SURFACE PENALIZATION OF SELF-INTERPENETRATION IN LINEAR AND NONLINEAR ELASTICITY

STEFAN KRÖMER AND JAN VALDMAN

ABSTRACT. We analyze a term penalizing surface self-penetration, as a soft constraint for models of hyperelastic materials to approximate the Ciarlet-Nečas condition (almost everywhere global invertibility of deformations). For a linear elastic energy subject to an additional local invertibility constraint, we prove that the penalized elastic functionals converge to the original functional subject to the Ciarlet-Nečas condition. The approach also works for nonlinear models of non-simple materials including a suitable higher order term in the elastic energy, without artificial local constraints. Numerical experiments illustrate our results for a self-contact problem in 3d.

1. INTRODUCTION

This article is a follow-up of [34], contributing to an ongoing effort of obtaining a mathematically rigorous computational approach for obtaining steady states or global energy minimizers in models of elastic solids (for general background, see [49, 16, 4], e.g.) featuring a global non-penetration constraint reflecting non-interpenetrability of matter. Here, we consider an approximation scheme "rigorous" if its solutions can be proved to converge in some sense to a solution of the original constrained problem. As we intend to focus on large deformations, we naturally have to study problems of self-contact (frictionless, which is the most simple case). Such problems are always inherently nonconvex, even if a linear elastic model is used for the local response of the material to stresses. As a consequence, many techniques developed for rigid substrate contact problems, including a formulation of the problem as a variational inequality, fail. Even models for one-dimensional rods in 2d already encounter some subtleties [35]. Moreover, results related to Lagrange multiplier theory are only available for elastic models of non-simple materials which offer higher order regularity [46, 45]. The study of critical points is therefore largely out of reach on the analytical level, and we are left with the more accessible theory for global energy

minimizers as pioneered by Ball for hyperelasticity [6]. On the level of numerical convergence results, we have to handle the possibility of non-uniqueness of solutions, and as a consequence, convergence to one of them can typically only be ensured for suitable subsequences of the approximate solutions. As we are in a variational framework, we will take advantage of the language of De Giogi's Γ-convergence to express this, cf. Remark 3.12 (or the slightly stronger Mosco-convergence).

On the level of the model, a non-penetration constraint translates to global injectivity of the deformation map $y : \Omega \to \mathbb{R}^d$ mapping the "reference configuration" $\Omega \subset \mathbb{R}^d$ to the deformed state (typically in dimension $d = 3$). In suitable spaces of orientation preserving deformations, almost everywhere global injectivity of y is equivalent to the well-known Ciarlet-Nečas condition [17]

$$\int_\Omega |\det(\nabla y)| \, \mathrm{d}x \le |y(\Omega)| . \tag{1.1}$$

As no rigorous and computationally feasible projection onto the constraint (1.1) is known, we approach it with a penalty method, roughly following [34]. For a class of nonlinear elastic models leading to deformations of bounded distortion, major progress has recently been made in [33], which uses a completely different penalty term which allows successful convergence analysis without any additional local constraints or higher order regularity as in [34].

In this paper, we will follow a more practically-minded path, combining a new penalty term supported only on the boundary (or the piece where expected contact) with a linear elastic model. It is introduced and analyzed in Section 2. One important feature of our penalty is that like its bulk variant in [34], it can fully prevent interpenetration as shown in Corollary 2.3. Its convergence in combination with elastic energies, also discretized, is discussed in Section 3. Compared to [34], restricting the penalization to the boundary allows a major reduction of computational cost, effectively reducing the dimension by two, as our penalty term is a nonlocal double integral. Here, notice that some degree of nonlocality is unavoidable – it simply reflects the nonlocal nature of global injectivity.

Of course, our choice of a linear elastic model is largely arbitrary. In fact, our analysis does not even exploit it beyond the fact that it avoids the determinant singularity of fully nonlinearly elastic energy densities. We prefer it as the main example here because we believe that linear elastic models are used in the vast majority of practical computations, especially in 3d, simply due to their much lower computational cost. Besides, in our concrete numerical experiments, this means that the performance gain from a more efficient interpenetration penalty becomes much more noticable.

Using a linear elastic model comes with the caveat that by itself, it is unable to enforce locally orientation preserving deformations, and so we have to supplement it with a local[1] locking-type constraint (3.6) which implies a local bi-Lipschitz property and thus prevents local loss of injectivity. As explained in greater detail in Subsection 3.2, we do not expect this local constraint to represent actual material properties; instead, our philosophy here is that if (3.6) (with generously chosen constants) is violated or active, deformation gradients deviate so far from the identity that the linear elastic model is too poor an approximation of reality, anyway. In particular, we do not try to enforce (3.6) in our numerical experiments, although it can be checked a posteriori. It mainly serves as a restriction necessary so that our convergence analysis is valid, see Theorem 2.1 and our main theoretical result, Theorem 3.5. At the same time, as further discussed below, it avoids a Lavrentiev phenomenon in context of finite element approximations. For comparison with [34], we also provide a convergence result featuring a nonlinear elastic energy for non-simple materials involving higher order derivatives, see Theorem 3.9. By results of [28, 34], this regularized energy implicitly enforces a local bi-Lipschitz property of deformations (3.7) acting as a suitable replacement for (3.6).

To keep technicalities to a minimum, the theoretical part does not involve partial Dirichlet boundary conditions or force terms, although typical examples of the latter are trivial to add. This is further discussed in Subsection 3.4.

We also point out that there is a rather straightforward and rigorous way to obtain a penalty term directly based on (1.1), see for instance [40]:

$$E_\varepsilon^{\mathrm{penalty}}(y) := \frac{1}{\varepsilon}\Big(\int_\Omega |\det(\nabla y)|\,\mathrm{d}x - |y(\Omega)| \Big), \quad \text{with } \varepsilon > 0 \text{ small.}$$

This approach seems to have major computational disadvantages however, including a very high computational cost and a non-smooth character. We are not aware of any practical implementation of it. In addition, unlike our penalty terms, it vanishes on all injective deformations and thus always leads to a little bit of interpenetration in computations if the approximated solution is in self-contact with non-vanishing surface contact forces. In particular, it probably cannot be generalized at all to a more efficient boundary variant, cf. Remark 2.4.

Apart from that, there are already many numerical approaches based on heuristical arguments, emphasizing performance but lacking a proof of convergence, see for instance [47, 2, 14, 23]. There also is a rich literature for energy terms with self-repulsive properties for curves or

---

[1]actually, it is slightly nonlocal, but with arbitrarily short range

surfaces, including numerical results, see e.g. [11, 9, 10, 13, 50]. These typically require higher regularity though, at least C^1 which is more than we want to impose here, as this can collide with a possible Lavrentiev phenomenon [26].

Generally, numerical approximation in the presence of constraints related to local invertibility or orientation preservation has to handled with care to avoid possible Lavrentiev phenomenona or related issues [43, 5]. In fact, even the density of finite element spaces in sets of admissible orientation-preserving deformations with finite energy is often nontrivial, as illustrated by the still open Ball-Evans problem [8]. We also refer to [21, 22, 31] for some results in dimension $d = 2$ and to [15] for a counterexample in Sobolev spaces with low integrability. To justify our assumptions in Theorem 3.5 in this regard, we explicitly show the existence of almost conforming finite elements for our constrained linear elastic model in Proposition 3.17. In fact, this is the reason for using the constraint (3.6) instead of a straightforward local bi-Lipschitz property like (3.7).

Our practical numerical experiments are presented in Section 4. There, we heavily exploit the linear elastic model using a Schur complement method, effectively pre-solving the linear elastic problem in the interior to transform the whole problem to one exclusively depending on boundary nodes. In practice, we can even use a prescribed subset of "non-penetration" boundary nodes, if we restrict the penalty energy to the the associated boundary part. In this way, a priori intuition about the expected contact set can be exploited to further reduce the cost of the computation.

2. Penalization terms for the Ciarlet-Nečas condition

The basic idea of a penalty method in variational context is to replace a given constraint by an additional "penalization" term in the energy. This term should roughly approximate a functional that vanishes where the constraint holds while it is infinite elsewhere. For asymptotics, it is natural to have this approximation governed by a penalization parameter, $\varepsilon > 0$ below, where $\varepsilon = 0$ at least formally corresponds to the case where the original constraint is again perfectly enforced.

Before we introduce a new penalization term $E_\varepsilon^{\partial\Omega}$ for the Ciarlet-Nečas condition acting only of the surface, let us recall the bulk version introduced in [34] which serves as a role model. It is given as follows:

$$E_\varepsilon^{CN}(y) := \frac{1}{\varepsilon^\beta} \int_\Omega d_{\varepsilon,y}^{CN}(x)\,\mathrm{d}x \quad \text{with}$$

$$d_{\varepsilon,y}^{CN}(x) := \int_\Omega \frac{1}{\varepsilon^d} \left[g(|\tilde{x} - x|) - g\Big(\frac{|y(\tilde{x}) - y(x)|}{\varepsilon}\Big) \right]^+ \mathrm{d}\tilde{x}, \tag{2.1}$$

where $[a]^+ := \max\{0, a\}$ denotes the positive part, $\beta > 0$ is a constant and

$$g : [0, \infty) \to [0, \infty) \text{ is a continuous,}$$
$$\text{strictly increasing function with } g(0) = 0. \tag{2.2}$$

For locally bi-Lipschitz deformations y, as $\varepsilon \to 0$, $E_\varepsilon^{\mathrm{CN}}(y)$ converges to zero whenever y satisfies the Ciarlet-Nečas condition (1.1) and is out of self-contact, and to $+\infty$ if (1.1) is violated. Borderline cases with self-contact are more subtle but also lead to the correct limit if y is replaced by an appropriate approximating sequence (a suitable recovery sequence in the spirit of Gamma convergence). For more details see [34, Thm. 3.3 and Thm. 4.6].

One of the disadvantages of $E_\varepsilon^{\mathrm{CN}}$ is its nonlocal nature and the ensuing computational complexity for numerical evaluation: Typically, a single evaluation of the double integral has a cost of the order of h^{-2d} elementary operations, where h is the grid size and d the dimension of the reference configuration Ω. While it is clear that the nonlocal nature of global invertibility will also be reflected in any associated penalty term, the computational cost can be reduced if we work with integration over the boundary $\partial\Omega$ instead of the full domain, effectively decreasing the dimension.

