GAMMA-CONVERGENT LDG METHOD FOR LARGE BENDING DEFORMATIONS OF BILAYER PLATES

ANDREA BONITO, RICARDO H. NOCHETTO, AND SHUO YANG

ABSTRACT. Bilayer plates are slender structures made of two thin layers of different materials. They react to environmental stimuli and undergo large bending deformations with relatively small actuation. The reduced model is a constrained minimization problem for the second fundamental form, with a given spontaneous curvature that encodes material properties, subject to an isometry constraint. We design a local discontinuous Galerkin (LDG) method which imposes a relaxed discrete isometry constraint and controls deformation gradients at barycenters of elements. We prove Γ-convergence of LDG, design a fully practical gradient flow, which gives rise to a linear scheme at every step, and show energy stability and control of the isometry defect. We extend the Γ-convergence analysis to piecewise quadratic creases. We also illustrate the performance of the LDG method with several insightful simulations of large deformations, one including a curved crease.

1. INTRODUCTION

Bilayer plates are slender structures made of two thin layers of different materials glued together. These layers react differently to non-mechanical stimuli, such as thermal, electrical, and chemical actuation [31, 44, 32]. Bilayer plates can undergo large bending deformations using a small amount of energy, which makes them appealing at small and large scales alike. Amongst the many and broad applications of bilayer materials in engineering and biomedical science, we list drug delivery vesicles [29, 45], cell encapsulation devices [46], sensors [35] and self-deployable sun sails [34].

We model bilayer plates as thin 3d hyper-elastic bodies as depicted in Fig. 1. Exploiting their relatively small thickness, two dimensional plate models for the mid-plane deformation $\mathbf{y}(\Omega)$, $\Omega \subset \mathbb{R}^2$, are derived and analyzed in [41, 42]; we also refer to [8] for a formal dimension reduction argument. The plates equilibria

are characterized as solutions to a nonlinear minimization problem with a noncon-vex constraint expressing the plates ability to bend without stretching or shearing. Therefore, distances within the midplane remain unchanged thereby resulting in isometric deformations.
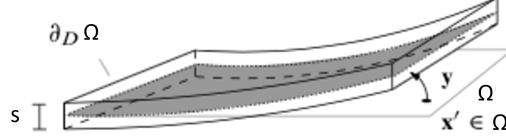

FIGURE 1. Bilayer plates: $\Omega \times (-s/2, s/2)$, $\Omega \subset \mathbb{R}^2$ is the mid-plane (bounded Lipschitz domain) and s is the thickness parameter. The sets $\Omega \times (-s/2, 0)$ and $\Omega \times (0, s/2)$ represent the two undeformed layers of different materials.

1.1. **Problem statement.** The plate deformation $\mathbf{y} : \Omega \to \mathbb{R}^3$ must belong to the following *admissible set* $\mathbb{A}$, which prevents shearing and stretching within the surface $\mathbf{y}(\Omega)$ and imposes possible boundary conditions:

(1) $$\mathbb{A} := \left\{ \mathbf{y} \in \left[H^2(\Omega) \right]^3 : \quad \mathrm{I}[\mathbf{y}] = I_2 \ \text{ in } \Omega, \quad \mathbf{y} = \boldsymbol{\varphi}, \ \nabla \mathbf{y} = \Phi \text{ on } \Gamma^D \right\},$$

where I_2 is the 2×2 identity matrix and

(2) $$\mathrm{I}[\mathbf{y}] := \nabla \mathbf{y}^T \nabla \mathbf{y}$$

is the first fundamental form of $\mathbf{y}(\Omega)$. We assume that $\Gamma^D \subset \partial\Omega$ is nonempty and open and $\boldsymbol{\varphi} \in \left[H^2(\Omega) \right]^3$ and $\Phi \in \left[H^1(\Omega) \right]^{3 \times 2}$ are given and are compatible with the isometry constraint, namely $\Phi = \nabla \boldsymbol{\varphi}$ and $\Phi^T \Phi = I_2$ on Γ_D; thus $\mathbb{A}$ is non-empty. Moreover, condition (2) entails that $\{\partial_i \mathbf{y}\}_{i=1}^2$ is an orthonormal basis of the tangent plane to $\mathbf{y}(\Omega)$ and its unit normal $\boldsymbol{\nu}$ can be written as

(3) $$\boldsymbol{\nu} := \frac{\partial_1 \mathbf{y} \times \partial_2 \mathbf{y}}{|\partial_1 \mathbf{y} \times \partial_2 \mathbf{y}|} = \partial_1 \mathbf{y} \times \partial_2 \mathbf{y}.$$

Although we will present simulations in Section 6 for both Dirichlet boundary conditions (i.e. $\Gamma^D \neq \varnothing$) and free boundary conditions (i.e. $\Gamma^D = \varnothing$), we focus our presentation on the former for convenience. We emphasize that the analysis of the latter follows from that in this paper. The modifications are in the spirit of [15], where we analyze the LDG method for prestrained plates with free boundary conditions. Consequently, we do not include details to avoid repetitions.

Equilibrium configurations of bilayer plates are solutions $\mathbf{y} \in \mathbb{A}$ of the following constrained minimization problem

(4) $$\min_{\mathbf{y} \in \mathbb{A}} E[\mathbf{y}] := \min_{\mathbf{y} \in \mathbb{A}} \frac{1}{2} \int_\Omega \left| \mathrm{II}[\mathbf{y}] - Z \right|^2,$$

where $\mathrm{II}[\mathbf{y}]$ is the second fundamental form of $\mathbf{y}(\Omega)$

(5) $$\mathrm{II}[\mathbf{y}] := \left(\partial_{ij} \mathbf{y} \cdot \boldsymbol{\nu} \right)_{ij=1}^2 = \left(\partial_{ij} \mathbf{y} \cdot (\partial_1 \mathbf{y} \times \partial_2 \mathbf{y}) \right)_{ij=1}^2,$$

and $Z \in [L^\infty(\Omega)]^{2 \times 2}$ is a *spontaneous curvature* which encodes the material properties of the bilayer plates. In fact, Z forces the plate $\mathbf{y}(\Omega)$ to bend so that $\mathrm{II}[\mathbf{y}]$ gets as close as possible to Z. If the material is homogenous and isotropic, then the spontaneous curvature is diagonal, i.e. $Z = \alpha I_2$ with a constant α depending on the materials parameters. In particular, when the two layers are identical, $Z = \mathbf{0}$ and the model reduces to a *single layer plate* [4, 18], which coincides with the classical (nonlinear) Kirchhoff plate theory.

Thanks to the isometry constraint $\mathrm{I}[\mathbf{y}] = I_2$, the energy functional $E[\mathbf{y}]$ can be further simplified. Recall that for isometries, there holds [5]

(6) $$\left|\mathrm{II}[\mathbf{y}]\right|^2 = \left|D^2\mathbf{y}\right|^2 = \left|\Delta\mathbf{y}\right|^2 = \left(\operatorname{tr}\mathrm{II}[\mathbf{y}]\right)^2,$$

whence expanding the square in (4) and using (5) and (6) yields

(7) $$E[\mathbf{y}] = \frac{1}{2}\int_\Omega \left|D^2\mathbf{y}\right|^2 - \sum_{i,j=1}^{2}\int_\Omega \partial_{ij}\mathbf{y}\cdot(\partial_1\mathbf{y}\times\partial_2\mathbf{y})Z_{ij} + \frac{1}{2}\int_\Omega |Z|^2.$$

Furthermore, since $\frac{1}{2}\int_\Omega |Z|^2$ does not depend on $\mathbf{y}$, minimizing the energy in (7) over $\mathbb{A}$ is equivalent to minimizing the reduced energy

(8) $$E[\mathbf{y}] := \frac{1}{2}\int_\Omega \left|D^2\mathbf{y}\right|^2 - \sum_{i,j=1}^{2}\int_\Omega \partial_{ij}\mathbf{y}\cdot(\partial_1\mathbf{y}\times\partial_2\mathbf{y})Z_{ij},$$

over $\mathbb{A}$; we keep the same notation for the energies in (7) and (8) for simplicity. The effect of the layers mismatch appears in the cubic term leading to a nonlinear Euler-Lagrange equation for the equilibrium deformation $\mathbf{y}$, namely

(9)
$$0 = \delta E[\mathbf{y}](\mathbf{v}) := \int_\Omega D^2\mathbf{y} : D^2\mathbf{v} - \sum_{i,j=1}^{2}\int_\Omega \partial_{ij}\mathbf{v}\cdot(\partial_1\mathbf{y}\times\partial_2\mathbf{y})Z_{ij}$$
$$- \sum_{i,j=1}^{2}\int_\Omega \partial_{ij}\mathbf{y}\cdot(\partial_1\mathbf{v}\times\partial_2\mathbf{y})Z_{ij} - \sum_{i,j=1}^{2}\int_\Omega \partial_{ij}\mathbf{y}\cdot(\partial_1\mathbf{y}\times\partial_2\mathbf{v})Z_{ij},$$

where $\mathbf{v} \in [H_0^2(\Omega)]^3$ is an arbitrary test function. For later use, we also introduce a notation for the single layer bending energy

(10) $$B[\mathbf{y}] := \frac{1}{2}\int_\Omega \left|D^2\mathbf{y}\right|^2$$

and the cubic term

(11) $$C[\mathbf{y}] := \sum_{i,j=1}^{2}\int_\Omega \partial_{ij}\mathbf{y}\cdot(\partial_1\mathbf{y}\times\partial_2\mathbf{y})Z_{ij},$$

so that

$$E[\mathbf{y}] = B[\mathbf{y}] - C[\mathbf{y}].$$

We emphasize that the cubic term C satisfies

$$|C[\mathbf{y}]| \leq \|\mathbf{y}\|_{H^2(\Omega)}\|\nabla\mathbf{y}\|_{L^2(\Omega)}\|\nabla\mathbf{y}\|_{L^\infty(\Omega)}\|Z\|_{L^\infty(\Omega)}$$

and $I[\mathbf{y}] = I_2$ implies $\|\nabla \mathbf{y}\|_{L^\infty(\Omega)} \lesssim \|I[\mathbf{y}]\|_{L^\infty(\Omega)} \lesssim 1$, whence

$$|C[\mathbf{y}]| \lesssim \|\mathbf{y}\|_{H^2(\Omega)}^2.$$

Since the discrete deformation $\mathbf{y}_h$ is piecewise polynomial, our numerical method cannot guarantee that $\mathbf{y}_h$ satisfies the isometry constraint $I[\mathbf{y}_h] = I_2$ everywhere in Ω. We choose to enforce a relaxed constraint solely at the barycenter of elements. This is a chief ingredient of our LDG method and is inspired by Bartels and Palus for Kirchhoff elements [10].

1.2. **Previous numerical methods.** There are several finite element methods available for the numerical simulation of bilayers plates [8, 7, 10, 17]. In all of them, the isometry constraint $I[\mathbf{y}] = I_2$ is linearized at $\mathbf{y}$

(12) $$L[\mathbf{v}; \mathbf{y}] := \nabla \mathbf{v}^T \nabla \mathbf{y} + \nabla \mathbf{y}^T \nabla \mathbf{v} = \mathbf{0},$$

and tangential variations $\mathbf{v}$ are evolved within a gradient flow that decreases the energy $E[\mathbf{y}]$ and is favored for its robustness.

The gradients of Kirchhoff finite elements are uniquely defined at the mesh vertices, which is where (12) is imposed in [7, 8]. The discrete gradient flow in [8] treats the cubic energy $C[\mathbf{y}]$ implicitly to get an energy decreasing scheme but requires the normalization (3) of the discrete normal, which renders the algorithm nonlinear. Discrete energies are shown to Γ-converge in [8]. In contrast, the scheme of [7] is linear and much more efficient, but stability and Γ-convergence are still open. Recently, Bartels and Palus [10] reformulated the discretization of $C[\mathbf{y}]$ making it fully explicit and the ensuing algorithm linear, and were also able to show an energy decreasing property for the explicit gradient flow with a mild time-step constraint and Γ-convergence of the discrete energies.

On the other hand, interior penalty discontinuous Galerkin (IPDG) finite element methods are proposed and studied in [17] because they require a lower polynomial degree (2 instead of 3), are easier to find in existing software platforms, are more flexible in imposing boundary conditions as well as the linearized isometry constraint (12), and are amenable to subdivisions containing curved boundaries which is crucial to deal with creases. The linearized constraint (12) is enforced in average on all elements of the subdivision. Furthermore, the cubic energy $C[\mathbf{y}]$ is treated explicitly at each step of the discrete gradient flow and the ensuing algorithm is linear. However, Γ-convergence and energy decreasing properties remain open problems.

We note that the bilayer model (7) reduces to single layer plates endowed with the bending energy $B[\mathbf{y}]$ for $\mathbf{y} \in \mathbb{A}$ provided the upper and lower layers are identical, i.e. $Z = 0$. We refer to [4, 18] for the design and analysis of Kirchhoff and IPDG methods in this simpler context.

1.3. **LDG-discretization and our contribution.** We propose a local discontinuous Galerkin (LDG) method for the approximation of the minimization problem (4) along the lines of [14, 15]. LDG method was originally introduced in [24], and further explored in [11, 21, 22, 25, 26]. Our discrete energy $E_h[\mathbf{y}_h]$ is obtained (up

to stabilization terms) by simply replacing the Hessian $D^2\mathbf{y}$ in (8) by a discrete Hessian $H_h[\mathbf{y}_h]$ (defined by (29) below), which is constructed and analyzed in [14, 15] in terms of the discontinuous Galerkin solution $\mathbf{y}_h$. This is conceptually simpler than IPDG methods, which are based on integration by parts and are harder to design for intricate nonlinear systems. In contrast to IPDG, LDG is also stable for any positive stabilization parameters, and exhibits better convergence properties at the expense of a slightly worse sparsity pattern [14, 15].

Our treatment of the cubic term hinges on the mid-point quadrature. If $\mathcal{T}_h$ is a mesh made of shape-regular triangles or quadrilaterals T with barycenter x_T, let

$$(13) \qquad C_h[\mathbf{y}_h] := \sum_{i,j=1}^{2} \sum_{T \in \mathcal{T}_h} |T| \Big(\overline{H}_h[\mathbf{y}_h]_{ij} \cdot (\partial_1 \mathbf{y}_h \times \partial_2 \mathbf{y}_h) Z_{ij} \Big)(x_T)$$

where $\overline{H}_h[\mathbf{y}_h] = \frac{1}{|T|} \int_T H_h[\mathbf{y}_h]$ for all $T \in \mathcal{T}_h$ is the piecewise constant reduced discrete Hessian. We also make the simplifying assumption that the spontaneous curvature Z in (4) is piecewise constant over all partitions $\mathcal{T}_h$, $h > 0$. Moreover, we control the isometry defect at barycenters, namely given a parameter $\delta > 0$ we impose

$$(14) \qquad D_h[\mathbf{y}_h](x_T) := \Big| \big[\nabla \mathbf{y}_h^T \nabla \mathbf{y}_h - I_2\big](x_T) \Big| \leq \delta \quad \forall\, T \in \mathcal{T}_h.$$

We enforce the Dirichlet condition upon augmenting the discrete energy $E_h[\mathbf{y}_h]$ via a Nitsche method. Therefore, we say that discrete functions satisfying (14) belong to the discrete admissible set $\mathbb{A}_{h,\delta}$, the discrete counterpart of $\mathbb{A}$ in (1). We prove that $\mathbb{A}_{h,\delta}$, is non-empty, and derive convergence of global minimizers $\mathbf{y}_h$ of E_h within $\mathbb{A}_{h,\delta}$, towards global minimizers $\mathbf{y}$ of (4) in the spirit of Γ-convergence.

It is worth pointing out that Γ-convergence does not give error estimates and that, except for [9] for linear plates with folding, we are not aware of such bounds for nonlinear plates undergoing large deformations. The main obstructions are: the energy is nonconvex; the isometry constraint is nonconvex; there might be multiple solutions; there is no regularity theory beyond the basic energy estimate $\mathbf{y} \in [H^2(\Omega)]^3$; mapping properties of the linearized Euler-Lagrange equation and isometry constraint, that govern the behavior of perturbations, have to be discovered and most likely will entail additional regularity of $\mathbf{y}$; no monotonicity argument is available because $\mathbf{y}$ is vector-valued. However, it is plausible that error estimates are valid for small perturbations of smooth branches of solutions. Proving error estimates is a challenging and important endeavor, but is far beyond the scope of this paper.

Solving the *nonconvex* discrete minimization counterpart of (4) is a highly nontrivial task. We resort to a discrete gradient flow that enforces the linearized isometry constraint (12) at the barycenters x_T of elements T

$$(15) \qquad L[\mathbf{v}_h; \mathbf{y}_h](x_T) := \big[\nabla \mathbf{v}_h^T \nabla \mathbf{y}_h + \nabla \mathbf{y}_h^T \nabla \mathbf{v}_h\big](x_T) = \mathbf{0} \quad \forall\, T \in \mathcal{T}_h,$$

and solve a discrete minimization problem for a tangential variation $\mathbf{v}_h$ of $\mathbf{y}_h$, in the sense (15), with the cubic term (13) treated explicitly. The latter is a clever idea, due to Bartels and Palus [10], that renders the problem linear at each step of the gradient flow; however, our approach is different from [10]. We show that this procedure is energy decreasing, convergent, and preserves the isometry defect (14)

provided δ is proportional to h, which entails a linear relation between the time step τ of the gradient flow and h. Moreover, we derive a (suboptimal) discrete inf-sup condition for the Lagrange multiplier approach to the linear constraint (15), which seems to be the first such result for this type of matrix constraint and is consistent with computations.

The rest of this article is organized as follows. Section 2 is about LDG. We introduce the (*broken*) finite element spaces in Section 2.2. We examine the discrete Hessian operator and its reduced counterpart in Subsection 2.3, together with their boundedness and convergence properties. In Subsections 2.4 and 2.5, we define the discrete problem and investigate consistency of the cubic discrete energy C_h. The proof of Γ-convergence of the discrete energy to the exact one is the content of Section 3, and its extension to a bilayer model with piecewise quadratic creases is included in Section 4. In Section 5, we introduce the gradient flow scheme used to solve the discrete problem, prove its conditional stability and show how the constraint violation (14) is controlled throughout the flow. Moreover, we derive a suboptimal inf-sup condition for (15) at each step of the flow. We present several insightful simulations in Section 6 to illustrate the performance of LDG, including folding across a curved crease.

2. LDG Discretization

2.1. **Subdivisions.** From now on, we assume that $\Omega \subset \mathbb{R}^2$ is a polygonal domain and denote by $\{\mathcal{T}_h\}_{h>0}$ a shape-regular sequence of conforming partitions of Ω made of either triangles or quadrilaterals T with diameter $h_T := \operatorname{diam}(T) \le h$. The set of edges $\mathcal{E}_h := \mathcal{E}_h^0 \cup \mathcal{E}_h^b$ is decomposed into the interior edges $\mathcal{E}_h^0$ and boundary edges $\mathcal{E}_h^b$. For $e \in \mathcal{E}_h$, we define $h_e := \operatorname{diam}(e)$ and note that $h_e \le h$, and thus

$$(16) \qquad\qquad h^{-1} \le h_e^{-1} \quad \forall e \in \mathcal{E}_h.$$

We assume a compatible representation of the Dirichlet boundary $\Gamma^D = \cup\{e : e \in \mathcal{E}_h^D\}$, and let $\mathcal{E}_h^a := \mathcal{E}_h^0 \cup \mathcal{E}_h^D$ be the set of *active edges* on which jumps and averages will be computed. The union of these edges gives rise to the corresponding skeletons of $\mathcal{T}_h$

$$(17) \qquad \Gamma_h^0 := \cup\{e : e \in \mathcal{E}_h^0\}, \quad \Gamma_h^D := \Gamma^D, \quad \Gamma_h^a := \Gamma_h^0 \cup \Gamma_h^D.$$

We use the notation $(\cdot, \cdot)_{L^2(\Omega)}$ and $(\cdot, \cdot)_{L^2(\Gamma_h^a)}$ to denote the L^2 inner products over Ω and Γ_h^a, and a similar notation for subsets of Ω and Γ_h^a. We denote by h a mesh density function, locally equivalent to h_T and h_e, and utilize it as a weight in the preceding norms. We often write $f \lesssim g$ to indicate that there exists a constant C independent of discretization parameters such that $f \le Cg$.

2.2. **Broken spaces and operators.** For an integer $r \ge 0$, we denote by $\mathbb{P}_r$ the space of polynomials of total degree at most r when the subdivision is made of triangles and by $\mathbb{Q}_r$ the space of polynomials of degree at most r in each variable when quadrilaterals are used. We also use the same notation, $\widehat{T}$, to denote either the unit triangle or the unite square depending on the type of subdivision used. We let $F_T : \widehat{T} \to T \in [\mathbb{Q}_1]^2$ be the generic map from the reference element to the physical element. It is affine only when the subdivision is made of triangles.

We fix the polynomial degree $k \geq 2$. The (*broken*) finite element space $\mathbb{V}_h^k$ to approximate each component of the deformation $\mathbf{y}$ reads

$$(18) \qquad \mathbb{V}_h^k := \left\{ v_h \in L^2(\Omega) : \; v_{h|T} \circ F_T \in \mathbb{Q}_k \quad \forall T \in \mathcal{T}_h \right\},$$

when the subdivision is made of quadrilaterals, and we replace $\mathbb{Q}_k$ by $\mathbb{P}_k$ if we have triangular elements. We define the broken gradient $\nabla_h v_h$ of $v_h \in \mathbb{V}_h^k$ to be the elementwise gradient, and use similar notation for other differential operators. For instance $D_h^2 v_h = \nabla_h \nabla_h v_h$ stands for the broken Hessian, and $\partial_i v_h := \partial_{i,h} v_h$ denotes the components $i = 1, 2$ of the broken gradient $\nabla_h v_h$.

We now introduce the jump and average operators. For every $e \in \mathcal{E}_h^0$, fix $\mathbf{n}_e$ to be one of the two unit normals to e (the choice is arbitrary but does not affect the formulation). For a boundary edge $e \in \mathcal{E}_h^b$, we set $\mathbf{n}_e = \mathbf{n}$, the outward unit normal vector to $\partial\Omega$. The jump of $v_h \in \mathbb{V}_h^k$ and $\nabla_h v_h$ across $e \in \mathcal{E}_h^0$ are given by

$$(19) \qquad [v_h]|_e := v_h^- - v_h^+, \quad [\nabla_h v_h]|_e := \nabla_h v_h^- - \nabla_h v_h^+,$$

where $v_h^\pm(\mathbf{x}) := \lim_{s \to 0^+} v_h(\mathbf{x} \pm s\mathbf{n}_e)$ for $\mathbf{x} \in e$. The jumps of a vector or matrix valued function are computed componentwise.

