27) commutes, then the discrete spaces form a *discrete subcomplex* of the relevant section of the Cosserat complex. The key for the remaining results is encoded in the continuity of the commuting linear maps $\pi^\star$ [9, 5]. Following those previous works, we recall the following definition.

DEFINITION 3.12 (Bounded cochain projection). *If the linear maps π^p and π^u satisfy the following properties:*
 1. *Both operators are projections, thus $\pi^p\pi^p = \pi^p$ and $\pi^u\pi^u = \pi^u$.*
 2. *Both operators are* uniformly bounded, *thus there exists a constant C, independent of h, such that*

$$(\|\pi^p p\|_{\mathscr{A}} + \|\mathcal{S}\pi^p p\|) \leq C(\|p\|_{\mathscr{A}} + \|\mathcal{S}p\|) \qquad \forall p \in H(\nabla\cdot;\mathscr{V}),$$
$$\|\pi^u u\| \leq C\|u\| \qquad \forall u \in L^2\mathscr{V}.$$

 3. *The operators commute:*

$$\mathcal{S}\pi^u = \pi^p\mathcal{S}.$$

Then we refer to the two linear maps $\pi^\star$ collectively as a bounded cochain projection.

Existence of the above bounded cochain projections, which we will verify below, is sufficient to obtain unique solvability and optimal approximation properties.

THEOREM 3.13 (Optimal approximation). *Let the spaces X_h^p and X_h^u be strongly coupled, and let there exist bounded cochain projections $\pi^\star$. Then problem (3.26) is uniquely solvable and, moreover, the solution of (3.26) satisfies standard approximation properties relative to the solution of (3.22):*

$$\|p - p_h\|_{\mathscr{A}} + \|\mathcal{S}p - \mathcal{S}p_h\| + \|u - u_h\|$$
$$(3.28) \qquad \leq C\left(\inf_{p' \in H(\nabla\cdot;\mathscr{M})}(\|p - p'\|_{\mathscr{A}} + \|\mathcal{S}(p - p')\|) + \inf_{u' \in L^2\mathscr{V}}(\|u - u'\|)\right)$$

where the constant C depends on the Poincaré constant of the section of the Cosserat complex given in Definition 3.10, as well as the uniform bound C on the bounded cochain projections stated in Definition 3.12.

Proof. See [4, Thm. 5.4 and 5.5]. □

THEOREM 3.14 (Strongly coupled MFE for Cosserat). *For a family of simplicial partitions $\mathfrak{F}_h$ of the domain Ω, and for a non-negative integer k, define*

$$(3.29) \qquad X_{h,k}^p := W_k(\mathfrak{F}_h)^3 \times W_{k+1}(\mathfrak{F}_h)^3, \qquad \text{with } W_k \in \{BDM_{k+1},\ RT_k\}$$
$$(3.30) \qquad X_{h,k}^u := P_{-k}(\mathfrak{F}_h)^3 \times P_{-(k+1)}(\mathfrak{F}_h)^3.$$

Then the following statements hold:
 1) *The spaces $X_{h,k}^p$ and $X_{h,k}^u$ are strongly coupled (Def. 3.10).*
 2) *There exist bounded cochain projections π^p and π^u (Def. 3.12).*
 3) *If the solution of equation (3.22) has sufficient regularity, then the following a priori convergence rates hold for $u = (u, r)$ and $p = (\sigma, \omega)$*

$$(3.31)$$
$$\|\sigma - \sigma_h\|_{H(\nabla\cdot;\mathbb{M})} + \|u - u_h\| + \|\omega - \omega_h\|_{H(\nabla\cdot;\mathbb{M})} + \|r - r_h\| = \mathcal{O}\left(h^{k+1}\right)$$

Proof. 1) Clearly, $X^p_{h,k} \subset H(\nabla \cdot; \mathcal{M})$ and $X^u_{h,k} \subset L^2 \mathcal{V}$, and any linear projection onto these spaces is a linear map. It remains to show that the range of $\mathcal{S}$ applied to $X^p_{h,k}$ is contained in $X^u_{h,k}$. Thus for any $(\sigma, \omega) \in X^p_{h,k}$, we need to show that

$$(3.32) \qquad \begin{pmatrix} -\nabla \cdot & 0 \\ S & -\nabla \cdot \end{pmatrix} \begin{pmatrix} \sigma \\ \omega \end{pmatrix} = \begin{pmatrix} -\nabla \cdot \sigma \\ S\sigma - \nabla \cdot \omega \end{pmatrix} \in \begin{pmatrix} P_{-k} (\mathfrak{F}_h)^3 \\ P_{-(k+1)} (\mathfrak{F}_h)^3 \end{pmatrix}$$

By the definition of $X^p_{h,k}$, we have that $\sigma \in W_k (\mathfrak{F}_h)^3$ and $\omega \in W_{k+1} (\mathfrak{F}_h)^3$, thus $\nabla \cdot \sigma \in P_{-k} (\mathfrak{F}_h)^3$ and $\nabla \cdot \omega \in P_{-(k+1)} (\mathfrak{F}_h)^3$ as differentiation reduces polynomial degree. It remains to show that $S\sigma \in P_{-(k+1)} (\mathfrak{F}_h)^3$. Note that $W_k (\mathfrak{F}_h) \subseteq P_{-(k+1)} (\mathfrak{F}_h)^{3 \times 3}$, and since S is a linear transformation with constant weights, it does not increase polynomial degree. In turn, we have $S\sigma \in P_{-(k+1)} (\mathfrak{F}_h)^3$, and thus the spaces are strongly coupled.

2) We first note that the existence of bounded cochain projections for the de Rham complex are well established, such as e.g. the Fortin operator and the L^2 projection (see e.g. [9, Prop. 2.5.2]). Let $\tilde{\pi}^p_{h,k}$ and $\tilde{\pi}^u_{h,k}$ be any such pair for the de Rham complex, i.e. such that

$$(3.33) \qquad \begin{CD} H(\nabla \cdot; \mathcal{M}) @>{\begin{pmatrix} -\nabla \cdot & 0 \\ 0 & -\nabla \cdot \end{pmatrix}}>> L^2 \mathcal{V} @>>> 0 \\ @V{\tilde{\pi}^p_{h,k}}VV @V{\tilde{\pi}^u_{h,k}}VV \\ X^p_{h,k} @>{\begin{pmatrix} -\nabla \cdot & 0 \\ 0 & -\nabla \cdot \end{pmatrix}}>> X^u_{h,k} @>>> 0 \end{CD}$$

is a commuting diagram, and $\tilde{\pi}^p_{h,k}$ and $\tilde{\pi}^u_{h,k}$ are uniformly bounded projections in $H(\nabla \cdot; \mathcal{M})$ and $L^2 \mathcal{V}$, respectively. To obtain a bounded cochain projection for the Cosserat complex, we exploit the isomorphism given in Lemma 3.4, and define

$$(3.34) \qquad \pi^p := \Phi \tilde{\pi}^p_{h,k} \Phi^{-1} \qquad\qquad \pi^u := \Phi \tilde{\pi}^u_{h,k} \Phi^{-1}$$

Following Definition 3.7, we first verify that these operators are projections. For π^u, consider any $(u, r) \in X^u_{h,k} = P_{-k} (\mathfrak{F}_h)^3 \times P_{-(k+1)} (\mathfrak{F}_h)^3$, and recall that U is a homogeneous linear function in x (cf. (3.3a)). Therefore $\Phi^{-1} (u, r) = (u, r + Uu) \in P_{-k} (\mathfrak{F}_h)^3 \times P_{-(k+1)} (\mathfrak{F}_h)^3 = X^u_{h,k}$. Now, since $\tilde{\pi}^u_{h,k}$ is a projection onto $X^u_{h,k}$, we derive

$$(3.35) \qquad \pi^u \begin{pmatrix} u \\ r \end{pmatrix} = \Phi \left(\tilde{\pi}^u_{h,k} \Phi^{-1} \begin{pmatrix} u \\ r \end{pmatrix} \right) = \Phi \left(\Phi^{-1} \begin{pmatrix} u \\ r \end{pmatrix} \right) = \begin{pmatrix} u \\ r \end{pmatrix}$$

and thus π^u is also a projection onto $X^u_{h,k}$.

To show that also π^p is a projection, we start by showing that that $\Phi^{-1} X^p_{h,k} \subseteq X^p_{h,k}$. We first consider the case $W_k = BDM_{k+1}$ and select any $(\sigma, \omega) \in BDM_{k+1} (\mathfrak{F}_h)^3 \times BDM_{k+2} (\mathfrak{F}_h)^3$. Then also $\Phi^{-1} (u, r) = (\sigma, \omega + U\sigma) \in BDM_{k+1} (\mathfrak{F}_h)^3 \times BDM_{k+2} (\mathfrak{F}_h)^3$, since the operator U increases polynomial degree, but does not reduce regularity.

Secondly for $W_k = RT_k$, consider $(\sigma, \omega) \in RT_k (\mathfrak{F}_h)^3 \times RT_{k+1} (\mathfrak{F}_h)^3$. Again we have $\Phi^{-1} (u, r) = (\sigma, \omega + U\sigma) \in RT_k (\mathfrak{F}_h)^3 \times RT_{k+1} (\mathfrak{F}_h)^3$ since the operator U is a linear combinations of products with coordinates, and by the definition of RT_k spaces given in Section 2.2, it holds that for any coordinate x_i, and $w \in RT_k (\mathfrak{F}_h)$

then $x_i w \in RT_{k+1}(\mathfrak{F}_h)$. Finally, a similar calculation as in (3.24) completes the verification that π^ρ is a projection onto $X^\rho_{h,k}$ for both cases of W_k.

From Lemma 3.4, we know that the operators Φ and Φ^{-1} are bounded (independent of h and k), and thus π^ρ and π^u inherit the boundedness of $\tilde{\pi}^\rho_{h,k}$ and $\tilde{\pi}^u_{h,k}$, respectively.

In order to verify commutativity, we first note that Lemma 3.4 allows us to explicitly write

$$
(3.36) \qquad \mathcal{S} = \Phi \begin{pmatrix} -\nabla \cdot & 0 \\ 0 & -\nabla \cdot \end{pmatrix} \Phi^{-1}
$$

Using this expression, we can easily verify the commutativity of the projection operators:

$$
(3.37) \quad \mathcal{S}\pi^\rho = \Phi \begin{pmatrix} -\nabla \cdot & 0 \\ 0 & -\nabla \cdot \end{pmatrix} \Phi^{-1} \Phi \tilde{\pi}^\rho_{h,k} \Phi^{-1} = \Phi \tilde{\pi}^u_{h,k} \begin{pmatrix} -\nabla \cdot & 0 \\ 0 & -\nabla \cdot \end{pmatrix} \Phi^{-1} = \pi^u \mathcal{S}
$$

We have thus established the existence of bounded cochain projections through an explicit construction.