As we will show, this is possible while retaining the main effect of the penalty term, using the following surface variant:

$$E_\varepsilon^{\partial\Omega}(y) := \frac{1}{\varepsilon^\beta} \int_{\partial\Omega} d_{\varepsilon,y}^{\partial\Omega}(x)\, \mathrm{d}\mathcal{H}^{d-1}(x) \quad \text{with}$$
$$d_{\varepsilon,y}^{\partial\Omega}(x) := \int_{\partial\Omega} \frac{1}{\varepsilon^{d-1}} P\left(g(|\tilde{x} - x|) - g\left(\frac{|y(\tilde{x}) - y(x)|}{\varepsilon}\right)\right) \mathrm{d}\mathcal{H}^{d-1}(\tilde{x}), \tag{2.3}$$

where $\beta > 0$, g satisfies (2.2) as before, $\mathcal{H}^{d-1}$ denotes the $(d-1)$-dimensional Hausdorff measure (the surface measure) and P is an approximation of the positive part $[\cdot]^+$ in the sense that

$$P : \mathbb{R} \to [0, \infty) \text{ is continuous with } [t-1]^+ \leq P(t) \leq [t]^+,$$
$$\text{and } P(t) > 0 \text{ for } t > 0. \tag{2.4}$$

In practice, we choose P and g as smooth approximations of the map $t \mapsto [t]^+$.

Illustrations of $d_{\varepsilon,y}^{CN}$ and $d_{\varepsilon,y}^{\partial\Omega}$ for a 2d deformation with self-penetration are given in Figure 1 and were generated by the complementary code of [34] .

2.1. **Analytic investigation of the penalty term.** We now analyze the behavior of $E_\varepsilon^{\partial\Omega}$ as $\varepsilon \to 0$.

(A) Undeformed domain.

(D) Deformed domain.

(B) Density $d_{\varepsilon,y}^{CN}(x)$ for $\epsilon = 1/2$.

(E) Density $d_{\varepsilon,y}^{CN}(x)$ for $\epsilon = 1/4$.

(C) Density $d_{\varepsilon,y}^{\partial\Omega}(x)$ for $\epsilon = 1/2$.

(F) Density $d_{\varepsilon,y}^{\partial\Omega}(x)$ for $\epsilon = 1/4$.
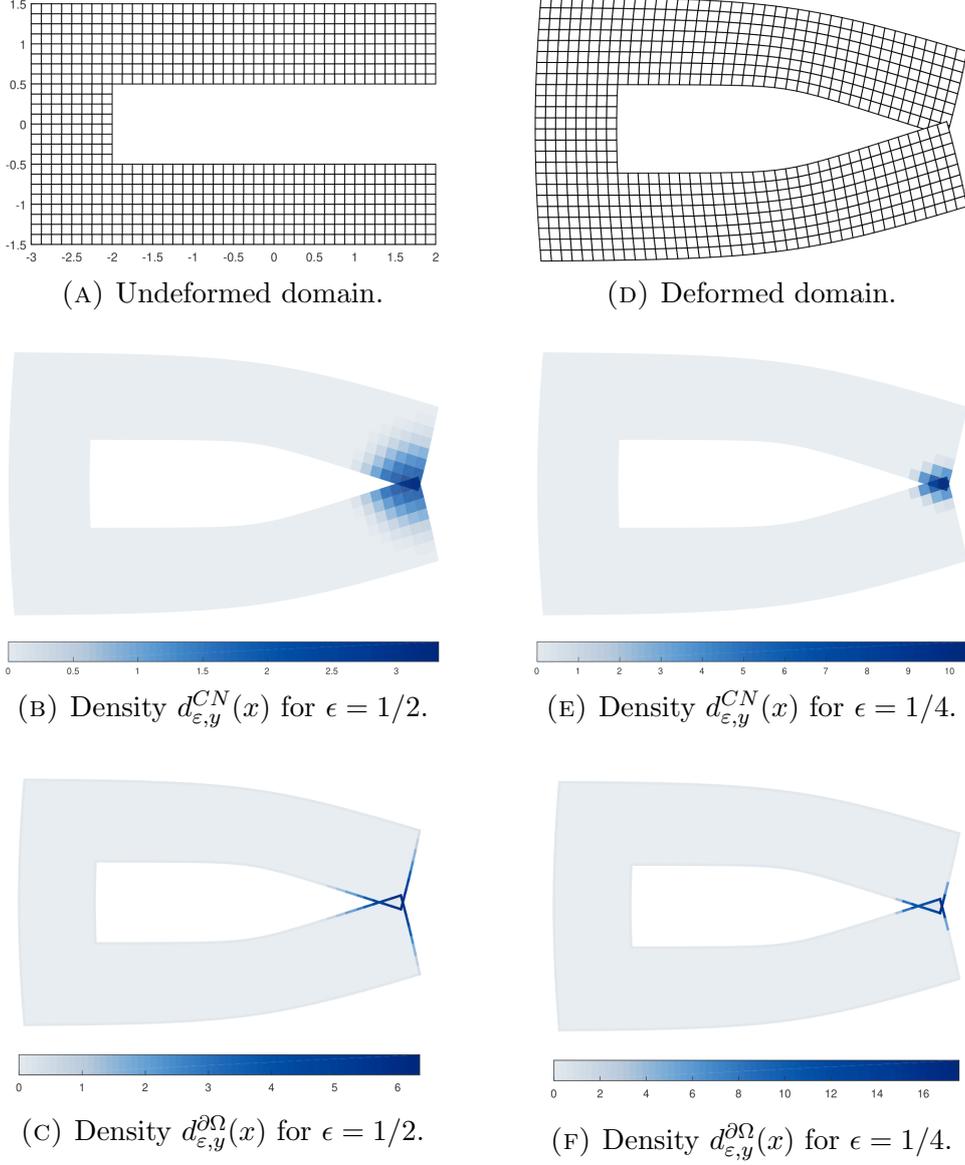
FIGURE 1. Pincers domain under given deformation.

Theorem 2.1 (Asymptotics of the surface penalty term (2.3)).
Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain, let $E_\varepsilon^{\partial\Omega}$ be the functional defined in (2.3) with $\beta > 0$, let g satisfy (2.2) and P satisfy (2.4). Moreover, suppose that $y \in C^0(\partial\Omega; \mathbb{R}^d)$ is locally bi-Lipschitz in the sense that

$$l\,|x_1 - x_2| \leq |y(x_1) - y(x_2)| \leq L\,|x_1 - x_2|$$
$$\text{for all } x_1, x_2 \in \partial\Omega \text{ with } |x_1 - x_2| < \varrho, \tag{2.5}$$

with some constants $\varrho, l, L > 0$. Finally, for $s \geq 0$, define the set

$$P_y(s) := \left\{ x \in \partial\Omega \mid \exists \tilde{x} \in \partial\Omega : |y(x) - y(\tilde{x})| \leq s \text{ and } |x - \tilde{x}| > \tfrac{\varrho}{2} \right\}.$$

Then there exist constants $r, R, a, A, \bar{\varepsilon} > 0$ only depending on d, Ω, ϱ, l, L, P and g such that

$$a\mathcal{H}^{d-1}(P_y(r\varepsilon)) \leq \varepsilon^\beta E_\varepsilon^{\partial\Omega}(y) \leq A\mathcal{H}^{d-1}(P_y(R\varepsilon)) \tag{2.6}$$

every $0 < \varepsilon \leq \bar{\varepsilon}$.

Remark 2.2. As a consequence of (2.5),

$$P_y(0) = \{x \in \partial\Omega \mid \#y^{-1}(y(x)) > 1\},$$

i.e., it is precisely the set where $y|_{\partial\Omega}$ fails to be injective.

Proof of Theorem 2.1. In essence, the proof is analogous to the one of [34, Theorem 3.3]. We here present it self-contained for the reader's convenience.

The crucial quantity to study is the (rescaled) density of the penalty term given by

$$J_{y,\varepsilon}(x) := \varepsilon^{d-1} d_{\varepsilon,y}^{\partial\Omega}(x)$$
$$= \int_{\partial\Omega} P\left(g(|\tilde{x} - x|) - g\left(\frac{|y(\tilde{x}) - y(x)|}{\varepsilon}\right)\right) \, d\mathcal{H}^{d-1}(\tilde{x}).$$

Due to the properties of g and P required in (2.2) and (2.4), for given $x \in \partial\Omega$, $J_{y,\varepsilon}(x)$ gives a non-zero contribution if and only if there is a point $\tilde{x}$ such that

$$\varepsilon |\tilde{x} - x| > |y(\tilde{x}) - y(x)|. \tag{2.7}$$

(By (2.5), this automatically means there is a set of positive measure of such points.) Due to the lower bound in (2.5), if ε is small enough, (2.7) never happens locally with $|\tilde{x} - x| < \varrho$.

Lower bound in (2.6)**:** Defining a small enough constant r based on the constants appearing in (2.5), more precisely,

$$r := \min\left\{\tfrac{\varrho}{4}, 1\right\},$$

the lower bound in (2.6) relies on the following observation: If $x \in P_y(r\varepsilon)$ and $\tilde{x}_0$ is an admissible choice in the definition of $P_y(r\varepsilon)$, i.e.,

$$|x - \tilde{x}_0| > \frac{\varrho}{2} \text{ and } |y(x) - y(\tilde{x}_0)| \leq r\varepsilon, \tag{2.8}$$

then (2.7) holds, $|x - \tilde{x}_0| \geq \varrho$ and for $\tilde{x} = \tilde{x}_0$,

$$g(|\tilde{x} - x|) - g\left(\frac{|y(\tilde{x}) - y(x)|}{\varepsilon}\right) \geq c_\varrho > 0, \tag{2.9}$$

with $c_\varrho := g(\varrho) - g\left(\frac{\varrho}{4}\right)$. In fact, to obtain (2.9) it is enough if r is replaced by the bigger constant $\frac{\varrho}{4}$ in (2.8). Combining this with (2.5), we see that for the smaller constant $\tilde{c}_\varrho := g\left(\frac{3}{4}\varrho\right) - g\left(\frac{\varrho}{2}\right)$, the set

$$G(x) = G(x, \tilde{x}_0) := \left\{\tilde{x} \in \partial\Omega \cap B_{\frac{\varrho}{4}}(\tilde{x}_0) \,\Big|\, (2.9) \text{ holds with } \tilde{c}_\varrho\right\}$$

contains all $\tilde{x} \in \partial\Omega$ with $|\tilde{x} - \tilde{x}_0| < \frac{r}{L+1}\varepsilon$ (as long as $\frac{r}{L+1}\varepsilon \leq \varrho$). Since $\partial\Omega$ is Lipschitz, this implies that $G(x)$ has a $(d-1)$-dimensional measure of the order of ε^{d-1} (or more) for small enough ε, independently of x and the choice of $\tilde{x}_0$. By (2.9) and (2.4), we conclude that for all $x \in P_y(r\varepsilon)$, $\varepsilon^{-(d-1)} J_{y,\varepsilon}(x)$ is bounded from below by a fixed constant, which gives the lower bound in (2.6).