In order to incorporate the Dirichlet boundary conditions $\mathbf{y} = \boldsymbol{\varphi}$, $\nabla\mathbf{y} = \Phi$ on Γ^D, we resort to a Nitsche's approach which does not impose essential restrictions on the discrete space $[\mathbb{V}_h^k]^3$ but rather modifies the discrete formulation by including boundary jumps defined for $\mathbf{v}_h \in [\mathbb{V}_h^k]^3$

$$(20) \qquad [\mathbf{v}_h]|_e := [\mathbf{v}_h]|_e(\boldsymbol{\varphi}) := \mathbf{v}_h - \boldsymbol{\varphi}, \quad [\nabla_h \mathbf{v}_h]|_e := [\nabla_h \mathbf{v}_h]|_e(\Phi) := \nabla_h \mathbf{v}_h - \Phi,$$

for all $e \in \mathcal{E}_h^D$. However, to simplify the notation, it is convenient to introduce the discrete set $\mathbb{V}_h^k(\boldsymbol{\varphi}, \Phi)$

$$(21) \qquad \mathbb{V}_h^k(\boldsymbol{\varphi}, \Phi) := \left\{ \mathbf{v}_h \in [\mathbb{V}_h^k]^3 : \; [\mathbf{v}_h]|_e, \, [\nabla_h \mathbf{v}_h]|_e \text{ given by (20) for all } e \in \mathcal{E}_h^D \right\},$$

which coincide with $[\mathbb{V}_h^k]^3$ but carries the notion of boundary jump (20) for its elements. We define the *average* of $v_h \in \mathbb{V}_h^k$ across an edge $e \in \mathcal{E}_h$ as

$$(22) \qquad \{v_h\}|_e := \begin{cases} \frac{1}{2}(v_h^+ + v_h^-) & e \in \mathcal{E}_h^0 \\ v_h^- & e \in \mathcal{E}_h^b, \end{cases}$$

and apply (22) componentwise to vector and matrix-valued functions.

We let $\langle \cdot, \cdot \rangle_{H_h^2(\Omega)}$ be the following mesh-dependent form defined, for any $\mathbf{v}_h, \mathbf{w}_h \in \mathbb{V}_h^k(\boldsymbol{\varphi}, \Phi)$, by

$$(23) \qquad \begin{aligned} \langle \mathbf{v}_h, \mathbf{w}_h \rangle_{H_h^2(\Omega)} := & \, (D_h^2 \mathbf{v}_h, D_h^2 \mathbf{w}_h)_{L^2(\Omega)} \\ & + (\mathrm{h}^{-1}[\nabla_h \mathbf{v}_h], [\nabla_h \mathbf{w}_h])_{L^2(\Gamma_h^a)} + (\mathrm{h}^{-3}[\mathbf{v}_h], [\mathbf{w}_h])_{L^2(\Gamma_h^a)}. \end{aligned}$$

We emphasize that (23) is not bilinear in $\mathbb{V}_h^k(\boldsymbol{\varphi}, \Phi)$ because of the presence of $(\boldsymbol{\varphi}, \Phi)$ in the boundary jump terms, unless $\boldsymbol{\varphi} = \mathbf{0}$, $\Phi = \mathbf{0}$. Moreover, we set

$$(24) \qquad \|\mathbf{v}_h\|_{H_h^2(\Omega)}^2 := \langle \mathbf{v}_h, \mathbf{v}_h \rangle_{H_h^2(\Omega)} \quad \forall \, \mathbf{v}_h \in \mathbb{V}_h^k(\boldsymbol{\varphi}, \Phi),$$

and observe the validity of the following Friedrichs-type inequality [18, (2.27)]

$$(25) \quad \|\mathbf{v}_h\|_{L^2(\Omega)} + \|\nabla_h \mathbf{v}_h\|_{L^2(\Omega)} \lesssim \|\mathbf{v}_h\|_{H_h^2(\Omega)} + \|\boldsymbol{\varphi}\|_{H^1(\Omega)} + \|\Phi\|_{H^1(\Omega)} \quad \forall\, \mathbf{v}_h \in \mathbb{V}_h^k(\boldsymbol{\varphi}, \Phi).$$

Once restricted to $\mathbb{V}_h^k(\mathbf{0}, \mathbf{0})$, the form $\langle \cdot, \cdot \rangle_{H_h^2(\Omega)}$ turns out to be a scalar product, according to (25), which corresponds to the discrete counterpart of $\langle \cdot, \cdot \rangle_{H^2(\Omega)}$.

2.3. **Discrete Hessians.** The central ingredient in the proposed LDG approximation is the reconstructed Hessian $H_h[\mathbf{y}_h] \in \left[L^2(\Omega)\right]^{3\times2\times2}$ defined in [14, 15]. Let $l_1, l_2 \geq 0$ be integers and consider two *local lifting operators* $r_e : [L^2(e)]^2 \to [\mathbb{V}_h^{l_1}]^{2\times2}$ and $b_e : L^2(e) \to [\mathbb{V}_h^{l_2}]^{2\times2}$ defined for $e \in \mathcal{E}_h^a$ by

$$(26) \qquad r_e(\boldsymbol{\phi}) \in [\mathbb{V}_h^{l_1}]^{2\times2} : \int_{\omega_e} r_e(\boldsymbol{\phi}) : \tau_h = \int_e \{\tau_h\}\, \mathbf{n}_e \cdot \boldsymbol{\phi} \quad \forall \tau_h \in [\mathbb{V}_h^{l_1}]^{2\times2},$$

$$(27) \qquad b_e(\phi) \in [\mathbb{V}_h^{l_2}]^{2\times2} : \int_{\omega_e} b_e(\phi) : \tau_h = \int_e \{\mathrm{div}\,\tau_h\} \cdot \mathbf{n}_e \phi \quad \forall \tau_h \in [\mathbb{V}_h^{l_2}]^{2\times2},$$

where ω_e is the union of the two elements of $\mathcal{T}_h$ sharing $e \in \Gamma_h^0$ or the element of $\mathcal{T}_h$ having $e \in \Gamma_h^b$ as part of its boundary. The definitions extend to $\left[[L^2(e)]^2\right]^3$ and $[L^2(e)]^3$ by component-wise application. The corresponding *global lifting operators* are then given by

$$(28) \qquad \begin{aligned} \mathcal{R}_h &:= \sum_{e\in\mathcal{E}_h^a} r_e : [L^2(\Gamma_h^a)]^2 \to [\mathbb{V}_h^{l_1}]^{2\times2}, \\ \mathcal{B}_h &:= \sum_{e\in\mathcal{E}_h^a} b_e : L^2(\Gamma_h^a) \to [\mathbb{V}_h^{l_2}]^{2\times2}. \end{aligned}$$

Their purpose is to lift inter-element information to the cells so that once added to the piecewise Hessian D_h^2, they constitute a weakly convergent approximation of the exact Hessian (see Lemma 1). In fact, we define the discrete Hessian operator $H_h : \mathbb{V}_h^k(\boldsymbol{\varphi}, \Phi) \to \left[L^2(\Omega)\right]^{3\times2\times2}$ by

$$(29) \qquad\qquad H_h[\mathbf{v}_h] := D_h^2 \mathbf{v}_h - \mathcal{R}_h([\nabla_h \mathbf{v}_h]) + \mathcal{B}_h([\mathbf{v}_h]).$$

We point out the implicit dependence on data $(\boldsymbol{\varphi}, \Phi)$ and that we will later compute $H_h[\mathbf{v}_h]$ for $\mathbf{v}_h \in \mathbb{V}_h^k(\mathbf{0}, \mathbf{0})$, i.e. $\boldsymbol{\varphi} = \mathbf{0}$, $\Phi = \mathbf{0}$, slightly abusing notation. Thanks to the relation between the edge and cell diameter (16), we have the following a priori upper bounds for lifting operators

$$(30) \qquad\qquad \|H_h[\mathbf{v}_h]\|_{L^2(\Omega)} \lesssim \|\mathbf{v}_h\|_{H_h^2(\Omega)}.$$

Moreover, we have the following properties of the discrete Hessian $H_h[\mathbf{v}_h]$.

Lemma 1 (weak convergence of H_h). *Let $k \geq 2$ and $\mathbf{v}_h \in \mathbb{V}_h^k(\boldsymbol{\varphi}, \Phi)$. If $\|\mathbf{v}_h\|_{H_h^2(\Omega)} \lesssim 1$ and $\mathbf{v}_h \to \mathbf{v} \in [H^2(\Omega)]^3$ in $[L^2(\Omega)]^3$ as $h \to 0$, then for any polynomial degree $l_1, l_2 \geq 0$ we have*

$$(31) \qquad\qquad H_h[\mathbf{v}_h] \rightharpoonup D^2 \mathbf{v} \quad in \left[L^2(\Omega)\right]^{3\times2\times2} \quad as\ h \to 0.$$

Proof. See [15, Lemma 2.4 and Appendix B]. □

Lemma 2 (strong convergence of H_h). *Let* $\mathbf{v} \in [H^2(\Omega)]^3$ *be any function such that* $\mathbf{v} = \boldsymbol{\varphi}$ *and* $\nabla\mathbf{v} = \Phi$ *on* Γ^D. *Moreover, let* $\mathbf{v}_h \in \mathbb{V}_h^k(\boldsymbol{\varphi}, \Phi)$ *satisfy*

$$(32) \qquad \|D^2\mathbf{v}_h\|_{L^2(T)} \lesssim \|\mathbf{v}\|_{H^2(T)} \ \forall T \in \mathcal{T}_h, \quad \sum_{T\in\mathcal{T}_h} \|\mathbf{v}_h - \mathbf{v}\|_{H^2(T)}^2 \to 0 \ as \ h \to 0^+.$$

Then for any polynomial degree $l_1, l_2 \geq 0$ *we have as* $h \to 0^+$

$$(33) \qquad H_h[\mathbf{v}_h] \to D^2\mathbf{v} \quad strongly \ in \ [L^2(\Omega)]^{3\times2\times2}.$$

Proof. This is a minor modification of [15, Lemma 2.5 and Appendix B], which assumes that $\mathbf{v}_h$ is the Lagrange interpolant of $\mathbf{v} \in H^2(\Omega)$. □

For later use, we now discuss properties of the *reduced discrete Hessian* defined as the local L^2 projection onto the space of piecewise constants, i.e.

$$(34) \qquad \overline{H}_h[\mathbf{v}_h]|_T := \frac{1}{|T|} \int_T H_h[\mathbf{v}_h] \quad \forall\, T \in \mathcal{T}_h.$$

We start with the stability of $\overline{H}_h[\mathbf{y}_h]$.

Lemma 3 (stability of $\overline{H}_h[\mathbf{v}_h]$). *For any* $\mathbf{v}_h \in \mathbb{V}_h^k(\boldsymbol{\varphi}, \Phi)$, *there holds*

$$(35) \qquad \|\overline{H}_h[\mathbf{v}_h]\|_{L^2(\Omega)} \leq c_{stab}\|\mathbf{v}_h\|_{H_h^2(\Omega)},$$

where the constant c_{stab} *is independent of* h.

Proof. This result is a direct consequence of the stability of the reconstructed Hessian (30) and the local L^2 projection (34). □

The reduced discrete Hessian is also weakly converging.

Lemma 4 (weak convergence of $\overline{H}_h[\mathbf{v}_h]$). *Let* $\mathbf{v}_h \in \mathbb{V}_h^k(\boldsymbol{\varphi}, \Phi)$ *be a sequence of discrete deformations satisfying* $\|\mathbf{v}_h\|_{H_h^2(\Omega)} \lesssim 1$ *for all* h *and such that* $\mathbf{v}_h \to \mathbf{v}$ *in* $[L^2(\Omega)]^3$ *for some* $\mathbf{v} \in [H^2(\Omega)]^3$. *Then,* $\overline{H}_h[\mathbf{v}_h]$ *converges weakly to* $D^2\mathbf{v}$ *in* $[L^2(\Omega)]^3$.

Proof. For any $\phi \in [C_0^\infty(\Omega)]^{3\times2\times2}$, we have

$$\int_\Omega \overline{H}_h[\mathbf{v}_h] : \phi = \sum_{T\in\mathcal{T}_h} \int_T H_h[\mathbf{v}_h] : \overline{\phi} = \sum_{T\in\mathcal{T}_h} \int_T H_h[\mathbf{v}_h] : \phi + H_h[\mathbf{v}_h] : (\overline{\phi} - \phi),$$

where $\overline{\phi} := \frac{1}{|T|} \int_T \phi$. Lemma 1 (weak convergence of $H_h[\mathbf{y}_h]$) implies

$$\int_\Omega H_h[\mathbf{v}_h] : \phi \to \int_\Omega D^2\mathbf{v} : \phi.$$

On the other hand, the uniform boundedness (30) and the assumption $\|\mathbf{v}_h\|_{H_h^2(\Omega)} \lesssim 1$ guarantee that

$$\left| \sum_{T\in\mathcal{T}_h} H_h[\mathbf{v}_h] : (\overline{\phi} - \phi) \right| \lesssim h\|H_h[\mathbf{v}_h]\|_{L^2(\Omega)}\|\nabla\phi\|_{L^2(\Omega)} \lesssim h\|\nabla\phi\|_{L^2(\Omega)} \to 0$$

as $h \to 0^+$. Combining these two estimates yields the desired result. □

2.4. **Discrete admissible set.** We introduce the discrete counterpart of the admissible set $\mathbb{A}$. Given a parameter $\delta > 0$ to be related later to h, we recall the discrete isometry defect $D_h[\mathbf{y}_h]$ from (14) and define the *discrete admissible set* $\mathbb{A}_{h,\delta}$ as

$$(36) \qquad \mathbb{A}_{h,\delta} := \left\{ \mathbf{y}_h \in \mathbb{V}_h^k(\boldsymbol{\varphi}, \Phi) : \quad D_h[\mathbf{y}_h](x_T) \le \delta \quad \forall T \in \mathcal{T}_h \right\},$$

where the polynomial degree is $k \ge 2$ and x_T is the barycenter of $T \in \mathcal{T}_h$. The Dirichlet boundary conditions are hidden within the definition (21) of $\mathbb{V}_h^k(\boldsymbol{\varphi}, \Phi)$ and imposed in the weak formulation; hence they do not contribute to any essential restriction in $\mathbb{A}_{h,\delta}$. The following two lemmas are simple consequences of (36).

Lemma 5 ($\mathbb{A}_{h,\delta}$ is non-empty)**.** *For all $h > 0$ there exists $\mathbf{y}_h \in \mathbb{V}_h^k(\boldsymbol{\varphi}, \Phi)$ such that $D_h[\mathbf{y}_h](x_T) = 0$ for all $T \in \mathcal{T}_h$.*

Proof. Let $\mathbf{y}_h(x) := x$ for $x \in \Omega$. We see that $\mathbf{y}_h \in [\mathbb{V}_h^k]^3$, and therefore $\mathbf{y}_h \in \mathbb{V}_h^k(\boldsymbol{\varphi}, \Phi)$ because the Dirichlet boundary conditions are not imposed *essentially* in the space $\mathbb{V}_h^k(\boldsymbol{\varphi}, \Phi)$ defined in (21). Moreover, $\mathrm{I}[\mathbf{y}_h](x_T) = I_2$, whence $D_h[\mathbf{y}_h](x_T) = 0$. $\qquad\square$

Note that this implies $\mathbb{A}_{h,\delta}$ is non-empty for any $\delta > 0$, because $\mathbb{A}_{h,0} \subset \mathbb{A}_{h,\delta}$. We postpone until Theorem 9 the hard question whether $\mathbb{A}_{h,\delta}$ is sufficiently rich to approximate $\mathbb{A}$: for any $\mathbf{y} \in \mathbb{A}$ there is $\mathbf{y}_h \in \mathbb{A}_{h,\delta}$ close to $\mathbf{y}$ in a suitable sense. The following lemma provides an estimate on the amount of local stretch and shear associated with functions in $\mathbb{A}_{h,\delta}$.

Lemma 6 (pointwise isometry constraint)**.** *If $\mathbf{y}_h \in \mathbb{A}_{h,\delta}$, then for all $T \in \mathcal{T}_h$ and $i = 1, 2$*

$$(37) \qquad 1 - \delta \le |\partial_i \mathbf{y}_h(x_T)|^2 \le 1 + \delta, \quad |\partial_1 \mathbf{y}_h(x_T) \cdot \partial_2 \mathbf{y}_h(x_T)| \le \delta.$$

Proof. From definition (36), we deduce that for any $i, j = 1, 2$

$$\left| \partial_i \mathbf{y}_h(x_T) \cdot \partial_j \mathbf{y}_h(x_T) - \delta_{ij} \right| \le \delta,$$

where δ_{ij} is the Kronecker delta. The assertion thus follows. $\qquad\square$

The pointwise control of isometry defect in (36) is inspired by the algorithms based on Kirchhoff finite elements developed in [8, 10], where this constraint is imposed at the element vertices. Dealing with element barycenters is novel in the context of DG methods in that previous schemes impose this constraint in average over elements [17, 15]. Having control at barycenters does not imply control of $\nabla_h \mathbf{y}_h$ anywhere else, and dictates the use of mid-point quadrature for the discretization of the cubic nonlinear energy C_h. We discuss this next.

2.5. **Discrete energy.** The LDG approximation of the energy $E[.]$ reads

$$(38) \qquad E_h[\mathbf{y}_h] := B_h[\mathbf{y}_h] - C_h[\mathbf{y}_h]$$

where $B_h[.]$ approximates the bending energy (10) and $C_h[.]$ approximates the cubic interaction energy in (11). The energy $B_h[\mathbf{y}_h]$ is defined by

$$(39) \qquad B_h[\mathbf{y}_h] := \frac{1}{2} \int_\Omega \left| H_h[\mathbf{y}_h] \right|^2 + S_h[\mathbf{y}_h],$$

where

$$(40) \qquad S_h[\mathbf{y}_h] := \gamma_1 \|\mathrm{h}^{-\frac{1}{2}}[\nabla_h \mathbf{y}_h]\|_{L^2(\Gamma_h^a)}^2 + \gamma_0 \|\mathrm{h}^{-\frac{3}{2}}[\mathbf{y}_h]\|_{L^2(\Gamma_h^a)}^2$$

is a stabilization term with parameters $\gamma_0, \gamma_1 > 0$, whereas $C_h[\mathbf{y}_h]$ is given by (13)

$$(41) \qquad C_h[\mathbf{y}_h] := \sum_{i,j=1}^{2} \sum_{T \in \mathcal{T}_h} |T|\Big(\overline{H}_h[\mathbf{y}_h]_{ij} \cdot (\partial_1 \mathbf{y}_h \times \partial_2 \mathbf{y}_h) Z_{ij}\Big)(x_T).$$

With these notations the discrete minimization problem reads

$$(42) \qquad \min_{\mathbf{y}_h \in \mathbb{A}_{h,\delta}} E_h[\mathbf{y}_h].$$

We devote the rest of this section to examine the cubic energy (41). Combining Lemma 6 (pointwise isometry constraint) with Lemma 3 (stability of $\overline{H}_h[\mathbf{y}_h]$) yields

$$\big|C_h[\mathbf{y}_h]\big| \lesssim (1 + \delta)\|\mathbf{y}_h\|_{H_h^2(\Omega)}\|Z\|_{L^2(\Omega)},$$

whence $\big|C_h[\mathbf{y}_h]\big|$ is uniformly bounded whenever $\|\mathbf{y}_h\|_{H_h^2(\Omega)}$ is. Another crucial aspect of (41) is the convergence of C_h towards the continuous energy C within the basic H^2-regularity framework. This requires dealing with the reduced Hessian $\overline{H}_h[\mathbf{y}_h]$ as we show next.

Lemma 7 (convergence of cubic energy). *Let Z be piecewise constant over $\mathcal{T}_h$. Let $\mathbf{y}_h \in \mathbb{A}_{h,\delta}$ be a sequence of discrete deformations satisfying*

$$(43) \qquad \|\mathbf{y}_h\|_{H_h^2(\Omega)} \lesssim 1 \quad \forall\, h > 0$$

and such that $\mathbf{y}_h \to \mathbf{y}$ in $[L^2(\Omega)]^3$, $\nabla_h \mathbf{y}_h \to \nabla \mathbf{y}$ in $[L^2(\Omega)]^{3\times 2}$ for $\mathbf{y} \in [H^2(\Omega) \cap W_\infty^1(\Omega)]^3$ as $h \to 0^+$. Then

$$(44) \qquad \lim_{h\to 0^+} C_h[\mathbf{y}_h] = C[\mathbf{y}].$$

Proof. For any $\epsilon > 0$, it suffices to show that

$$(45) \qquad \limsup_{h\to 0^+} \big|C[\mathbf{y}] - C_h[\mathbf{y}_h]\big| \lesssim \epsilon.$$

We need a regularization argument to deal with the effect of quadrature. Since Ω is Lipschitz we can regularize $\mathbf{y}$, say by convolution, in such a manner that the approximate deformation $\mathbf{y}^\epsilon \in [H^3(\Omega)]^3$ satisfies

$$(46) \qquad \|\mathbf{y}^\epsilon\|_{H^2(\Omega)} + \|\mathbf{y}^\epsilon\|_{W_\infty^1(\Omega)} \lesssim 1, \qquad \|\mathbf{y} - \mathbf{y}^\epsilon\|_{H^2(\Omega)} \lesssim \epsilon;$$

we recall the convention that constants hidden in $\lesssim$ are independent of h and ϵ. We point out that this procedure is simpler than the regularization due to Hornung [30] in that $\mathbf{y}^\epsilon$ need not be an isometry. We first observe that the energies $C[\mathbf{y}]$ and $C[\mathbf{y}^\epsilon]$ can be made arbitrarily close because

$$\big|C[\mathbf{y}] - C[\mathbf{y}^\epsilon]\big| \lesssim \|\mathbf{y} - \mathbf{y}^\epsilon\|_{H^2(\Omega)}\|\partial_1 \mathbf{y}\|_{L^2(\Omega)}\|\partial_2 \mathbf{y}\|_{L^\infty(\Omega)}\|Z\|_{L^\infty(\Omega)}$$
$$+ \|\mathbf{y}^\epsilon\|_{H^2(\Omega)}\|\mathbf{y} - \mathbf{y}^\epsilon\|_{H^1(\Omega)}\big(\|\partial_2 \mathbf{y}\|_{L^\infty(\Omega)} + \|\partial_1 \mathbf{y}^\epsilon\|_{L^\infty(\Omega)}\big)\|Z\|_{L^\infty(\Omega)} \lesssim \epsilon.$$

We next write $C_h[\mathbf{y}_h] - C[\mathbf{y}^\epsilon] = \sum_{i,j=1}^2 \sum_{T \in \mathcal{T}_h} R_1(T) + R_2(T) + R_3(T)$, where

$$R_1(T) := \int_T (\overline{H}_h[\mathbf{y}_h]_{ij} - \partial_{ij}\mathbf{y}^\epsilon) \cdot (\partial_1 \mathbf{y}^\epsilon \times \partial_2 \mathbf{y}^\epsilon) Z_{ij},$$

$$R_2(T) := |T| \big[\overline{H}_h[\mathbf{y}_h]_{ij} \cdot (\partial_1 \mathbf{y}_h \times \partial_2 \mathbf{y}_h - \partial_1 \mathbf{y}^\epsilon \times \partial_2 \mathbf{y}^\epsilon) Z_{ij} \big](x_T),$$

$$R_3(T) := |T| \big[\overline{H}_h[\mathbf{y}_h]_{ij} \cdot (\partial_1 \mathbf{y}^\epsilon \times \partial_2 \mathbf{y}^\epsilon) Z_{ij} \big](x_T) - \int_T \overline{H}_h[\mathbf{y}_h]_{ij} \cdot (\partial_1 \mathbf{y}^\epsilon \times \partial_2 \mathbf{y}^\epsilon) Z_{ij},$$

and disregard the non critical dependence on $i, j = 1, 2$ in the notation. Lemma 4 (weak convergence of $\overline{H}_h[\mathbf{y}_h]$) in conjunction with (46) implies that

$$\limsup_{h \to 0^+} \sum_{i,j=1}^2 \sum_{T \in \mathcal{T}_h} |R_1(T)| \lesssim \epsilon.$$