3) is a corollary of Theorem 3.13, and the error rates follow from the standard approximation properties of the spaces [9], and the equivalence between the $\mathscr{A}$-weighted and unweighted norms on $L^2\mathcal{M}$ (a direct consequence of Assumptions 2.1 and 2.2). In particular, we have as a consequence of Theorem 3.13 that:

$$
\begin{aligned}
& \|\sigma - \sigma_h\|_{A_\sigma} + \|\nabla \cdot (\sigma - \sigma_h)\| + \|u - u_h\| + \|\omega - \omega_h\|_{A_\omega} \\
(3.38) \qquad & + \|S(\sigma - \sigma_h) - \nabla \cdot (\omega - \omega_h)\| + \|r - r_h\| = \mathcal{O}\left(h^{k+1}\right)
\end{aligned}
$$

The stress norms satisfy

$$
\|\sigma - \sigma_h\| \le \left(A^-_\sigma\right)^{-\frac{1}{2}} \|\sigma - \sigma_h\|_{A_\sigma} = \mathcal{O}\left(h^{k+1}\right)
$$
$$
\|\omega - \omega_h\| \le \left(A^-_\omega\right)^{-\frac{1}{2}} \|\omega - \omega_h\|_{A_\omega} = \mathcal{O}\left(h^{k+1}\right)
$$

For the compound norm, we can calculate:

$$
\begin{aligned}
\|\nabla \cdot (\omega - \omega_h)\| & \le \|S(\sigma - \sigma_h) - \nabla \cdot (\omega - \omega_h)\| + \|S(\sigma - \sigma_h)\| \\
& \le \|S(\sigma - \sigma_h) - \nabla \cdot (\omega - \omega_h)\| + \|\sigma - \sigma_h\|\|S\|
\end{aligned}
$$

Now, since $\|S\|$ is finite, it follows from equation (3.38) that

$$
(3.39) \qquad \|\nabla \cdot (\omega - \omega_h)\| \le \mathcal{O}\left(h^{k+1}\right)
$$

from which (3.31) follows. □

Higher convergence rates can to various degrees be obtained for all four variables. We present two different approaches that lead to complementary results in the following two theorems.

THEOREM 3.15 (Improved convergence for BDM_{k+1}). *Let $W_k = BDM_{k+1}$, and assume that the true solution has sufficient regularity. Then the following improved convergence rates hold:*

$$
(3.40) \qquad \|\sigma - \sigma_h\| + \|\omega - \omega_h\|_{H(\nabla \cdot;\mathbb{M})} + \|r - r_h\| = \mathcal{O}\left(h^{k+2}\right)
$$

Proof. We recall the improved L^2 estimates available for mixed methods [3, Theorem 3.11]:

$$(3.41) \qquad \|\wp - \wp_h\| \leq C \left(\inf_{p' \in X_h^\wp} \|\wp - \wp'\| + h \inf_{p' \in X_h^\wp} \|\mathcal{S}(\wp - \wp')\| \right)$$

The approximation properties of $X_h^\wp$ and X_h^u depend on the regularity of the solution, and will be dominated by the lower-order spaces for displacement and stress. Concretely, $\wp_h \in X_{h,k}^\wp$ implies that $\sigma_h \in BDM_{k+1}(\mathfrak{F}_h)^3$, thus we obtain from [9, Prop. 2.5.4] that

$$(3.42) \quad \inf_{p' \in X_h^\wp} \|\wp - \wp'\| \leq Ch^{k+2} \|\sigma\|_{k+2}, \quad \inf_{p' \in X_h^\wp} \|\mathcal{S}(\wp - \wp')\| \leq Ch^{k+1} \|\nabla \cdot \sigma\|_{k+1}$$

Thus for sufficiently regular σ, we have from (3.41) and (3.42) that:

$$(3.43) \qquad \|\sigma - \sigma_h\| + \|\omega - \omega_h\| = \mathcal{O}\left(h^{k+2}\right)$$

Consider now the error equation for the rotation and couple stress:

$$(3.44a) \qquad (e_\omega, \omega')_{\mathscr{A}} - (e_r, \nabla \cdot \omega') = 0$$
$$(3.44b) \qquad (\nabla \cdot e_\omega, r') = (Se_\sigma, r')$$

Since (3.44) is a Laplace-type elliptic equation, it follows from standard theory that the solution (e_ω, e_r) is bounded and satisfies

$$(3.45) \qquad \|\omega - \omega_h\|_{H(\nabla\cdot; M)} + \|r - r_h\| \leq C \|Se_\sigma\| = \mathcal{O}\left(h^{k+2}\right),$$

where the last equality follows from (3.42). The theorem follows by combining (3.42) and (3.45). □

For the second analysis of improved convergence rates, applicable also to $W_k = RT_k$, we consider a duality argument. Let $\varpi_{h,k}$ denote the L^2-projection onto $P_{-k}(\mathfrak{F}_h)^3$, and define $\varphi = (\varphi_u, \varphi_r)$ as the solution to the Cosserat equations (2.15) with right-hand side

$$(3.46) \qquad f_u = \varpi_{h,k}(u - u_h) \quad \text{and} \quad f_r = \varpi_{h,k+1}(r - r_h).$$

Now, for sufficiently regular domains, φ will enjoy higher regularity than the minimum regularity implied by the differential operators (see e.g. [18]). Then the following theorem applies.

THEOREM 3.16 (Improved convergence of rotation and displacement). *Let $W_k \in \{BDM_{k+1}, RT_k\}$, and φ as defined above be H^2 regular in the sense that*

$$(3.47) \qquad \left\|\mathscr{A}^{-1}\mathcal{S}^*\varphi\right\|_1 + \|\varphi\|_2 \lesssim \|\varpi_h e_u\|$$

Then the rotation r_h and displacement u_h satisfy the optimal rates

$$(3.48) \qquad \|r - r_h\| + \|\varpi_{h,k}u - u_h\| = \mathcal{O}\left(h^{k+2}\right).$$

Proof. The definition of φ implies that $\mathcal{S}\mathscr{A}^{-1}\mathcal{S}^*\varphi = \varpi_h e_u$ with $e_u := (e_u, e_r) = (u - u_h, r - r_h)$ and the L^2-projection $\varpi_h e_u = (\varpi_{h,k}e_u, \varpi_{h,k+1}e_r)$. Corollary 3.7 ensures that the solution is stable with

$$(3.49) \qquad \|\varphi\| + \|\mathcal{S}^*\varphi\| \leq C \|\varpi_h e_u\|$$

Introducing also $e_p := (e_\sigma, e_\omega) = (\sigma - \sigma_h, \omega - \omega_h)$, the error equations for the Cosserat approximation are obtained by subtracting (3.26) from (3.22):

(3.50a) $$(e_p, p')_{\mathscr{A}} - (e_u, \mathscr{S}p') = 0$$

(3.50b) $$(\mathscr{S}e_p, u') = 0$$

which holds for all $(p', u') \in X_h^p \times X_h^u$.

Choosing now $p' = \pi_h \mathscr{A}^{-1} \mathscr{S}^* \varphi$, this implies that $\mathscr{S}p' = \varpi_h e_u$, thus we obtain from (3.50a) that

(3.51) $$\|\varpi_h e_u\|^2 = (e_u, \varpi_h e_u) = (e_u, \mathscr{S}\pi_h \mathscr{A}^{-1} \mathscr{S}^* \varphi) = (e_p, \pi_h \mathscr{A}^{-1} \mathscr{S}^* \varphi)_{\mathscr{A}}$$

We now bound the right-hand side term:

$$
\begin{aligned}
\left(e_p, \pi_h \mathscr{A}^{-1} \mathscr{S}^* \varphi\right)_{\mathscr{A}} &= - \left(e_p, (1 - \pi_h) \mathscr{A}^{-1} \mathscr{S}^* \varphi\right)_{\mathscr{A}} + \left(e_p, \mathscr{A}^{-1} \mathscr{S}^* \varphi\right)_{\mathscr{A}} \\
&\lesssim \|e_p\| \|(1 - \pi_h) \mathscr{A}^{-1} \mathscr{S}^* \varphi\| + (e_p, \mathscr{S}^* \varphi) \\
&\leq h\|e_p\| \|\mathscr{A}^{-1} \mathscr{S}^* \varphi\|_1 + (\mathscr{S}e_p, (1 - \pi_h)\varphi) + (\mathscr{S}e_p, \pi_h \varphi) \\
&\leq h\|e_p\| \|\mathscr{A}^{-1} \mathscr{S}^* \varphi\|_1 + h^2 \|\mathscr{S}e_p\| \|\varphi\|_2 + 0
\end{aligned}
$$
(3.52)

Here we have used equation (3.50b) to eliminate the last term, and moreover the fact that since φ is H^2-regular, then [9]:

(3.53) $$\|(1 - \pi_h)\varphi\| \leq h^2 \|\varphi\|_2 \qquad \|(1 - \pi_h) \mathscr{A}^{-1} \mathscr{S}^* \varphi\| \leq h \|\mathscr{A}^{-1} \mathscr{S}^* \varphi\|_1$$

Combining Equations (3.51) and (3.52), and using inequalities (3.49) and (3.47), we obtain

(3.54) $$\|\varpi_{h,k} e_u\| + \|\varpi_{h,k+1} e_r\| \leq C \|\varpi_h e_u\| = \mathscr{O}\left(h^{k+2}\right)$$

However, for the rotations, we also note that for sufficiently smooth solutions r, it holds that:

(3.55) $$\|\varpi_{h,k+1} r - r\| = \mathscr{O}\left(h^{k+2}\right)$$

Thus

$$\|e_r\| = \|r - \varpi_{h,k+1} r + \varpi_{h,k+1} r - r_h\| \leq \|\varpi_{h,k+1} r - r\| + \|\varpi_{h,k+1} e_r\| = \mathscr{O}\left(h^{k+2}\right)$$

This proves the theorem.   □

Remark 3.17 (Extensions). The proofs of the above lemmas and theorems are based on rather general arguments. As such, we expect that it should be possible to reproduce the above constructions to obtain MFE discretizations of Hodge-Laplace equations for any part of the Cosserat complex. While we do not presently know of a physically relevant motivation for Hodge-Laplace equations on the $\mathscr{M}$-spaces of the complex, Corollary 3.3 shows another manifestation of the Cosserat equations on the left side of the Cosserat complex. The counterpart of Corollary 3.7 would be the variational system stated in (3.24). The conforming finite element spaces for equations (3.24) in the context of the Cosserat complex is then to use $p_h \in \mathbb{R}^6$ and standard Lagrange finite elements for u_h, with the space of displacements having a polynomial degree one higher than the rotations. The analysis of two-field finite elements methods for Cosserat materials can be conducted by more direct arguments (standard Galerkin theory for quadratic minimization problems, see e.g. [11]), and have been applied previously in literature [24].