Upper bound in (2.6): We exploit that the integrand of $J_{y,\varepsilon}(x)$ is always bounded from above:

$$g(|\tilde{x} - x|) - g\Big(\frac{|y(\tilde{x}) - y(x)|}{\varepsilon}\Big) \leq g(|\tilde{x} - x|) \leq g(\operatorname{diam}\Omega),$$

where $\operatorname{diam}\Omega := \sup_{x_1,x_2\in\Omega} |x_1 - x_2|$.

The integrand of $J_{y,\varepsilon}(x)$ vanishes if (2.7) does not hold, in particular if $\varepsilon < l$ and $|\tilde{x} - x| < \varrho$ due to lower bound in (2.5). Hence, we obtain that

$$J_{y,\varepsilon}(x) \leq C_1 \mathcal{H}^{d-1}(Q_y(\varepsilon, x)) \tag{2.10}$$

with

$$Q_y(\varepsilon, x) := \Big\{ \tilde{x} \in \partial\Omega \,\Big|\, |y(\tilde{x}) - y(x)| < \varepsilon \operatorname{diam}\Omega \text{ and } |\tilde{x} - x| \geq \frac{\varrho}{2} \Big\}$$

and the constant $C_1 := \sup_{0 \leq t \leq g(\operatorname{diam}\Omega)} P(t)$. As another consequence of the lower bound in (2.5), we know that for each $\tilde{x} \in Q_y(\varepsilon, \varrho, x)$,

$$z \notin Q_y(\varepsilon, x) \quad \text{for all } z \in \partial\Omega \text{ with } \varrho \geq |z - \tilde{x}| \geq \frac{2\operatorname{diam}\Omega}{l}\varepsilon.$$

This means that $Q_y(\varepsilon, x)$ can be covered by a finite number (depending only on ϱ and Ω) of balls of radius $\frac{2\operatorname{diam}\Omega}{l}\varepsilon$. Consequently,

$$\mathcal{H}^{d-1}\big(Q_y(\varepsilon, x)\big) \leq C_2 \varepsilon^{d-1},$$

where the constant $C_2 = C_2(\varrho, l, \Omega) > 0$ also compensates the fact the $\partial\Omega$ locally is Lipschitz but not necessarily flat. With $R := \operatorname{diam}\Omega$ and $A := C_1 C_2 > 0$, we conclude that

$$A\mathcal{H}^{d-1}(P_y(R\varepsilon)) = \int_{\{x\in\partial\Omega \,|\, Q_y(\varepsilon,x)\neq\emptyset\}} C_1 C_2 \,\mathrm{d}\mathcal{H}^{d-1}(x)$$

$$\geq \int_{\partial\Omega} \frac{C_1}{\varepsilon^{d-1}} \mathcal{H}^{d-1}\big(Q_y(\varepsilon, x)\big) \,\mathrm{d}\mathcal{H}^{d-1}(x).$$

Together with (2.10), this yields the upper bound in (2.6). $\qquad\square$

The lower bound in (2.6) immediately implies that if a deformation y has bounded energy including the penalty term for all small $\varepsilon > 0$, then $\mathcal{H}^{d-1}(P_y(0)) = 0$. With a more refined argument, we can even obtain that $P_y(0) = \emptyset$, i.e., full invertibility of $y|_{\partial\Omega}$, as long as $\beta > d - 1$:

Corollary 2.3 (Boundary invertibility for finite penalization). *Suppose that the assumptions of Theorem 2.1 hold and let $C > 0$. If $\beta > d - 1$ in (2.3), then there exists a constant $0 < \tilde{\varepsilon} \leq \bar{\varepsilon}$ which only depends on β, C, d, Ω, ϱ, l, L and g, such that for all $\varepsilon < \tilde{\varepsilon}$ and all $y \in C^0(\partial\Omega; \mathbb{R}^d)$ satisfying* (2.5),

$$E_\varepsilon^{\partial\Omega}(y) \leq C \quad \text{implies that } y|_{\partial\Omega} \text{ is injective.} \tag{2.11}$$

Proof. Suppose that $y|_{\partial\Omega}$ is not injective. Then there is a pair of points $x_1, x_2 \in \partial\Omega$ with $x_1 \neq x_2$ and $y(x_1) = y(x_2)$. Since y is locally bi-Lipschitz by (2.5), there are neighborhoods of each of these two points in $\partial\Omega$ which are fully contained in $P_y(r\varepsilon)$, and each neighborhood locally covers $\partial\Omega$ around x_i for a radius of the order of ε (or more). The surface measure of such a neighborhood is therefore of the order of ε^{d-1} (or more). Hence, up to a positive multiplicative constant, $\mathcal{H}^{d-1}(P_y(r\varepsilon)) \geq \varepsilon^{d-1}$. Due to the lower bound in (2.6), this contradicts our given energy bound, the premise of (2.11). $\qquad\square$

Remark 2.4. Unlike the bulk variant (2.1), (2.3) cannot be expected to correctly reproduce global invertibility as a limiting condition as $\varepsilon \to 0$ for all $\beta > 0$. Moreover, on the boundary, it is not helpful to have global invertibility merely a.e.. Roughly speaking, this is related to the same issue which prevents a straightforward modification of the Ciarlet-Nečas condition (1.1) to surface integrals. Such a formally analogous variant of (1.1) on the boundary is given by

$$\int_{\partial\Omega} \left[\det\left(n \otimes n + (\nabla_t y)^\top \nabla_t y \right) \right]^{\frac{1}{2}} \mathrm{d}\mathcal{H}^{d-1}(x) = \mathcal{H}^{d-1}\big(y(\partial\Omega)\big), \quad (2.12)$$

where ∇_t denotes the tangential gradient, i.e., $\nabla_t y = (\nabla y)P_t$ with the orthogonal projection $P_t = \mathbb{1} - n \otimes n$ onto the tangent space orthogonal to the outer normal n. However, (2.12) only rules out self-intersections of the boundary on $(d-1)$-dimensional sets, which is practically useless because generic self-intersection on the boundary happens on sets of dimension $d-2$, sets of measure zero with respect to the surface measure $\mathcal{H}^{d-1}$.

Theorem 2.1 also shows that the penalty term eventually vanishes for deformations without self-contact on the boundary:

Corollary 2.5. *In the situation of Theorem 2.1, let $\varepsilon \leq \bar{\varepsilon}$ and suppose in addition that y is more than a distance of $R\varepsilon$ away from any non-local self-contact, i.e.,*

$$|y(x_1) - y(x_2)| > R\varepsilon \quad \text{for all } x_1, x_2 \in \partial\Omega \text{ with } |x_1 - x_2| > \tfrac{\varrho}{2} \quad (2.13)$$

with $R = \operatorname{diam}\Omega$. Then $E_\varepsilon^{\partial\Omega}(y) = 0$.

Proof. This is a direct consequence of the upper bound in (2.6) and the definition of $P_y(R\varepsilon)$: (2.13) implies that $P_y(R\varepsilon) = \emptyset$. $\qquad\square$

2.2. **On the assumed local bi-Lipschitz property.** For the preceding analysis, deformations are only admissible if they satisfy the local bi-Lipschitz property (2.5) on the boundary. While this does restrict the applicability of our results, (2.5) can be justified in certain realistic scenarios, in fact even the more restrictive (3.7) on all of Ω.

First, as explained below, one can study linear elastic models while imposing (3.6) or (3.7) as a constraint, since both are stable under

pointwise convergence of y. The drawback is that solutions obtained in such a framework are clearly not physical if the constraint is active. However, our theory does not require any particular choice for the local bi-Lipschitz constants l, L, so that we can admit any fixed deformation y which is not too far from the identity in $W^{1,\infty}$. These are the deformations for which a small strain (linear elastic) model is most likely to provide decent approximations. Numerically, both can be checked a posteriori relatively easily as long as boundary self-contact is prevented, see Remark 4.5. Still, it would be nice to weaken (2.5) if possible, because Linear Elasticity can be rigorously justified as a Gamma-limit of suitable nonlinear models where the deformation gradients are required to be close to the identity matrix only in, say, L^2 [19, 1] (not in L^∞ as suggested by (3.6)). For further related results and a discussion of the case without Dirichlet boundary conditions see Remark 3.11.

Secondly, for nonlinear elastic energies resisting extreme compression, another justification of (3.7) can be found if suitable higher order terms are included in the nonlinear elastic energy (non-simple materials): A version of the Inverse Mapping Theorem, summarized in Lemma 2.7 below, then guarantees (3.7). For this argument, the extra regularity of deformations provided by the second gradient term in $E_{\varepsilon,\sigma}$ in Subsection 3.3 below, together with suitable properties of the nonlinear elastic energy density, cf. (3.12) and (3.14), guarantees that in any set of states with bounded energy, $J := \det \nabla y$ is uniformly positive, bounded away from zero by a result of [28]:

Lemma 2.6 (cf. Lemma 4.1 in [28]). *Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain, $\alpha > 0$ and $q \geq d/\alpha$. Then for every $C_1, C_2 > 0$ there exists a constant $\delta = \delta(C_1, C_2, \Omega, \alpha, q, d) > 0$ such that*

$$\int_\Omega J^{-q} \, dx \leq C_1 \quad \text{implies that} \quad J \geq \delta > 0 \text{ on } \Omega$$

for any $J \in C^\alpha(\Omega)$ with $\|J\|_{C^\alpha(\Omega)} \leq C_2$ and $J > 0$ a.e. in Ω.

In particular, if $y \in W^{2,s}$ with $s > d$, we have that $y \in C^{1,\alpha}$ by embedding, where $\alpha = (s - d)/s$. Consequently, $J := \det \nabla y \in C^\alpha$, and we can apply Lemma 2.6 as long as $q \geq d/\alpha = sd/(s-d)$, which is assumed in (3.14). In this way, (2.14) in the following lemma can be justified.