For R_2, we note that

$$\big| (\partial_1 \mathbf{y}_h \times \partial_2 \mathbf{y}_h - \partial_1 \mathbf{y}^\epsilon \times \partial_2 \mathbf{y}^\epsilon)(x_T) \big| \le \big| \nabla(\mathbf{y}_h - \mathbf{y}^\epsilon)(x_T) \big| \big(|\nabla \mathbf{y}_h(x_T)| + |\nabla \mathbf{y}^\epsilon(x_T)| \big).$$

By Lemma 6 (pointwise isometry constraint), the fact that $\mathbf{y}_h \in \mathbb{A}_{h,\delta}$ and (46), we have the uniform bound $|\nabla \mathbf{y}_h(x_T)| + |\nabla \mathbf{y}^\epsilon(x_T)| \lesssim 1$ for all x_T. If $I_h \nabla \mathbf{y}^\epsilon$ indicates the standard $\mathbb{P}_1$-Lagrange interpolant of $\nabla \mathbf{y}^\epsilon$, applying approximating properties of I_h together with an inverse inequality for polynomials, we conclude

$$\big| \nabla(\mathbf{y}_h - \mathbf{y}^\epsilon)(x_T) \big| \le \big| (\nabla \mathbf{y}_h - I_h \nabla \mathbf{y}^\epsilon)(x_T) \big| + \big| (I_h \nabla \mathbf{y}^\epsilon - \nabla \mathbf{y}^\epsilon)(x_T) \big|$$
$$\lesssim h_T^{-1} \| \nabla \mathbf{y}_h - I_h \nabla \mathbf{y}^\epsilon \|_{L^2(T)} + h_T \| D^3 \mathbf{y}^\epsilon \|_{L^2(T)}.$$

We next add and subtract $\nabla \mathbf{y}^\epsilon$ in the first term of the right-hand side and apply again an interpolation estimate of I_h to derive

$$|T|^{1/2} \big| \nabla(\mathbf{y}_h - \mathbf{y}^\epsilon)(x_T) \big| \lesssim \| \nabla(\mathbf{y}_h - \mathbf{y}^\epsilon) \|_{L^2(T)} + h_T^2 \| D^3 \mathbf{y}^\epsilon \|_{L^2(T)}.$$

Moreover, since $|T|^{1/2} \big| \overline{H}_h[\mathbf{y}_h]_{ij}(x_T) \big| = \| \overline{H}_h[\mathbf{y}_h]_{ij} \|_{L^2(T)}$ because $\overline{H}_h[\mathbf{y}_h]$ is piecewise constant, we obtain

$$|R_2(T)| \lesssim \| \overline{H}_h[\mathbf{y}_h]_{ij} \|_{L^2(T)} \big(\| \nabla(\mathbf{y}_h - \mathbf{y}) \|_{L^2(T)} + \| \nabla(\mathbf{y} - \mathbf{y}^\epsilon) \|_{L^2(T)} + h_T^2 \| D^3 \mathbf{y}^\epsilon \|_{L^2(T)} \big),$$

where the hidden constant is proportional to $\|Z\|_{L^\infty(\Omega)}$. After summing over elements, Lemma 3 (stability of $\overline{H}_h[\mathbf{y}_h]$), together with the assumption $\nabla_h \mathbf{y}_h \to \nabla \mathbf{y}$ in $[L^2(\Omega)]^{3 \times 2}$, (43) and (46), yields

$$\limsup_{h \to 0^+} \sum_{i,j=1}^2 \sum_{T \in \mathcal{T}_h} |R_2(T)| \lesssim \epsilon.$$

It remains to deal with R_3 which entails the effect of quadrature. Since Z and $\overline{H}_h[\mathbf{y}_h]$ are constant in T, which is the chief reason for utilizing the reduced discrete Hessian, we can equivalently rewrite $R_3(T)$ as follows:

$$R_3(T) = \overline{H}_h[\mathbf{y}_h]_{ij} Z_{ij} \int_T (\mathbf{f}(x_T) - \mathbf{f})$$

with $\mathbf{f} = \partial_1 \mathbf{y}^\epsilon \times \partial_2 \mathbf{y}^\epsilon$. The Bramble-Hilbert Lemma, in conjunction with the Sobolev embedding $W_1^2(T) \subset C(\overline{T})$ (cf. [20, Lemma 4.3.4]), implies the existence of a linear

polynomial $\mathbf{p} \in [\mathbb{P}_1(T)]^3$ such that $\|\mathbf{f} - \mathbf{p}\|_{L^\infty(T)} \lesssim \|D^2\mathbf{f}\|_{L^1(T)}$. Since the mid-point quadrature is exact for linears, we deduce

$$\left| \int_T (\mathbf{f}(x_T) - \mathbf{f}) \right| = \left| \int_T \{ (\mathbf{f} - \mathbf{p})(x_T) + (\mathbf{p} - \mathbf{f}) \} \right| \le 2|T| \|\mathbf{f} - \mathbf{p}\|_{L^\infty(T)} \lesssim h_T^2 \|D^2\mathbf{f}\|_{L^1(T)}.$$

Moreover, invoking (46),

$$\|D^2\mathbf{f}\|_{L^1(T)} \lesssim \|D^3\mathbf{y}^\epsilon\|_{L^2(T)} \|\nabla\mathbf{y}^\epsilon\|_{L^2(T)} + \|D^2\mathbf{y}^\epsilon\|_{L^2(T)}^2 \lesssim \|\mathbf{y}^\epsilon\|_{H^3(T)}.$$

Inserting this back into $R_3(T)$ and adding we end up with

$$\limsup_{h \to 0^+} \sum_{i,j=1}^2 \sum_{T \in \mathcal{T}_h} |R_3(T)| \lesssim \limsup_{h \to 0^+} \left(h \|\overline{H}_h[\mathbf{y}_h]\|_{L^2(\Omega)} \right) \|\mathbf{y}^\epsilon\|_{H^3(\Omega)} = 0,$$

because of Lemma 3. Altogether, we arrive at

$$\limsup_{h \to 0^+} \left| C_h[\mathbf{y}_h] - C[\mathbf{y}^\epsilon] \right| \lesssim \epsilon$$

which implies the desired estimate (45). $\qquad\square$

It is worth realizing the role of the reduced discrete Hessian $\overline{H}_h[\mathbf{y}_h]$ in the preceding proof, namely that it factors out the integral defining $R_3(T)$. If we had used the discrete Hessian $H_h[\mathbf{y}_h]$ instead, then there would have been a term of the form $h_T^2 \|D^2 H_h[\mathbf{y}_h]\|_{L^2(T)}$ that could only be handled via an inverse inequality within the H^2-regularity setting. This in turn would have gotten rid of the factor h_T^2 and the proof of (44) would have failed.

3. Γ-CONVERGENCE

The reduced energy (8) consists of a bending energy $B[\mathbf{y}]$ and a cubic term $C[\mathbf{y}]$, and so does its discrete counterpart (38), namely $B_h[\mathbf{y}_h]$ and $C_h[\mathbf{y}_h]$. Compactness and Γ-convergence of the bending energy part, being similar to the single layer model, could be deduced from the results in [15]. For instance, we have that for any $\gamma_0, \gamma_1 > 0$, there exists a constant c_{coer} such that [15, (37) and (38)]

$$(47) \qquad c_{coer}^{-1} \|\mathbf{y}_h\|_{H_h^2(\Omega)}^2 \le B_h[\mathbf{y}_h] \le c_{cont} \|\mathbf{y}_h\|_{H_h^2(\Omega)}^2 \qquad \forall \mathbf{y}_h \in \mathbb{V}_h^k(\boldsymbol{\varphi}, \Phi),$$

and the constant $c_{coer} \to \infty$ if either γ_0 or $\gamma_1 \to 0^+$. In spite of that, [15] enforces the isometry constraint in average and constructs the recovery sequence needed for Γ-convergence via standard nodal interpolation. Therefore, the analysis below incorporates new ideas which do not follow from [15].

We start with the equicoercivity of energy E_h. The difficulty is dealing with C_h.

Lemma 8 (coercivity of total energy). *Let $\delta > 0$ and $\mathbf{y}_h \in \mathbb{A}_{h,\delta}$. There exists a constant $\tilde{c}_{coer} > 0$ independent of δ, but depending on the given data Z and $\mathcal{T}_h$ only through its shape regularity constant, such that*

$$(48) \qquad (2c_{coer})^{-1} \|\mathbf{y}_h\|_{H_h^2(\Omega)}^2 \le E_h[\mathbf{y}_h] + \tilde{c}_{coer}(1 + \delta)^2.$$

Proof. We write $B_h = E_h + C_h$ and employ (47) for B_h to obtain

$$c_{coer}^{-1}\|\mathbf{y}_h\|_{H_h^2(\Omega)}^2 \le E_h[\mathbf{y}_h] + C_h[\mathbf{y}_h].$$

It remains to estimate the cubic term $C_h[\mathbf{y}_h]$. Combining Lemma 6 (pointwise isometry constraint) with the Cauchy-Schwarz inequality yields

$$C_h[\mathbf{y}_h] \le \sum_{i,j=1}^{2} \sum_{T \in \mathcal{T}_h} |T| |\overline{H}_h[\mathbf{y}_h]_{ij} \cdot (\partial_1\mathbf{y}_h \times \partial_2\mathbf{y}_h)Z_{ij}|(x_T)$$

$$\le \sum_{i,j=1}^{2} \sum_{T \in \mathcal{T}_h} |T|^{\frac{1}{2}} \|\overline{H}_h[\mathbf{y}_h]_{ij}\|_{L^2(T)} |\partial_1\mathbf{y}_h(x_T)| |\partial_2\mathbf{y}_h(x_T)| \|Z\|_{L^\infty(T)}$$

$$\le 2(1+\delta)\|Z\|_{L^\infty(\Omega)} |\Omega|^{\frac{1}{2}} \|\overline{H}_h[\mathbf{y}_h]\|_{L^2(\Omega)}.$$

Invoking Lemma 3 (stability of $\overline{H}_h[\mathbf{y}_h]$) and Young's inequality yields

$$(49) \qquad \frac{1}{2c_{coer}}\|\mathbf{y}_h\|_{H_h^2(\Omega)}^2 \le E_h[\mathbf{y}_h] + 2c_{coer}c_{stab}^2|\Omega|\|Z\|_{L^\infty(\Omega)}^2(1+\delta)^2,$$

which is the desired estimate (48) with $\tilde{c}_{coer} = 2c_{coer}c_{stab}^2|\Omega|\|Z\|_{L^\infty(\Omega)}^2$.   □

We now prove a compactness result and Γ-convergence of E_h towards E, which consists of a $\liminf$ and a $\limsup$ property.

Theorem 9 (compactness and Γ-convergence). *Let $\delta = \delta(h) \to 0$ as $h \to 0^+$. Then*

(i) *Compactness: Let $\mathbf{y}_h \in \mathbb{A}_{h,\delta}$ be a sequence such that $E_h[\mathbf{y}_h]$ is uniformly bounded in h. Then there exists $\mathbf{y} \in \mathbb{A}$ such that $\mathbf{y}_h \to \mathbf{y}$ in $[L^2(\Omega)]^3$ and $\nabla_h\mathbf{y}_h \to \nabla\mathbf{y}$ in $[L^2(\Omega)]^{3\times2}$ for a subsequence (not relabeled).*

(ii) *Lim-inf property: If $\mathbf{y}_h \in \mathbb{A}_{h,\delta}$ satisfies $\mathbf{y}_h \to \mathbf{y}$ in $[L^2(\Omega)]^3$ and $\nabla_h\mathbf{y}_h \to \nabla\mathbf{y}$ in $[L^2(\Omega)]^{3\times2}$ where $\mathbf{y} \in [H^2(\Omega) \cap W_\infty^1(\Omega)]^3$, then $E[\mathbf{y}] \le \liminf_{h \to 0^+} E_h[\mathbf{y}_h]$.*

(iii) *Lim-sup property: For any $\mathbf{y} \in \mathbb{A}$ there exists $\mathbf{y}_h \in \mathbb{A}_{h,\delta}$ such that $\mathbf{y}_h \to \mathbf{y}$ in $[L^2(\Omega)]^3$ and $E[\mathbf{y}] \ge \limsup_{h \to 0^+} E_h[\mathbf{y}_h]$.*

Proof. We prove properties (i),(ii) and (iii) separately.

(i) Compactness. Lemma 8 (coercivity of total energy) and (25) imply

$$\|\mathbf{y}_h\|_{L^2(\Omega)} + \|\nabla_h\mathbf{y}_h\|_{L^2(\Omega)} + \|\mathbf{y}_h\|_{H_h^2(\Omega)} \lesssim 1.$$

Proceeding as in [18, Proposition 5.2], there exists $\mathbf{y} \in [H^2(\Omega)]^3$ satisfying the Dirichlet boundary conditions in (1) and $\mathbf{y}_h \to \mathbf{y}$ in $[L^2(\Omega)]^3$, $\nabla_h\mathbf{y}_h \to \nabla\mathbf{y}$ in $[L^2(\Omega)]^{3\times2}$.

It just remains to prove the isometry constraint $\mathrm{I}[\mathbf{y}] = I_2$ a.e. in Ω. To this end, recall that $\mathrm{I}[\mathbf{y}_h] = \nabla_h\mathbf{y}_h^T\nabla_h\mathbf{y}_h$, let $T \in \mathcal{T}_h$ and note that

$$\|\mathrm{I}[\mathbf{y}_h] - I_2\|_{L^1(T)} \le \|\mathrm{I}[\mathbf{y}_h] - \mathrm{I}[\mathbf{y}_h](x_T)\|_{L^1(T)} + |T| |\mathrm{I}[\mathbf{y}_h](x_T) - I_2|$$

$$\lesssim h_T\|D_h^2\mathbf{y}_h\|_{L^2(T)}\|\nabla_h\mathbf{y}_h\|_{L^2(T)} + \delta|T|,$$

because $\mathbf{y}_h \in \mathbb{A}_{h,\delta}$ whence $D_h[\mathbf{y}_h](x_T) = \big|\mathrm{I}[\mathbf{y}_h](x_T) - I_2\big| \le \delta$. Adding over T and employing the uniform boundedness of $\|D_h^2 \mathbf{y}_h\|_{L^2(\Omega)}$ and $\|\nabla_h \mathbf{y}_h\|_{L^2(\Omega)}$ results in

$$\|\mathrm{I}[\mathbf{y}_h] - I_2\|_{L^1(\Omega)} \lesssim h + \delta \to 0 \quad \text{as } h \to 0^+.$$

On the other hand, we see that

$$\mathrm{I}[\mathbf{y}_h] - \mathrm{I}[\mathbf{y}] = \nabla_h(\mathbf{y}_h - \mathbf{y})^T \nabla_h \mathbf{y}_h + \nabla \mathbf{y}^T \nabla_h(\mathbf{y}_h - \mathbf{y})$$

implies

$$\|\mathrm{I}[\mathbf{y}_h] - \mathrm{I}[\mathbf{y}]\|_{L^1(\Omega)} \le \big(\|\nabla \mathbf{y}\|_{L^2(\Omega)} + \|\nabla_h \mathbf{y}_h\|_{L^2(\Omega)}\big) \|\nabla_h \mathbf{y}_h - \nabla \mathbf{y}\|_{L^2(\Omega)} \to 0,$$

as $h \to 0^+$ because $\|\nabla_h \mathbf{y}_h\|_{L^2(\Omega)} \lesssim 1$. This and the triangle inequality lead to $\|\mathrm{I}[\mathbf{y}] - I_2\|_{L^1(\Omega)} = 0$ and consequently $\mathrm{I}[\mathbf{y}] = I_2$ a.e. in Ω, as desired.

(ii) lim-inf property. In view of Lemma 1 (weak convergence of H_h), we deduce $H_h[\mathbf{y}_h] \rightharpoonup D^2\mathbf{y}$ in $\big[L^2(\Omega)\big]^{3\times 2\times 2}$. The lower-semicontinuity of the L^2-norm under weak-limits together with the fact that the stabilization terms in $B_h[\mathbf{y}_h]$ are positive guarantee that

$$B[\mathbf{y}] = \frac{1}{2} \int_\Omega |D^2\mathbf{y}|^2 \le \liminf_{h\to 0^+} B_h[\mathbf{y}_h].$$

In addition, Lemma 7 (convergence of cubic energy) yields $\lim_{h\to 0^+} C_h[\mathbf{y}_h] = C[\mathbf{y}]$, and altogether gives $E[\mathbf{y}] \le \liminf_{h\to 0^+} E_h[\mathbf{y}_h]$ as asserted.

(iii) lim-sup property. The difficulty to construct a recovery sequence $\mathbf{y}_h \in \mathbb{A}_{h,\delta}$ is that the regularity $\mathbf{y} \in [H^2(\Omega) \cap W_\infty^1(\Omega)]^3$ is borderline to define pointwise values of $\nabla \mathbf{y}$ and thus enforce the isometry defect $D_h[\mathbf{y}_h](x_T)$ at every element barycenter x_T. Hence, we invoke the regularization procedure of P. Hornung [30]: given an isometry $\mathbf{y} \in [H^2(\Omega)]^3$ and $\epsilon > 0$, there exists an *isometry* $\mathbf{y}^\epsilon \in [H^3(\Omega)]^3$ such that

(50) $$\|\mathbf{y} - \mathbf{y}^\epsilon\|_{H^2(\Omega)} \lesssim \epsilon, \qquad \|D^2\mathbf{y}^\epsilon\|_{L^2(\Omega)} \lesssim \|D^2\mathbf{y}\|_{L^2(\Omega)}.$$

As usual, the constants hidden in the symbol $\lesssim$ are independent of h and ϵ. We now set $\mathbf{y}_h := R_h[\mathbf{y}^\epsilon]$, where the recovery operator $R_h : [H^3(\Omega)]^3 \to [\mathbb{V}_h^k]^3$ is the following quadratic Taylor expansion about x_T for every $T \in \mathcal{T}_h$

(51) $$R_h[\mathbf{w}](x) := \mathbf{w}(x_T) + \nabla \mathbf{w}(x_T)(x - x_T) + \frac{1}{2}(x - x_T)^T Q_T[\mathbf{w}](x - x_T) \quad \forall\, x \in T,$$

where $Q_T[\mathbf{w}] := \frac{1}{|T|} \int_T D^2\mathbf{w}$. Note that $\nabla \mathbf{y}_h(x_T) = \nabla \mathbf{y}^\epsilon(x_T)$ and $D_h[\mathbf{y}_h](x_T) = 0$, whence $\mathbf{y}_h \in \mathbb{A}_{h,0} \subset \mathbb{A}_{h,\delta}$. We next show the two convergence properties of $\mathbf{y}_h$ in (ii).

Since $R_h\big|_T$ is invariant over the space $\mathbb{P}_1$ of polynomials of degree ≤ 1, we have

$$\mathbf{w} - R_h[\mathbf{w}] = (\mathbf{w} - \mathbf{p}) - R_h[\mathbf{w} - \mathbf{p}] \quad \forall\, \mathbf{p} \in [\mathbb{P}_1]^3.$$

Therefore, combining the stability in $W_\infty^1(T)$ of the linear part of R_h with the Bramble-Hilbert lemma and the property $\|Q_T[\mathbf{w}]\|_{L^2(T)} \le |\mathbf{w}|_{H^2(T)}$, we deduce

$$\|\mathbf{w} - R_h[\mathbf{w}]\|_{H^1(T)} \lesssim h_T \|\nabla(\mathbf{w} - \mathbf{p})\|_{W_\infty^1(T)} + h_T \|Q_T[\mathbf{w}]\|_{L^2(T)}$$
$$\lesssim h_T^2 \|\mathbf{w}\|_{H^3(T)} + h_T |\mathbf{w}|_{H^2(T)} \lesssim h_T \|\mathbf{w}\|_{H^3(T)}.$$

Notice the presence of the full H^3-norm on the right-hand side of the above estimate, which accounts for possible subdivisions made of quadrilaterals [23, 27, 18]. We next square and add over $T \in \mathcal{T}_h$ to obtain

$$\|\mathbf{w} - R_h[\mathbf{w}]\|_{L^2(\Omega)} + \|\nabla\mathbf{w} - \nabla_h R_h[\mathbf{w}]\|_{L^2(\Omega)} \lesssim h\|\mathbf{w}\|_{H^3(\Omega)}.$$

This estimate for $\mathbf{w} = \mathbf{y}^\epsilon$, in conjunction with (50), yields

$$\|\mathbf{y} - \mathbf{y}_h\|_{L^2(\Omega)} + \|\nabla\mathbf{y} - \nabla_h\mathbf{y}_h\|_{L^2(\Omega)} \lesssim \epsilon + h\|\mathbf{y}^\epsilon\|_{H^3(\Omega)},$$

whence $\|\mathbf{y} - \mathbf{y}_h\|_{L^2(\Omega)} \lesssim \epsilon$ provided h is sufficiently small so that $h\|\mathbf{y}^\epsilon\|_{H^3(\Omega)} \leq \epsilon$. This shows the asserted convergence $\mathbf{y}_h \to \mathbf{y}$ in $[L^2(\Omega)]^2$ because ϵ is arbitrary.

It remains to show the convergence $E_h[\mathbf{y}_h] \to E[\mathbf{y}]$ as $h \to 0^+$, which in turn implies the desired lim-sup property. Since $D^2\mathbf{y}_h = Q_T[\mathbf{y}^\epsilon]$, we infer that

$$\|D_h^2\mathbf{y}_h\|_{L^2(\Omega)}^2 = \sum_{T\in\mathcal{T}_h} \|Q_T[\mathbf{y}^\epsilon]\|_{L^2(T)}^2 \leq \sum_{T\in\mathcal{T}_h} \|D^2\mathbf{y}^\epsilon\|_{L^2(T)}^2 \lesssim \|D^2\mathbf{y}\|_{L^2(\Omega)}^2,$$

according to (50). Moreover,

$$\|D_h^2\mathbf{y}_h - D^2\mathbf{y}\|_{L^2(\Omega)}^2 = \sum_{T\in\mathcal{T}_h} \|Q_T[\mathbf{y}^\epsilon] - D^2\mathbf{y}\|_{L^2(T)}^2$$

$$\leq \sum_{T\in\mathcal{T}_h} \|Q_T[\mathbf{y}^\epsilon] - D^2\mathbf{y}^\epsilon\|_{L^2(T)}^2 + \|D^2\mathbf{y}^\epsilon - D^2\mathbf{y}\|_{L^2(T)}^2 \lesssim h^2\|\mathbf{y}^\epsilon\|_{H^3(\Omega)}^2 + \epsilon^2$$

shows that $D_h^2\mathbf{y}_h \to D^2\mathbf{y}$ and Lemma 2 (strong convergence of H_h) gives

$$H_h[\mathbf{y}_h] \to D^2\mathbf{y} \quad \text{strongly in} \quad [L^2(\Omega)]^{3\times2\times2}.$$

An argument similar to [15, Appendix B and C], invoking the trace inequality, yields

$$S_h[\mathbf{y}_h] \lesssim \sum_{T\in\mathcal{T}_h} \|\mathbf{y} - \mathbf{y}_h\|_{H^2(T)}^2 \to 0, \quad \text{as } h \to 0^+$$

for the stabilization energy $S_h[\mathbf{y}_h]$ in (40) and implies convergence of the bending energy B_h in (39), namely $\lim_{h\to0^+} B_h[\mathbf{y}_h] = B[\mathbf{y}]$. Finally, in view of the preceding discussion, we see that the assumptions of Lemma 7 (convergence of the cubic energy) are valid, whence Lemma 7 implies $C_h[\mathbf{y}_h] \to C[\mathbf{y}]$ and completes the proof.   □

The construction of the recovery sequence in Theorem 9 (compactness and Γ-convergence) is closely related to Lemma 7 (convergence of the cubic energy) and illustrates the crucial interplay between enforcing the isometry defect $D_h[\mathbf{y}_h]$ at barycenters and the mid-point quadrature rule in the cubic energy $C_h[\mathbf{y}_h]$. This, however, limits the accuracy of LDG to that of lowest polynomial degree $k = 2$. We leave the design of an LDG method with formal higher accuracy $k > 2$ open.