Remark 3.18 (Lack of stability with respect to A_σ^- and A_ω^-). Assumption 2.2 ensures that A_ω^- and A_ω^+ are bounded from above and below, respectively, and this is an essential component of the proof of the stability both of the conforming MFE method defined in Theorem 3.14, as well as two-field finite element methods for Cosserat equations discussed in Remark 3.17. This has the practical implication that the conforming method of Theorem 3.14 will not be stable for degeneracy to normal elastic material laws (for which $A_\omega^- \to \infty$).

4. Analysis and MFE via Ladyzhenskaya–Babuška–Brezzi theory. In this section, we wish to dispense with Assumption 2.2, which was a prerequisite for the development in Section 3. As such, we will lose access to the direct identification of the Cosserat equations with the Cosserat complex, since we can no longer expect $(p, p')_{\mathscr{A}}$ to be a continuous bilinear form, and thus it fails to be an inner product on $L^2\mathscr{M}$.

Therefore, we will analyze the weak variational form of (2.15) directly, as stated in (3.22). Thereafter, we will again consider MFE discretization as in (3.26), however, this time we will have different conditions on the spaces, which results in a different discretization.

Throughout this section, both Assumptions 2.1 and 2.3 will be assumed. Moreover, we emphasize that Remark 2.6 applies, that is to say that for isotropic materials, the constants appearing in the various proofs will be independent of the coefficient λ_σ.

4.1. The degenerate linearized Cosserat equations. We begin by restating equations (2.12) subject to the substitution $A_\omega \to \epsilon^{-2}\tilde{A}_\omega$: Find $(\sigma, \omega, u, r) \in \mathbb{M} \times \mathbb{M} \times \mathbb{V}_0 \times \mathbb{V}_0$ such that

$$(4.1) \qquad \begin{pmatrix} A_\sigma & & -\nabla & -S^* \\ & \epsilon^{-2}\tilde{A}_\omega & & -\nabla \\ -\nabla\cdot & & & \\ S & -\nabla\cdot & & \end{pmatrix} \begin{pmatrix} \sigma \\ \omega \\ u \\ r \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ f_u \\ f_r \end{pmatrix}$$

We recognize that the obstacle to showing well-posedness is the degeneracy of the diagonal block $\epsilon^{-2}\tilde{A}_\omega$, and therefore make the change of variables $\tilde{\omega} = \epsilon^{-1}\omega$, after which we obtain the problem: Find $(\sigma, \tilde{\omega}, u, r) \in \mathbb{M} \times \mathbb{M} \times \mathbb{V}_0 \times \mathbb{V}_0$ such that

$$(4.2) \qquad \begin{pmatrix} A_\sigma & & -\nabla & -S^* \\ & \tilde{A}_\omega & & -\epsilon\nabla \\ -\nabla\cdot & & & \\ S & -\nabla\cdot\epsilon & & \end{pmatrix} \begin{pmatrix} \sigma \\ \tilde{\omega} \\ u \\ r \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ f_u \\ f_r \end{pmatrix}$$

This motivates defining the scaled block operators

$$(4.3) \qquad \tilde{\mathscr{A}} = \begin{pmatrix} A_\sigma & \\ & \tilde{A}_\omega \end{pmatrix} \qquad\qquad \mathscr{S}_\epsilon = \begin{pmatrix} -\nabla\cdot & \\ S & -\nabla\cdot\epsilon \end{pmatrix}$$

and we can write for $(\sigma, \tilde{\omega}) = \tilde{p} \in \mathscr{M}$ and $(u, r) = u \in \mathscr{V}_0$:

$$(4.4) \qquad \begin{pmatrix} \tilde{\mathscr{A}} & -\mathscr{S}_{\epsilon,0}^* \\ \mathscr{S}_\epsilon & \end{pmatrix} \begin{pmatrix} \tilde{p} \\ u \end{pmatrix} = \begin{pmatrix} 0 \\ \ell \end{pmatrix}$$

Where $\mathscr{S}_{\epsilon,0}^*$ is the adjoint operator satisfying $(\mathscr{S}_{\epsilon,0}^* u, p) = (u, \mathscr{S}_\epsilon p)$ for all $(u, p) \in \mathscr{V}_0 \times \mathscr{M}$

4.2. Well-posedness of the degenerate linearized Cosserat equations.
For the well-posedness analysis, we recognize that the space $H(\nabla\cdot;\mathbb{M})$ requires too much regularity of the couple stress when $\epsilon = 0$. Thus, we begin by introducing the weighted space

$$(4.5) \qquad H(\nabla \cdot \epsilon; \mathbb{M}) = \left\{ \omega \in L^2\mathbb{M} \mid \nabla \cdot (\epsilon\omega) \in L^2\mathbb{V} \right\}$$

Based on this space, we define the semi-weighted Hilbert space for $\mathscr{M}$, which we denote $H_\epsilon\mathscr{M}$ for simplicity, as

$$(4.6) \qquad H_\epsilon\mathscr{M} = H(\nabla\cdot;\mathbb{M}) \times H(\nabla \cdot \epsilon; \mathbb{M})$$

These spaces are naturally endowed with the norms:

$$(4.7) \quad \|\omega\|_{H(\nabla\cdot\epsilon)}^2 := \|\omega\|^2 + \|\nabla \cdot (\epsilon\omega)\|^2, \quad \|(\sigma,\omega)\|_{H_\epsilon\mathscr{M}}^2 := \|\sigma\|_{H(\nabla\cdot)}^2 + \|\omega\|_{H(\nabla\cdot\epsilon)}^2$$

With this choice of spaces, we establish sufficient conditions for well-posedness of (4.4) in the following three lemmas:

LEMMA 4.1 (Continuity). *The bilinear forms $\left(\tilde{\mathscr{A}} p, r\right)$ and $(\mathscr{S}_\epsilon p, u)$ are continuous with continuity constants independent of ϵ.*

Proof. Let $p = (\sigma,\ \omega) \in H_\epsilon\mathscr{M}$ and $r = (\tau, \rho) \in H_\epsilon\mathscr{M}$. We first note the following trivial bound:

$$(4.8) \qquad \|p\|_{H_\epsilon\mathscr{M}}^2 = \|\sigma\|_{H(\nabla\cdot)}^2 + \|\omega\|_{H(\nabla\cdot\epsilon)}^2 \geq \|\sigma\|^2 + \|\omega\|^2 = \|p\|^2$$

From Assumptions 2.1 and 2.3, we now obtain the continuity:

$$\left(\tilde{\mathscr{A}} p, r\right) \leq A_\sigma^+ \|\sigma\|\|\tau\| + A_\omega^+ \|\omega\|\|\rho\| \leq \max\left(A_\sigma^+, A_\omega^+\right) \|p\|\|r\|$$
$$(4.9) \qquad\qquad\qquad \leq \max\left(A_\sigma^+, A_\omega^+\right) \|p\|_{H_\epsilon\mathscr{M}} \|r\|_{H_\epsilon\mathscr{M}}$$

Similarly, for $(\mathscr{S}_\epsilon p, u)$ with $p = (\sigma,\ \omega) \in H_\epsilon\mathscr{M}$ and $u = (u, r) \in L^2\mathscr{V}$, we obtain the following bound:

$$(4.10) \qquad (\mathscr{S}_\epsilon p, u) \leq \|\nabla \cdot \sigma\|\|u\| + \|S\sigma\|\|r\| + \|\nabla \cdot (\epsilon\omega)\|\|r\| \leq C\|p\|_{H_\epsilon\mathscr{M}}\|u\| \qquad \square$$

LEMMA 4.2 (Coercivity). *The bilinear form $\left(\tilde{\mathscr{A}} p, r\right)$ is coercive on the null-space of $\mathscr{S}_\epsilon$ in $H_\epsilon\mathscr{M}$, with a coercivity constant independent of ϵ.*

Proof. Consider any $p = (\sigma,\omega) \in H_\epsilon\mathscr{M}$ with $\mathscr{S}_\epsilon p = 0$. Then both $\nabla \cdot \sigma = 0$ and $\nabla \cdot \epsilon\omega = S\sigma$. Thus

$$(4.11) \qquad \|p\|_{H_\epsilon\mathscr{M}}^2 = \|\sigma\|_{H(\nabla\cdot)}^2 + \|\omega\|_{H(\nabla\cdot\epsilon)}^2 = \|\sigma\|^2 + \|\omega\|^2 + \|S\sigma\|^2 \leq C\|p\|^2$$

Where the constant C only depends on the bounded operator S. From Assumptions 2.1 and 2.3, we then obtain coercivity:

$$\left(\tilde{\mathscr{A}} p, p\right) \geq A_\sigma^- \|\sigma\|^2 + A_\omega^- \|\omega\|^2 \geq \min\left(A_\sigma^-, A_\omega^-\right)\|p\|^2 \geq C^{-1} \min\left(A_\sigma^-, A_\omega^-\right)\|p\|_{H_\epsilon\mathscr{M}}^2 \quad \square$$

LEMMA 4.3 (LBB: inf-sup). *The following inf-sup condition holds, with β independent of ϵ:*

$$(4.12) \qquad \inf_{u \in L^2\mathscr{V}} \sup_{p \in H_\epsilon\mathscr{M}} \frac{(\mathscr{S}_\epsilon p, u)}{\|u\|\|p\|_{H_\epsilon\mathscr{M}}} \geq \beta > 0$$

Proof. We first spell out the inf-sup statement for $u = (u, r)$ and $p = (\sigma, \omega)$:

$$(4.13) \qquad \inf_{(u,r) \in L^2 \mathscr{V}} \sup_{\substack{\sigma \in H(\nabla \cdot; \mathbb{M}) \\ \omega \in H(\nabla \cdot \epsilon; \mathbb{M})}} \frac{-(\nabla \cdot \sigma, u) + (S\sigma, r) - (\nabla \cdot (\epsilon \omega), r)}{(\|u\|^2 + \|r\|^2)^{\frac{1}{2}} (\|\sigma\|_{H(\nabla \cdot)} + \|\omega\|_{H(\nabla \cdot \epsilon)})^{\frac{1}{2}}}$$

We can obtain a lower bound on expression (4.13) by considering in the supremum the subspace of $H_\epsilon \mathscr{M}$ where $\omega = 0$, i.e. we have the bound

$$\inf_{(u,r) \in L^2 \mathscr{V}} \sup_{\substack{\sigma \in H(\nabla \cdot; \mathbb{M}) \\ \omega \in H(\nabla \cdot \epsilon; \mathbb{M})}} \frac{-(\nabla \cdot \sigma, u) + (S\sigma, r) - (\nabla \cdot (\epsilon \omega), r)}{(\|u\|^2 + \|r\|^2)^{\frac{1}{2}} (\|\sigma\|_{H(\nabla \cdot)} + \|\omega\|_{H(\nabla \cdot \epsilon)})^{\frac{1}{2}}}$$

$$(4.14) \qquad \geq \inf_{(u,r) \in L^2 \mathscr{V}} \sup_{\sigma \in H(\nabla \cdot; \mathbb{M})} \frac{-(\nabla \cdot \sigma, u) + (S\sigma, r)}{(\|u\|^2 + \|r\|^2)^{\frac{1}{2}} \|\sigma\|_{H(\nabla \cdot)}}$$

However, the right-hand side is known from the study of equations (2.11) to be bounded from below by a positive constant (see [2, 9]). Since ϵ does not appear in the right-hand side of (4.14), the lower bound is independent of ϵ, thus the lemma holds. $\qquad \square$

COROLLARY 4.4 (Well-posedness). *Let the weak mixed variational form of the Cosserat equations* (4.4) *be stated as follows: For $\ell \in L^2 \mathscr{V}$, find $(\tilde{p}, u) \in H_\epsilon \mathscr{M} \times L^2 \mathscr{V}$ such that:*

$$(4.15a) \qquad (\tilde{\mathscr{A}} \tilde{p}, p') - (u, \mathcal{S}_\epsilon p') = 0$$

$$(4.15b) \qquad (\mathcal{S}_\epsilon \tilde{p}, u') = (\ell, u')$$

for all $(p', u') \in H_\epsilon \mathscr{M} \times L^2 \mathscr{V}$. Then equations (4.15) *have a unique solution, satisfying*

$$(4.16) \qquad \|\tilde{p}\|_{H_\epsilon \mathscr{M}} + \|u\| \leq C \|\ell\|$$

Where the constant C is finite and independent of ϵ.