Lemma 2.7 (Lemma 3.6 in [34]). *Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain with local Lipschitz constants bounded by a fixed $L_\Omega > 0$, let $\alpha, \delta, M_1, M_2 > 0$ and let $y \in C^{1,\alpha}(\Omega; \mathbb{R}^d)$ such that*

$$\det \nabla y \geq \delta > 0 \text{ and } |\nabla y| \leq M_1 \text{ on } \Omega \text{ and } \|\nabla y\|_{C^\alpha(\Omega)} \leq M_2. \quad (2.14)$$

Then there exists a $\varrho > 0$ which only depends on δ, M_1, M_2, α and Ω such that for every $\bar{x} \in \mathbb{R}^d$, y is injective on $\Omega(\bar{x}, \frac{\varrho}{2}) := B_{\varrho/2}(\bar{x}) \cap \Omega$.

Moreover, y is bi-Lipschitz on $\Omega(\bar{x}, \frac{\varrho}{2})$ for any $\bar{x}$, i.e., y satisfies (3.7) on Ω (and thus on $\overline{\Omega}$ by continuity), where the constants $l, L > 0$ can be explicitly chosen as

$$l := \frac{1}{2}\frac{\delta}{M_1^{d-1}}, \quad L := M_1\sqrt{L_\Omega^2 + 1}.$$

2.3. **From invertibility on the boundary to invertibility everywhere.** It is not difficult to construct examples of deformations exhibiting interior self-penetration while preserving invertibility at the boundary. However, by a result of [32] generalizing [7] (see also [29] for another recent extension), for topologically simple domains no such example is possible in the class of locally orientation preserving deformations. For the precise statement given in Theorem 2.9 below, we recall the concept of deformations that are *approximately invertible on the boundary* used in [32]:

Definition 2.8 (AIB). Let $\Omega \subset \mathbb{R}^d$ be open and bounded, and $y : \bar{\Omega} \to \mathbb{R}^d$ with $y \in C(\partial\Omega; \mathbb{R}^d)$. We say that y is *approximately invertible on the boundary*, or, shortly, $y \in$ AIB, if there exists a sequence $(\varphi_k)_{k\in\mathbb{N}} \subset C(\partial\Omega; \mathbb{R}^d)$ such that $\varphi_k : \partial\Omega \to \varphi_k(\partial\Omega)$ is invertible for each k and $\varphi_k \to y$ uniformly on $\partial\Omega$.

Theorem 2.9 ([32, Corollary 6.5 and Remark 6.3]). *Let $\Omega \subset \mathbb{R}^d$ be a bounded bounded Lipschitz domain such that $\mathbb{R}^d \setminus \partial\Omega$ has exactly two connected components. If $p > d$ and $y \in W^{1,p}(\Omega; \mathbb{R}^d) \cap$ AIB with $\det\nabla y > 0$ a.e. in Ω, then y satisfies the Ciarlet-Nečas condition (1.1).*

Remark 2.10. If $\Omega \subset \mathbb{R}^d$ is a bounded Lipschitz domain, $p > d$ and $(y_k) \subset W^{1,p}(\Omega; \mathbb{R}^d)$ is a bounded sequence, we automatically have that $y_k \in C(\bar{\Omega}; \mathbb{R}^d)$ by embedding. If, in addition, $\varphi_k := y_k|_{\partial\Omega}$ is invertible for each k, for example due to a bound on the surface penalty term as in Corollary 2.3, then any limit y of a subsequence weakly converging in $W^{1,p}$ belongs to AIB.

Remark 2.11. A variant of Theorem 2.9, only valid for piecewise affine maps on a tetrahedral mesh which are invertible on the boundary, is given by [2, Thm. 3]. An important step in its proof, namely the calculation of the global degree of the deformation map, is not well explained in [2], though. In the more general framework of Theorem 2.9, the additional topological assumption on the reference configuration Ω (M in [2]) plays a crucial role at exactly this point (cf. [32, Thm. 4.2 and Appendix B]), in combination with the orientation preserving property of the full deformation (not just on the boundary). In particular, for the results of [32] it is not enough to have the deformed boundary $y(\partial\Omega)$ ($\Psi(\partial M)$ in [2]) coincide with the boundary of some domain (in fact, the latter does not even imply that the reference configuration is connected). To what extent the restriction to piecewise affine maps can

be exploited to avoid topological assumptions on the reference config-
uration is not clear to us from the proof in [2].

3. Convergence of energies

We will show here that in the limit as the penalty parameter ε converges
to zero, elastic energies augmented with the surface penalty term will
reproduce the original energy with global injectivity added as a con-
straint in form of the Ciarlet-Nečas condition (1.1). As mentioned
before, due to the assumptions of Theorem 2.1, our analysis is lim-
ited to cases where deformations are known to be locally bi-Lipschitz.
We discuss two such scenarios below, first tailored to linear elastic
energies subject to a local constraint (Theorem 3.5), and then to an
unconstrained nonlinear elastic energy containing a regularizing term
of higher order (Theorem 3.9).

3.1. **Auxiliary results: domain shrinking.** In both scenarios, we
will need the following two technical lemmas which were also implicitly
used in [34] and play a similar role in [33]. For the proofs of the
theorems below, the case $\Gamma = \emptyset$ in Lemma 3.1 and Lemma 3.2 is
enough, but we prefer to present a slightly more general version here
which is designed to handle additional Dirichlet boundary conditions
on Γ, cf. Remark 3.10.

Lemma 3.1 (domain shrinking). *Let $\Omega \subset \mathbb{R}^d$ be a bounded Lips-
chitz domain, and $\Gamma \subset \partial\Omega$ a closed set. Then there exist a decreas-
ing sequence of closed sets Γ_j with $\Gamma \subset \Gamma_{j+1} \subset \Gamma_j \subset \partial\Omega$ such that
$\bigcap_{j\in\mathbb{N}} \Gamma_j = \Gamma$ and a sequence of C^∞-diffeomorphisms*

$$\Psi_j : \overline{\Omega} \to \Psi_j(\overline{\Omega}) \subset\subset \Omega \cup \Gamma_j \quad with \quad \Psi_j|_\Gamma = \mathrm{id}\,|_\Gamma$$

*such that as $j \to \infty$, $\Psi_j \to \mathrm{id}$ in $C^m(\overline{\Omega};\mathbb{R}^d)$ for all $m \in \mathbb{N}$. If $\Gamma = \emptyset$,
we may choose $\Gamma_j = \emptyset$, too.*

Lemma 3.2 (composition with domain shrinking is continuous).
*Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain, $k \in \mathbb{N}_0$, $1 \le p < \infty$ and
$f \in W^{k,p}(\Omega;\mathbb{R}^n)$. With the maps Ψ_j of Lemma 3.1, we then have that
$f \circ \Psi_j \to f$ in $W^{k,p}(\Omega;\mathbb{R}^n)$.*

Proof of Lemma 3.1. If Ω is strictly star-shaped with respect to a
point $x_0 \in \Omega$ and $\Gamma = \emptyset$, one may take $\Psi_j(x) := x_0 + \frac{j-1}{j}(x - x_0)$. For
the general case, we combine local constructions near the boundary
using a smooth decomposition of unity: First, choose functions

$$\eta_i \in C^\infty(\mathbb{R}^d;[0,1]) \text{ such that } \eta_i(x) = \begin{cases} 0 & \text{if } x \in \Gamma, \\ 1 & \text{if } \mathrm{dist}\,(x;\Gamma) \ge \frac{1}{i}, \end{cases}$$

(with the understanding that $\eta_i \equiv 1$ if $\Gamma = \emptyset$) and define

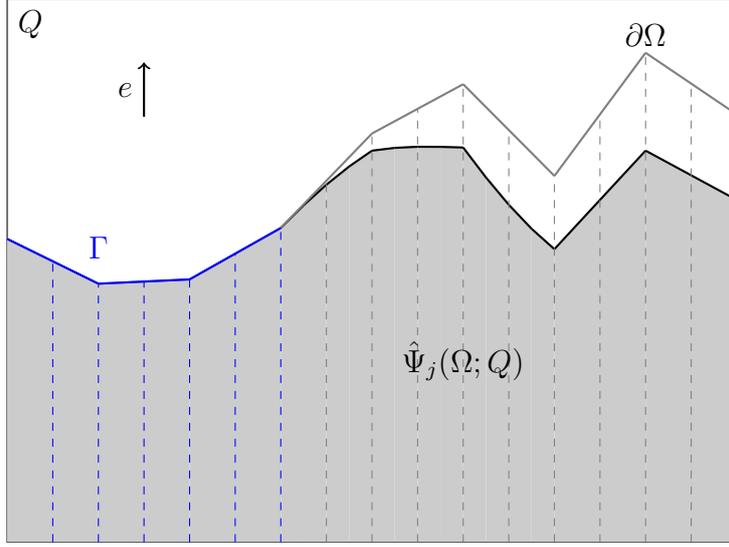$$\Gamma_i := \{x \in \partial\Omega \mid \eta_n(x) = 0 \text{ for } n = 1, \ldots, i\}.$$

FIGURE 2. Locally shrinking Ω with factor $\frac{j-1}{j} = \frac{4}{5}$.

If, locally in some open cube Q, Ω is a Lipschitz subgraph, i.e.,

$$\Omega \cap Q = \{x \in Q \mid x \cdot e < f(x')\} \text{ and } \partial\Omega \cap Q = \{x' + ef(x') \mid x \in Q\},$$

where e is a unit vector orthogonal to one of the faces of Q, $x' := x - (x \cdot e)e$ and f is a real-valued Lipschitz function, we define for $x \in Q$

$$\hat{\Psi}_j(x; Q) := (1 - \eta_{i(j)}(x))x + \eta_{i(j)}(x)\big(x' + \alpha_0 e + \frac{j-1}{j}(e \cdot x - \alpha_0)e\big),$$

$$\text{where } \alpha_0 := \inf_{x \in Q} e \cdot x \text{ and } i(j) \in \mathbb{N} \text{ will be fixed later.}$$

Notice that $\hat{\Psi}_j(\cdot; Q)$ keeps Γ fixed, but as soon as as we are more than $\frac{1}{i(j)}$ away from Γ, it pulls the local boundary piece $\partial\Omega \cap Q$ "down" (in direction $-e$) into the original domain while leaving the "lower" face of Q fixed.