Corollary 10 (convergence of global minimizers). *If $\mathbf{y}_h \in \mathbb{A}_{h,\delta}$ is an almost global minimizer of E_h in the sense that*

$$E_h[\mathbf{y}_h] \leq \inf_{\mathbf{w}_h\in\mathbb{A}_{h,\delta}} E_h[\mathbf{w}_h] + \sigma$$

where $\sigma, \delta \to 0$ as $h \to 0^+$, then $\{\mathbf{y}_h\}_{h>0}$ is precompact in $[L^2(\Omega)]^3$ and every cluster point $\mathbf{y}$ belongs to $\mathbb{A}$ and is a global minimizer of E, namely $E[\mathbf{y}] = \inf_{\mathbf{w} \in \mathbb{A}} E[\mathbf{w}]$. Moreover, up to a subsequence (not relabeled) the energies converge

$$E[\mathbf{y}] = \lim_{h \to 0^+} E_h[\mathbf{y}_h].$$

We omit the proof of Corollary 10, which readily follows from Theorem 9 (compactness and Γ-convergence), and refer instead to [4, 5, 8, 18] for details.

4. BILAYER MODEL WITH CREASES

Bartels, Bonito and Hornung have recently developed a reduced single layer model that allows for folding across creases [6]. The resulting two-dimensional model hinges on a general hyperelastic material description with appropriate scaling conditions on the energy, and consists of a piecewise nonlinear Kirchhoff plate bending model with a continuity condition at the creases. For a prescribed Lipschitz curve $\mathcal{C}$ intersecting the boundary of Ω transversally, the modified bending energy of [6] reads

$$\widetilde{B}[\mathbf{y}] := \frac{1}{2} \int_{\Omega \smallsetminus \mathcal{C}} |\mathrm{II}[\mathbf{y}]|^2 = \frac{1}{2} \int_{\Omega \smallsetminus \mathcal{C}} |D^2 \mathbf{y}|^2$$

for deformations $\mathbf{y} \in [H^2(\Omega \smallsetminus \mathcal{C}) \cap W^{1,\infty}(\Omega)]^3$ satisfying the isometry constraint $\mathrm{I}[\mathbf{y}] = I_2$ along with possible boundary conditions. Properly designed creases allow for flapping mechanisms upon actuation at the boundary which are of interest in engineering and medicine [6].

In this section we explore a similar modification of the elastic energy (4)

(52) $$\widetilde{E}[\mathbf{y}] := \frac{1}{2} \int_{\Omega \smallsetminus \mathcal{C}} |\mathrm{II}[\mathbf{y}] - Z|^2,$$

but without justification from 3d hyperelasticity. Therefore, we leave open the question whether this energy is the appropriate Γ-limit for bilayer materials. We also modify the admissible set to be

$$\widetilde{\mathbb{A}} := \left\{ \mathbf{y} \in [H^2(\Omega \smallsetminus \mathcal{C}) \cap W^1_\infty(\Omega)]^3 : \quad \mathrm{I}[\mathbf{y}] = I_2 \ \text{ in } \Omega, \quad \mathbf{y} = \boldsymbol{\varphi}, \ \nabla \mathbf{y} = \Phi \text{ on } \Gamma^D \right\}.$$

Our goal is, instead, to investigate the relation between (52) and its fully discrete counterpart, and demonstrate computationally the crucial role of spontaneous curvature Z to produce plate folding without actuation via boundary conditions.

We extend our LDG method to account for creases as in [6]. We consider isoparametric partitions $\mathcal{T}_h$ made of possibly curved elements, i.e. the mapping F_T used to define the finite element space $\mathbb{V}^k_h$ locally is $[\mathbb{Q}_2]^2$ instead of $[\mathbb{Q}_1]^2$ (or $[\mathbb{P}_2]^2$ instead of $[\mathbb{P}_1]^2$). We further assume that the crease $\mathcal{C}$ is exactly matched by $\mathcal{T}_h$:

(53) $$\mathcal{C} \text{ is made of piecewise quadratic edges } e_1, ..., e_J \in \mathcal{E}_h.$$

This geometric assumption is restrictive but instrumental for the theory below. Dealing with more general creases $\mathcal{C}$, just interpolated by $\mathcal{E}_h$, is important and the subject of current research; we refer to [6, Section 4.4] for some discussion.

The distributional derivative of $\mathbf{y} \in [H^2(\Omega \smallsetminus \mathcal{C}) \cap W^1_\infty(\Omega)]^3$ reads

$$D^2 \mathbf{y} = \widetilde{D}^2 \mathbf{y} + [\nabla \mathbf{y}] \otimes \mathbf{n} \, \delta_{\mathcal{C}},$$

where $\widetilde{D}^2\mathbf{y}$ stands for the absolutely continuous part of $D^2\mathbf{y}$, or restriction of $D^2\mathbf{y}$ to $\Omega \smallsetminus \mathcal{C}$ that happens to be L^2, while $[\nabla\mathbf{y}] \otimes \mathbf{n}\,\delta_{\mathcal{C}}$ is the singular part supported on $\mathcal{C}$ and $\mathbf{n}$ is a unit normal vector to $\mathcal{C}$. The first issue to tackle is the construction of a discrete Hessian $\widetilde{H}_h[\mathbf{y}_h]$ that allows for folding across $\mathcal{C}$ and mimics $\widetilde{D}^2\mathbf{y}$. As in [6], we replace the global lift $\mathcal{R}_h$ in (28) by

$$\widetilde{\mathcal{R}}_h := \sum_{e \in \mathcal{E}_h^a \smallsetminus \{e_1,\dots,e_J\}} r_e,$$

where $\{e_j\}_{j=1}^J$ are defined in (53), and let the *modified discrete Hessian* be

$$\widetilde{H}_h[\mathbf{y}_h] := D_h^2\mathbf{y}_h - \widetilde{\mathcal{R}}_h([\nabla_h\mathbf{y}_h]) + \mathcal{B}_h([\mathbf{y}_h]).$$

We likewise replace (23) by the modified mesh-dependent form $\langle\cdot,\cdot\rangle_{\widetilde{H}_h^2}$

$$\langle\mathbf{v}_h, \mathbf{w}_h\rangle_{\widetilde{H}_h^2} := (D_h^2\mathbf{v}_h, D_h^2\mathbf{w}_h)_{L^2(\Omega)}$$
$$+ (\mathrm{h}^{-1}[\nabla_h\mathbf{v}_h], [\nabla_h\mathbf{w}_h])_{L^2(\Gamma_h^a\smallsetminus\mathcal{C})} + (\mathrm{h}^{-3}[\mathbf{v}_h], [\mathbf{w}_h])_{L^2(\Gamma_h^a)}.$$

In essence, the ability for the plates to fold freely across $\mathcal{C}$ is reflected in the absence of all the contributions related to $[\nabla\mathbf{y}_h]$ across $\mathcal{C}$. This is the key to the following lemma whose proof follows along the lines of [15, Appendix B] and is thus omitted.

Lemma 11 (convergence of $\widetilde{H}_h$). *Let the crease $\mathcal{C}$ satisfy* (53). *Then there holds*

(i) *Weak convergence: If $k \geq 2$ and $\mathbf{v}_h \in \mathbb{V}_h^k(\boldsymbol{\varphi}, \Phi)$ satisfies $\|\mathbf{v}_h\|_{H_h^2(\Omega)} \lesssim 1$ and $\mathbf{v}_h \to \mathbf{v} \in [H^2(\Omega \smallsetminus \mathcal{C}) \cap H^1(\Omega)]^3$ in $[L^2(\Omega)]^3$ as $h \to 0^+$, then we have*

$$\widetilde{H}_h[\mathbf{v}_h] \rightharpoonup \widetilde{D}^2\mathbf{v} \quad in \; [L^2(\Omega)]^{3\times2\times2} \quad as \; h \to 0^+.$$

(ii) *Strong convergence: Let $\mathbf{v} \in [H^2(\Omega \smallsetminus \mathcal{C})]^3$ be any function such that $\mathbf{v} = \boldsymbol{\varphi}$ and $\nabla\mathbf{v} = \Phi$ on Γ^D. Moreover, let $\mathbf{v}_h \in \mathbb{V}_h^k(\boldsymbol{\varphi}, \Phi)$ satisfy*

$$\|D^2\mathbf{v}_h\|_{L^2(T)} \lesssim \|\mathbf{v}\|_{H^2(T)} \; \forall T \in \mathcal{T}_h, \quad \sum_{T\in\mathcal{T}_h} \|\mathbf{v}_h - \mathbf{v}\|_{H^2(T)}^2 \to 0 \; as \; h \to 0^+.$$

Then we have as $h \to 0^+$

$$\widetilde{H}_h[\mathbf{v}_h] \to \widetilde{D}^2\mathbf{v} \quad strongly \; in \; [L^2(\Omega)]^{3\times2\times2}.$$

We are now ready to introduce the LDG approximation of $\widetilde{E}[\mathbf{y}]$ in (52), namely

$$\widetilde{E}_h[\mathbf{y}_h] := \widetilde{B}_h[\mathbf{y}_h] + \widetilde{C}_h[\mathbf{y}_h],$$

where

$$\widetilde{B}_h[\mathbf{y}_h] := \frac{1}{2} \int_\Omega |\widetilde{H}_h[\mathbf{y}_h]|^2 + \gamma_1 \|\mathrm{h}^{-\frac{1}{2}}[\nabla_h\mathbf{y}_h]\|_{L^2(\Gamma_h^a\smallsetminus\mathcal{C})}^2 + \gamma_0 \|\mathrm{h}^{-\frac{3}{2}}[\mathbf{y}_h]\|_{L^2(\Gamma_h^a)}^2,$$

and

$$\widetilde{C}_h[\mathbf{y}_h] := \sum_{i,j=1}^2 \sum_{T\in\mathcal{T}_h} |T|\,\overline{H}_h[\mathbf{y}_h]_{ij} \cdot (\partial_1\mathbf{y}_h \times \partial_2\mathbf{y}_h)(x_T)\, Z_{ij},$$

with $\overline{H}_h[\mathbf{y}_h]|_T := \frac{1}{|T|}\int_T \widetilde{H}_h[\mathbf{y}_h]$. Lemmas 3 and 4 are valid for $\overline{H}_h[\mathbf{y}_h]$, as well as Lemma 7 (convergence of cubic energy) and Lemma 8 (coercivity of total energy).

It remains to examine the convergence of the discrete global minimizers towards the continuous global minimizers. Assume that the crease $\mathcal{C}$ splits Ω into two disjoint sets Ω_1 and Ω_2. Since Hornung's regularization procedure [30] cannot guarantee general Dirichlet boundary conditions, it is not clear how to regularize in Ω_1 and Ω_2 functions that belong to $[H^2(\Omega \smallsetminus \mathcal{C}) \cap W_\infty^1(\Omega)]^3$ and yet maintain the location of the crease $\mathcal{C}$, namely obtain an isometry in $[H^3(\Omega \smallsetminus \mathcal{C}) \cap W_\infty^1(\Omega)]^3$. Another obstruction stems from the use of curved elements necessarily for the subdivisions to match the crease. When using polynomial mappings from the reference to the physical elements, the resulting finite element functions are not necessarily polynomial in the physical element, thereby ruling out the construction of the recovery sequence proposed to guarantee the $\lim\sup$ property; see Theorem 9(ii).

We circumvent these issues by requiring slightly more smoothness on *one* of the global minimizers $\mathbf{y}$, which in turn allows for a different, more generic, construction of its recovery sequence. Because the additional regularity cannot be derived from our Γ-convergence theory, we assume the existence of a global minimizer $\mathbf{y}^* \in \widetilde{\mathbb{A}}$ of $\widetilde{E}$ with the following property

$$(54) \qquad\qquad \mathbf{y}^*|_{\Omega_i} \in C^1(\overline{\Omega}_i), \qquad i = 1, 2.$$

Note that the above assumption is consistent with practical configurations. We also point out that this regularity assumption and the fact that the subdivision matches the crease entail the existence of a modulus of smoothness ω so that

$$(55) \qquad\qquad |\nabla \mathbf{y}^*(x) - \nabla \mathbf{y}^*(z)| \leq \omega(h_T) \quad \forall\, x, z \in T, \ \forall\, T \in \mathcal{T}_h,$$

with $\omega(s) \to 0$ as $s \to 0^+$.

The construction of the recovery sequence for deformations satisfying the additional regularity (54) is then based on a piecewise averaged Taylor polynomial. The latter does not preserve the isometry constraint pointwise but (54) allows for control of the isometry defect.

Before embarking on the proofs, we recall a useful result on the averaged Taylor polynomial [20] defined on the reference element $\widehat{T}$ (see Section 2.2). Until the end of this section, we consider the case when the reference element is a square and $\mathbb{Q}_k$ finite element functions are used. The case where $\widehat{T}$ is the unit simplex is somewhat simpler and can be dealt with similarly. Let $\widehat{B}$ be a ball centered at the barycenter of $\widehat{T}$ such that its closure is contained in $\widehat{T}$ and $\widehat{\zeta}$ be a cut-off function with unit mass supported on the closure of $\widehat{B}$. For $\widehat{w} \in L^1(\widehat{T})$ let

$$(56) \qquad Q[\widehat{w}](\hat{x}) := \sum_{|\alpha|_\infty \leq 2} \int_{\widehat{B}} \frac{1}{\alpha!} \widehat{D}^\alpha \widehat{w}(\hat{z})(\hat{x} - \hat{z})^\alpha \widehat{\zeta}(\hat{z}) d\hat{z} \in \mathbb{Q}_2,$$

be the averaged Taylor polynomial where $\alpha := (\alpha_1, \alpha_2)$ is a multi-index with non-negative integers α_1, α_2 and $|\alpha|_\infty := \max\{\alpha_1, \alpha_2\}$. We recall the following useful properties of Q and refer to [20] for additional details: Q preserves $\mathbb{Q}_2$ on $\widehat{T}$

$$(57) \qquad\qquad Q[\widehat{p}] = \widehat{p}, \qquad \widehat{p} \in \mathbb{Q}_2,$$

is stable

$$(58) \qquad \|Q[\widehat{w}]\|_{W_\infty^k(\widehat{T})} \lesssim \|\widehat{w}\|_{L^1(\widehat{B})}, \qquad \forall k \in \mathbb{N}$$

and convergent

$$(59) \qquad |\widehat{w} - Q[\widehat{w}]|_{H^k(\widehat{T})} \lesssim \left(\sum_{i=1}^{2} \|\partial_{\widehat{x}_i}^3 \widehat{w}\|_{L^2(\widehat{T})}^2 \right)^{1/2}, \qquad 0 \le k < 3.$$

We next discuss estimates for isoparametric mappings $F_T : \widehat{T} \to T$ between the reference element $\widehat{T}$ and $T \in \mathcal{T}_h$ so that $F_T \in [\mathbb{Q}_2]^2$. They establish relationship between norms on $\widehat{T}$ and T, as well as provide an interpolation estimate in $[\mathbb{V}_h^2]^3$. In fact, for $v \in H^2(T)$ and $\widehat{v} = v \circ F_T \in H^2(\widehat{T})$, there holds

$$(60) \qquad \|\widehat{v}\|_{L^2(\widehat{T})} \approx h_T^{-1} \|v\|_{L^2(T)}, \quad \|\widehat{v}\|_{L^\infty(\widehat{T})} \approx \|v\|_{L^\infty(T)}$$

$$(61) \qquad \|\widehat{\nabla}\widehat{v}\|_{L^2(\widehat{T})} \approx \|\nabla v\|_{L^2(T)}, \quad \|\widehat{\nabla}\widehat{v}\|_{L^\infty(\widehat{T})} \approx h_T \|\nabla v\|_{L^\infty(T)},$$

$$(62) \qquad \|\widehat{D}^2\widehat{v}\|_{L^2(\widehat{T})} \lesssim h_T \|D^2 v\|_{L^2(T)} + \|\nabla v\|_{L^2(T)},$$

$$(63) \qquad \|D^2 v\|_{L^2(T)} \lesssim h_T^{-1} \|\widehat{v}\|_{H^2(\widehat{T})},$$

whence for $m = 0, 1, 2$

$$(64) \qquad |v|_{H^m(T)} \lesssim h_T^{1-m} \|\widehat{v}\|_{H^2(\widehat{T})}.$$

Moreover if $v \in H^3(T)$, we further obtain

$$(65) \qquad \|\widehat{D}^3\widehat{v}\|_{L^2(\widehat{T})} \lesssim h_T^2 \|D^3 v\|_{L^2(T)} + h_T \|D^2 v\|_{L^2(T)} + \|\nabla v\|_{L^2(T)}.$$

Note that the first four estimates are discussed and proved in [18, Appendix], while one can extend the proof of (62) to show (65). Additionally, as in [18, Lemma A.4], the local Lagrange interpolant $I_h w \in [\mathbb{V}_h^2]^3$ for any $w \in H^3(T)$ satisfies the estimate

$$(66) \qquad |w - I_h w|_{H^m(T)} \lesssim h^{3-m} \|v\|_{H^3(T)},$$

for $0 \le m \le 3$.

The next lemma describes the modified $\limsup$ property which hinges on the averaged Taylor polynomial (56); compare with Theorem 9(iii).

Lemma 12 ($\limsup$ property with creases)**.** *Let $\mathbf{y}^* \in \widetilde{\mathbb{A}}$ satisfy the regularity assumption (54) and let ω be the modulus of smoothness in (55). There is a constant c and $\mathbf{y}_h^* \in \mathbb{A}_{h,c\omega(h)}$ such that $\mathbf{y}_h^* \to \mathbf{y}^*$ in $[L^2(\Omega)]^3$ and $\lim_{h \to 0^+} \widetilde{E}_h[\mathbf{y}_h] = \widetilde{E}[\mathbf{y}^*]$.*

Proof. As usual, the hat symbol denotes quantities defined on the reference element $\widehat{T}$. Let $\mathbf{y}_h^* \in [\mathbb{V}_h^2]^3$ be defined locally by

$$\mathbf{y}_h^*|_T := \widehat{\mathbf{y}}_h^*|_T \circ F_T^{-1}, \quad \widehat{\mathbf{y}}_h^*|_T := Q[\mathbf{y}^* \circ F_T] \quad \forall T \in \mathcal{T}_h,$$

where Q is given in (56) and is applied component-wise. Note that by construction we indeed have $\mathbf{y}_h^* \in [\mathbb{V}_h^2]^3$. The rest of the proof consists of 4 steps.

Step (i): isometry defect. We claim the intermediate estimates

(67) $$\|\nabla \mathbf{y}_h^*\|_{L^\infty(T)} \lesssim \|\nabla \mathbf{y}^*\|_{L^\infty(T)}, \quad \|\nabla(\mathbf{y}^* - \mathbf{y}_h^*)\|_{L^\infty(T)} \lesssim \omega(h_T) \quad \forall\, T \in \mathcal{T}_h.$$

To show the first estimate, we use (61) and (57) to write

$$\|\nabla \mathbf{y}_h^*\|_{L^\infty(T)} \lesssim h_T^{-1} \|\widehat{\nabla}\widehat{\mathbf{y}}_h^*\|_{L^\infty(\widehat{T})} \lesssim h_T^{-1} \|\widehat{\nabla}(\widehat{\mathbf{y}}_h^* - \mathbf{c})\|_{L^\infty(\widehat{T})} = h_T^{-1} \|\widehat{\nabla}Q[\widehat{\mathbf{y}}^* - \mathbf{c}]\|_{L^\infty(\widehat{T})},$$

where $\mathbf{c} := |T|^{-1} \int_T \mathbf{y}^* \in \mathbb{R}^3$ is the average of $\mathbf{y}^*$. We then employ the stability (58) of Q, together with (60) and Poincáre inequality, to deduce

$$\|\nabla \mathbf{y}_h^*\|_{L^\infty(T)} \lesssim h_T^{-1} \|\widehat{\mathbf{y}}^* - \mathbf{c}\|_{L^\infty(\widehat{T})} \lesssim h_T^{-1} \|\mathbf{y}^* - \mathbf{c}\|_{L^\infty(T)} \lesssim \|\nabla \mathbf{y}^*\|_{L^\infty(T)},$$

which is the first estimate in (67).

To prove the second estimate in (67), we first notice that for any $\mathbf{p} \in [\mathbb{P}_1]^3$ we have $\widehat{\mathbf{p}} := \mathbf{p} \circ F_T \in [\mathbb{Q}_2]^3$. Since $\widehat{\mathbf{p}} = Q[\widehat{\mathbf{p}}]$, according to (57), we proceed as before, but now using $\widehat{\mathbf{p}} \in \mathbb{Q}_2$ instead of the constant $\mathbf{c}$ along with (61), to write

$$\|\nabla(\mathbf{y}^* - \mathbf{y}_h^*)\|_{L^\infty(T)} \lesssim h_T^{-1} \left(\|\widehat{\nabla}(\widehat{\mathbf{y}}^* - \widehat{\mathbf{p}})\|_{L^\infty(\widehat{T})} + \|\widehat{\nabla}Q[\widehat{\mathbf{y}}^* - \widehat{\mathbf{p}}]\|_{L^\infty(\widehat{T})} \right)$$

(68) $$\lesssim h_T^{-1} \|\widehat{\mathbf{y}}^* - \widehat{\mathbf{p}}\|_{W_\infty^1(\widehat{T})}$$

$$\lesssim h_T^{-1} \|\mathbf{y}^* - \mathbf{p}\|_{L^\infty(T)} + \|\nabla(\mathbf{y}^* - \mathbf{p})\|_{L^\infty(T)}.$$

We next choose $\mathbf{y}^*$ to take advantage of the piecewise smoothness (54), namely

$$\mathbf{p}(x) := \mathbf{y}^*(x_T) + \nabla \mathbf{y}^*(x_T)(x - x_T).$$

The property (55) of ω implies

$$\|\nabla(\mathbf{y}^* - \mathbf{p})\|_{L^\infty(T)} \leq \omega(h_T)$$

and combined with $\mathbf{y}^*(x) - \mathbf{y}^*(x_T) = \nabla \mathbf{y}^*(\xi)(x - x_T)$ for some $\xi \in T$, gives

$$\|\mathbf{y}^* - \mathbf{p}\|_{L^\infty(T)} \leq h_T\,\omega(h_T).$$

Inserting these estimates in (68) yields the second estimate in (67).