Proof. As a consequence of Lemmas 4.1, 4.2, and 4.3, standard theory applies (see e.g. [9, Thm. 4.2.3]), and the constant C only depends on the constants in the cited lemmas, which are independent of ϵ. $\qquad \square$

4.3. Weakly coupled mixed finite element approximation. Our objective is now to obtain a finite-dimensional approximation to equations (4.15). That is to say, for a family of finite-dimensional spaces $X_h^p \subset H(\nabla \cdot; \mathscr{V})$ and $X_h^u \subset L^2 \mathscr{V}$, both indexed by h, we approximate equations (4.15) by the following finite-dimensional system: For $\ell \in L^2 \mathscr{V}$, find $(\tilde{p}_h, u_h) \in X_h^p \times X_h^u$ such that:

$$(4.17a) \qquad (\tilde{\mathscr{A}} \tilde{p}_h, p_h') - (u_h, \mathcal{S}_\epsilon p_h') = 0 \qquad \qquad \forall p_h' \in X_h^p$$

$$(4.17b) \qquad (\mathcal{S}_\epsilon \tilde{p}_h, u_h') = (\ell, u_h') \qquad \qquad \forall u_h' \in X_h^u.$$

We will in this section consider a pair of *weakly coupled spaces*, in the sense that we will not require Definition 3.10 to hold. In particular, we will allow the range of $\mathcal{S}_\epsilon$ to be outside X_h^u. Nevertheless, we will still consider finite-dimensional subspaces of $H(\nabla \cdot; \mathscr{V})$ and $L^2 \mathscr{V}$, thus the discretization is conforming in the sense of function spaces.

In the proof of Lemma 4.3 we exploited the relationship between the Cosserat equations and those of linearized elasticity, in the sense that despite the somewhat different structure, they share the same inf-sup constant. This observation motivates the following lemma:

LEMMA 4.5 (Stability from elasticity). *Let a quadruplet of finite-dimensional spaces X_h^i, for $i \in \{\sigma, \omega, u, r\}$, be such that:*

$$(4.18) \qquad X_h^\sigma \subset H(\nabla\cdot; \mathbb{M}), \qquad X_h^\omega \subset H(\nabla \cdot \epsilon; \mathbb{M}), \qquad X_h^u \subset L^2\mathbb{V}, \qquad X_h^r \subset L^2\mathbb{V}$$

and let the spaces X_h^σ, X_h^u and X_h^r satisfy the discrete LBB condition for linearized elasticity with weak symmetry:

$$(4.19) \qquad \inf_{(u,r) \in X_h^u \times X_h^r} \sup_{\sigma \in X_h^\sigma} \frac{-(\nabla \cdot \sigma, u) + (S\sigma, r)}{(\|u\|^2 + \|r\|^2)^{\frac{1}{2}} \|\sigma\|_{H(\nabla\cdot)}} \geq \beta_h \geq \beta_0 > 0$$

Then the following LBB condition holds for $X_h^u := X_h^\sigma \times X_h^\omega$ and $X_h^p := X_h^u \times X_h^r$:

$$(4.20) \qquad \inf_{u \in X_h^u} \sup_{p \in X_h^p} \frac{(\mathcal{S}_\epsilon p, u)}{\|u\| \|p\|_{H_\epsilon \mathcal{M}}} \geq \beta_0 > 0,$$

with β_0 independent of ϵ.

Proof. The same arguments as in the proof of Lemma 4.3 hold. $\square$

The significance of Lemma 4.5 is that any mixed finite element method for elasticity with weakly imposed symmetry will be applicable as a building block for a weakly coupled mixed finite element method for the Cosserat equations. The literature on these methods is reviewed in detail in [9, Chapter 9]. Here, we will for concreteness consider the triplet proposed in [2], consisting of BDM elements for the stress, and equal-order discontinous Galerkin elements for displacement and rotation.

THEOREM 4.6 (Weakly coupled MFE for Cosserat equations). *For a family of simplicial partitions $\mathfrak{F}_h$ of the domain Ω, and for a non-negative integer k, define*

$$(4.21a) \qquad X_{h,k}^p := BDM_{k+1}(\mathfrak{F}_h)^3 \times W_k(\mathfrak{F}_h)^3, \qquad with\ W_k \in \{BDM_{k+1},\ RT_k\}$$

$$(4.21b) \qquad X_{h,k}^u := P_{-k}(\mathfrak{F}_h)^3 \times P_{-k}(\mathfrak{F}_h)^3.$$

and consider these spaces together with problem (4.17). Then the following statements hold:

1) *The stated choice of X_h^σ, X_h^u and X_h^r satisfies the assumptions of Lemma 4.5.*
2) *The resulting approximation is stable and convergent, independent of ϵ, satisfying*

$$\|\sigma - \sigma_h\|_{H(\nabla\cdot)} + \|\tilde{\omega} - \tilde{\omega}_h\|_{H(\nabla\cdot\epsilon)} + \|u - u_h\| + \|r - r_h\|$$
$$\leq C \left(\inf_{\sigma' \in X_h^\sigma} \|\sigma - \tau\|_{H(\nabla\cdot)} + \inf_{\tilde{\omega}' \in X_h^\omega} \|\tilde{\omega} - \tilde{\omega}'\|_{H(\nabla\cdot\epsilon)} \right.$$
$$\left. + \inf_{u' \in X_h^u} \|u - u'\| + \inf_{r' \in X_h^r} \|r - r'\| \right)$$

3) *If the solution of (4.15) has sufficient regularity, then the following optimal a priori convergence rates hold for $u = (u, r)$ and $p = (\sigma, \tilde{\omega})$:*

$$(4.22a) \quad \|\sigma - \sigma_h\|_{H(\nabla\cdot)} + \|\tilde{\omega} - \tilde{\omega}_h\|_{H(\nabla\cdot\epsilon)} + \|u - u_h\| + \|r - r_h\| = \mathcal{O}\left(h^{k+1}\right)$$

In terms of non-scaled variables, this translates to:

$$\|\sigma - \sigma_h\|_{H(\nabla\cdot)} + \|\epsilon^{-1}(\omega - \omega_h)\| + \|\nabla \cdot (\omega - \omega_h)\|$$
$$(4.22b) \qquad\qquad\qquad + \|u - u_h\| + \|r - r_h\| = \mathcal{O}\left(h^{k+1}\right)$$

Proof. 1) This is exactly [5, Theorem 11.4].

2) We need in addition to the inf-sup condition from Lemma 4.5, the coercivity and continuity of $(\tilde{\mathscr{A}}\mathcal{p}, \mathcal{r})$ for this choice of spaces. This is established in Lemma 4.7 below, and thus again we can invoke [9, Thm. 4.2.3 and 5.2.2], which ensures stability and the approximation property [2].

3) This follows from 2), together with the approximation properties of the spaces [9]. We only need to show that the approximation properties of X_h^ω hold in the weighted norm $\|\cdot\|_{H(\nabla\cdot\epsilon)}$. However, a product rule and Assumption 2.3 imply that this norm is weaker:

$$(4.23) \qquad \|\nabla \cdot \epsilon\tilde{\omega}\| \leq \|\epsilon\nabla \cdot \tilde{\omega}\| + \|(\nabla\epsilon)\cdot\tilde{\omega}\| \leq C\|\tilde{\omega}\|_{H(\nabla\cdot)}$$

In turn, the approximation properties of $X_h^\omega \subset H(\nabla\cdot; \mathbb{M})$ are directly inherited. □

The choice of weakly coupled spaces implies that the coercivity of $\tilde{\mathscr{A}}$ on the kernel of $\mathcal{S}_\epsilon$ is not an immediate consequence of Lemma 4.2. We therefore show this in a separate lemma.

LEMMA 4.7 (Discrete coercivity). *With the choice of finite-dimensional sub-spaces $X_{h,k}^p \subset H_\epsilon\mathcal{M}$ and $X_{h,k}^u \subset L^2\mathcal{V}$ given in Theorem 4.6, the bilinear form $\tilde{\mathscr{A}}$ is coercive on the null-space $\mathcal{S}_\epsilon$ on $X_{h,k}^u$, with coercivity constant independent of ϵ.*

Proof. Let $\mathcal{p} = (\sigma, \omega) \in X_{h,k}^p$ such that

$$(4.24) \qquad (\mathcal{S}_\epsilon\mathcal{p}, \mathcal{u}') = 0, \qquad\qquad \forall \mathcal{u}' \in X_{h,k}^u$$

We now note that $\nabla \cdot X_{h,k}^\sigma \subseteq X_{h,k}^u$, so that we may set $\mathcal{u}' = (\nabla\cdot\sigma, 0)$ in (4.24) and conclude $\nabla\cdot\sigma = 0$. Secondly, we set $\mathcal{u}' = (0, \epsilon_h\nabla\cdot u)$ in (4.24), where $\epsilon_h \in P_0(\mathfrak{F}_h)$ is a constant on each element equal to the maximum value of ϵ on that element. Thus $\epsilon_h\nabla\cdot\omega \in X_{h,k}^r$, and (4.24) evaluates to:

$$(4.25) \qquad (S\sigma + \nabla\cdot(\epsilon\omega), \epsilon_h\nabla\cdot\omega) = 0$$