Clearly, $\hat{\Psi}_j(\cdot; Q)$ is of class C^∞ on Q. Moreover, we can choose $i(j) \to \infty$ as $j \to \infty$, slowly enough so that still $\big\|\nabla^m \eta_{i(j)}\big\|_{L^\infty} \leq C(m)j^{\frac{1}{2}}$, for each $j, m \in \mathbb{N}$ with constants $C(m) > 0$ independent of j. Thus $\big|1 - \frac{j-1}{j}\big| \big\|\nabla^m \eta_{i(j)}\big\|_{L^\infty} \leq C(m)j^{-\frac{1}{2}} \to 0$ as $j \to \infty$, for each m, and as a consequence, we obtain that $\hat{\Psi}_j(\cdot; Q) \to \mathrm{id}\,|_Q$ in C^m. Since $\partial\Omega$ can be covered by finitely many such cubes, we can write $\overline{\Omega} \subset Q_0 \cup \bigcup_{k=1}^n Q_k$ with some open interior set $Q_0 \subset\subset \Omega$. For a smooth decomposition of unity $1 = \sum_{k=0}^n \varphi_k$ subordinate to this covering of Ω (i.e., φ_k smooth, non-negative and compactly supported in Q_k),

$$\Psi_j(x) := \varphi_0(x)x + \sum_{k=1}^n \varphi_k(x)\hat{\Psi}_j(x; Q_k)$$

now has the asserted properties, with $\Gamma_j := \Gamma_{i(j)}$. $\qquad\square$

Proof of Lemma 3.2. We only provide a proof for the case $k = 1$, which will include the argument for $k = 0$. For $k \geq 2$, the assertion follows inductively. It suffices to show that as $j \to \infty$, $\partial_n[f \circ \Psi_j - f] \to 0$ in L^p, for each partial derivative ∂_n, $n = 1, \ldots, d$. By the chain rule,

$$\partial_n[f \circ \Psi_j - f] = [(\nabla f) \circ \Psi_j]\partial_n \Psi_j - \partial_n f$$
$$= \Big([(\nabla f) \circ \Psi_j]\partial_n \Psi_j - (\partial_n f) \circ \Psi_j\Big) + \Big((\partial_n f) \circ \Psi_j - \partial_n f\Big) \tag{3.1}$$

The first term above converges to zero in L^p since $(\nabla f)e_n = \partial_n f$ for the n-th unit vector e_n, and $\partial_n \Psi_j \to \partial_n \mathrm{id} = e_n$ uniformly. The convergence of the second term correspond to our assertion for the case $k = 0$, with $\tilde{f} := \partial_n f \in L^p$. It can be proved in the same way as the well-known continuity of the shift in L^p: If $\tilde{f}$ is smooth and can be extended to a smooth function on $\mathbb{R}^d$, we have

$$\|\tilde{f} \circ \Psi_j - \tilde{f}\|_{L^p(\Omega;\mathbb{R}^d)} \leq \|\nabla \tilde{f}\|_{L^\infty(\mathbb{R}^d;\mathbb{R}^{d\times d})} \|\Psi_j - \mathrm{id}\|_{L^p(\Omega;\mathbb{R}^d)} \xrightarrow[j\to\infty]{} 0. \tag{3.2}$$

The general case follows by approximation of $\tilde{f}$ in L^p with such smooth functions, by first extending $\tilde{f}$ by zero to all of $\mathbb{R}^d$, and then mollifying. Here, notice that for the mollified function, $\|\nabla \tilde{f}\|_{L^\infty}$ in (3.2) is unbounded in general as a function of the mollification parameter, but one can always choose the latter to converge slow enough with respect to j so that (3.2) still holds. $\qquad\square$

3.2. **Linear elasticity with a constraint guaranteeing local invertibility.** Linear elasticity, a very popular model due to its computational efficiency, constitutes a standard quadratic approximation of nonlinear elastic energies for the case where ∇y is close to the identity matrix I. Still, this potentially allows y to deviate from the identity quite a lot, depending on the size of the reference domain Ω, so that a global self-contact problem remains meaningful. In this model, we work with an elastic part of the energy of the form

$$E_{\mathrm{lin}}^{\mathrm{el}}(y) = \int_\Omega Q(\nabla y) \, \mathrm{d}x. \tag{3.3}$$

The prototypical example for the density Q is a quadratic function

$$Q(\nabla y) = \mu |e|^2 + (\lambda/2) \mathrm{tr}^2(e) \tag{3.4}$$

of the symmetrized small strain tensor

$$e = \frac{1}{2}[(\nabla y - I)^\top + (\nabla y - I)]$$

with the Lamé material parameters

$$\mu > 0 \quad \text{and} \quad \lambda > -\frac{2}{d}\mu. \tag{3.5}$$

In particular, Q as given in (3.4) is strictly convex in e for precisely the range of λ, μ stated above, and the existence and uniqueness of solutions (minimizers or critical points) is guaranteed for suitable boundary conditions and added linear potentials corresponding to external forces, essentially using Korn's inequality [44]. While Q as defined above corresponds to an isotropic material, we could just as well allow anisotropic cases. For instance, Q can be also be chosen as any symmetric quadratic form which is a strictly convex function of e. In fact, we do not even need that Q is quadratic, (3.10) below suffices.

This model is well understood, see, e.g., [16, Sec. 6.3] for the case $d = 3$, and asymptotically justified as a limit of nonlinear elasticity [19], provided that Dirichlet conditions are imposed on a part of the boundary. Pure traction problems are also possible but more subtle, see Remark 3.11. Note however that such asymptotic results does not automatically provide insight on self-contact problems. In fact, it is not hard to see that a constraint like the Ciarlet-Nečas condition would simply vanish in the passage to a linearized limit model unless the identity (or another rotation that we linearize at) is already in self-contact. Concerning the latter, the only related result so far available which rigorously justifies a linearized contact condition is given in [3] and essentially limited to dimension 2. As we intend to look at scenarios where the deformation is not really asymptotically small, but just has deformation gradients moderately close to the identity, we still aim to reproduce global injectivity as a constraint, without trying to linearize it.

An obvious issue of the model is that by itself, the linear elastic energy $E_{\text{lin}}^{\text{el}}$ does not enforce invertibility of the deformation y, neither locally nor globally. We will therefore augment it with additional constraints that prevent interpenetration of matter at least locally, to make sure that the global self-contact problem remains meaningful and can be tackled with penalization. More precisely, we will admit only locally orientation preserving deformations (i.e., $\det \nabla y > 0$ a.e. in Ω) also satisfying the constraints

$$\begin{aligned}
&l \left| x_2 - x_1 \right|^2 \leq (y(x_2) - y(x_1)) \cdot (x_2 - x_1), \\
&\left| y(x_2) - y(x_1) \right| \leq L \left| x_2 - x_1 \right|, \\
&\text{for a.e. } x_1, x_2 \in \Omega \text{ with } |x_1 - x_2| < \varrho,
\end{aligned} \tag{3.6}$$

where $\varrho, l, L > 0$ are given, fixed constants. In particular, (3.6) implies that y is uniformly locally bi-Lipschitz:

$$\begin{aligned}
&l \left| x_1 - x_2 \right| \leq \left| y(x_1) - y(x_2) \right| \leq L \left| x_1 - x_2 \right| \\
&\quad \text{for a.e. } x_1, x_2 \in \Omega \text{ with } |x_1 - x_2| < \varrho.
\end{aligned} \tag{3.7}$$

Consequently, $|\nabla y| \leq L$ and $|(\nabla y)^{-1}| \leq l^{-1}$ in operator norm, and for locally orientation preserving deformations y, $\det \nabla y \geq l^d > 0$. Our

constraints thus contain the local invertibility constraint studied in [25] as well as the locking constraints of [12]. Such invertibility properties of ∇y by themselves would not suffice for our purposes, however, because they do not provide enough control near the boundary. By contrast, (3.6) also prevents local self-intersections near the boundary that one could otherwise create by closing outer angles. In a nonlinear elastic settings, additional constraints like (3.6) can possibly be avoided as shown in [33], based on the theory of functions of finite distortion [30]. However, it is doubtful whether the approach of [33] would work with our penalty term here.

While further restrictions on l and L are not required for our theory (as long as $L > l$ so that some admissible y exist), for the model it is reasonable to have $l < 1 < L$ and close, while ϱ can be very small. With this, (3.6) implies that ∇y cannot be too far from the identity, with the distance controlled by l and L. From this point of view, one can also consider (3.6) as a way to specify the range for which we consider the linear elastic model an acceptable approximation of the nonlinear case. Besides local invertibility, (3.6) also provides additional regularity of deformations so that the pointwise definition of $\det \nabla y$ is meaningful and the Ciarlet-Nečas condition (1.1) implies global invertibility a.e.. The natural space $W^{1,2}$ would be too weak for these purposes in dimension $d \geq 3$. We refer to [27] for some results with minimal requirements.

The full internal energy with an added nonlocal surface term penalizing self-interpenetration on the boundary now reads

$$E_\varepsilon(y) = \begin{cases} E_{\text{lin}}^{\text{el}}(y) + E_\varepsilon^{\partial\Omega}(y) & \text{if } y \in \mathcal{Y}, \\ +\infty & \text{else}, \end{cases}$$

where for some fixed parameters $\varrho, l, L > 0$,

$$\mathcal{Y} = \mathcal{Y}(\varrho; l, L) := \left\{ y \in W^{1,2}(\Omega; \mathbb{R}^d) \mid (3.6) \text{ holds} \right\}$$

Notice that

$$\mathcal{Y} \subset \text{BiLip}(\Omega; \varrho; l, L) := \{ y \in W^{1,2}(\Omega; \mathbb{R}^d) \mid (3.7) \text{ holds}\},$$

$\mathcal{Y} \subset W^{1,\infty}$ and for all $y \in \mathcal{Y}$, we automatically have that $\det \nabla y \geq l^d$. Below, we will also need the subset of $\mathcal{Y}$ where the constraint (3.6) is inactive in the sense that it holds with slightly stronger constants, i.e,

$$\hat{\mathcal{Y}} = \hat{\mathcal{Y}}(\varrho; l, L) := \bigcup_{t>1} \mathcal{Y}_t, \quad \text{where } \mathcal{Y}_t := \mathcal{Y}(\varrho; lt, L/t) \subset \mathcal{Y}(\varrho; l, L).$$

We claim that as $\varepsilon \to 0$, E_ε converges in a suitable sense to

$$E(y) := \begin{cases} E_{\text{lin}}^{\text{el}}(y) & \text{if } y \in \mathcal{Y} \text{ and (1.1) holds}, \\ +\infty & \text{else}, \end{cases} \tag{3.8}$$

the original linear elastic energy with the Ciarlet-Nečas condition (1.1) added as a constraint.