The estimate on the isometry defect $\|(\nabla \mathbf{y}_h^*)^T \nabla \mathbf{y}_h^* - I_2\|_{L^\infty(T)}$ follows directly from the intermediate estimates (67) and the assumption $\mathrm{I}[\mathbf{y}^*] = I_2$

$$\|(\nabla \mathbf{y}_h^*)^T \nabla \mathbf{y}_h^* - I_2\|_{L^\infty(T)} = \|(\nabla \mathbf{y}_h^*)^T \nabla \mathbf{y}_h^* - (\nabla \mathbf{y}^*)^T \nabla \mathbf{y}^*\|_{L^\infty(T)}$$

$$\leq (\|\nabla \mathbf{y}^*\|_{L^\infty(T)} + \|\nabla \mathbf{y}_h^*\|_{L^\infty(T)})\|\nabla(\mathbf{y}_h^* - \mathbf{y}^*)\|_{L^\infty(T)} \leq c\,\omega(h_T).$$

for a constant c independent of the discretization parameters; hence $\mathbf{y}_h^* \in \mathbb{A}_{h,c\omega(h)}$.

Step (ii): Broken H^2- Stability. For $\mathbf{p} \in [\mathbb{P}_1]^3$, we set $\widehat{\mathbf{p}} := \mathbf{p} \circ F_T \in [\mathbb{Q}_2]^3$ to get

$$\|D^2 \mathbf{y}_h^*\|_{L^2(T)} = \|D^2(\mathbf{y}_h^* - \mathbf{p})\|_{L^2(T)} \lesssim h_T^{-1} \|\widehat{\mathbf{y}}_h^* - \widehat{\mathbf{p}}\|_{H^2(\widehat{T})},$$

in view of (63). Thanks to the invariance (57) and stability (58) of Q, we obtain

$$\|D^2 \mathbf{y}_h^*\|_{L^2(T)} \lesssim h_T^{-1} \|Q[\widehat{\mathbf{y}}^* - \widehat{\mathbf{p}}]\|_{H^2(\widehat{T})} \lesssim h_T^{-1} \|\widehat{\mathbf{y}}^* - \widehat{\mathbf{p}}\|_{L^2(\widehat{T})}.$$

This, together with (60) and a standard interpolation estimate on T, yields

$$\|D^2 \mathbf{y}_h^*\|_{L^2(T)} \lesssim h_T^{-2} \|\mathbf{y}^* - \mathbf{p}\|_{L^2(T)} \lesssim \|D^2 \mathbf{y}^*\|_{L^2(T)},$$

which is the desired stability estimate.

Step (iii): H^2-Convergence. We exploit a density argument. For any $\epsilon > 0$, there exists $\mathbf{y}^\epsilon$ so that $\mathbf{y}^\epsilon|_{\Omega_i} \in [H^3(\Omega_i)]^3$ and $\|\mathbf{y}^* - \mathbf{y}^\epsilon\|_{H^2(\Omega_i)} \le \epsilon$ for $i = 1, 2$; $\mathbf{y}^\epsilon$ may not be an isometry and not be continuous across $\mathcal{C}$. Exploiting the fact that $\mathcal{C}$ is exactly matched by $\mathcal{T}_h$, we split

(69)
$$\sum_{T \in \mathcal{T}_h} \|\mathbf{y}^* - \mathbf{y}_h^*\|_{H^2(T)}^2 \lesssim \sum_{i=1}^2 \|\mathbf{y}^* - \mathbf{y}^\epsilon\|_{H^2(\Omega_i)}^2 + \sum_{T \in \mathcal{T}_h} \|\mathbf{y}^\epsilon - \mathbf{y}_h^\epsilon\|_{H^2(T)}^2 + \sum_{T \in \mathcal{T}_h} \|\mathbf{y}_h^\epsilon - \mathbf{y}_h^*\|_{H^2(T)}^2$$

with $\mathbf{y}_h^\epsilon = \widehat{\mathbf{y}}_h^\epsilon \circ F_T^{-1}$ and $\widehat{\mathbf{y}}_h^\epsilon := Q[\mathbf{y}^\epsilon \circ F_T]$, and estimate each of the three terms separately. The first term is obviously bounded by ϵ^2.

For the second term, we let $m = 0, 1, 2$ and combine (59) with (64) to arrive at

$$|\mathbf{y}_h^\epsilon - \mathbf{y}^\epsilon|_{H^m(T)} \lesssim h_T^{1-m} \left(\sum_{i=1}^2 \|\partial_{\widehat{x}_i}^3 \widehat{\mathbf{y}}^\epsilon\|_{L^2(\widehat{T})}^2 \right)^{1/2} \lesssim h_T^{1-m} \left(\sum_{i=1}^2 \|\partial_{\widehat{x}_i}^3 (\widehat{\mathbf{y}}^\epsilon - \widehat{I_h \mathbf{y}^\epsilon})\|_{L^2(\widehat{T})}^2 \right)^{1/2},$$

where I_h is the local Lagrange interpolant onto $[\mathbb{V}_h^2]^3$ and $\widehat{I_h \mathbf{y}^\epsilon} := I_h \mathbf{y}^\epsilon \circ F_T^{-1} \in [\mathbb{Q}_2]^3$. As a consequence, using the estimate (65) to map back to the physical element T, and applying the error estimate (66) for I_h to the ensuing terms, we deduce

$$|\mathbf{y}_h^\epsilon - \mathbf{y}^\epsilon|_{H^m(T)} \lesssim h_T^{3-m} \|\mathbf{y}^\epsilon\|_{H^3(T)},$$

for $m = 0, 1, 2$. We thus conclude

$$\sum_{T \in \mathcal{T}_h} \|\mathbf{y}^\epsilon - \mathbf{y}_h^\epsilon\|_{H^2(T)}^2 \lesssim h^2 \sum_{i=1}^2 \|\mathbf{y}^\epsilon\|_{H^3(\Omega_i)}^2.$$

It remains to estimate the third term $\|\mathbf{y}_h^* - \mathbf{y}_h^\epsilon\|_{H^2(T)}$. To deal with each term $|\mathbf{y}_h^* - \mathbf{y}_h^\epsilon|_{H^m(T)}$ for $m = 0, 1, 2$, we let $\mathbf{p}_{m-1} \in \mathbb{P}_{m-1}$ to be chosen later with the convention that $\mathbf{p}_{-1} = \mathbf{0}$, and use the invariance $Q[\mathbf{p} \circ F_T] = \mathbf{p} \circ F_T \in [\mathbb{Q}_2]^3$ for any $\mathbf{p} \in [\mathbb{P}_1]^3$. Combining (64) with the stability (58) of Q yields

$$|\mathbf{y}_h^* - \mathbf{y}_h^\epsilon|_{H^m(T)} = |\mathbf{y}_h^* - \mathbf{y}_h^\epsilon - \mathbf{p}_{m-1}|_{H^m(T)} \lesssim h_T^{1-m} \|Q[\widehat{\mathbf{y}}^* - \widehat{\mathbf{y}}^\epsilon - \mathbf{p}_{m-1} \circ F_T]\|_{H^2(\widehat{T})}$$
$$\lesssim h_T^{1-m} \|\widehat{\mathbf{y}}^* - \widehat{\mathbf{y}}^\epsilon - \mathbf{p}_{m-1} \circ F_T\|_{L_2(\widehat{T})} \lesssim h_T^{-m} \|\mathbf{y}^* - \mathbf{y}^\epsilon - \mathbf{p}_{m-1}\|_{L^2(T)}$$
$$\lesssim \|\mathbf{y}^* - \mathbf{y}^\epsilon\|_{H^2(T)},$$

provided $\mathbf{p}_{m-1}$ is an averaged Taylor polynomial of $\mathbf{y}^* - \mathbf{y}^\epsilon$. This in turn implies

$$\sum_{T \in \mathcal{T}_h} \|\mathbf{y}_h^\epsilon - \mathbf{y}_h^*\|_{H^2(T)}^2 \lesssim \sum_{i=1}^2 \|\mathbf{y}^* - \mathbf{y}^\epsilon\|_{H^2(\Omega_i)}^2 \lesssim \epsilon^2.$$

Therefore, gathering the estimates for the three terms in (69) we obtain

$$\sum_{T \in \mathcal{T}_h} \|\mathbf{y}^* - \mathbf{y}_h^*\|_{H^2(T)}^2 \lesssim \epsilon^2 + h^2 \sum_{i=1}^2 \|\mathbf{y}^\epsilon\|_{H^3(\Omega_i)}^2 \lesssim \epsilon^2,$$

provided $h\|\mathbf{y}^\epsilon\|_{H^3(\Omega_i)} \le \epsilon$ for h sufficiently small and $i = 1, 2$. Since ϵ is arbitrary, we deduce

$$\sum_{T \in \mathcal{T}_h} \|\mathbf{y}^* - \mathbf{y}_h^*\|_{H^2(T)}^2 \to 0,$$

and in particular $\mathbf{y}_h^* \to \mathbf{y}^*$ in $[L^2(\Omega)]^3$, as $h \to 0^+$.

Step (iv): Convergence of $\widetilde{E}_h[\mathbf{y}_h^*]$. Steps (iii) and (iv) show that the conditions in Lemma 11(ii) (strong convergence of $\widetilde{H}_h$) are fulfilled, whence $\widetilde{H}_h[\mathbf{y}_h^*] \to \widetilde{D}^2\mathbf{y}^*$ strongly in $[L^2(\Omega)]^{3\times2\times2}$. Consequently, convergence of $\widetilde{E}_h[\mathbf{y}_h^*]$ towards $\widetilde{E}[\mathbf{y}^*]$ reduces to the argument given in Theorem 9 (ii) and is not repeated here.   □

The next theorem guarantees convergence of discrete global minimizers towards exact global minimizers, but it is not a standard Γ-convergence result because we assume (54) for one global minimizer. Other minimizers may fail to satisfy (54).

Theorem 13 (convergence of global discrete minimizers with creases). *Assume that a global minimizer $\mathbf{y}^* \in \widetilde{\mathbb{A}}$ of $\widetilde{E}$ satisfies the additional regularity (54). Let $\delta = \delta(h) \geq c\omega(h)$ with c the constant in Lemma 12 and ω the modulus of smoothness in (55). If $\mathbf{y}_h \in \mathbb{A}_{h,\delta}$ is an almost global minimizer of $\widetilde{E}_h$ in the sense that*

$$(70) \qquad \widetilde{E}_h[\mathbf{y}_h] \leq \inf_{\mathbf{w}_h \in \mathbb{A}_{h,\delta}} \widetilde{E}_h[\mathbf{w}_h] + \sigma$$

where $\sigma, \delta \to 0$ as $h \to 0^+$, then $\{\mathbf{y}_h\}_{h>0}$ is precompact in $[L^2(\Omega)]^3$ and every cluster point $\mathbf{y}$ belongs to $\widetilde{\mathbb{A}}$ and is a global minimizer of $\widetilde{E}$, namely $\widetilde{E}[\mathbf{y}] = \inf_{\mathbf{w}\in\widetilde{\mathbb{A}}} \widetilde{E}[\mathbf{w}]$. Moreover, up to a subsequence (not relabeled)

$$(71) \qquad \widetilde{E}[\mathbf{y}] = \lim_{h\to0^+} \widetilde{E}_h[\mathbf{y}_h].$$

Proof. The lim inf property follows along the lines of Theorem 9 (i) because it is based on Lemmas 8 and 7, which remain valid in this context, as well as Lemma 11 (i) (weak convergence of $\widetilde{H}_h$) instead of Lemma 1 (weak convergence of H_h) and the weak lower semicontinuity of the L^2-norm. Therefore, there is $\mathbf{y} \in \widetilde{\mathbb{A}}$ such that (up to a subsequence not relabelled) $\mathbf{y}_h \to \mathbf{y}$ in $[L^2(\Omega)]^3$ and

$$\widetilde{E}[\mathbf{y}] \leq \liminf_{h\to0^+} \widetilde{E}_h[\mathbf{y}_h].$$

To show that $\mathbf{y}$ is a global minimizer of $\widetilde{E}$ we resort to the extra regularity (54) of the global minimizer $\mathbf{y}^* \in \widetilde{\mathbb{A}}$ of $\widetilde{E}$. Let $\{\mathbf{y}_h^*\}_{h>0} \subset [L^2(\Omega)]^3$ be the sequence provided by Lemma 12 (lim sup property with creases), which satisfies

$$\mathbf{y}_h^* \in \mathbb{A}_{h,c\omega(h)}, \quad \widetilde{E}_h[\mathbf{y}_h^*] \to \widetilde{E}[\mathbf{y}^*]$$

as $h \to 0^+$. In view of (70) and $\mathbf{y}_h \in \mathbb{A}_{h,\delta(h)}$, we end up with

$$\widetilde{E}[\mathbf{y}] \leq \liminf_{h\to0^+} \widetilde{E}_h[\mathbf{y}_h] \leq \limsup_{h\to0^+} \left(\widetilde{E}_h[\mathbf{y}_h^*] + \sigma\right) = \widetilde{E}[\mathbf{y}^*] = \inf_{\mathbf{w}\in\widetilde{A}} \widetilde{E}[\mathbf{w}].$$

Therefore, $\mathbf{y}$ is indeed a global minimizer of $\widetilde{E}$ and (71) is valid.   □

5. Discrete gradient flow

Solving the minimization problem (42) is a nontrivial task because it entails enforcing the *nonconvex constraint* $D_h[\mathbf{y}_h](x_T) \leq \delta$ at element barycenters x_T. We

now develop a discrete gradient flow with respect to the H_h^2 metric (23) that linearizes the isometry constraint according to (15). We refer to [4, 7, 8, 14, 15, 18, 17] and especially to S. Bartels and Ch. Palus [10] for similar gradient flows.

We start recalling the notion of linearized isometry constraint for $\mathbf{v}_h, \mathbf{y}_h \in [\mathbb{V}_h^k]^3$

$$(72) \qquad L[\mathbf{v}_h; \mathbf{y}_h](x_T) = \left[\nabla \mathbf{v}_h^T \nabla \mathbf{y}_h + \nabla \mathbf{y}_h^T \nabla \mathbf{v}_h \right](x_T) \quad \forall\, T \in \mathcal{T}_h$$

and defining a tangent space associated with the isometry constraint for any $\mathbf{y}_h \in \mathbb{A}_{h,\delta}$

$$(73) \qquad \mathcal{F}_h(\mathbf{y}_h) := \left\{ \mathbf{v}_h \in \mathbb{V}_h^k(\mathbf{0}, \mathbf{0}) : \quad L[\mathbf{v}_h; \mathbf{y}_h](x_T) = 0 \quad \forall\, T \in \mathcal{T}_h \right\}.$$

Given $\mathbf{y}_h^0 \in \mathbb{A}_{h,0}$ (i.e, $\mathbf{y}_h^0$ satisfies the isometry constraint $\mathrm{I}[\mathbf{y}_h^0](x_T) = I_2$ at each barycenter x_T), the discrete gradient flow consists of seeking recursively $\delta \mathbf{y}_h^{n+1} := \mathbf{y}_n^{n+1} - \mathbf{y}_h^n \in \mathcal{F}_h(\mathbf{y}_h^n)$ such that

$$(74) \quad \frac{1}{\tau} \left\langle \delta \mathbf{y}_h^{n+1}, \mathbf{v}_h \right\rangle_{H_h^2(\Omega)} + a_h(\delta \mathbf{y}_h^{n+1}, \mathbf{v}_h) = -a_h(\mathbf{y}_h^n, \mathbf{v}_h) + \ell[\mathbf{y}_h^n](\mathbf{v}_h) \quad \forall\, \mathbf{v}_h \in \mathcal{F}_h(\mathbf{y}_h^n).$$

Here $\tau > 0$ is a pseudo time step and a_h is the bilinear form corresponding to the variational derivative of the bending energy $B_h[\mathbf{y}_h]$ defined in (39), i.e.

$$(75) \qquad \begin{aligned} a_h(\mathbf{y}_h, \mathbf{v}_h) &:= \int_\Omega H_h[\mathbf{y}_h] : H_h[\mathbf{v}_h] \\ &\quad + \gamma_1(\mathrm{h}^{-1}[\nabla_h \mathbf{y}_h], [\nabla_h \mathbf{v}_h])_{L^2(\Gamma_h^a)} + \gamma_0(\mathrm{h}^{-3}[\mathbf{y}_h], [\mathbf{v}_h])_{L^2(\Gamma_h^a)}. \end{aligned}$$

The linear form $\ell[\mathbf{y}_h^n](\mathbf{v}_h)$ on $\mathbf{v}_h$ is the first variation of the cubic energy $C_h[\mathbf{y}_h^n]$, defined in (41), along the direction of the test function $\mathbf{v}_h$ and is given by

$$\begin{aligned} \ell[\mathbf{y}_h^n](\mathbf{v}_h) &:= \sum_{i,j=1}^2 \sum_{T \in \mathcal{T}_h} |T| \, \overline{H}_h[\mathbf{v}_h]_{ij} \cdot (\partial_1 \mathbf{y}_h^n \times \partial_2 \mathbf{y}_h^n)(x_T)\, Z_{ij} \\ &\quad + \sum_{i,j=1}^2 \sum_{T \in \mathcal{T}_h} |T| \, \overline{H}_h[\mathbf{y}_h^n]_{ij} \cdot (\partial_1 \mathbf{v}_h \times \partial_2 \mathbf{y}_h^n)(x_T)\, Z_{ij} \\ &\quad + \sum_{i,j=1}^2 \sum_{T \in \mathcal{T}_h} |T| \, \overline{H}_h[\mathbf{y}_h^n]_{ij} \cdot (\partial_1 \mathbf{y}_h^n \times \partial_2 \mathbf{v}_h)(x_T)\, Z_{ij}; \end{aligned}$$

recall that both $\overline{H}_h[\mathbf{v}_h]$ and Z are piecewise constant on $\mathcal{T}_h$. The explicit treatment of $\mathbf{y}_h^n$ in $\ell[\mathbf{y}_h^n](\mathbf{v}_h)$ is similar to the scheme proposed and analyzed by S. Bartels and Ch. Palus [10]. We note that (74) is the discrete Euler-Lagrange equation for the augmented energy $E_h[\mathbf{y}_h^{n+1}] + \frac{1}{2\tau} \|\mathbf{y}_h^{n+1} - \mathbf{y}_h^n\|_{H_h^2(\Omega)}^2$, except that the nonlinear terms corresponding to the cubic energy $C_h[\mathbf{y}_h^{n+1}]$ are linearized by $\ell[\mathbf{y}_h^n](\mathbf{v}_h)$. We refer to the nonlinear continuous Euler-Lagrange equation (9) for a comparison.

For later use, we note that Lemma 3 (stability of $\overline{H}_h[\mathbf{v}_h]$) yields

$$|\ell[\mathbf{y}_h](\mathbf{v}_h)| \lesssim \sqrt{1+\delta} \|Z\|_{L^\infty(\Omega)} \left(\|\nabla_h \mathbf{v}_h\|_{L^2(\Omega)} \|\mathbf{y}_h\|_{H_h^2(\Omega)} + \|\nabla_h \mathbf{y}_h\|_{L^2(\Omega)} \|\mathbf{v}_h\|_{H_h^2(\Omega)} \right),$$

provided $\mathbf{y}_h \in \mathbb{A}_{h,\delta}$ because $|\nabla \mathbf{y}_h(x_T)| \le \sqrt{2(1+\delta)}$ from Lemma 6 (pointwise isometry constraint) and the inverse inequality $|\nabla \mathbf{v}_h(x_T)| \lesssim h_T^{-1} \|\nabla \mathbf{v}_h\|_{L^2(T)}$ applies. In

addition, we rewrite the Friedrichs inequality (25) as follows

$$\|\nabla_h \mathbf{y}_h\|_{L^2(\Omega)} \lesssim \|\mathbf{y}_h\|_{H_h^2(\Omega)} + C_{\boldsymbol{\varphi},\Phi}, \qquad \forall \mathbf{w}_h \in \mathbb{V}_h^k(\boldsymbol{\varphi}, \Phi), \tag{76}$$

where $C_{\boldsymbol{\varphi},\Phi} = \|\boldsymbol{\varphi}\|_{H^1(\Omega)} + \|\Phi\|_{H^1(\Omega)}$. From these estimates we deduce the existence of a constant c_{nl} such that for $\mathbf{y}_h \in \mathbb{A}_{h,\delta}$ and $\mathbf{v}_h \in \mathbb{V}_h^k(\mathbf{0},\mathbf{0})$ we have

$$\left|\ell[\mathbf{y}_h](\mathbf{v}_h)\right| \le c_{nl}\sqrt{1+\delta}\|Z\|_{L^\infty(\Omega)}\big(\|\mathbf{y}_h\|_{H_h^2(\Omega)} + C_{\boldsymbol{\varphi},\Phi}\big)\|\mathbf{v}_h\|_{H_h^2(\Omega)}. \tag{77}$$

5.1. **Energy stability and admissibility.** We discuss in this section the energy reduction property of the gradient flow and, although the isometry constraint $D_h[\mathbf{y}_h^n](x_T) = 0$ is relaxed and linearized in the iterative scheme, the deviation of $D_h[\mathbf{y}_h^n](x_T)$ from 0 is controlled by a parameter $\delta > 0$ provided τ is sufficiently small. These results rely on a discrete inverse inequality on finite dimensional subsets.

Lemma 14 (discrete Sobolev inequality). *Let $W_h \subset \Pi_{T \in \mathcal{T}_h} H^1(T)$ be a finite element space subordinated to the partition $\mathcal{T}_h$. For any $w_h \in W_h$ there holds*

$$\|w_h\|_{L^\infty(\Omega)} \lesssim \big(1 + |\log h_{\min}|\big)^{\frac{1}{2}}\Big(\|w_h\|_{L^2(\Omega)} + \|\nabla_h w_h\|_{L^2(\Omega)} + \|\mathrm{h}^{-\frac{1}{2}}[w_h]\|_{L^2(\Gamma_h^0)}\Big), \tag{78}$$

where $h_{\min} := \min_{T \in \mathcal{T}_h} h_T$.

Proof. We denote by $\Pi_h : \Pi_{T \in \mathcal{T}_h} H^1(T) \to \mathbb{V}_h^k \cap H^1(\Omega)$ the smoothing operator from [16, 18] and recall that it satisfies

$$\|\nabla \Pi_h w_h\|_{L^2(\Omega)} + \|\mathrm{h}^{-1}(w_h - \Pi_h w_h)\|_{L^2(\Omega)} \lesssim \|\nabla_h w_h\|_{L^2(\Omega)} + \|\mathrm{h}^{-\frac{1}{2}}[w_h]\|_{L^2(\Gamma_h^0)},$$

and

$$\|\Pi_h w_h\|_{L^2(\Omega)} \lesssim \|w_h\|_{L^2(\Omega)}.$$

Therefore, combining the triangle and inverse inequalities for $w_h \in W_h$ implies

$$\|w_h\|_{L^\infty(\Omega)} \lesssim \|w_h - \Pi_h w_h\|_{L^\infty(\Omega)} + \|\Pi_h w_h\|_{L^\infty(\Omega)}$$

$$\lesssim \|\mathrm{h}^{-1}(w_h - \Pi_h w_h)\|_{L^2(\Omega)} + \big(1 + |\log h_{\min}|\big)^{\frac{1}{2}}\|\Pi_h w_h\|_{H^1(\Omega)},$$

in view of the following discrete Sobolev inequality in 2d [19, 20]

$$\|\Pi_h w_h\|_{L^\infty(\Omega)} \lesssim \big(1 + |\log h_{\min}|\big)^{\frac{1}{2}}\|\Pi_h w_h\|_{H^1(\Omega)}.$$

This leads to the assertion upon applying the preceding estimates for Π_h. $\square$

We are now in a position to prove the main result of this section, namely that the gradient flow is energy decreasing and controls the isometry defect.