Due to Assumption 2.3, we have sufficient regularity to apply the chain rule, and thus (4.25) is rewritten as

$$(4.26) \qquad (\epsilon\nabla\cdot\omega, \epsilon_h\nabla\cdot\omega) + (S\sigma + \omega\nabla\epsilon, \epsilon_h\nabla\cdot\omega) = 0.$$

We now calculate

$$(4.27) \qquad \begin{aligned} \|\sqrt{\epsilon\epsilon_h}\nabla\cdot\omega\|^2 = (\epsilon\nabla\cdot\omega, \epsilon_h\nabla\cdot\omega) &= -(S\sigma + \omega\nabla\epsilon, \epsilon_h\nabla\cdot\omega) \\ &\leq \|S\sigma + \omega\nabla\epsilon\|\|\epsilon_h\nabla\cdot\omega\| \end{aligned}$$

To proceed, we note that $\nabla\cdot\omega \in P_{-k}(\mathfrak{F}_h)^3$. Thus Lemma 4.8 below ensures that a finite C_k exists with which we rewrite (4.27) as

$$(4.28)$$
$$\|\epsilon_h\nabla\cdot\omega\| \leq C_k\|\sqrt{\epsilon\epsilon_h}\nabla\cdot\omega\| \leq C_k\|S\sigma + \omega\nabla\epsilon\|\frac{\|\epsilon_h\nabla\cdot\omega\|}{\|\sqrt{\epsilon\epsilon_h}\nabla\cdot\omega\|} \leq C_k^2\|S\sigma + \omega\nabla\epsilon\|$$

We now establish the following bound for $\not{p} = (\sigma, \omega) \in X_{h,k}^{\not{p}}$ satisfying (4.24):

$$\left(\|\not{p}\|_{H_\epsilon \mathscr{M}}^2 - \|\not{p}\|^2 \right)^{\frac{1}{2}} = \|\omega \nabla \epsilon + \epsilon \nabla \cdot \omega\|$$

$$\leq \|\omega \nabla \epsilon\| + \|\epsilon \nabla \cdot \omega\|$$

$$\leq \|\omega \nabla \epsilon\| + \|\epsilon_h \nabla \cdot \omega\|$$

$$\leq \|\omega \nabla \epsilon\| + C_k^2 \|S\sigma + \omega \nabla \epsilon\|$$

$$\leq \left(1 + C_k^2 \right) \|\omega \nabla \epsilon\| + C_k^2 \|S\sigma\|$$

$$\leq \left(1 + C_k^2 \right) C_\epsilon \|\omega\| + C_k^2 \|S\| \|\sigma\|$$

(4.29)
$$\leq \max \left(\left(1 + C_k^2 \right) C_\epsilon, C_k^2 \|S\| \right) \sqrt{2} \|\not{p}\|$$

Squaring both sides leads us to

(4.30)
$$\|\not{p}\|_{H_\epsilon \mathscr{M}}^2 \leq C \|\not{p}\|^2$$

where the constant C does not depend on ϵ, but only on the assumed upper bound on its gradient C_ϵ. Coercivity now follows by the same calculation as in Lemma 4.2.□

LEMMA 4.8. *For $k \geq 0$, ϵ satisfying Assumption 2.3, and $\epsilon_h \in P_0(\mathfrak{F}_h)$ the maximum value of ϵ on each simplex $\Delta \in \mathfrak{F}_h$, then a constant C_k exists such that:*

(4.31)
$$\sup_{u \in P_{-k}(\mathfrak{F}_h)^3} \frac{\|\epsilon_h u\|}{\|\sqrt{\epsilon \epsilon_h} u\|} \leq C_k < \infty$$

In particular, we have $C_0 = 2$ and $C_k = \sqrt{5}$ for $k \geq 1$.

Proof. To investigate this supremum, we take its square and rewrite as follows:

(4.32)
$$\sup_{u \in P_{-k}(\mathfrak{F}_h)^3} \frac{\|\epsilon_h u\|^2}{\|\sqrt{\epsilon \epsilon_h} u\|^2} = \sup_{u \in P_{-k}(\mathfrak{F}_h)^3} \frac{\int_\Omega \epsilon_h^2 |u|^2 dV}{\int_\Omega \epsilon \epsilon_h |u|^2 dV}$$

Recall now that by Assumption 2.3, $\epsilon \in P_1(\mathfrak{E})$, and consider the case where h is sufficiently small such that the triangulation $\mathfrak{F}_h$ resolves $\mathfrak{E}$. In turn, $\epsilon \in P_1(\mathfrak{F}_h)$. Therefore, on a single 3-simplex $\Delta \in \mathfrak{F}_h$, ϵ is a non-negative, linear function and the following mean value theorem applies

(4.33)
$$\frac{1}{4} \int_\Delta \epsilon_h dV \leq \int_\Delta \epsilon dV$$

Let ϖ_k denote the L^2 projection onto $P_{-k}(\mathfrak{F}_h)$. Using (4.33) with the fact that $\epsilon_h \in P_0(\mathfrak{F}_h)$ and $\epsilon \in P_1(\mathfrak{F}_h)$, we obtain

(4.34)
$$\frac{\int_\Omega \epsilon_h^2 |u|^2 dV}{\int_\Omega \epsilon \epsilon_h |u|^2 dV} = \frac{\int_\Omega \epsilon_h^2 \varpi_0(|u|^2) \, dV}{\int_\Omega \epsilon \epsilon_h \varpi_1(|u|^2) \, dV} \leq 4 \frac{\int_\Omega \epsilon \epsilon_h \varpi_0(|u|^2) \, dV}{\int_\Omega \epsilon \epsilon_h \varpi_1(|u|^2) \, dV}$$

For the case $k = 0$, we have that u is piecewise constant and thus $\varpi_1(|u|^2) = \varpi_0(|u|^2) = |u|^2$. Counteracting the square taken at the beginning of the proof, the combination of (4.32) and (4.34) gives us $C_0 = \sqrt{4} = 2$.

We continue with $k \geq 1$. Due to the projections ϖ_k, the supremum over $u \in P_{-k}(\mathfrak{F}_h)^3$ can be rewritten as a supremum over $\phi \in P_{-1}(\mathfrak{F}_h)$ with $\phi \geq 0$:

$$\sup_{u \in P_{-k}(\mathfrak{F}_h)} \frac{\int_\Omega \epsilon\epsilon_h \varpi_0(|u|^2) \, dV}{\int_\Omega \epsilon\epsilon_h \varpi_1(|u|^2) \, dV} = \sup_{\substack{\phi \in P_{-1}(\mathfrak{F}_h) \\ \phi \geq 0}} \frac{\int_\Omega \epsilon\epsilon_h \varpi_0(\phi) \, dV}{\int_\Omega \epsilon\epsilon_h \phi \, dV}$$

$$(4.35) \qquad\qquad = \sup_{\substack{\phi \in P_{-1}(\mathfrak{F}_h) \\ \phi \geq 0}} \frac{\int_\Omega \varpi_0(\epsilon)\epsilon_h \varpi_0(\phi) \, dV}{\int_\Omega \epsilon\epsilon_h \phi \, dV}$$

Next, we consider a single simplex $\Delta \in \mathfrak{F}_h$ and let ϵ_i denote the value of ϵ at vertex $v_i \in \Delta$. Using the fact that both ϵ and ϕ are nonnegative, linear functions, we use the definition of the local mass matrix of $P_1(\Delta)$ in 3D to compute:

$$\int_\Delta \varpi_0(\epsilon)\varpi_0(\phi) \, dV = |\Delta| \frac{\sum_i \epsilon_i}{4} \frac{\sum_i \phi_i}{4} \leq \frac{5}{4} \left(|\Delta| \frac{\sum_i \epsilon_i \sum_i \phi_i + \sum_i \epsilon_i\phi_i}{20} \right)$$

$$(4.36) \qquad\qquad = \frac{5}{4} \int_\Delta \epsilon\phi \, dV.$$

Scaling both sides by ϵ_h and summing over all Δ, we combine (4.36) with (4.32), (4.34), and (4.35) to obtain the upper bound:

$$(4.37) \qquad \sup_{u \in P_{-k}(\mathfrak{F}_h)^3} \frac{\|\epsilon_h u\|^2}{\|\sqrt{\epsilon\epsilon_h} u\|^2} \leq 4 \sup_{\substack{\phi \in P_{-1}(\mathfrak{F}_h) \\ \phi \geq 0}} \frac{\int_\Omega \varpi_0(\epsilon)\epsilon_h\varpi_0(\phi) \, dV}{\int_\Omega \epsilon\epsilon_h\phi \, dV} \leq 4 \left(\frac{5}{4} \right) = C_k^2$$

In turn, we obtain $C_k = \sqrt{5}$ for $k \geq 1$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

Remark 4.9 (Improved convergence). It is natural to ask whether improved convergence results similar to Theorem 3.15 and 3.16 can be developed for the weakly coupled discretizations. Inspecting the proofs of these theorems, it is clear that the improved convergence results can not in general be expected to be parameter robust with respect to ϵ for the weakly coupled method. We will therefore not detail improved convergence results in this section, but summarize the observed improved convergence rates obtained in Section 5.

Remark 4.10 (Computational cost). The weakly coupled discretization given in Theorem 4.6 is nominally a cheaper method than the strongly coupled discretization in Theorem 3.14 for a given element degree k. This is due to the fact that lower-order spaces are employed for the rotation variables, thus more evenly balancing the computational cost between displacements and rotations, leading to a lower total number of degrees of freedom. On the other hand, the strongly coupled BDM method enjoys some form of improved convergence rates of all variables, and is thus comparatively speaking a $k+1$-order method. The choice of most efficient discretization for a given problem is thus a balance of considerations.

Remark 4.11 (Elasticity). Finally, we emphasize that by construction, Theorem 4.6 contains a stable discretization of linear elasticity, coinciding with the so-called Arnold-Falk-Winther elements [2]. Thus, if $\epsilon = 0$ on any nondegenerate subdomain $\Omega_0 \subseteq \Omega$, then for any choice of finite-dimensional sub-space X_h^ω the discrete solution is $\tilde{\omega}_h = 0$ om Ω_0. Moreover, the remaining variables still satisfy all properties shown in Theorem 4.6. Finally, the elastic discretization is robust in the limit of incompressible materials [2, 9], a property that is inherited by the method when applied to Cosserat materials.

5. Numerical verification. We verify the performance of the MFE discretizations proposed above in terms of convergence rates and robustness. We consider three examples. In the first and second examples, we assess the convergence and robustness relative to the material properties of the couple stress and linear stress, respectively. In the third example, we consider a choice of material parameters such that a pure linear elastic medium is modeled in parts of the domain, thus verifying the applicability of the weakly coupled MFE discretization.

Results are presented for the discretization with strong coupling detailed in Theorem 3.14, with both choices of couple stress space, i.e. $W_k = RT_k$ and $W_k = BDM_{k+1}$. We abbreviate this discretization as "SC-RT" and "SC-BDM". Results are also presented for the discretization with weak coupling detailed in Theorem 4.6, also with both choices of couple-stress space, which are abbreviated as "WC-RT" and "WC-BDM".