Remark 3.3. The set of admissible deformations $\mathcal{Y}$ is closed with respect to weak convergence in $W^{1,2}$, see the proof of Theorem 3.5 (i) below. It is therefore compatible with direct methods for the existence of minimizers. In particular, E always has a minimizer obtained as a $W^{1,2}$-weak limit of a minimizing sequence. Here, notice that since $\mathcal{Y}$ is a bounded subset of $W^{1,\infty}$ (possibly up to rigid translations that can be removed, cf. Remark 3.16), neither additional coercivity properties nor Korn's inequality are needed. Moreover, $W^{1,2}$-weak convergence of a sequence in $\mathcal{Y}$ automatically implies $W^{1,p}$-weak convergence for all $1 \leq p < \infty$. Using $p > d$, we see that the Ciarler-Nečas conditon (1.1) is preserved in the limit, too [17].

Remark 3.4. From a purely theoretical perspective, we could also replace $\mathcal{Y}$ with the larger set of orientation preserving maps satisfying the weaker condition (3.7), the only property we really exploit. However, (3.6) is better for numerical results because as we will see, it admits (approximately) conforming finite elements. This seems to be unclear for (3.7), where, as far as we known, it is unknown whether or not a Lavrentiev phenomenon could occur in our context, at least in dimension $d \geq 3$. For $d = 2$, see [21] for a related approximation result.

As before, discrete Galerkin-type approximations can be included by further restricting E_ε. Let $h > 0$ (typically a mesh size) and let Y_h be an associated finite dimensional subspace of $W^{1,2}(\Omega; \mathbb{R}^d)$ such that the approximation error satisfies

$$\mathcal{E}_h(y) := \sup_{\tau > 1} \Big(\inf_{y_h \in Y_h \cap \mathcal{Y}(\varrho; \tilde{l}/\tau, \tilde{L}\tau)} \|y - y_h\|_{W^{1,2}} \Big) \underset{h \to 0}{\longrightarrow} 0 \tag{3.9}$$

for all $\tilde{l}, \tilde{L}$ with $0 < l < \tilde{l} < \tilde{L} \leq L$ and all $y \in \tilde{\mathcal{Y}}(\varrho; \tilde{l}, \tilde{L})$.

In fact, it is enough to have the the above for $(\tilde{l}, \tilde{L})$ in a small neighborhood of (l, L).

The corresponding finite dimensional approximation of E_ε is defined by

$$E_\varepsilon^h(y) := \begin{cases} E_{\text{lin}}^{\text{el}}(y) + E_\varepsilon^{\partial\Omega}(y) & \text{if } y \in Y_h \cap \mathcal{Y}, \\ +\infty & \text{else}, \end{cases}$$

In view of the rather strong restriction (3.7), we can actually work with a much more general class of densities Q, not necessarily quadratic:

$$Q : \mathbb{R}^{d \times d} \to \mathbb{R} \text{ is continuous and quasiconvex.} \tag{3.10}$$

As a consequence of (3.10), $y \mapsto E_{\text{lin}}^{\text{el}}(y) = \int_\Omega Q(x, \nabla y)\,\mathrm{d}x$ is $W^{1,\infty}$-weak* sequentially lower semicontinuous (see [18, Theorem 8.4], e.g.), and also $W^{1,p}$-strongly continuous on $\mathcal{Y}$ (or any other bounded subset

of $W^{1,\infty}$) for arbitrary $1 \leq p < \infty$. Concerning $E_{\mathrm{lin}}^{\mathrm{el}}$, this is all we will need below.

Our main result is the following.

Theorem 3.5. *Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain such that $\mathbb{R}^d \setminus \partial\Omega$ has only two connected components, and let $E_{\mathrm{lin}}^{\mathrm{el}}$ (see (3.3), (3.10)), $E_{\varepsilon}^{\partial\Omega}$ (see (2.3), (2.2), (2.4)), $\mathcal{Y}$, $\hat{\mathcal{Y}}$, E and E_{ε}^{h} be given as above. In addition, assume that (3.9) holds and that the constant β in $E_{\varepsilon}^{\partial\Omega}$ satisfies $\beta > d - 1$. For every sequence $(h(k), \varepsilon(k)) \in (0, \infty)^2$, $k \in \mathbb{N}$, with $h(k) \to 0$ and $\varepsilon(k) \to 0$ as $k \to \infty$, we then have the following properties for all $y \in W^{1,2}(\Omega; \mathbb{R}^d)$:*

(i) For every sequence $y_k \rightharpoonup y$ in $W^{1,2}$ (weakly),

$$\liminf_{k \to \infty} E_{\varepsilon(k)}^{h(k)}(y_k) \geq E(y);$$

$(ii)_1$ if $y \in \hat{\mathcal{Y}}$ or y does not satisfy (1.1), then there exists a sequence $y_k \to y$ in $W^{1,2}$ (strongly) such that

$$\lim_{k \to \infty} E_{\varepsilon(k)}^{h(k)}(y_k) = E(y);$$

$(ii)_2$ if $y \in \mathcal{Y} \setminus \hat{\mathcal{Y}}$ and y satisfies (1.1), then for every $\tau > 1$ there exists a sequence $y_k \in Y_{h(k)}$ with $y_k \to y$ in $W^{1,2}$ (strongly),

$$y_k \in \mathcal{Y}(\varrho, l/\tau, L\tau) \text{ and } E_{\mathrm{lin}}^{\mathrm{el}}(y_k) + E_{\varepsilon}^{\partial\Omega}(y_k) \xrightarrow[k \to \infty]{} E_{\mathrm{lin}}^{\mathrm{el}}(y).$$

This also remains true for the case $h = 0$ if we define $E_{\varepsilon}^{0} := E_{\varepsilon}$, and the assumption (3.9) can be dropped in this case.

Remark 3.6. Finite elements fully conforming with $\mathcal{Y}$, i.e., (3.9) with fixed $\tau = 1$ instead of relaxing the constraint constants using $\tau > 1$, seem to be hard to obtain if not impossible. In Proposition 3.17 at the end of this section, (3.9) is proved for piecewise affine elements on simplicial meshes.

Remark 3.7. Theorem 3.5 $(ii)_2$ is weaker than one might like, because our construction only yields the constraint (3.6) for y_k in slightly relaxed form, with constants modified by a factor $\tau > 1$. As a consequence, it is too weak to ensure finite $E_{\varepsilon}^{h(k)}(y_k)$. On the technical level, we do not know how to handle the scenario where (1.1) holds with self-contact on the surface while the deformation maximally stretches and compresses locally nearby, say, one end forming a slim-necked peg which, after deformation, is maximally compressed and stuck in a superficial, maximally stretched opening somewhere else on the body. In such an example, it is possible that E_{ε} is infinite in a whole neighborhood of y for all $\varepsilon > 0$, and the correct value of the limit functional at such a y should read $E(y) = +\infty$, not $E(y) = E_{\mathrm{lin}}^{\mathrm{el}}(y)$ as we defined E.

Proof of Theorem 3.5. **(i) "Lower bound":** Let $y_k \rightharpoonup y$ be $k \to \infty$, weakly in $W^{1,2}$. Passing to a subsequence if necessary, we may

assume that $e_0 := \liminf E^{h(k)}_{\varepsilon(k)}(y_k)$ is a limit. Moreover, we may assume that $e_0 < \infty$, because otherwise the asserted lower bound is trivial. As a consequence, $y_k \in \mathcal{Y}$ for all sufficiently large k, because $E^{h(k)}_{\varepsilon(k)} = +\infty$ on $W^{1,2} \setminus \mathcal{Y}$. In particular, (y_k) is bounded in $W^{1,\infty}$ and we also have that $y_k \rightharpoonup^* y$ in $W^{1,\infty}$. By compact embedding, again passing to a subsequence if necessary, we have $y_k \to y$ uniformly on $\overline{\Omega}$. In particular, the limit y also satisfies (3.6), whence $y \in \mathcal{Y}$.

With $K := e_0 + 1$, we also have that $E^{h(k)}_{\sigma,\varepsilon(k)}(y_k) \leq K$ for all k sufficiently large. As a consequence of this energy bound and the fact that $E^{\mathrm{el}}_{\mathrm{lin}}$ is bounded from below on $\mathcal{Y}$, $E^{\partial\Omega}_{\varepsilon(k)}(y_k)$ is bounded, so that y_k is invertible on $\partial\Omega$ by Corollary 2.3 for all large enough k. Hence, $y \in \mathrm{AIB}$ (cf. Definition 2.8 and Remark 2.10) and y satisfies the Ciarlet-Nečas condition (1.1) due to Theorem 2.9. As $E^{\mathrm{el}}_{\mathrm{lin}}$ is weakly* sequentially lower semicontinuous in $W^{1,\infty}$, we infer that

$$\liminf E^{h(k)}_{\varepsilon(k)}(y_k) \geq \liminf E^{\mathrm{el}}_{\mathrm{lin}}(y_k) \geq E^{\mathrm{el}}_{\mathrm{lin}}(y) = E(y).$$

(ii)$_1$ recovery sequence for inactive constraint: If (1.1) does not hold, we may choose $y_k \equiv y$. Otherwise, $y : \Omega \to y(\Omega)$ is a homeomorphism, as a locally bi-Lipschitz map satisfying (1.1). To obtain an approximation of y with a controllable contribution in $E^{\partial\Omega}_{\varepsilon(k)}$, we first create a small gap all around the boundary, using the smooth injective maps $\Psi_j : \bar{\Omega} \to \Omega$ close to the identity from Lemma 3.1 which shrink the Lipschitz domain Ω into itself. $y \circ \Psi_j$ fully avoids self-contact: $y \circ \Psi_j(\overline{\Omega})$ is compactly contained in the open set $y(\Omega)$ and $y \circ \Psi_j$ is still locally bi-Lipschitz. In particular, $E^{\partial\Omega}_\varepsilon(y \circ \Psi_j) = 0$ for all $\varepsilon > 0$ small enough by Corollary 2.5. Moreover, $y \in \mathcal{Y}(\varrho; lt, L/t)$ for some $t > 1$, and thus $y \circ \Psi_j \in \hat{\mathcal{Y}}$ for all large enough j because $\Psi_j \to \mathrm{id}$ in C^1. In addition, $y \circ \Psi_j \to y$ in $W^{1,2}$ by Lemma 3.2. For each j, we approximate $y \circ \Psi_j$ with suitable finite elements $y_{j,k} \in Y_{h(k)} \cap \hat{\mathcal{Y}}$ according to (3.9). This leads to a sequence $y_{j,k} \in Y_{h(k)} \cap \mathcal{Y}$ such that as $k \to \infty$, $y_{j,k} \to y \circ \Psi_j$ in $W^{1,2}$, and for all k large enough (depending on j) $E^{\partial\Omega}_{\varepsilon(k)}(y_{j,k}) = 0$. Moreover, $\lim_j \lim_k E^{\mathrm{el}}_{\mathrm{lin}}(y_{j,k}) = \lim_j E^{\mathrm{el}}_{\mathrm{lin}}(y \circ \Psi_j) = E^{\mathrm{el}}_{\mathrm{lin}}(y)$ by continuity of $E^{\mathrm{el}}_{\mathrm{lin}}$ on $\mathcal{Y}$ with respect to the strong topology of $W^{1,2}$. With a suitable diagonal sequence $y_k := y_{j(k),k}$ with $j(k) \to \infty$ (slow enough), $y_k \to y$ in $W^{1,2}$ and $E^{h(k)}_{\varepsilon(k)}(y_k) = E^{\mathrm{el}}_{\mathrm{lin}}(y_k) \to E^{\mathrm{el}}_{\mathrm{lin}}(y)$ as asserted.