Theorem 15 (properties of gradient flow). *Let $\{\mathbf{y}_h^0\}_{h>0} \subset \mathbb{A}_{h,0}$ satisfy $E_h[\mathbf{y}_h^0] \le c_0$ with c_0 a constant independent of h and let all subdivisions $\mathcal{T}_h$ be such that $|\log h_{\min}| \ge 1$. Let N be the number of iterations of the gradient flow and τ be its pseudo-time step. There exists a constant $\alpha_1 = \alpha_1(\boldsymbol{\varphi}, \Phi, Z) > 0$ independent of h and N such that if $\tau \le (2\alpha_1|\log h_{\min}|)^{-1}$, then the energy $E_h[\mathbf{y}_h^N]$ satisfies*

$$E_h[\mathbf{y}_h^N] + \frac{1}{2\tau}\sum_{n=0}^{N-1}\|\delta\mathbf{y}_h^{n+1}\|_{H_h^2(\Omega)}^2 \le E_h[\mathbf{y}_h^0]. \tag{79}$$

In addition, there are constants $\alpha_2 = \alpha_2(\boldsymbol{\varphi}, \Phi, Z) > 0$ and $\alpha_3 > 0$, both independent of h and N, such that the isometry defect $D_h[\mathbf{y}_h^N]$ satisfies

$$(80) \qquad D_h[\mathbf{y}_h^N](x_T) \le \alpha_3 \tau |\log h_{\min}| \big(E_h[\mathbf{y}_h^0] + \alpha_2\big) \quad \forall\, T \in \mathcal{T}_h.$$

Proof. We proceed by induction. We first note that estimates (79) and (80) hold trivially for $N = 0$ and $\mathbf{y}_h^0 \in \mathbb{A}_{h,0}$. Therefore, we assume that (79) and (80) are valid for $N = 1, ..., M$ with positive constants $\alpha_1, \alpha_2, \alpha_3$ to be specified below and prove the validity of the same estimates for $N = M + 1$ with the same constants $\alpha_1, \alpha_2, \alpha_3$.

We split the proof into four steps with the following roadmap. After deriving an intermediate estimate in Step (i), we prove (79) in Step (ii) and (80) in Step (iii) under suitable restrictions on α_i, $i = 1, 2, 3$. In Step (iv), we show that it is always possible to find values of these parameters satisfying the desired restrictions. In this proof, the generic constants hidden in the symbol "$\lesssim$" are not only independent of h but also of τ, M and α_i, $i = 1, 2, 3$.

Step (i): intermediate estimate. We take $\mathbf{v}_h = \delta\mathbf{y}_h^{M+1} \in \mathcal{F}_h(\mathbf{y}_h)$ in (74) for $n = M$, use the elementary relation $2b(b-a) = (b-a)^2 + b^2 - a^2$ and discard the positive term $(b-a)^2 = a_h(\delta\mathbf{y}_h^{M+1}, \delta\mathbf{y}_h^{M+1})$ to write

$$(81) \qquad \|\delta\mathbf{y}_h^{M+1}\|_{H_h^2(\Omega)}^2 + \frac{\tau}{2} a_h\big(\mathbf{y}_h^{M+1}, \mathbf{y}_h^{M+1}\big) - \frac{\tau}{2} a_h\big(\mathbf{y}_h^M, \mathbf{y}_h^M\big) \le \tau \ell\big[\mathbf{y}_h^M\big](\delta\mathbf{y}_h^{M+1}).$$

Using (80) with $N = M$ (induction assumption), together with the restriction on τ and the uniform bound $E_h[\mathbf{y}_h^0] \le c_0$, we obtain

$$(82) \qquad \big|D_h[\mathbf{y}_h^M](x_T)\big| \le \frac{\alpha_3}{2\alpha_1}\big(c_0 + \alpha_2\big) = \delta.$$

To simplify the expressions below, we let $\alpha = (\alpha_i)_{i=1}^3$ and $c_\alpha^2 = 1 + \delta$, whence

$$(83) \qquad c_\alpha := \sqrt{\frac{\alpha_3}{2\alpha_1}\big(c_0 + \alpha_2\big) + 1}.$$

Estimate (82) shows that $\mathbf{y}_h^M \in \mathbb{A}_{h,\delta}$ with $\delta = c_\alpha^2 - 1$ which in turn implies

$$|\partial_i \mathbf{y}_h^M(x_T)| \le c_\alpha, \qquad i = 1, 2, \quad T \in \mathcal{T}_h,$$

according to Lemma 6 (pointwise isometry constraint). Substituting into (77) yields

$$\big|\ell[\mathbf{y}_h^M](\delta\mathbf{y}_h^{M+1})\big| \le c_{nl} c_\alpha \|Z\|_{L^\infty(\Omega)} \|\delta\mathbf{y}_h^{M+1}\|_{H_h^2(\Omega)}\big(\|\mathbf{y}_h^M\|_{H_h^2(\Omega)} + C_{\boldsymbol{\varphi},\Phi}\big).$$

Inserting this back into (81) and using Young's inequality to absorb the term $\|\delta\mathbf{y}_h^{k+1}\|_{H_h^2(\Omega)}^2$ in the left hand side of (81), gives the estimate

$$(84) \qquad \begin{aligned} \frac{1}{2}\|\delta\mathbf{y}_h^{M+1}\|_{H_h^2(\Omega)}^2 &+ \frac{\tau}{2} a_h\big(\mathbf{y}_h^{M+1}, \mathbf{y}_h^{M+1}\big) \\ &\le \frac{\tau}{2} a_h\big(\mathbf{y}_h^M, \mathbf{y}_h^M\big) + \tau^2 c_{nl}^2 c_\alpha^2 \|Z\|_{L^\infty(\Omega)}^2 \big(\|\mathbf{y}_h^M\|_{H_h^2(\Omega)}^2 + C_{\boldsymbol{\varphi},\Phi}^2\big). \end{aligned}$$

We next improve upon (84) by deriving a uniform bound for the right-hand side. According to (82), the isometry defect is controlled by $\delta = c_\alpha^2 - 1$. Moreover, (79)

for $N = M$ (induction assumption) implies that $E_h[\mathbf{y}_h^M] \leq E_h[\mathbf{y}_h^0] \leq c_0$. Hence, the coercivity estimate (48) reads

$$(85) \qquad\qquad (2c_{coer})^{-1}\|\mathbf{y}_h^M\|_{H_h^2(\Omega)}^2 \leq c_0 + \tilde{c}_{coer}c_\alpha^4.$$

Estimate (47) can be rewritten in terms of the bilinear form a_h as follows

$$(86) \qquad c_{coer}^{-1}\|\mathbf{v}_h\|_{H_h^2(\Omega)}^2 \leq \frac{1}{2}a_h(\mathbf{v}_h, \mathbf{v}_h) \leq c_{cont}\|\mathbf{v}_h\|_{H_h^2(\Omega)}^2 \qquad \forall \mathbf{v}_h \in \mathbb{V}_h^k(\boldsymbol{\varphi}, \Phi).$$

Since $|\log h_{\min}| \geq 1$, τ satisfies $\tau \leq \frac{1}{2\alpha_1} \leq 1$ provided $\alpha_1 \geq \frac{1}{2}$, which is our first restriction on α_1. Combining this with (85) and (86) with $\mathbf{v}_h = \mathbf{y}_h^M$, and replacing back into (84), gives the desired intermediate estimate

$$(87) \qquad\qquad \frac{1}{2\tau}\|\delta\mathbf{y}_h^{M+1}\|_{H_h^2(\Omega)}^2 + \frac{1}{2}a_h(\mathbf{y}_h^{M+1}, \mathbf{y}_h^{M+1}) \leq \psi_1(c_\alpha)$$

where $\psi_1(c_\alpha) \geq 0$ is a positive increasing function of its argument c_α but whose specific expression is irrelevant except that it is independent of h, M and depends on $\alpha = (\alpha_i)_{i=1}^3$ only through the variable c_α rather than separately on each α_i.

Step (ii): proof of (79) *for $N = M+1$.* In view of (81), and telescopic cancellation, this requires dealing with the cubic term $\ell[\mathbf{y}_h^M](\delta\mathbf{y}_h^{M+1})$. Using the identity

$$a^{M+1}b^{M+1}c^{M+1} - a^M b^M c^M$$
$$= (a^{M+1} - a^M)b^{M+1}c^{M+1} + a^M(b^{M+1} - b^M)c^{M+1} + a^M b^M(c^{M+1} - c^M),$$

we deduce

$$(a^{M+1} - a^M)b^M c^M + a^M(b^{M+1} - b^M)c^M + a^M b^M(c^{M+1} - c^M)$$
$$= a^{M+1}b^{M+1}c^{M+1} - a^M b^M c^M - (a^{M+1} - a^M)(b^{M+1} - b^M)c^{M+1}$$
$$- (a^{M+1} - a^M)b^M(c^{M+1} - c^M) - a^M(b^{M+1} - b^M)(c^{M+1} - c^M),$$

and rewrite $\ell[\mathbf{y}_h^M](\delta\mathbf{y}_h^{M+1})$ as follows:

$$\ell[\mathbf{y}_h^M](\delta\mathbf{y}_h^{M+1}) = \sum_{i,j=1}^2 \sum_{T\in\mathcal{T}_h} |T|\overline{H}_h[\mathbf{y}_h^{M+1}]_{ij} \cdot (\partial_1\mathbf{y}_h^{M+1} \times \partial_2\mathbf{y}_h^{M+1})(x_T)\, Z_{ij}$$
$$- \sum_{i,j=1}^2 \sum_{T\in\mathcal{T}_h} |T|\overline{H}_h[\mathbf{y}_h^M]_{ij} \cdot (\partial_1\mathbf{y}_h^M \times \partial_2\mathbf{y}_h^M)(x_T)\, Z_{ij}$$
$$- \sum_{i,j=1}^2 \sum_{T\in\mathcal{T}_h} |T|\overline{H}_h[\delta\mathbf{y}_h^{M+1}]_{ij} \cdot (\partial_1\delta\mathbf{y}_h^{M+1} \times \partial_2\mathbf{y}_h^{M+1})(x_T)\, Z_{ij}$$
$$- \sum_{i,j=1}^2 \sum_{T\in\mathcal{T}_h} |T|\overline{H}_h[\delta\mathbf{y}_h^{M+1}]_{ij} \cdot (\partial_1\mathbf{y}_h^M \times \partial_2\delta\mathbf{y}_h^{M+1})(x_T)\, Z_{ij}$$
$$- \sum_{i,j=1}^2 \sum_{T\in\mathcal{T}_h} |T|\overline{H}_h[\mathbf{y}_h^M]_{ij} \cdot (\partial_1\delta\mathbf{y}_h^{M+1} \times \partial_2\delta\mathbf{y}_h^{M+1})(x_T)\, Z_{ij}.$$

We note that the first two terms are exactly the cubic energies $C_h[\mathbf{y}_h^{M+1}]$ and $C_h[\mathbf{y}_h^M]$ and together with the bending energies $B_h[\mathbf{y}_h^{M+1}] = \frac{1}{2}a_h(\mathbf{y}_h^{M+1}, \mathbf{y}_h^{M+1})$ and

$B_h[\mathbf{y}_h^M] = \frac{1}{2}a_h(\mathbf{y}_h^M, \mathbf{y}_h^M)$ in (81) give rise to the full energies $E_h[\mathbf{y}_h^{M+1}]$ and $E_h[\mathbf{y}_h^M]$ in (79). In contrast, the last three terms must be estimated and absorbed into the remaining term $\|\delta\mathbf{y}_h^{M+1}\|^2_{H_h^2(\Omega)}$ in (79). To this end, we combine the Friedrichs inequality (76) for $\mathbf{y}_h = \mathbf{y}_h^{M+1}, \mathbf{y}_h^M \in \mathbb{V}_h^k(\boldsymbol{\varphi}, \Phi)$ and $\mathbf{y}_h = \delta\mathbf{y}_h^{M+1} \in \mathbb{V}_h^k(\mathbf{0}, \mathbf{0})$, and Lemma 3 (stability of $\overline{H}_h[\mathbf{v}_h]$), to obtain

$$\|\delta\mathbf{y}_h^{M+1}\|^2_{H_h^2(\Omega)} + \tau E_h[\mathbf{y}_h^{M+1}] - \tau E_h[\mathbf{y}_h^M]$$
$$\lesssim \tau\|\delta\mathbf{y}_h^{M+1}\|_{H_h^2(\Omega)}\|\nabla\delta\mathbf{y}_h^{M+1}\|_{L^\infty(\Omega)}\Big(\|\mathbf{y}_h^{M+1}\|_{H_h^2(\Omega)} + \|\mathbf{y}_h^M\|_{H_h^2(\Omega)} + C_{\boldsymbol{\varphi},\Phi}\Big),$$

where the symbol $\lesssim$ hides $\|Z\|_{L^\infty(\Omega)}$. To estimate the L^∞-norm on the right-hand side, we resort to Lemma 14 (discrete Sobolev inequality)

$$\|\delta\mathbf{y}_h^{M+1}\|^2_{H_h^2(\Omega)} + \tau E_h[\mathbf{y}_h^{M+1}] - \tau E_h[\mathbf{y}_h^M]$$
$$\leq c_Z\tau|\log h_{\min}|\|\delta\mathbf{y}_h^{M+1}\|^2_{H_h^2(\Omega)}\Big(\|\mathbf{y}_h^{M+1}\|_{H_h^2(\Omega)} + \|\mathbf{y}_h^M\|_{H_h^2(\Omega)} + C_{\boldsymbol{\varphi},\Phi}\Big),$$

because $|\log h_{\min}| \geq 1$, with constant c_Z depending on $\|Z\|_{L^\infty(\Omega)}$ and the constants hidden in (76) and (78). Moreover, the coercivity estimate (47) of B_h, written now as

$$\|\mathbf{y}_h^{M+1}\|^2_{H_h^2(\Omega)} \leq c_{coer}B_h[\mathbf{y}_h^{M+1}] = \frac{c_{coer}}{2}a_h(\mathbf{y}_h^{M+1}, \mathbf{y}_h^{M+1}) \leq c_{coer}\psi_1(c_\alpha)$$

according to (87), together with (85) guarantees that

$$\|\mathbf{y}_h^{M+1}\|_{H_h^2(\Omega)} + \|\mathbf{y}_h^M\|_{H_h^2(\Omega)} + C_{\boldsymbol{\varphi},\Phi} \leq \psi_2(c_\alpha),$$

where $\psi_2(c_\alpha)$ is a positive increasing function of the argument c_α which is independent of h, M, and the individual parameters $(\alpha_i)_{i=1}^3$. Substituting back yields

$$\|\delta\mathbf{y}_h^{M+1}\|^2_{H_h^2(\Omega)} + \tau E_h[\mathbf{y}_h^{M+1}] - \tau E_h[\mathbf{y}_h^M] \leq c_Z\tau|\log h_{\min}|\psi_2(c_\alpha)\|\delta\mathbf{y}_h^{M+1}\|^2_{H_h^2(\Omega)}.$$

Consequently, since $\tau \leq (2\alpha_1|\log h_{\min}|)^{-1}$, it remains to choose $(\alpha_i)_{i=1}^3$ so that

$$\psi_2(c_\alpha) \leq \alpha_1 c_Z^{-1},$$

to derive the desired estimate (79) for $N = M + 1$. The validity of this estimate will be justified in Step (iv).

Step (iii): proof of (80) *for* $N = M + 1$. Since $\delta\mathbf{y}_h^{n+1} \in \mathcal{F}_h(\mathbf{y}_h^n)$, expanding $\mathrm{I}[\mathbf{y}_h^{n+1}](x_T) = [(\nabla\mathbf{y}_h^{n+1})^T\nabla\mathbf{y}_h^{n+1}](x_T)$ and using the definition (73) of $\mathcal{F}_h(\mathbf{y}_h^n)$ yields

(88) $D_h[\mathbf{y}_h^{n+1}](x_T) \leq D_h[\mathbf{y}_h^n](x_T) + \big|\mathrm{I}[\delta\mathbf{y}_h^{n+1}](x_T)\big| \quad \forall T \in \mathcal{T}_h.$

Applying Lemma 14 (discrete Sobolev inequality), followed by the discrete Friedrichs inequality (76) to estimate $\|\nabla_h\delta\mathbf{y}_h^{n+1}\|_{L^2(\Omega)}$, implies

(89) $\big|\mathrm{I}[\delta\mathbf{y}_h^{n+1}](x_T)\big| \leq \|\nabla_h\delta\mathbf{y}_h^{n+1}\|^2_{L^\infty(T)} \leq c_F|\log h_{\min}|\|\delta\mathbf{y}_h^{n+1}\|^2_{H_h^2(\Omega)}$

where $c_F > 0$ is a constant that combines constants hidden in (78) and (76), because $|\log(h_{\min})| \geq 1$ and $\delta\mathbf{y}_h^{n+1} \in \mathbb{V}_h^k(\mathbf{0}, \mathbf{0})$. Summing over $0 \leq n \leq M$, and using telescopic

cancellation along with $\mathbf{y}_h^0 \in \mathbb{A}_{h,0}$, yields

$$(90) \qquad D_h[\mathbf{y}_h^{M+1}](x_T) \le c_F |\log h_{\min}| \sum_{n=0}^{M} \|\delta \mathbf{y}_h^{n+1}\|_{H_h^2(\Omega)}^2.$$

Exploiting the energy decay (79), proved for $N = M + 1$ in Step (ii), gives

$$(91) \qquad \sum_{n=0}^{M} \|\delta \mathbf{y}_h^{n+1}\|_{H_h^2(\Omega)}^2 \le 2\tau \big(E_h[\mathbf{y}_h^0] - E_h[\mathbf{y}_h^{M+1}] \big).$$

We now need a lower bound for the energy $E_h[\mathbf{y}_h^{M+1}]$, which is a consequence of (48) provided $\mathbf{y}_h^{M+1} \in \mathbb{A}_{h,\epsilon}$ for some $\epsilon > 0$. To this end, we resort again to (88). We first bound the second term on the right-hand side upon combining the intermediate estimate (87) for $\|\delta \mathbf{y}_h^{M+1}\|_{H_h^2(\Omega)}$ with (89)

$$\big| \mathbb{I}[\delta \mathbf{y}_h^{M+1}](x_T) \big| \le 2\tau c_F |\log h_{\min}| \psi_1(c_\alpha) \le c_F \alpha_1^{-1} \psi_1(c_\alpha),$$

because $\tau \le (2\alpha_1 |\log h_{\min}|)^{-1}$. Using this bound in (88), along with the fact that $\mathbf{y}_h^M \in \mathbb{A}_{h,\delta}$ for $\delta = c_\alpha^2 - 1$ according to the induction assumption (82), implies

$$\big| D_h[\mathbf{y}_h^{M+1}] \big| (x_T) \le c_\alpha^2 - 1 + c_F \alpha_1^{-1} \psi_1(c_\alpha) =: \epsilon,$$

whence $\mathbf{y}_h^{M+1} \in \mathbb{A}_{h,\epsilon}$. Inserting $E_h[\mathbf{y}_h^{M+1}] \ge -\tilde{c}_{coer}(1 + \epsilon)^2$ from (48) into (91) gives

$$\sum_{n=0}^{M} \|\delta \mathbf{y}_h^{n+1}\|_{H_h^2(\Omega)}^2 \le 2\tau \Big(E_h[\mathbf{y}_h^0] + \tilde{c}_{coer}\big(c_\alpha^2 + c_F \alpha_1^{-1} \psi_1(c_\alpha) \big)^2 \Big).$$

Returning to (90), we arrive at

$$D_h[\mathbf{y}_h^{M+1}](x_T) \le 2c_F \tau |\log h_{\min}| \Big(E_h[\mathbf{y}_h^0] + \tilde{c}_{coer}\big(c_\alpha^2 + c_F \alpha_1^{-1} \psi_1(c_\alpha) \big)^2 \Big),$$

and we emphasize that c_F is a constant independent of h, M and $(\alpha_i)_{i=1}^3$. The desired control on the isometry defect (80) is thus guaranteed provided

$$\alpha_3 \ge 2c_F, \qquad \alpha_2 \ge \tilde{c}_{coer}\big(c_\alpha^2 + c_F \alpha_1^{-1} \psi_1(c_\alpha) \big)^2.$$

Step (iv): choice of parameters. $\alpha = (\alpha_i)_{i=1}^3$ must satisfy

$$\alpha_1 \ge \frac{1}{2}, \quad \psi_2(c_\alpha) \le \alpha_1 c_Z^{-1}, \quad \alpha_3 \ge 2c_F, \quad \alpha_2 \ge \tilde{c}_{coer}\big(c_\alpha^2 + c_F \alpha_1^{-1} \psi_1(c_\alpha) \big)^2 =: \psi_3(\alpha_1, c_\alpha),$$

where c_α is defined in (83) and the functions ψ_1, ψ_2 are positive and increasing in their arguments. One admissible set of parameters is

$$\alpha_3 = 2c_F, \quad \alpha_2 = \alpha_1^{\frac{1}{2}},$$

with $\alpha_1 \ge \frac{1}{2}$ sufficiently large. In fact, we note that as $\alpha_1 \to \infty$

$$c_\alpha \downarrow 1, \quad \psi_1(c_\alpha) \downarrow \psi_1(1) \ge 0, \quad \psi_2(c_\alpha) \downarrow \psi_2(1) \ge 0, \quad \psi_3(\alpha_1, c_\alpha) \downarrow \tilde{c}_{coer},$$

and the condition $\alpha_1 \ge \max\big\{ \frac{1}{2}, c_Z \psi_2(c_\alpha), \psi_3(\alpha_1, c_\alpha)^2 \big\}$ admits a solution. This completes the induction argument. $\qquad \square$

It is worth realizing that the ℓ^∞-control of the isometry defect (80) implies that $\mathbf{y}_h \in \mathbb{A}_{h,\delta}$ provided τ is so small that

$$\alpha_3 \tau |\log h_{\min}| \left(E_h[\mathbf{y}_h^0] + \alpha_2 \right) \le \delta,$$

where $\mathbb{A}_{h,\delta}$ is defined in (36). This property is novel in the context of DG approximations [14, 15, 17, 18, 39], but is inspired by a similar one at element vertices shown by S. Bartels and Ch. Palus for Kirchhoff elements [10]. It is responsible for the explicit treatment of the cubic term in (74), which in turn converts (74) into a *linear system* to solve for $\delta\mathbf{y}_h^{n+1}$. The fact that $H^2(\Omega)$ does not embed in $W_\infty^1(\Omega)$ in two dimensions, but is borderline instead, explains the critical nature of the estimates (79) and (80). The discrete H^2-metric of the gradient flow (74), combined with Lemma 14 (discrete Sobolev inequality), makes it possible to exploit this borderline structure discretely at the expense of a log term. No weaker metric for the gradient flow than H^2 would allow for ℓ^∞-control of the isometry defect.