Throughout the section, we will consider an isotropic medium according to Example 2.5, with material parameters:

$$(5.1) \qquad \mu_\sigma = 1, \qquad\qquad \mu_\omega = 1, \qquad\qquad \lambda_\omega = 0.$$

The values of λ_σ and ϵ, will be varied to validate robustness, and are thus specified in the various subsections. Nevertheless, we note that for $0 \leq \lambda_\sigma < \infty$, these material parameters clearly satisfy Assumption 2.1, and further satisfy Assumption 2.2 if $\epsilon > 0$. Finally, for spatially varying (piecewise linear) $\epsilon \in [0, 1]$, Assumption 2.3 holds.

When assessing the convergence of the methods, the norms given in Theorems 3.14 and 4.6 are used for the strongly and weakly coupled discretizations, respectively. When assessing the improved convergence of the methods, we use method-specific norms, consolidating both the theoretical results (Theorems 3.15 and 3.16) as well as the numerical results. For the strongly coupled methods, the norms in which we measure improved convergence are given by:

$$(5.2a) \qquad \|(p, u)\|^*_{\mathrm{SC-RT}} = \|\omega\| + \{\|r\| + \|\varpi_{h,k} u\|\}$$

$$(5.2b) \qquad \|(p, u)\|^*_{\mathrm{SC-BDM}} = \left\{ \|\sigma\| + \|\omega\|_{H(\nabla\cdot;\mathbb{M})} + \|r\| + \|\varpi_{h,k} u\| \right\}$$

Here the braces around terms in the norm indicates that improved convergence is proved in the cited theorems, while for the remaining term improved convergence is seen numerically, but has not been proved. For the weakly coupled methods, the improved convergence is not in general ϵ-robust, and we therefore specify the norms of improved convergence in the relevant sections.

For all cases, we consider as domain a unit cube domain, with manufactured solutions in $C_0^\infty \mathcal{V}$. In Sections 5.1 and 5.2, a regular triangulation is employed as illustrated in Figure 1. In section 5.3, the triangulation is adapted to resolve the structure of $\epsilon(x)$.

It is important to remark that due to the saddle-point structure of the formulation, efficient preconditioning is essential for the numerical performance of the implementation. To this end, we have implemented the operator-based norm-equivalent preconditioning [21], adapting the approach detailed for the Biot equations [7]. The details of the preconditioner are given in Appendix A. The preconditioner is robust with respect to the material and discretization parameters considered herein.

Python scripts were employed to generate all the numerical values and figures in this study. For each numerical example, the table data is generated by a dedicated script. For reproducing numerical results and figures, a Docker image was created, serialized, and made available for download [15].
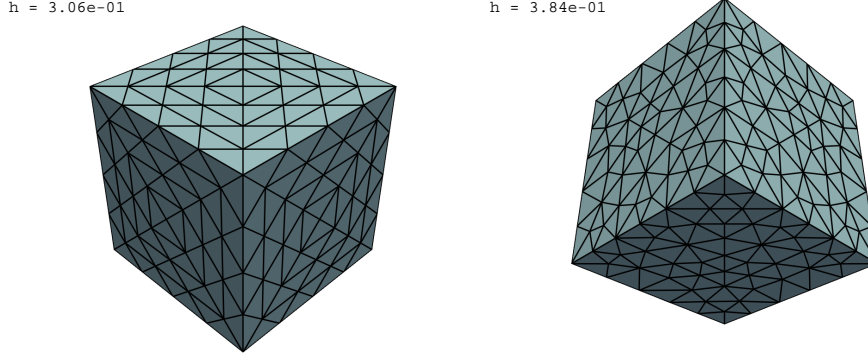
h = 3.06e-01   h = 3.84e-01

Fig. 1. *Examples of grid used in the numerical validation. Left: Grid for Section 5.1 and 5.2. Right: Grid for Section 5.3.*

5.1. Verification of expected convergence rates and stability for $\epsilon > 0$. In this section, we validate the stated *a priori* convergence rates of both the strongly and weakly coupled MFE methods detailed in Theorem 3.14 and 4.6. In this subsection, we fix $\lambda_\sigma = 1$, while $\epsilon > 0$ will be considered constant in space. We present results for the two lowest-order spaces, $k = 0, 1$, for both the weakly and strongly coupled MFE approximations.

We consider the unit cube domain $\Omega = [0,1]^3$, with an exact solution $(u, r) \in C_0^\infty \mathscr{V}$ defined by:

$$\text{(5.3a)} \qquad u(x) = \sum_{i=1}^{3} \sin\left(\pi x_i\right) \left(1 - x_{i+1}\right) x_{i+1} \left(1 - x_{i-1}\right) x_{i-1} \boldsymbol{e}_i$$

$$\text{(5.3b)} \qquad r(x) = \sum_{i=1}^{3} \left(1 - x_i\right) x_i \sin\left(\pi x_{i+1}\right) \sin\left(\pi x_{i-1}\right) \boldsymbol{e}_i$$

where as usual the indices are understood modulo 3, and $\boldsymbol{e}_i = \nabla x_i$ is the unit vector in the i-th coordinate direction. The solution is chosen to have high regularity, satisfy zero Dirichlet boundary conditions, and contain a mix of polynomial and geometric terms to reduce the chance of spurious super-convergence phenomena.

The convergence results for all methods are shown in Figure 2, for both $k = 0$ and $k = 1$. Considering first the results with $\epsilon = 1$, we note that all methods converge according to the expected rates. Moreover, considering the results for $\epsilon = \left\{10^{-2}, 10^{-4}\right\}$, we see that the weakly coupled method is robust for varying ϵ, while the strongly coupled method SC-RT is as expected not robust, and a large error is observed on coarse grids. The strongly coupled method SC-BDM is also not robust from a theoretical perspective, however, within the range of ϵ considered here, it performs well in practice, and no loss of accuracy is seen at any grid level. Keeping in mind that $A_\omega \sim \epsilon^{-2}$, this suggests that the SC-BDM method is suitable for a wide range of relative strengths between the stress and couple-stress parameters.

We also report improved convergence results in Figure 3. For the strongly coupled discretizations, the norms plotted are given in equations (5.2). For the weakly coupled methods, we report the results for only the error in the projected displacement:

$$\text{(5.4)} \quad \|(p - p_h, u - u_h)\|^*_{\text{WC-RT}} = \|(p - p_h, u - u_h)\|^*_{\text{WC-BDM}} = \|\varpi_{h,k} u - u_h\|$$

As expected, the weakly coupled methods enjoy good properties of improved convergence, and indeed for the SC-BDM method, all variables enjoy some form of improved convergence, which appears to ϵ-robust in the range considered herein, so that the method can essentially be considered one degree higher than its nominal rate. The improved convergence of the SC-RT method is also according to expectation, but is not ϵ-robust in magnitude. The SC-RT method also enjoys an improved convergence in the couple-stress, which is not completely unexpected from literature, but which is known to be difficult to prove in general [8]. Finally, our numerical results reveal that the weakly coupled methods only enjoy improved convergence properties with respect to the projected displacement.

Overall, the observed convergence of the methods is summarized as follows:

OBSERVATION 5.1 (Numerical convergence and robustness for spatially constant ϵ). *Under the conditions of Assumption 2.1 and 2.2, both the strongly coupled and weakly coupled MFE methods are convergent, satisfying the rates stated in Theorems 3.14, 3.15, 3.16 and 4.6. For the weakly coupled method, the observed convergence rates are fully robust with respect ϵ, as expected from theory. For the methods with strong coupling, a sensitivity to small values of ϵ is observed, however, the SC-BDM method appears to be more robust than SC-RT in practice. Furthermore, improved convergence rates are observed numerically for additional variables as specified in equations (5.2) and (5.4).*
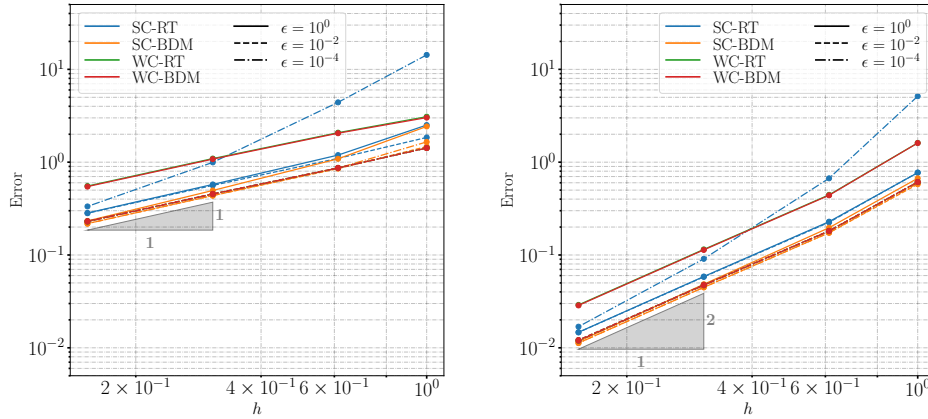

FIG. 2. *Convergence data for the methods as summarized in Theorem 3.14 and 4.6, for spatially constant $\epsilon \in \left\{1, 10^{-2}, 10^{-4}\right\}$ and $\lambda_\sigma = 1$. The error is calculated according to equations (3.31) and (4.22a), for the strongly coupled and weakly coupled methods, respectively. Left panel shows convergence results for $k = 0$, right panel for $k = 1$.*

5.2. Verification of expected convergence rates and stability for nearly incompressible materials.
In this section, we verify the robustness of the strongly and weakly coupled MFE methods for in the limit of nearly incompressible materials, as outlined in Remark 2.6. In this subsection, we fix $\epsilon = 1$, while λ_σ will be considered a parameter constant space. As in the previous section, we present results for the two-lowest-order spaces, $k = 0, 1$, for both the weakly and strongly coupled MFE approximations.
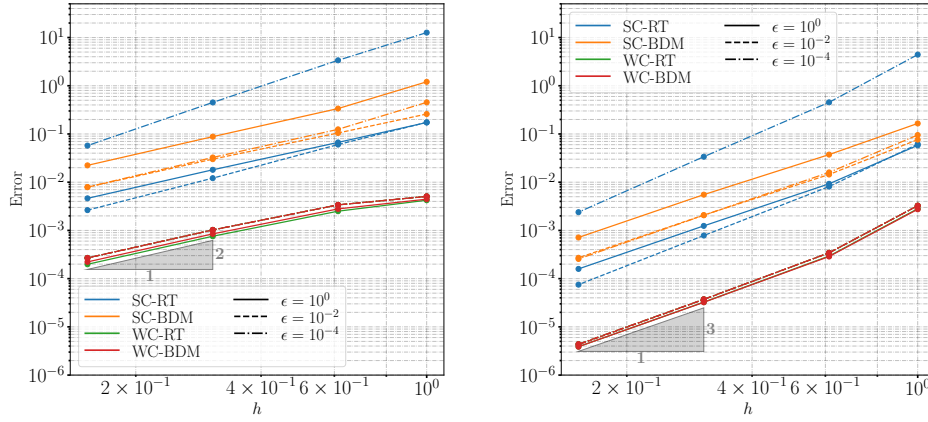
Fig. 3. *Observed improved convergence rates seen using the norms stated in equations* (5.2) *and* (5.4) *for spatially constant* $\epsilon \in \left\{1, 10^{-2}, 10^{-4}\right\}$ *and* $\lambda_\sigma = 1$. *Left panel shows convergence results for* $k = 0$, *right panel for* $k = 1$.