(ii)$_2$ recovery sequence with weaker constraint constants: This is fully analogous to $(ii)_1$. We now use (3.9) with $\tilde{l} = l$ and $\tilde{L} = L$. $\square$

3.3. **Nonlinear elasticity with higher order terms.** As an alternative model, we briefly revisit the scenario already studied in [34]. There, artificial constraints like (3.6) or (3.7) are not imposed. Instead, a higher order term is added to a nonlinear elastic energy, leading to a model of a so-called non-simple material. The local bi-Lipschitz

property required for our analysis of the boundary penalty term in Theorem 2.1 will now be obtained as consequence of an energy bound. For $y \in W^{1,p}(\Omega; \mathbb{R}^d)$, consider the penalized energy given by

$$E_{\varepsilon,\sigma}(y) = \begin{cases} E^{el}(y) + E_{\sigma}^{reg}(y) + E_{\varepsilon}^{\partial\Omega}(y) & \text{if } y \in W^{2,s}, \\ +\infty & \text{else,} \end{cases}$$

in the limit as $\varepsilon \to 0$. We will see that it converges to

$$E_{\sigma}(y) = \begin{cases} E^{el}(y) + E_{\sigma}^{reg}(y) & \text{if } y \in W^{2,s} \text{ and } (1.1) \text{ holds,} \\ +\infty & \text{else,} \end{cases}$$

the original energy which includes the Ciarlet-Nečas condition (1.1) as a built-in constraint. Here,

$$E^{el}(y) = \int_{\Omega} W(\nabla y) \, \mathrm{d}x.$$

where

$$W : \mathbb{R}^{d \times d} \to \mathbb{R} \cup \{+\infty\} \text{ is continuous} \qquad (3.11)$$

Moreover, for all $F \in \mathbb{R}^{d \times d}$,

$$\begin{aligned} +\infty &= W(F) & \text{if } \det F \leq 0, \\ +\infty &> W(F) \geq c_1 \left(|F|^p + (\det F)^{-q} \right) - c_2 & \text{if } \det F > 0, \end{aligned} \qquad (3.12)$$

with constants $q > d$ (which is necessary for (3.14) below), $c_1 > 0$ and $c_2 \geq 0$. In addition, we assume that W is polyconvex, i.e.,

$$W(F) = \Psi(m(F)), \text{ with a convex function } \Psi, \qquad (3.13)$$

where $m(F) \in \mathbb{R}^{n(d)}$, $n(d) := \sum_{k=1}^{d} \binom{d}{k}^2$, denotes the collection of all minors of F, i.e., all $k \times k$ sub-determinants with $1 \leq k \leq d$. This means that for $d = 2$, $m(F) = (F, \det F) \in \mathbb{R}^5$ and for $d = 3$, $m(F) = (F, \operatorname{cof} F, \det F) \in \mathbb{R}^{19}$. Here, $\operatorname{cof} F \in \mathbb{R}^{d \times d}$ denotes cofactor matrix so that $F^{-1} = (\operatorname{cof} F)^T (\det F)^{-1}$ whenever F is invertible.

It would also be possible to use a more general W with explicit dependence on x or approximate it with truncated, everywhere finite integrands W_{κ} (which are safer for numerical purposes), subsequently considering the simultaneous limit $(\kappa, \varepsilon) \to (0, 0)$ as in [34]. As our numerical experiment are not conducted in this framework, however, we will not further discuss these generalizations here.

Remark 3.8. Due to [6, 17], E^{el} always has a minimizer y^* in $W^{1,p}(\Omega; \mathbb{R}^d)$ even if partial Dirichlet boundary conditions or compact force terms are added. Like all states with finite energy, it must satisfy $\det \nabla y^* > 0$ a.e. in Ω.

With the penalty proposed here, classical nonlinear elasticity seems to be out of reach for our analysis concerning the approximation of invertibility with surface penalization without additional constraints. We thus further modify the elastic energy by adding a regularizing term. Accordingly, we regularize E^{el} by adding the higher order term

$$E_\sigma^{reg}(y) := \sigma \int_\Omega \left| D^2 y \right|^s \, \mathrm{d}x,$$

with a fixed parameter $\sigma > 0$. Altogether, the exponents are assumed to satisfy

$$p > d, \quad s > d, \quad q > \frac{sd}{s-d}. \tag{3.14}$$

Here, as in [34], the latter ensures that q, s are admissible for the result of [28] summarized in Lemma 2.6, so that we can obtain a uniform lower bound for $\det \nabla y$.

Discrete Galerkin-type approximations can also be included. For that, let $h > 0$ (typically a mesh size) and let Y_h denote an associated finite dimensional subspace of $(W^{2,s} \cap W^{1,p})(\Omega; \mathbb{R}^d)$ (typically $Y_h \subset W^{2,\infty}$) such that for each y, the approximation error $\mathcal{E}(y; h)$ satisfies

$$\mathcal{E}(y; h) := \inf_{y_h \in Y_h} \left(\{ \| y - y_h \|_{W^{2,s} \cap W^{1,p}} \right) \xrightarrow[h \to 0]{} 0. \tag{3.15}$$

The corresponding finite dimensional approximations of $E_{\varepsilon,\sigma}$ are

$$E_{\varepsilon,\sigma}^h(y) := \begin{cases} E^{el}(y) + E_\sigma^{reg}(y) + E_\varepsilon^{\partial\Omega}(y) & \text{if } y \in Y_h, \\ +\infty & \text{else,} \end{cases}$$

For this model, we have convergence of $E_{\varepsilon,\sigma}^h$ to E_σ in the following sense:

Theorem 3.9. *Let $\sigma > 0$ and $\beta > d - 1$ be fixed, let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain such that $\mathbb{R}^d \setminus \partial\Omega$ has only two connected components, and assume that (3.11)–(3.14) hold. For every sequence $(h(k), \varepsilon(k)) \in (0, \infty)^2$, $k \in \mathbb{N}$, with $h(k) \to 0$ and $\varepsilon(k) \to 0$ as $k \to \infty$, we have the following two properties for all $y \in W^{2,s}(\Omega; \mathbb{R}^d)$:*

(i) *For every sequence $y_k \rightharpoonup y$ in $W^{2,s}$ (weakly),*

$$\liminf_{k \to \infty} E_{\varepsilon(k),\sigma}^{h(k)}(y_k) \geq E_\sigma(y);$$

(ii) *there exists a sequence $y_k \to y$ in $W^{2,s}$ (strongly) such that*

$$\lim_{k \to \infty} E_{\varepsilon(k),\sigma}^{h(k)}(y_k) = E_\sigma(y).$$

This also holds for the case $h = 0$, if we define $E_{\varepsilon,\sigma}^0 := E_{\varepsilon,\sigma}$.

Proof. We provide only a proof for the case including Galerkin approximations with $h(k) > 0$, $h(k) \to 0$. The case $h = 0$ is similar and even slightly simpler.

(i) "Lower bound": Let $y_k \rightharpoonup y$ as $k \to \infty$, weakly in $W^{2,s}$. By compact embedding, this implies that $y_k \to y$ strongly in $W^{1,\infty}$. Passing to a suitable subsequence (not relabeled), we may assume that $e_0 := \liminf E_{\sigma,\varepsilon(k)}^{h(k)}(y_k) = \lim E_{\sigma,\varepsilon(n)}^{h(n)}(y_k)$. In addition, we may assume that $e_0 < +\infty$ because otherwise there is nothing to show. With $K := e_0 + 1$, we have $E_{\sigma,\varepsilon(k)}^{h(k)}(y_k) \leq K$ for all k sufficiently large.

As a consequence of the energy bound, (2.14) holds for $y = y_k$. More precisely, the bounds on ∇y_k follow by embedding and the fact that $E_\sigma^{reg}(y_k)$ and $E^{el}(y_k)$ control $\|D^2 y_k\|_{L^s}$ and $\|\nabla y_k\|_{L^p}$, respectively. The lower bound on $J := \det \nabla y_k$ is provided by Lemma 2.6 combined with our assumptions (3.12) and (3.14) on W. Given $\delta > 0$ and M_1 as in (2.14), (3.11) and (3.12) imply that

$$W \text{ is uniformly continuous on } \left\{ F \in \mathbb{R}^{d \times d} \;\middle|\; \begin{matrix} |F| \leq M_1, \\ \det F \geq \delta \end{matrix} \right\}. \qquad (3.16)$$

As $\nabla y_k \to \nabla y$ in L^∞, we infer that

$$E^{el}(y_k) = \int_\Omega W(\nabla y_k) \, \mathrm{d}x \xrightarrow[k\to\infty]{} \int_\Omega W(\nabla y) \, \mathrm{d}x = E^{el}(y). \qquad (3.17)$$

Moreover, by the weak lower semicontinuity of the convex functional E_σ^{reg},

$$\liminf_{k\to\infty} E_\sigma^{reg}(y_k) \geq E_\sigma^{reg}(y). \qquad (3.18)$$

In addition to (3.17) and (3.18), it also trivially holds that $E_{\varepsilon(k)}^{\partial\Omega} \geq 0$. Thus, we conclude that $\liminf E_{\sigma,\varepsilon(k)}^{h(k)}(y_k) \geq E_\sigma(y)$ as asserted, provided that y satisfies the Ciarlet-Nečas condition. The latter follows from Theorem 2.9, Remark 2.10 and Corollary 2.3.