5.2. **Lagrange multipliers for the isometry constraint.** We enforce tangential variations $\delta\mathbf{y}_h^{n+1} \in \mathcal{F}_h(\mathbf{y}_h^n)$ at each step of the gradient flow using Lagrange multipliers within the space of symmetric piecewise constant tensors

$$\Lambda_h := \left\{ \lambda_h : \Omega \to \mathbb{R}^{2\times 2} : \ \lambda_h^T = \lambda_h, \ \lambda_h \in \left[\mathbb{V}_h^0 \right]^{2\times 2} \right\}.$$

For any $\mathbf{w}_h \in \mathbb{V}_h^k(\boldsymbol{\varphi}, \Phi)$, we define the bilinear form $b_h(\mathbf{w}_h; \cdot, \cdot)$ on $\mathbb{V}_h^k(\mathbf{0}, \mathbf{0}) \times \Lambda_h$ as

$$(92) \qquad b_h(\mathbf{w}_h; \mathbf{v}_h, \boldsymbol{\mu}_h) := \sum_{T \in \mathcal{T}_h} |T| \, L[\mathbf{v}_h; \mathbf{w}_h](x_T) : \boldsymbol{\mu}_h,$$

where the linearized isometry constraint $L[\mathbf{v}_h; \mathbf{w}_h]$ is given in (72). Note that b_h is continuous with a continuity constant uniform in h

$$(93) \qquad |b_h(\mathbf{w}_h; \mathbf{v}_h, \boldsymbol{\mu}_h)| \lesssim \|\mathbf{w}_h\|_{H_h^2(\Omega)} \|\mathbf{v}_h\|_{H_h^2(\Omega)} \|\mu_h\|_{L^2(\Omega)},$$

thanks to the inverse inequality $|L[\mathbf{v}_h; \mathbf{w}_h](x_T)| \lesssim h_T^{-1} \|L[\mathbf{v}_h; \mathbf{w}_h]\|_{L^2(T)}$ and the discrete Sobolev inequality $\|\nabla_h \mathbf{w}_h\|_{L^4(\Omega)} \lesssim \|\mathbf{w}_h\|_{H_h^2(\Omega)}$ valid for all $\mathbf{w}_h \in [\mathbb{V}_h^k]^3$, see [15, (6.9)]. We also observe that $b_h(\mathbf{w}_h; \mathbf{v}_h, \boldsymbol{\mu}_h) = 0$ for all $\boldsymbol{\mu}_h \in \Lambda_h$ implies $\mathbf{v}_h \in \mathcal{F}_h(\mathbf{w}_h)$ according to (73). Therefore, in each step of the gradient flow augmented with the linearized metric constraint, we seek $(\delta\mathbf{y}_h^{n+1}, \boldsymbol{\lambda}_h^{n+1}) \in \mathbb{V}_h^k(\mathbf{0}, \mathbf{0}) \times \Lambda_h$ such that

$$(94) \qquad \begin{aligned} \tau^{-1}(\delta\mathbf{y}_h^{n+1}, \mathbf{v}_h)_{H_h^2(\Omega)} &+ a_h(\delta\mathbf{y}_h^{n+1}, \mathbf{v}_h) + b_h(\mathbf{y}_h^n; \mathbf{v}_h, \boldsymbol{\lambda}_h^{n+1}) \\ &+ b_h(\mathbf{y}_h^n; \delta\mathbf{y}_h^{n+1}, \boldsymbol{\mu}_h) = \ell[\mathbf{y}_h^n](\mathbf{v}_h) - a_h(\mathbf{y}_h^n, \mathbf{v}_h), \end{aligned}$$

for all $(\mathbf{v}_h, \boldsymbol{\mu}_h) \in \mathbb{V}_h^k(\mathbf{0}, \mathbf{0}) \times \Lambda_h$.

The proposed strategy is summarized in Algorithm 1.

It is worth pointing out that utilizing Lagrange multipliers is ubiquitous to enforce linearized metric constraints [4, 7, 8, 10, 14, 15, 18, 17]. In particular, the system (94) is solved using the *Schur complement approach*, whose performance depends on the inf-sup stability of b_h; see e.g. [39, 12] and refer to Section 6.1 for additional details on the practical implementation. Unfortunately, there are no results available in the literature guaranteeing a uniform inf-sup not even for the continuous problem.

Algorithm 1: (discrete-H^2 gradient flow with Lagrange multipliers)

Given a pseudo-time step $\tau > 0$ and a target tolerance *tol*;
Choose an initial guess $\mathbf{y}_h^0 \in \mathbb{A}_{h,0}$;
while $\tau^{-1}\big|E_h[\mathbf{y}_h^{n+1}] - E_h[\mathbf{y}_h^n]\big| > tol$ **do**
  Solve (94) for $(\delta\mathbf{y}_h^{n+1}, \boldsymbol{\lambda}_h^{n+1}) \in \mathbb{V}_h^k(\mathbf{0},\mathbf{0}) \times \Lambda_h$;
  Update $\mathbf{y}_h^{n+1} = \mathbf{y}_h^n + \delta\mathbf{y}_h^{n+1}$;
end

In this section, we make a first step towards a better understanding of the situation in that we derive a sub-optimal estimate of the discrete inf-sup constant. We start with a linear algebra lemma.

Lemma 16 (solvability of a matrix equation). *Given a 2×2 symmetric matrix C and a full-rank 3×2 matrix B, there exists a 3×2 matrix A that solves the equation*

$$(95) \qquad (A^T B + B^T A) : C = |C|^2$$

and satisfies $|A| \le \frac{|C|}{2\sigma_2(B)}$, where $|\cdot|$ denotes the Frobenius norm of matrices and $\sigma_2(B) > 0$ is the smallest singular value of B.

Proof. Using the cyclic properties of the trace operator yields

$$A^T B : C = \operatorname{tr}(B^T A C) = \operatorname{tr}(C A^T B) = \operatorname{tr}(A^T B C) = B^T A : C = A : BC,$$

whence (95) is equivalent to

$$A : BC = \frac{1}{2}|C|^2.$$

Let $B = U\Sigma V^T$ be the singular value decomposition of B, where $U \in \mathbb{R}^{3\times3}$ and $V = \mathbb{R}^{2\times2}$ are orthogonal matrices, and $\Sigma = [\sigma_1(B),0;0,\sigma_2(B);0,0] \in \mathbb{R}^{3\times2}$ carries the singular values $\sigma_1(B) \ge \sigma_2(B) \ge 0$ of B. Since B is full-rank, we deduce that $\sigma_2(B) > 0$ and

$$|BC|^2 = |U\Sigma V^T C|^2 = |\Sigma\overline{C}|^2 \ge \sigma_2(B)^2|\overline{C}|^2 = \sigma_2(B)^2|C|^2,$$

where $\overline{C} = V^T C$ and thus $|\overline{C}| = |C|$. We can now assume that $C \ne 0$, for otherwise $A = 0$ solves (95). We then realize that $|BC| > 0$ and

$$A = \frac{(BC)|C|^2}{2|BC|^2}$$

is clearly a solution to (95) as well as

$$|A| = \frac{|C|^2}{2|BC|} \le \frac{|C|}{2\sigma_2(B)},$$

which is the desired estimate $|A| \lesssim |C|$. $\qquad\square$

The following sub-optimal estimate of the discrete inf-sup constant is a consequence of the previous lemma. Since only the gradient of $\mathbf{v}_h \in \mathbb{V}_h^k(\mathbf{0},\mathbf{0})$ appears in (92), but the underlying norm of $\mathbb{V}_h^k(\mathbf{0},\mathbf{0})$ is the discrete H^2-norm, it seems natural to consider a negative Sobolev norm of order -1 for the space of Lagrange multipliers Λ_h. However, the fact that $\nabla_h \mathbf{v}_h$ is discontinuous makes it problematic to pair it with a distribution in a negative Sobolev space of order -1. This leads to the embedding of Λ_h into $[L^2(\Omega)]^{2\times2}$, which is somehow responsible for suboptimality.

Theorem 17 (discrete inf-sup constant). *For any $n \geq 0$ and $\mathbf{y}_h^n \in \mathbb{A}_{h,\delta}$, there exists a constant β independent of n and h such that $\beta_h = \beta h_{\min} > 0$ satisfies*

$$(96) \qquad \inf_{\boldsymbol{\mu}_h \in \Lambda_h} \sup_{\mathbf{v}_h \in \mathbb{V}_h^k(\mathbf{0},\mathbf{0})} \frac{b_h(\mathbf{y}_h^n; \mathbf{v}_h, \boldsymbol{\mu}_h)}{\|\mathbf{v}_h\|_{H_h^2(\Omega)} \|\boldsymbol{\mu}_h\|_{L^2(\Omega)}} \geq \beta_h.$$

Proof. We proceed in two steps: we first construct a suitable $\mathbf{v}_h$ and next show (96).

Step 1: Construction of $\mathbf{v}_h$. Given $\boldsymbol{\mu}_h \in \Lambda_h$, let $\boldsymbol{\mu}_{h,T} = \boldsymbol{\mu}_h|_T$ be the constant symmetric 2×2 restriction of $\boldsymbol{\mu}_h$ to any element $T \in \mathcal{T}_h$. Thanks to Lemma 16 (solvability of a matrix equation), there exists a 3×2 constant matrix A_T such that

$$L[A_T; \mathbf{y}_h^n](x_T) : \boldsymbol{\mu}_{h,T} = (A_T^T \nabla \mathbf{y}_h^n(x_T) + \nabla \mathbf{y}_h^n(x_T)^T A_T) : \boldsymbol{\mu}_{h,T} = |\boldsymbol{\mu}_{h,T}|^2,$$

and $|A_T| \leq \frac{|\boldsymbol{\mu}_{h,T}|}{2\sigma_{\min}(\nabla \mathbf{y}_h^n(x_T))}$. Let $\lambda_{\min} : \mathbb{M}^{2\times2} \to \mathbb{R}$ be the smallest eigenvalue function defined over the space of symmetric matrices $\mathbb{M}^{2\times2}$ into $\mathbb{R}$, which turns out to be continuous with respect to any norm. In particular, because $\mathbf{y}_h^n \in \mathbb{A}_{h,\delta}$ we have

$$D_h[\mathbf{y}_h^n](x_T) = \big|\mathrm{I}[\mathbf{y}_h^n](x_T) - I_2\big| \leq \delta.$$

and there is a constant c independent of h and n so that $\big|\lambda_{\min}\big(\mathrm{I}[\mathbf{y}_h^n](x_T) - I_2\big)\big| \leq c\delta$, or

$$\lambda_{\min}\big(\mathrm{I}[\nabla \mathbf{y}_h^n](x_T)\big) \geq 1 - c\delta,$$

Consequently, for δ sufficiently small we deduce

$$\sigma_{\min}\big((\nabla \mathbf{y}_h^n(x_T)\big) = \big(\lambda_{\min}\big(\mathrm{I}[\mathbf{y}_h^n](x_T)\big)\big)^{\frac{1}{2}} \geq \big(1 - c\delta\big)^{\frac{1}{2}},$$

is bounded away from 0 and we have $|A_T| \leq \frac{|\mu_{h,T}|}{2(1-c\delta)^{\frac{1}{2}}}$. We finally define $\mathbf{v}_h(\mathbf{x})\big|_T :=$ $A_T(\mathbf{x} - \mathbf{x}_T)$ on each $T \in \mathcal{T}_h$, where $\mathbf{x}_T$ is the barycenter of T, and observe that $\mathbf{v}_h \in [\mathbb{V}_h^k]^3$ for $k \geq 2$ and $\nabla \mathbf{v}_h|_T = A_T$.

Step 2: Discrete inf-sup property. We first compute

$$b_h(\mathbf{y}_h^n; \mathbf{v}_h, \boldsymbol{\mu}_h) = \sum_{T \in \mathcal{T}_h} |T| \, L[\mathbf{v}_h; \mathbf{y}_h^n](x_T) : \boldsymbol{\mu}_{h,T} = \sum_{T \in \mathcal{T}_h} |\boldsymbol{\mu}_{h,T}|^2 |T| = \|\boldsymbol{\mu}_h\|_{L^2(\Omega)}^2.$$

Since $D_h^2 \mathbf{v}_h = 0$ for $\mathbf{v}_h$ piecewise linear, combining a trace inequality with the Poincaré inequality on each element T gives

$$\|\mathbf{v}_h\|_{H_h^2(\Omega)}^2 = \sum_{e \in \mathcal{E}_h^a} \|\mathrm{h}^{-3/2}[\mathbf{v}_h]\|_{L^2(e)}^2 + \|\mathrm{h}^{-1/2}[\nabla \mathbf{v}_h]\|_{L^2(e)}^2$$

$$\lesssim \sum_{T \in \mathcal{T}_h} \mathrm{h}^{-4}\|\mathbf{v}_h\|_{L^2(T)}^2 + \mathrm{h}^{-2}\|\nabla \mathbf{v}_h\|_{L^2(T)}^2 \lesssim \sum_{T \in \mathcal{T}_h} \mathrm{h}^{-2}\|\nabla \mathbf{v}_h\|_{L^2(T)}^2$$

due to the the fact that $\mathbf{v}_h \in \mathbb{V}_h^k(\mathbf{0}, \mathbf{0})$ has vanishing mean value on T. Therefore,

$$(97) \qquad \|\mathbf{v}_h\|_{H_h^2(\Omega)}^2 \lesssim \sum_{T \in \mathcal{T}_h} \mathrm{h}^{-2} \int_T |A_T|^2 \lesssim h_{\min}^{-2}(1 - c\delta)^{-1} \|\boldsymbol{\mu}_h\|_{L^2(\Omega)}^2.$$

In summary, we have shown that for every $\boldsymbol{\mu}_h \in \Lambda_h$, there exists $\mathbf{v}_h \in \mathbb{V}_h^k(\mathbf{0}, \mathbf{0})$ such that $b_h(\mathbf{y}_h^n; \mathbf{v}_h, \boldsymbol{\mu}_h) = \|\boldsymbol{\mu}_h\|_{L^2(\Omega)}^2$ and $\|\mathbf{v}_h\|_{H_h^2(\Omega)} \lesssim h_{\min}^{-1} \|\boldsymbol{\mu}_h\|_{L^2(\Omega)}$. This is the desired inf-sup condition in disguised. $\qquad\square$

6. Numerical experiments

In this section we present several numerical experiments, some motivated by computations [8, 7, 10, 17] and other by lab experiments [1, 31, 37, 33, 36, 38, 40, 43]. We carry out simulations with several *spontaneous curvature* matrices Z and both *Dirichlet and free boundary conditions*, so as to capture a variety of insightful configurations exhibiting large bending deformations. We consider the effect of different aspect ratios of rectangular domains. We also explore properties of a novel model inspired by [6], which allows folding across curved creases (bilayer origami). Our numerical simulations illustrate the computational performance of our algorithm.

6.1. **Implementation.** We start with a few comments on the implementation of the gradient flow (94) and Algorithm 1.

Saddle-point structure. We resort to a *Schur complement method* to solve the discrete problem (94). We refer to [14] for full implementation details of a similar linear algebra structure, but emphasize here how Theorem 17 (inf-sup stability) guarantees solvability and affects the solver efficiency in the spirit of [12, Lemma 3.1].

To make explicit the Schur complement matrix and deduce its condition number, we denote by $\{\boldsymbol{\varphi}_i\}_{i=1}^N$ a basis for $\mathbb{V}_h^k(\mathbf{0}, \mathbf{0})$ and by $\{\boldsymbol{\psi}_i\}_{i=1}^M$ an orthonormal basis for Λ_h. The matrix representations of the bilinear forms $A_h(\cdot, \cdot) := \tau^{-1}(\cdot, \cdot)_{H_h^2(\Omega)} + a_h(\cdot, \cdot)$ and $b_h(\mathbf{y}_h^n; \cdot, \cdot)$ used to define the gradient flow (94) are thus given by

$$\mathbf{A} := \big(A_h(\boldsymbol{\varphi}_j, \boldsymbol{\varphi}_i)\big)_{i,j=1}^N, \qquad \mathbf{B}_n := \big(b_h(\mathbf{y}_h^n; \boldsymbol{\varphi}_j, \boldsymbol{\psi}_i)\big)_{i=1,j=1}^{M,N}.$$

With this notation, the Schur complement matrix reads $\mathbf{S}_n := \mathbf{B}_n \mathbf{A}^{-1} \mathbf{B}_n^T$ and satisfies

$$(\mathbf{S}_n \mathbf{m}, \mathbf{m}) = (\mathbf{A}^{-1/2} \mathbf{B}_n^T \mathbf{m}, \mathbf{A}^{-1/2} \mathbf{B}_n^T \mathbf{m}) = \sup_{\mathbf{w} \in \mathbb{R}^N} \left(\frac{(\mathbf{w}, \mathbf{A}^{-1/2} \mathbf{B}_n^T \mathbf{m})}{\|\mathbf{w}\|_2} \right)^2$$

$$= \sup_{\mathbf{v} \in \mathbb{R}^N} \left(\frac{(\mathbf{v}, \mathbf{B}_n^T \mathbf{m})}{\|\mathbf{A}^{1/2} \mathbf{v}\|_2} \right)^2 = \sup_{\mathbf{v}_h \in \mathbb{V}_h^k(\mathbf{0}, \mathbf{0})} \frac{b_h(\mathbf{y}_h^n; \mathbf{v}_h, \boldsymbol{\mu}_h)^2}{A_h(\mathbf{v}_h, \mathbf{v}_h)},$$

where $\boldsymbol{\mu}_h := \sum_{i=1}^M m_i \boldsymbol{\psi}_i \in \Lambda_h$, $\mathbf{v}_h := \sum_{j=1}^N v_j \boldsymbol{\varphi}_j \in \mathbb{V}_h^k(\mathbf{0}, \mathbf{0})$ with $\mathbf{m} = (m_i)_{i=1}^M$, $\mathbf{v} = (v_j)_{j=1}^N$, and $\|\cdot\|_2 = (\cdot, \cdot)^{1/2}$ is the Euclidean norm in $\mathbb{R}^N$. On one hand, the continuity (93) of b_h and the coercivity estimate (47) for $B_h[\mathbf{v}_h] = \frac{1}{2} a_h(\mathbf{v}_h, \mathbf{v}_h)$ yield

$$(\mathbf{S}_n \mathbf{m}, \mathbf{m}) \lesssim \tau \|\mathbf{y}_h^n\|_{H_h^2(\Omega)}^2 \|\boldsymbol{\mu}_h\|_{L^2(\Omega)}^2 \lesssim \tau \|\boldsymbol{\mu}_h\|_{L^2(\Omega)}^2 = \tau \|\mathbf{m}\|_2^2,$$

because $\|\mathbf{y}_h^n\|_{H_h^2(\Omega)} \lesssim 1$ in view of the energy stability (79) satisfied by $\mathbf{y}_h^n$ and the coercivity of total energy (48). On the other hand, the inf-sup stability (96) and the continuity estimate (47) for $B_h[\mathbf{v}_h] = \frac{1}{2}a_h(\mathbf{v}_h, \mathbf{v}_h)$ imply

$$\tau h_{\min}^2 \|\mathbf{m}\|_2^2 = \tau h_{\min}^2 \|\boldsymbol{\mu}_h\|_{L^2(\Omega)}^2 \lesssim (\mathbf{S}_n \mathbf{m}, \mathbf{m}).$$

Combining these two inequalities yields an estimate for the condition number of $\mathbf{S}_n$

$$(98) \qquad \kappa(\mathbf{S}_n) := \max_{\mathbf{m} \in \mathbb{R}^M} \frac{(\mathbf{S}_n \mathbf{m}, \mathbf{m})}{\|\mathbf{m}\|_2^2} \left(\min_{\mathbf{m} \in \mathbb{R}^M} \frac{(\mathbf{S}_n \mathbf{m}, \mathbf{m})}{\|\mathbf{m}\|_2^2} \right)^{-1} \lesssim h_{\min}^{-2}.$$

Estimate (98) shows that the saddle-point system is invertible but ill-conditioned. We use a conjugate gradient (CG) iterative solver for the numerical experiments below. Classical convergence theory for CG asserts that the number of iterations to achieve a desired accuracy is of order $\sqrt{\kappa(\mathbf{S}_n)}$ [28, Theorem 3.1.1]. Our numerical experiments reveal that the number of iterations needed in the CG solver roughly behaves like $h_{\min}^{-1}$, which is consistent with (98).

We emphasize that solving the linear system (94) by the Schur complement method for several steps of the gradient flow remains the bottle neck in terms of computing time. We leave the design of suitable preconditioners open.

Assembly. Since the scalar product $\langle \cdot, \cdot \rangle_{H_h^2(\Omega)}$ and bilinear form a_h do not change in the course of the gradient flow, we assemble them once for all before the main loop. In contrast, we assemble the bilinear form $b_h(\mathbf{y}_h^n; \cdot, \cdot)$ and right hand side $\ell[\mathbf{y}_h^n](.)$ at each step of the loop as they depend on the previous iterate $\mathbf{y}_h^n$. Computing the discrete Hessian $H_h[\mathbf{y}_h]$ is the most expensive part in the assembly process, as it requires solving the linear systems (26) and (27) for lifting operators. In order to save computing time, we find the discrete Hessian of each basis function at the beginning of the simulation and store its values for later use; this pre-processing drastically decreases the assembly time.

Software and data. We implement our LDG method within the software platform deal.ii [3] and visualize the outcome with paraview [2]. For all the simulations, we fix the polynomial degree k of the deformation $\mathbf{y}_h$ and the two liftings l_1, l_2 of the discrete Hessian $H_h[\mathbf{y}_h]$, as well as the stabilization parameters γ_1, γ_2 to be

$$k = l_1 = l_2 = 2, \quad \gamma_0 = \gamma_1 = 1.$$

Recall that LDG is stable for any positive choice of parameters γ_1, γ_2, which contrasts with IPDG that requires γ_1, γ_2 large for stability purposes [18, 17].

In the following numerical simulations, we consider both clamped Dirichlet ($\Gamma_D \neq \varnothing$) and free boundary conditions ($\Gamma_D = \varnothing$). For the latter, the discrete equation (74) is no longer well-defined. To fix the system kernel, we add an L^2-term to the metric $(\cdot, \cdot)_{H_h^2(\Omega)}$, while all other implementation aspects are similar to the case $\Gamma_D \neq \varnothing$. We refer to [14] for implementation details of free boundary conditions.

In either situation, a natural choice of initial deformation is that of a flat plate

$$\mathbf{y}_h^0(x_1, x_2) = (x_1, x_2, 0) \quad \forall (x_1, x_2) \in \Omega,$$

and satisfies clamped boundary conditions and the isometry constraint $I[\mathbf{y}_h^0] = I_2$ everywhere. This is much simpler than prestrained plates [14, 15], which require preprocessing of both boundary condition and metric constraint to construct suitable initial deformations for LDG to start.

6.2. **Clamped plate: Isotropic curvature.** We consider a rectangular plate $\Omega = (-5, 5) \times (-2, 2)$, clamped on the side $\{-5\} \times [-2, 2]$, with isotropic spontaneous curvature $Z = I_2$. The deformation with minimal energy corresponds to a cylinder of radius 1 and energy 20 [41]. This is confirmed by our simulations in Fig. 2, which displays iterations of the discrete gradient flow with 1024 elements (30720 dofs), $\tau = 5 \times 10^{-3}$ and $tol = 10^{-4}$ in Algorithm 1, as in [8].

We notice that surface self-intersecting develop during the relaxation dynamics of Algorithm 1; this is similar to [10] but different from [17]. Moreover, it takes fewer iterations for Algorithm 1 to reach the cylindrical equilibrium configuration, with the same or even smaller time step τ, than FEMs in [8, 10, 17].