We again consider the unit cube domain $\Omega = [0,1]^3$. However, when considering incompressible materials, it is important to note that the solution will satisfy $\nabla \cdot u = 0$ in the limit of $\lambda_\sigma \to 0$. This property does not hold for the expression in (5.3a), therefore we will in this section use the following expression for the exact solution of $u \in C_0^\infty \mathbb{V}$:

$$(5.5) \qquad u(x) = \nabla \times \left((\boldsymbol{e}_2 + \boldsymbol{e}_3) \prod_{i=1}^{3} \sin^2 (\pi x_i) \right)$$

Clearly, $u(x)$ defined in (5.5) will satisfy $\nabla \cdot u = 0$, and moreover, it is easy to verify that it satisfies the Dirichlet boundary conditions due to the presence of quadratic sine functions.

The results are shown as in the previous section both for convergence in all variables (Figure 3). The improved convergence rates (Figure 4) are as before evaluated for the strongly coupled discretizations in the norms stated in (5.2). For the weakly coupled methods according, we use the norms:

$$(5.6\text{a}) \qquad \|(\boldsymbol{p}, \boldsymbol{u})\|^*_{\text{WC}-\text{RT}} = \|\sigma\| + h^{1/2}\|\tilde{\omega}\|_{H(\nabla\cdot\epsilon)} + \|\varpi_{h,k}r\| + \|\varpi_{h,k}u\|$$

$$(5.6\text{b}) \qquad \|(\boldsymbol{p}, \boldsymbol{u})\|^*_{\text{WC}-\text{BDM}} = \|\sigma\| + \|\tilde{\omega}\| + \|\varpi_{h,k}r\| + \|\varpi_{h,k}u\|$$

Note the scaling of $h^{1/2}$ in front of the second term in the definition of the WC-RT norm. This reflects that we observe somewhere between a half and a full order of improved convergence for this variable.

We observe that all discretizations are fully robust with respect to variations in $\lambda_\sigma \in \left\{10^0, 10^2, 10^4\right\}$, and indeed the lines for the different values of λ_σ can hardly be distinguished in the plots. It is also noteworthy that for non-degenerate ϵ, the weakly coupled discretizations also enjoys almost as favorable convergence properties as their strongly coupled counterparts.

OBSERVATION 5.2 (Robustness for incompressibile materials). *Under the conditions of Assumption 2.1 and 2.2, both the strongly coupled and weakly coupled MFE*

methods are convergent, satisfying the rates stated in Theorems 3.14, 3.15, 3.16 and 4.6. Additional improved convergence rates are observed according to equations (5.2) and (5.6). All results are fully robust with respect to λ_σ, as expected from theory.
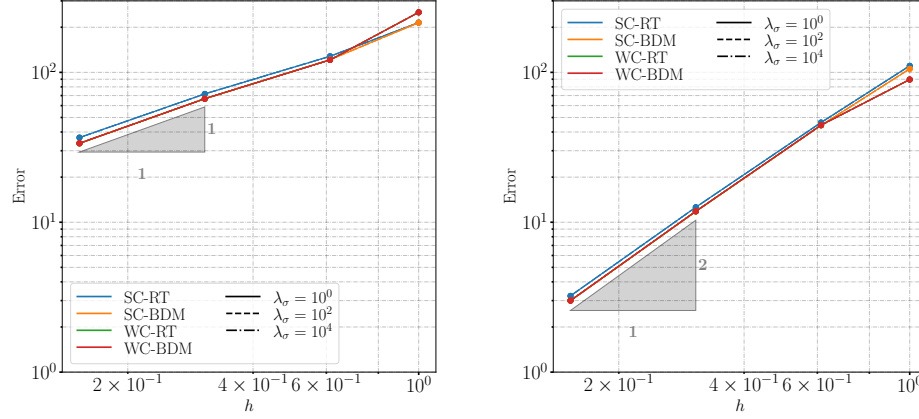

FIG. 4. *Convergence data for the methods as summarized in Theorem 3.14 and 4.6, for spatially constant $\lambda_\sigma \in \{1, 10^2, 10^4\}$ and $\epsilon = 1$. The error is calculated according to equations (3.31) and (4.22a), for the strongly coupled and weakly coupled methods, respectively. Left panel shows convergence results for $k = 0$, right panel for $k = 1$. Note that as the methods are fully robust with respect to λ_σ, the lines are overlapping within the resolution of the figure.*
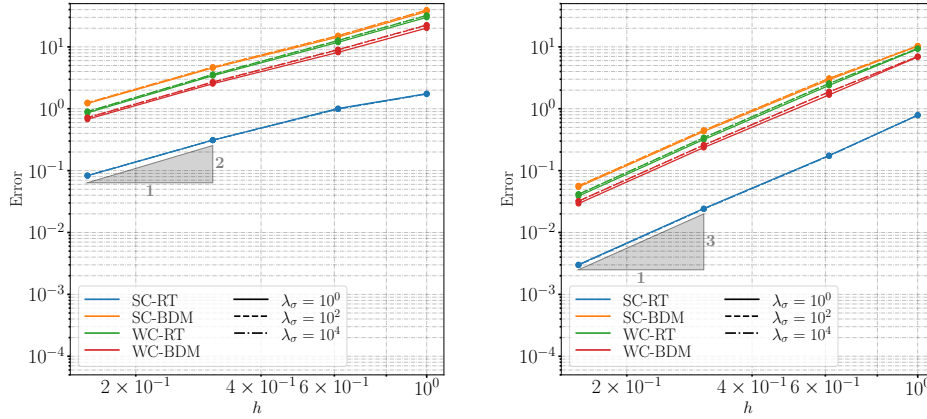

FIG. 5. *Super-convergence data for the strongly coupled methods as summarized in Theorem 3.14. for spatially constant $\lambda_\sigma \in \{1, 10^2, 10^4\}$ and $\epsilon = 1$. Super-convergence appears in the rotation and couple-stress variables, and is calculated according to (3.40). Left panel shows convergence results for $k = 0$, right panel for $k = 1$. Note that as the methods are fully robust with respect to λ_σ, the lines are overlapping within the resolution of the figure.*

5.3. Verification of robustness for degenerate ϵ. Theorem 4.6 ensures that the stability and approximation properties of the weakly coupled MFE discretization

is preserved in the presence of problems where regions of the contain fully degenerate Cosserat materials, i.e. regions of the domain where a standard linear elastic model is employed. In this final example, we therefore return to the same domain and manufactured solution as in Section 5.1, with $\lambda_\sigma = 1$. However, in this section we consider a spatially varying $\epsilon(x)$, with a nontrivial region where $\epsilon(x) = 0$. This thereby allows us to showcase the applicability of the weakly coupled MFE discretization to such composite materials, where regions of the material are a regular elastic material, and simultaneously verify the stability of the method in the presence of variability of ϵ, including the region where $\epsilon = 0$.

Concretely, we define the following spatially variable $\epsilon(x)$:

$$(5.7) \qquad \epsilon(x) = \min \left(1, \max \left(0, \max_{i=1\ldots3} (3x_i - 1) \right) \right)$$

See Figure 6 for a 2D illustration of ϵ restricted to any plane defined by $x_3 \in [0, 1/3]$ (left), and for a view of the solution of the couple stress variable (right). We note that by this construction, $\epsilon \in C^0$ and $0 \le \epsilon \le 1$. Moreover, ϵ is piecewise linear with $|\nabla \epsilon| \le 3$ almost everywhere and thereby satisfies the conditions of Assumption 2.3. Note also that $\epsilon = 0$ in the cube close to the origin defined by $\max_{i=1\ldots3} (x_i) \le 1/3$, thus in this region the constitutive model corresponds to linear elasticity.

The strongly coupled discretizations are not applicable to this problem, as $\|A_\omega\|$ is unbounded. Correspondingly, we only consider the weakly coupled discretizations of Theorem 4.6, with the results being presented in Figure 7. The error of the WC-RT and WC-BDM methods are indistinguishable in the plot, as the couple stress error is (as expected) not dominating for this problem. We observe that the presence of a degenerate region does not affect the convergence of the method, and the theoretically expected rates are achieved. It is important to emphasize that this result reflects the careful definition of norms, since in the elastic region the regularity of the rotations will in general only be $r \in L^2\mathbb{V}$. We summarize this as follows.

OBSERVATION 5.3 (Robustness in transition to linear elasticity). *Under the conditions of Assumption 2.1 and 2.3, the weakly coupled MFE discretizations are convergent and stable in the presence of degenerate regions of the domain where the Cosserat model reduces to linear elasticity.*
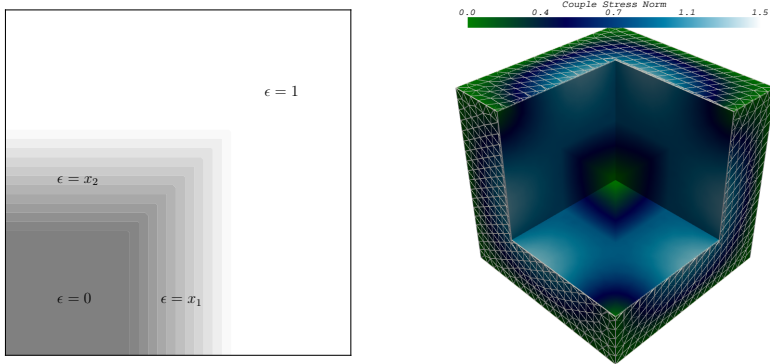

FIG. 6. *Left: Plot of spatially varying ϵ used in Section 5.3, as restricted to a plane defined by e.g. $x_i = 0$. Right: Plot of the (numerical) solution for the couple stress variable, seen facing the origin along the $x_1 = x_2 = x_3$ line. The blue square in the center of the cut-out indicates that the $\tilde{\omega}_h = 0$ in the region of pure elasticity.*
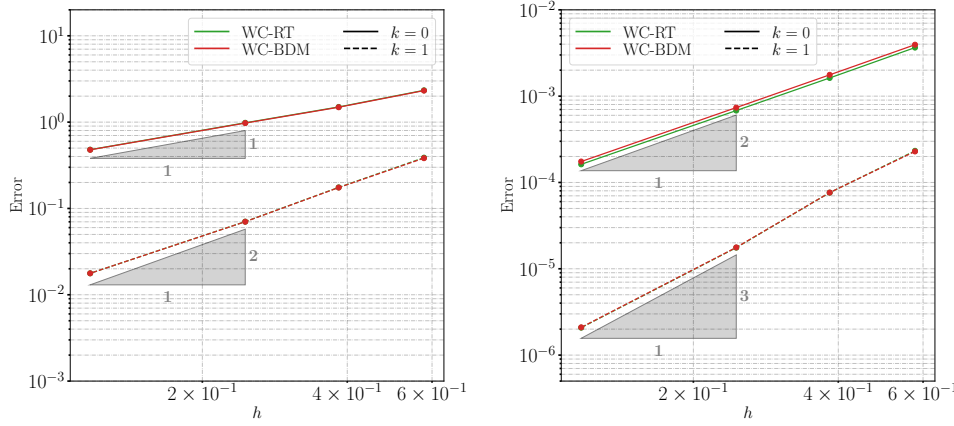
FIG. 7. *Left: Convergence data for the weakly coupled discretizations of Theorem 4.6 for spatially varying ϵ, and $k \in \{0, 1\}$. The error is calculated according to equations (4.22a). Right: Results showing improved convergence rates according to the norms defined in equation (5.6).*

Reproducible data and figures. All data and figures presented herein can be reproduced using a Docker image that has been created and made publicly available [15]. Source code is included in the Docker image for use, modification, and redistribution.