Moreover, we keep the time step $\tau = 5 \times 10^{-3}$ fixed and consider two quasi-uniform meshes with 256 and 1024 elements. We obtain bending energies $E_h = 16.8627$ and $E_h = 17.8038$ (16% and 11% relative error) respectively, which exhibit smaller errors than the corresponding ones $E_h = 15.961$ and $E_h = 16.544$ with the Kirchhoff FEM of [8] for the same mesh-size and time step. In addition, the energy error compares favorably with the new Kirchhoff FEM in [10], which computes with $Z = 2.5I_2$ and produces a 36% relative error even with a finer mesh of 5120 triangular elements.
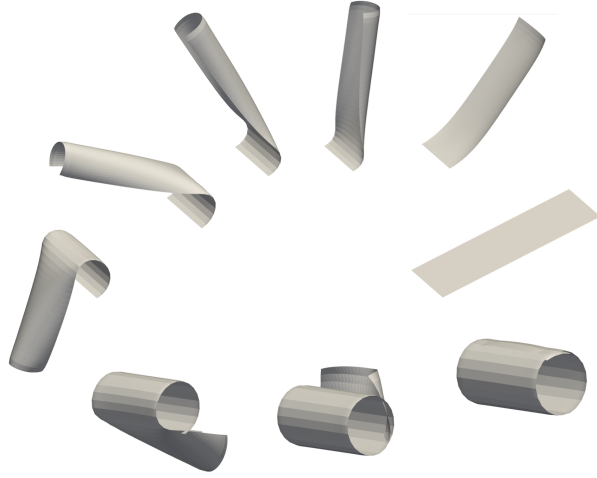

FIGURE 2. Isotropic curvature: Relaxation dynamics of Algorithm 1 towards the cylinder equilibrium shape of a clamped rectangular plate with the isotropic spontaneous curvature $Z = I$. The bilayer plate is depicted at times $0, 50, 1000, 9000, 18000, 36050, 48100, 56050, 72100$ of the gradient flow (counter-clockwise).

6.3. **Free plate: Anisotropic curvature.** We now explore a cigar-type configuration motivated by experiments [31] and computations [8, 17]. The plate is again the rectangle $\Omega = (-5, 5) \times (-2, 2)$, but now we impose no boundary condition (free boundary) along with the anisotropic spontaneous curvature

$$(99) \qquad Z = \begin{bmatrix} 3 & -2 \\ -2 & 3 \end{bmatrix}.$$

We observe that the eigenpairs of Z are $(1, [1, 1]^T)$ and $(5, [1, -1]^T)$. We thus expect that the plate deforms at -45 degrees with respect to the Cartesian axes in a symmetric way and eventually reaches a cigar-like configuration, as in [17]. We confirm this in Fig. 3, that displays computations with 1024 elements (30720 dofs) and $\tau = 5 \times 10^{-3}$. The final energy is $E_h = 46.3898$. Remarkably, Algorithm 1 takes fewer iterations to reach the equilibrium configuration than [17].
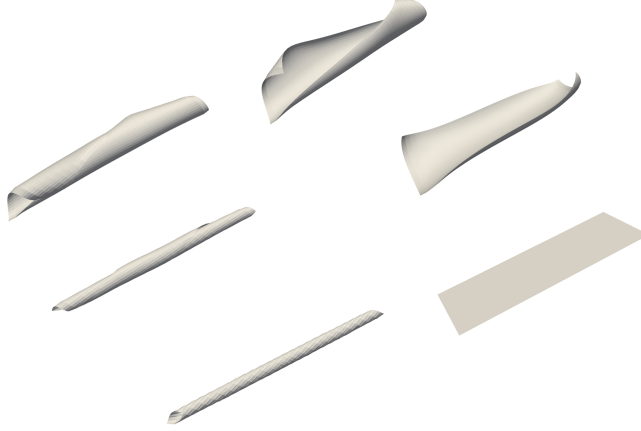

FIGURE 3. Anisotropic curvature: Relaxation dynamics of Algorithm 1 towards the cigar-type equilibrium of a free rectangular plate with the anisotropic spontaneous curvature of (99). The bilayer plate is depicted at times $0, 50, 200, 1000, 10000, 30000$ of the gradient (counter-clockwise).

6.4. **Free plate: Helix shape.** We present a helix-type shape motivated by a DNA-like configuration [43]. We consider a high aspect ratio plate $\Omega = (-8, 8) \times (-0.5, 0.5)$, with free boundary condition and anisotropic spontaneous curvature

$$(100) \qquad Z = \begin{bmatrix} 1 & -3/2 \\ -3/2 & 1 \end{bmatrix}.$$

We point out that the eigenpairs of Z are $(-\frac{1}{2}, [1, 1]^T)$ and $(\frac{5}{2}, [1, -1]^T)$, which again correspond to principal directions that form an angle of 45 degrees with the coordinate axes. This, together with eigenvalues of opposite sign and high aspect ratio, leads to a deformation that resembles the twisting of DNA molecules, as in [17]. We display several snapshots of the relaxation dynamics of Algorithm 1 in

FIGURE 4. Anisotropic indefinite curvature: Relaxation dynamics of Algorithm 1 with spontaneous curvature (100) towards a DNA-like equilibrium configuration of a free rectangular strip with large aspect ratio. The bilayer plate is depicted at times $0, 100, 200, 1000, 4000, 12600$ of the gradient flow (left to right).

Fig. 4. The simulation is carried out with 1024 elements and $\tau = 10^{-2}$, and yields a final energy $E_h = 3.2507$. Moreover, it again takes fewer iterations for LDG to reach the equilibrium configuration than the DG method of [17].

6.5. **Climate responsive architectures.** Bilayer devices can be used to control the temperature or moisture inside a room. The HygroSkin project [37, 38, 40, 33, 36] exploits this technology by designing visually appealing humidity responsive apertures to a pavilion. Heat and moisture are thus dynamically controlled without any high-tech equipment owing to the dominant orientation of fibers in plywood.

To simulate this device with our bilayer model, we consider an equilateral triangle with side length 1 and vertices $(0,0)$, $(1,0)$ and $(\frac{1}{2}, \frac{\sqrt{3}}{2})$. The actual climate responsive device consists of 6 of these triangular shapes suitably rotated and arranged together as to form a flat regular hexagon, with the exterior edge of each triangle clamped; we refer to Fig. 5. To mimic the effect of different relative humidity values, we choose several anisotropic spontaneous curvatures

$$(101) \qquad\qquad Z = \begin{pmatrix} 0 & 0 \\ 0 & \alpha \end{pmatrix}, \qquad \text{with } \alpha = 0, 1, 2, 3, 4, 5,$$

for the triangle with exterior edge parallel to the x-axis and suitably rotated for the other triangles. This matrix favors bending along the y-axis exclusively.

Upon actuation, the climate device automatically opens as depicted in Fig. 6. The matching of the computed (left) and actual (right) equilibrium shapes in Fig. 6
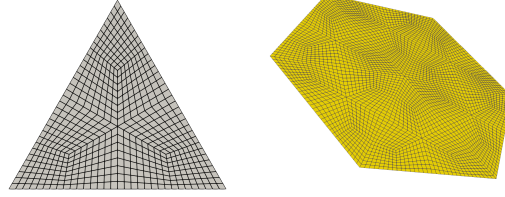
FIGURE 5. Climate responsive device. The undeformed plate is made of 6 equilateral triangles that together form a regular hexagon (right). Finite element partition of each triangle into trapezoids (left).

is quite remarkable for a model with just one parameter α within Z. We run this simulation with time step $\tau = 1$ and stopping tolerance $tol = 10^{-4}$.
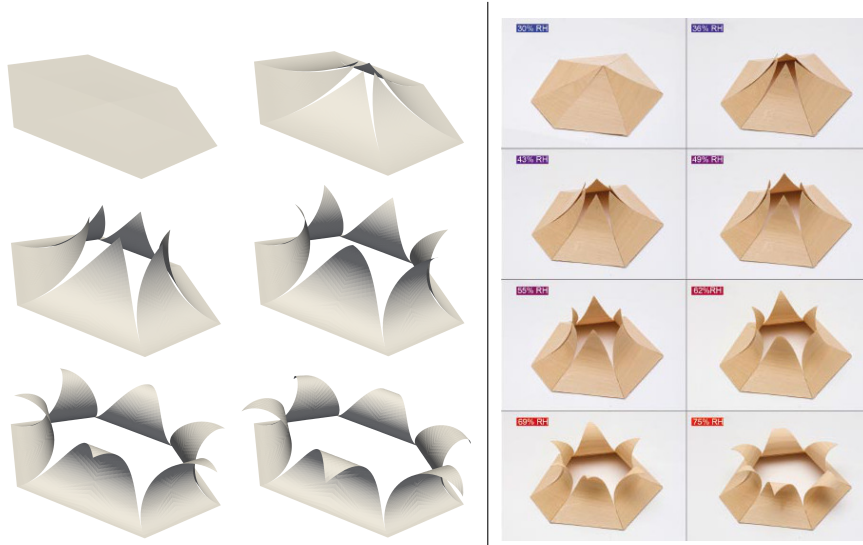

FIGURE 6. Climate responsive device. (Left) Approximate deformation for different values of spontaneous curvature (101) with parameter $\alpha = 0, 1, 2, 3, 4, 5$. (from left to right and top to bottom) (Right) Experimental deformations of a device made of plywood. Picture taken from [36] (courtesy of Prof. Achim Menges); see also [37]. The matching is remarkable.

6.6. **Folding Model: Bilayer Origami.** We finally explore computationally the combined effect of spontaneous curvature, as driving mechanism, and folding across a preassigned crease. The corresponding bilayer model and LDG method are discussed in Section 4. We consider below the setting from [6, Section 5.2] and refer to [13] for additional numerical simulations.
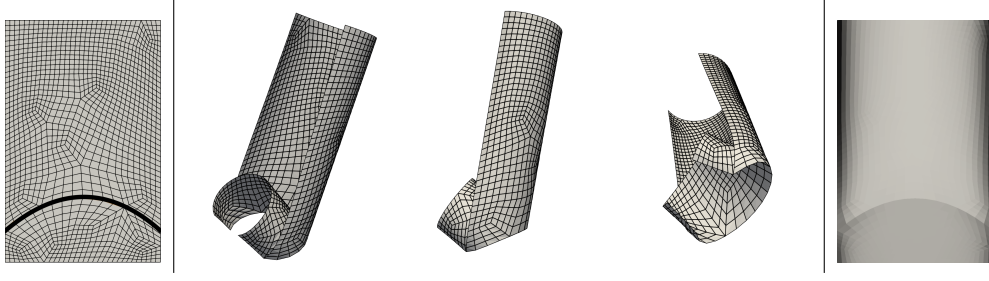
FIGURE 7. Bilayer origami: Flapping mechanism generated by folding of a bilayer plate across a crease. (Left) Conforming subdivision of Ω with quadratic crease. (Middle) Different perspectives of the resulting very large deformation. (Right) Isometry defect $D_h[\mathbf{y}_h](x)$ ranging from 3.0×10^{-5} (white) to 3.9×10^{-2} (black).

The computational domain is a rectangle $\Omega = (0, 9.6) \times (0, 15)$ and the folding crease is a quadratic curve $\mathcal{C}$ passing through the points $(0, 2)$, $(9.6, 2)$, and $(4.8, 6)$, which can be exactly represented by the isoparametric mesh $\mathcal{T}_h$; see Fig. 7. In order to generate a configuration similar to the flapping mechanism in [6], which is obtained by compression of the lateral boundary, we set the spontaneous curvatures

$$Z = \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}, \quad Z = \begin{pmatrix} 0 & 0 \\ 0 & -\frac{1}{2} \end{pmatrix},$$

below the folding arc and above of it, respectively, and do not impose any boundary condition. The resulting equilibrium shape is displayed in Fig. 7 along with the isometry defect $D_h[\mathbf{y}_h](x)$ at equilibrium; it ranges from 3.0×10^{-5} to 3.9×10^{-2}. We point out the crucial role played by the sign of principal curvatures $\lambda = \frac{1}{2}, -\frac{1}{2}$ corresponding to the same coordinate eigendirection: bending of the lower and upper plates occurs in opposite directions which gives rise to folding across the crease and yields a rather large compatible deformation.

7. CONCLUSIONS

In this article, we present a new LDG method for large bending isometric deformations of bilayer plates. We summarize our contributions in this section.

1. *LDG discretization.* It consists of replacing the Hessian $D^2\mathbf{y}$ by a reconstructed Hessian $H_h[\mathbf{y}_h]$ in the bending energy $B_h[\mathbf{y}_h]$, and by a reduced (piecewise constant) discrete Hessian $\overline{H}_h[\mathbf{y}_h]$ in the cubic energy $C_h[\mathbf{y}_h]$, which encodes the interaction with spontaneous curvature. We use the mid-point quadrature to integrate $C_h[\mathbf{y}_h]$.

2. *Relaxed isometry constraint.* This allows for a slight violation of the isometry constraint $\mathrm{I}[\mathbf{y}_h] = I_2$ while providing control of the ℓ^∞-norm of the isometry defect $D_h[\mathbf{y}_h] = \left| \mathrm{I}[\mathbf{y}_h] - I_2 \right|$ at element barycenters. This turns out to be a significant improvement over previous DG methods that enforce such defect as sum of averages over elements [14, 15, 18, 17].

3. *Γ-convergence.* The key novelty of the Γ-convergence of discrete energies is the construction of the recovery sequence of any admissible deformation $\mathbf{y} \in [H^2(\Omega) \cap W^1_\infty(\Omega)]^3$. It hinges on a quadratic Taylor expansion at element barycenters of a suitable regularization of $\mathbf{y}$, and exploits that both the reduced quadrature of $C_h[\mathbf{y}_h]$ and the isometry defect $D_h[\mathbf{y}_h]$ are imposed at element barycenters.

4. *Bilayer model with foldings.* We extend the LDG method to deal with a piece-wise quadratic crease $\mathcal{C}$ and prove its convergence. The construction of a recovery sequence for one absolute minimizer $\mathbf{y}^* \in [H^2(\Omega \backslash \mathcal{C}) \cap W^1_\infty(\Omega)]^3$ requires the slightly stronger assumption that $\mathbf{y}^*$ is C^1 in each subdomain created by $\mathcal{C}$.

5. *Fully linear solver.* We design a semi-implicit discrete gradient flow that treats $B_h[\mathbf{y}_h]$ implicitly and $C_h[\mathbf{y}_h]$ explicitly. This leads to linear problems at each step. The scheme retains the key property of being energy diminishing and controls the isometry defect provided the fictitious time step τ satisfies a mild constraint.

6. *Sub-optimal discrete inf-sup.* As is customary in the literature [4, 7, 8, 10, 14, 15, 18, 17], we rely on Lagrange multipliers to enforce the linearized isometry constraint. We prove a sub-optimal inf-sup condition for the resulting saddle-point system, which seems to be the first such result for these type of problems.

7. *Simulations.* We present several insightful numerical experiments with large isometric deformations, including a climate responsive device and the folding of a plate across a quadratic crease that yields a bilayer origami as equilibrium shape. We also document the size of the isometry defect $D_h[\mathbf{y}_h]$ for the latter.

References

[1] Alben, S., Balakrisnan, B., and Smela, E. Edge effects determine the direction of bilayer bending. *Nano letters 11*, 6 (2011), 2280–2285.

[2] Ayachit, U. *The ParaView Guide: A Parallel Visualization Application.* Kitware, Inc., USA, 2015.

[3] Bangerth, W., Hartmann, R., and Kanschat, G. deal.II – a general purpose object oriented finite element library. *ACM Trans. Math. Softw. 33*, 4 (2007), 24/1–24/27.

[4] Bartels, S. Approximation of large bending isometries with discrete kirchhoff triangles. *SIAM Journal on Numerical Analysis 51*, 1 (2013), 516–525.

[5] Bartels, S. *Numerical methods for nonlinear partial differential equations*, vol. 47. Springer, 2015.

[6] Bartels, S., Bonito, A., and Hornung, P. Modeling and simulation of thin sheet folding. *Interfaces Free Bound. 24*, 4 (2022), 459–485.

[7] Bartels, S., Bonito, A., Muliana, A. H., and Nochetto, R. H. Modeling and simulation of thermally actuated bilayer plates. *Journal of Computational Physics 354* (2018), 512–528.

[8] Bartels, S., Bonito, A., and Nochetto, R. H. Bilayer plates: Model reduction, Γ-convergent finite element approximation, and discrete gradient flow. *Communications on Pure and Applied Mathematics 70*, 3 (2017), 547–589.

[9] Bartels, S., Bonito, A., and Tscherner, P. Error estimates for a linear folding model. *IMA Journal of Numerical Analysis* (02 2023), (to appear).

[10] Bartels, S., and Palus, C. Stable gradient flow discretizations for simulating bilayer plate bending with isometry and obstacle constraints. *IMA Journal of Numerical Analysis 42*, 3 (2022), 1903–1928.

[11] BASSI, F., REBAY, S., MARIOTTI, G., PEDINOTTI, S., AND SAVINI, M. A high-order accurate discontinuous finite element method for inviscid and viscous turbomachinery flows. In *Proceedings of the 2nd European Conference on Turbomachinery Fluid Dynamics and Thermodynamics* (1997), Antwerpen, Belgium, pp. 99–109.

[12] BERRONE, S., BONITO, A., STEVENSON, R., AND VERANI, M. An optimal adaptive fictitious domain method. *Mathematics of Computation 88*, 319 (2019), 2101–2134.

[13] BONITO, A., GUIGNARD, D., AND MORVANT, A. Numerical approximations of thin structure deformations. *Comptes Rendus. Mécanique 351*, S1 (2023), 1–37.

[14] BONITO, A., GUIGNARD, D., NOCHETTO, R. H., AND YANG, S. LDG approximation of large deformations of prestrained plates. *Journal of Computational Physics 448* (2022), 110719.

[15] BONITO, A., GUIGNARD, D., NOCHETTO, R. H., AND YANG, S. Numerical analysis of the LDG method for large deformations of prestrained plates. *IMA J. Numer. Anal. 43*, 2 (2023), 627–662.

[16] BONITO, A., AND NOCHETTO, R. Quasi-optimal convergence rate of an adaptive discontinuous Galerkin method. *SIAM Journal on Numerical Analysis 48*, 2 (2010), 734–771.

[17] BONITO, A., NOCHETTO, R. H., AND NTOGKAS, D. Discontinuous Galerkin approach to large bending deformation of a bilayer plate with isometry constraint. *Journal of Computational Physics 423* (2020), 109785.

[18] BONITO, A., NOCHETTO, R. H., AND NTOGKAS, D. DG approach to large bending plate deformations with isometry constraint. *Mathematical Models and Methods in Applied Sciences 31*, 01 (2021), 133–175.

[19] BRAMBLE, J. H., PASCIAK, J. E., AND SCHATZ, A. H. The construction of preconditioners for elliptic problems by substructuring. i. *Mathematics of Computation 47*, 175 (1986), 103–134.

[20] BRENNER, S., AND SCOTT, R. *The mathematical theory of finite element methods*, vol. 15. Springer Science & Business Media, 2007.

[21] BREZZI, F., MANZINI, G., MARINI, D., PIETRA, P., AND RUSSO, A. Discontinuous finite elements for diffusion problems. *Atti Convegno in onore di F. Brioschi (Milano 1997), Istituto Lombardo, Accademia di Scienze e Lettere 1999* (1999), 197–217.

[22] BREZZI, F., MANZINI, G., MARINI, D., PIETRA, P., AND RUSSO, A. Discontinuous Galerkin approximations for elliptic problems. *Numerical Methods for Partial Differential Equations: An International Journal 16*, 4 (2000), 365–378.

[23] CIARLET, P. G., AND RAVIART, P.-A. Interpolation theory over curved elements, with applications to finite element methods. *Computer Methods in Applied Mechanics and Engineering 1*, 2 (1972), 217–249.

[24] COCKBURN, B., AND SHU, C.-W. The local discontinuous Galerkin method for time-dependent convection-diffusion systems. *SIAM Journal on Numerical Analysis 35*, 6 (1998), 2440–2463.

[25] DI PIETRO, D., AND ERN, A. Discrete functional analysis tools for discontinuous Galerkin methods with application to the incompressible navier-stokes equations. *Mathematics of Computation 79*, 271 (2010), 1303–1330.

[26] DI PIETRO, D. A., AND ERN, A. *Mathematical aspects of discontinuous Galerkin methods*, vol. 69. Springer Science & Business Media, 2011.

[27] ERN, A., AND GUERMOND, J.-L. *Finite elements I: Approximation and interpolation*, vol. 72. Springer Nature, 2021.

[28] GREENBAUM, A. *Iterative methods for solving linear systems*. SIAM, 1997.

[29] GUAN, J., HE, H., HANSFORD, D. J., AND LEE, L. J. Self-folding of three-dimensional hydrogel microstructures. *The Journal of Physical Chemistry B 109*, 49 (2005), 23134–23137.

[30] HORNUNG, P. Approximating $W^{2,2}$ isometric immersions. *Comptes Rendus Mathematique 346*, 3-4 (2008), 189–192.

[31] JANBAZ, S., HEDAYATI, R., AND ZADPOOR, A. Programming the shape-shifting of flat soft matter: from self-rolling/self-twisting materials to self-folding origami. *Materials Horizons 3*, 6 (2016), 536–547.

[32] Kim, D.-H., and Rogers, J. A. Stretchable electronics: materials strategies and devices. *Advanced materials 20*, 24 (2008), 4887–4892.

[33] Krieg, O. D. Hygroskin–meteorosensitive pavilion. In *Advancing Wood Architecture*. Routledge, 2016, pp. 125–140.

[34] Love, M., Zink, P., Stroud, R., Bye, D., Rizk, S., and White, D. Demonstration of morphing technology through ground and wind tunnel tests. In *48th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference* (2007), p. 1729.

[35] Mano, J. F. Stimuli-responsive polymeric systems for biomedical applications. *Advanced Engineering Materials 10*, 6 (2008), 515–527.

[36] Menges, A. HygroSkin: Meteorosensitive pavilion. `http://www.achimmenges.net/?p=5612`, Last accessed on 2023-01-07.

[37] Menges, A., and Reichert, S. Material capacity: embedded responsiveness. *Architectural Design 82*, 2 (2012), 52–59.

[38] Menges, A., and Reichert, S. Performative wood: physically programming the responsive architecture of the hygroscope and hygroskin projects. *Architectural Design 85*, 5 (2015), 66–73.

[39] Ntogkas, D. *Non-linear geometric PDEs: algorithms, numerical analysis and computation.* PhD thesis, University of Maryland, 2018.

[40] Reichert, S., Menges, A., and Correa, D. Meteorosensitive architecture: Biomimetic building skins based on materially embedded and hygroscopically enabled responsiveness. *Computer-Aided Design 60* (2015), 50–69.

[41] Schmidt, B. Minimal energy configurations of strained multi-layers. *Calculus of Variations and Partial Differential Equations 30*, 4 (2007), 477–497.

[42] Schmidt, B. Plate theory for stressed heterogeneous multilayers of finite bending energy. *Journal de mathématiques pures et appliquées 88*, 1 (2007), 107–122.

[43] Simpson, B., Nunnery, G., Tannenbaum, R., and Kalaitzidou, K. Capture/release ability of thermo-responsive polymer particles. *Journal of Materials Chemistry 20*, 17 (2010), 3496–3501.

[44] Sodhi, J., and Rao, I. Modeling the mechanics of light activated shape memory polymers. *International Journal of Engineering Science 48*, 11 (2010), 1576–1589.

[45] Stoychev, G., Puretskiy, N., and Ionov, L. Self-folding all-polymer thermoresponsive microcapsules. *Soft Matter 7*, 7 (2011), 3277–3279.

[46] Stoychev, G., Zakharchenko, S., Turcaud, S., Dunlop, J. W., and Ionov, L. Shape-programmed folding of stimuli-responsive polymer bilayers. *ACS nano 6*, 5 (2012), 3925–3934.