Appendix A. Norm-equivalent preconditioning of saddle-point systems.

In the case of high-order elements, especially in the strongly coupled variant, the use of direct solvers becomes infeasible due to excessive memory consumption. Thus, we use iterative solution methods with a suitable preconditioner. Following [7, 21], we construct the preconditioner according to the Riesz map associated with the Hilbert spaces in which the problems were analyzed.

We recall the weighted norms from Theorem 3.14 for which the well-posedness of the strongly coupled methods is shown, which we here denote by the subscript P:

$$\|(\boldsymbol{p}, \boldsymbol{u})\|_P^2 = \|\sigma_h\|_{A_\sigma}^2 + \|\nabla \cdot \sigma_h\|^2 + \|\omega_h\|_{A_\omega}^2 + \|\nabla \cdot \omega_h\|^2 + \|u_h\|^2 + \|r_h\|^2$$

Similarly, for the weakly coupled methods, we use the norm from Theorem 4.6, for simplicity also denoted by the subscript P:

$$\|(\boldsymbol{p}, \boldsymbol{u})\|_P^2 = \|\sigma_h\|_{A_\sigma}^2 + \|\nabla \cdot \sigma_h\|^2 + \|\omega_h\|_{\tilde{A}_\omega}^2 + \|\nabla \cdot (\epsilon \omega_h)\|^2 + \|u_h\|^2 + \|r_h\|^2$$

Let now P be the linear operator associated with the corresponding inner product, i.e. such that $\langle P(\boldsymbol{p}, \boldsymbol{u}), (\boldsymbol{p}, \boldsymbol{u})\rangle = \|(\boldsymbol{p}, \boldsymbol{u})\|_P^2$. This operator takes the following block

diagonal form:

$$
\text{(A.1)} \qquad\qquad P = \begin{pmatrix} P_\sigma & & & \\ & P_\omega & & \\ & & P_u & \\ & & & P_r \end{pmatrix}.
$$

Following the framework of [21], we choose the preconditioner to be an application of the inverse of P. Since the well-posedness of the problem was shown with constants independent of material and discretization parameters, the preconditioner will be robust with respect to these parameters.

For the examples of Sections 5.1 to 5.3, we use a P-preconditioned minimal residual method (MinRes) until an error tolerance of 10^{-9} is achieved. For all combinations of refinement, polynomial order, and material parameters, the number of iterations is bounded, as shown in figures 8 and 9 for Sections 5.1 and 5.3, respectively. In particular we note that while for the SC-RT method on some grids the solver requires additional iterations to converge, there is no indication of a trend of increase in iteration count with grid refinement.
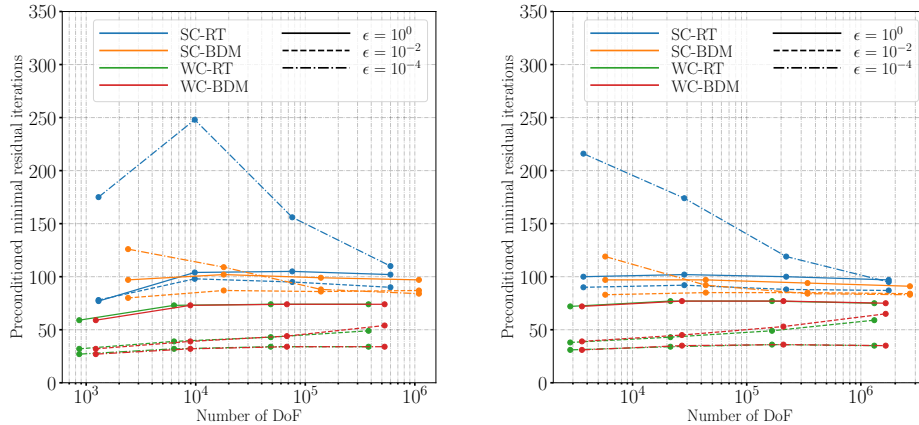

Fig. 8. *Number of preconditioned minimal residual iterations for the example presented in Section 5.1, shown as a function of the number of degrees of freedom. For the left and right panels, k is equal to 0 and 1 respectively.*

REFERENCES

[1] I. Ambartsumyan, E. Khattatov, J. M. Nordbotten, and I. Yotov, *A multipoint stress mixed finite element method for elasticity on simplicial grids*, SIAM Journal on Numerical Analysis, 58 (2020), pp. 630–656.

[2] D. Arnold, R. Falk, and R. Winther, *Mixed finite element methods for linear elasticity with weakly imposed symmetry*, Mathematics of Computation, 76 (2007), pp. 1699–1723.

[3] D. Arnold, R. Falk, and R. Winther, *Finite element exterior calculus: from hodge theory to numerical stability*, Bulletin of the American mathematical society, 47 (2010), pp. 281–354.

[4] D. N. Arnold, *Finite element exterior calculus*, Philadelphia: Society for Industrial and Applied Mathematics, 2018.

[5] D. N. Arnold, R. S. Falk, and R. Winther, *Finite element exterior calculus, homological techniques, and applications*, Acta numerica, 15 (2006), pp. 1–155.
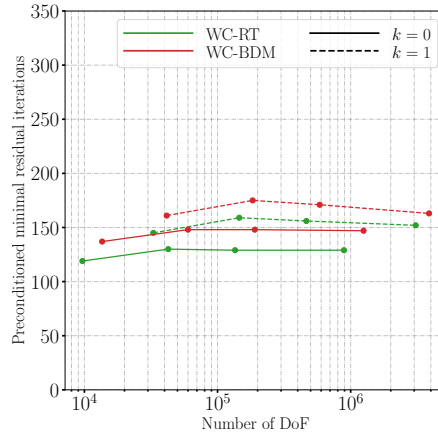
Fig. 9. *Number of preconditioned minimal residual iterations for the example presented in Section 5.3*

[6] D. N. Arnold and R. Winther, *Mixed finite elements for elasticity*, Numerische Mathematik, 92 (2002), pp. 401–419.

[7] T. Bærland, J. J. Lee, K.-A. Mardal, and R. Winther, *Weakly imposed symmetry and robust preconditioners for Biot's consolidation model*, Computational Methods in Applied Mathematics, 17 (2017), pp. 377–396.

[8] R. E. Bank and Y. Li, *Superconvergent recovery of Raviart–Thomas mixed finite elements on triangular grids*, Journal of Scientific Computing, 81 (2019), pp. 1882–1905.

[9] D. Boffi and F. Brezzi, *Mixed finite element methods and applications*, Springer, 2013.

[10] R. Bott and L. W. Tu, *Differential forms in Algebraic Topology*, New York: Springer Verlag, 1982.

[11] D. Braess, *Finite elements*, Cambridge University Press, 2007.

[12] J. Brüning and M. Lesch, *Hilbert complexes*, Journal of Functional Analysis, 108 (1992), pp. 88–132.

[13] E. M. P. Cosserat and F. Cosserat, *Théorie des corps déformables*, A. Hermann et fils, 1909.

[14] R. de Borst, *A generalisation of J2-flow theory for polar continua*, Computer Methods in Applied Mechanics and Engineering, 103 (1993), pp. 347–362.

[15] O. Duran, J. M. Nordbotten, and W. Boon, *Docker image: Mixed finite element methods for linear cosserat equations*, mar 2024, https://doi.org/10.5281/zenodo.10808503.

[16] M. Eastwood, *A complex from linear elasticity*, in Proceedings of the 19th Winter School "Geometry and Physics", Circolo Matematico di Palermo, 2000, pp. 23–29.

[17] W. Ehlers and S. Bidier, *Cosserat media*, Encyclopedia of Continuum Mechanics, (2020), pp. 436–446.

[18] P. Grisvard, *Elliptic problems in nonsmooth domains*, SIAM, 2011.

[19] M. E. Gurtin, *An Introduction to Continuum Mechanics*, San Diego: Academic Press, 1982.

[20] I. Hlaváček and M. Hlaváček, *On the existence and uniqueness of solution and some variational principles in linear theories of elasticity with couple-stresses. i: Cosserat continuum*, Aplikace matematiky, 14 (1969), pp. 387–410.

[21] K.-A. Mardal and R. Winther, *Preconditioning discretizations of systems of partial differential equations*, Numerical Linear Algebra with Applications, 18 (2011), pp. 1–40.

[22] P. Neff and J. Jeong, *A new paradigm: the linear isotropic Cosserat model with conformally invariant curvature energy*, Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik (ZAMM), 89 (2009), pp. 107–122.

[23] G. K. Pedersen, *Analysis now*, Springer Science & Business Media, 1989.

[24] A. Riahi and J. H. Curran, *Full 3d finite element Cosserat formulation with application in layered structures*, Applied mathematical modelling, 33 (2009), pp. 3450–3464.

[25] M. Spivak, *Calculus on Manifolds*, Reading, Massachusetts: Addison-Wesley Publishing Company, 1965.

[26] R. Temam and A. Miranville, *Mathematical modeling in continuum mechanics*, Cambridge
      University Press, 2000.
[27] C. Truesdell and W. Noll, *The non-linear field theories of mechanics*, Springer, 1965.
[28] A. Čap and K. Hu, *Bgg sequences with weak regularity and applications*, Foundations of
      Computational Mathematics, (2023), https://doi.org/10.1007/s10208-023-09608-9.