(ii) Existence of a strongly converging recovery sequence: We may assume that $E_\sigma(y) < +\infty$, because otherwise $E_{\varepsilon(k),\sigma}^{h(k)}(y) \to +\infty = E_\sigma(y)$. Hence, y satisfies (1.1). As $s > d$, y is also C^1. Moreover, the fact that $E_\sigma^{reg}(y) + E^{el}(y) < +\infty$ implies that $\inf \det \nabla y > 0$ by Lemma 2.6 (s, $\alpha := \frac{s-d}{s}$ and q are admissible for this result due to (3.14)). We infer that $y(\Omega)$ is open and $y : \Omega \to y(\Omega)$ is a homeomorphism that is locally uniformly bi-Lipschitz by Lemma 2.7. The main remaining difficulty is the possibility that y exhibits self-contact on the boundary. To handle this, we proceed as in the proof of Theorem 3.5 $(ii)_1$. First create a small gap around the boundary, using $y \circ \Psi_j$ with smooth injective maps $\Psi_j : \bar\Omega \to \Omega$ close to the identity from Lemma 3.1 that shrink the Lipschitz domain Ω into itself. The composition $y \circ \Psi_j(\overline{\Omega})$ is compactly contained in the open set $y(\Omega)$, and $y \circ \Psi_j$ is still locally bi-Lipschitz. Moreover, $y \circ \Psi_j \to y$ in $W^{2,s} \cap W^{1,p}$ by Lemma 3.2. Further approximations of $y \circ \Psi_j$ with finite elements in Y_h based on (3.15) can now be made while maintaining a safe distance from self-contact or loss of local invertibility. In particular, since

$\liminf_j \inf \det \nabla(y \circ \Psi_j) > 0$, we avoid the singularity of W and E^{el} be-
haves continuously along our sequence as in (i). We obtain $y_{j,k} \in Y_{h(k)}$
such that as $k \to \infty$, $y_{j,k} \to y \circ \Psi_j$ in $W^{2,s}$, $E^{\partial\Omega}_{\varepsilon(k)}(y_{j,k}) = 0$ for all
sufficiently large k (using Corollary 2.5) and $E^{h(k)}_{\varepsilon(k),\sigma}(y_{j,k}) \to E_\sigma(y \circ \Psi_j)$.
A suitable diagonal sequence $y_k := y_{j(k),k} \to y$ with $j(k) \to \infty$ (slow
enough) now yields the assertion. $\square$

3.4. **Remarks on the theoretical results.**

Remark 3.10 (Generalizations: force terms and boundary conditions).
Both Theorem 3.9 and Theorem 3.5 can be easily generalized by adding
a term to energy which is weakly lower semicontinuous and strongly
continuous in the relevant space, i.e., in $W^{2,s}$ for the former and in
$W^{1,2}$ for the latter. This comprises potentials associated to many typ-
ical force terms, including those used in our numerical experiments of
Section 4. Generalization including boundary conditions, say, a Dirich-
let condition like $y = \mathrm{id}$ on a closed subset Γ_D of $\partial\Omega$, would also make
sense, as long as the boundary condition is compatible with the con-
straints and stays away from self-contact. The proofs of the theorems
can be extended to cover this case: While the lower bound (i) in The-
orem 3.9 and Theorem 3.5 is not affected at all, constructions for (ii)
have to be adjusted to respect additional boundary conditions, but
Lemma 3.1 with $\Gamma := \Gamma_D$ is suitable for this purpose. However, the
second and bigger problem is hidden in the assumption (3.9), i.e., the
density of suitable finite elements, now with a Dirichlet condition on
Γ_D added to the definition of $\mathcal{Y}$. It is not clear if Proposition 3.17 can
be extended to this case.

Remark 3.11 (Pure traction problems). In our numerical experiments,
we do impose a Dirichlet condition on a part of the boundary to avoid
problems with coercivity. If such boundary conditions are completely
dropped, this leads to so-called pure traction problems, where the defor-
mation is nowhere fixed but subject to additional (conservative) body
and surface forces. The rigorous asymptotic derivation of linear from
nonlinear elasticity is subtle in such a scenario [36, 37, 39, 38]: The
forces have to be suitably equilibrated to avoid energies which are not
even bounded from below along rigid translations, and it is not always
obvious which rigid motion is preferred by the forces as the natural
point to linearize at.
The approaches of [36, 37, 38] and [39], respectively, differ in the way
the displacement is defined from a given deformation map y. If we
assume for simplicity that we linearize at the identity, then apart from
rescaling for the small strain limit, the options are either the standard
displacement $u = y - \mathrm{id}$ [36, 37, 39] or the renormalized displacement
$\tilde{u} = R_y^\top(y - (R_y x + c_y))$ with respect to the y-dependent optimal "ref-
erence configuration" $R_y x + c_y$, the rigid motion which minimizes $\tilde{u}$ in

$W^{1,2}$ [39, (1.3)]. The latter can avoid the technical condition of "compatibility" (of forces) [36, (2.25)] in context of compactness (see also [38] for a deeper discussion). In addition, both approaches suggest extending the linear elastic limit model, by introducing admissible states consisting of a pair (u, W_0), the displacement and an antisymmetric matrix W_0 which represents an infinitesimal rotation and enters the energy as a correction. However, global minimizers in the limit models typically can be obtained with $W_0 = 0$, see [36, Corollary 4.2] and [39, p.5], respectively.

For our purposes in Subsection 3.2, this means we have to be careful how we should interpret our linear elastic model as an approximation of a nonlinear elastic pure traction problem with moderate strains. Clearly, we need that the latter was already properly rotated so that the optimal rotation to linearize at is given by the identity. (If compatibility does not hold, the optimal rotation is not uniquely determined by the forces and not visible in the linearized model!) Fortunately, using $u = y - \mathrm{id}$ as before is reasonable even if we choose to follow the point of view of [39], because this does match their construction of the recovery sequence (the proof of the "upper bound" of [39, Thm. 5.2]) when approximating a global minimizer with $W_0 = 0$ (and the optimal rotation normalized to $R_0 = \mathbb{1}$). Be warned that more general forces or additional constraints can potentially further complicate the picture.

Remark 3.12 (Γ-convergence). Combined, (i) and (ii) in Theorem 3.9 are equivalent to Mosco convergence [41] of $E^{h(k)}_{\varepsilon(k),\sigma}$ to E_σ, which is stronger than $\Gamma(W^{2,s}$-weak)-convergence (see, e.g., [20]), since it requires the existence of a *strongly* converging "recovery sequence" in (ii). Among other things, the limit functional is always unique if it exists in the sense that (i) and (ii) hold.

Remark 3.13 (Convergence of discrete minimizers). (i) and (ii) in Theorem 3.9 imply that a sequence of minimizers of $E^{h(k)}_{\varepsilon(k),\sigma}$ always has a subsequence which weakly converges to a minimizer of E_σ. In the linear elastic setting of Theorem 3.5, this is very similar, except that (ii)$_1$ and (ii)$_2$ combined still do not fully cover the borderline case $y \in \mathcal{Y} \setminus \hat{\mathcal{Y}}$ where the local constraint is active. Any $y^* \in \hat{\mathcal{Y}}$ that arises a weak limit in $W^{1,2}$ of a sequence (y_k) of minimizers of $E^{h(k)}_{\varepsilon(k)}$ is automatically a minimizer of E in $\mathcal{Y}$, but we do not know what happens if $y^* \in \mathcal{Y} \setminus \hat{\mathcal{Y}}$.

Remark 3.14 (Possible non-uniqueness of minimizers). When we compare Theorem 3.9 and Theorem 3.5 to classical numerical convergence results, (i) and (ii) in a sense play the role of stability and consistency, respectively. As explained in more detail in the previous remark, we only get (weak) convergence of a sequence of discrete minimizers up to a subsequence, though. In our scenarios, much more cannot be expected because in general, global minimizers do not have to be unique, as for

instance the classical example of the buckling rod shows in the nonlinear elastic setting. Moreover, even in the linear elastic setting, the nonlinear constraints can break uniqueness. One such example was observed in [24] for a local determinant constraint. The global invertibility constraint (1.1) apparently can break uniqueness, too, for instance in our numerical pincer example, where, when pressed enough, the pincers naturally have two symmetric ways of sliding past each other. As far as we know, no analytical results on this kind of nonlocally driven bifurcation scenario are available so far, though.

Remark 3.15 (Further errors due to numerical integration). As defined, $E_{\varepsilon,\sigma}^h$ and E_ε^h are assumed to be exact on their associated finite element space. In practice, additional approximations are usually needed at this point. In order to not break the analysis above, corresponding additional errors terms should converge to zero as $(h,\varepsilon) \to 0$ along any sequence of states with bounded penalized energies. As a rule of the thumb, this forces a scaling regime where the mesh is fine enough with respect to ε, i.e., $h << \varepsilon$.

Remark 3.16 (Existence of penalized minimizers). By standard applications of the direct method, for fixed ε and h, we can always get the existence of minimizers for the penalized energies, both for the nonlinear elastic $E_{\varepsilon,\sigma}$, $E_{\varepsilon,\sigma}^h$ and the linear elastic E_ε, E_ε^h. In particular, the penalty term $E_\varepsilon^{\partial\Omega}$ is continuous in $L^1(\partial\Omega;\mathbb{R}^d)$, a space into which the trace embeds compactly in both of our models (for deformations with bounded energy). Moreover, the constraints of E_ε, E_ε^h built into $\mathcal{Y}$ are stable under weak convergence in $W^{1,2}$ on sets of bounded energy, cf. the proof of Theorem 3.5 (i). We also point out that without additional terms or boundary conditions, all energies are translation invariant: constant vectors can be added to y without changing the energy. In addition, the linear elastic energy is invariant with respect to addition of linear transformations with vanishing symmetric part of the matrix. Nevertheless, coercivity of the energy can be recovered by working in subspaces of $W^{2,s}(\Omega;\mathbb{R}^d)$ or $W^{1,2}(\Omega;\mathbb{R}^d)$ that remove these symmetries, say, by fixing appropriate averages. This extra step is not needed if added constraints, boundary conditions or terms in the energy already fix or control the otherwise free constants.

3.5. **Conforming finite elements for the constrained linear elasticity.** We now show that asymptotically conforming finite elements as assumed in (3.9) for Theorem 3.5 actually exist.

Proposition 3.17. *Suppose that the bounded domain $\Omega \subset \mathbb{R}^d$ has a polygonal boundary and thus can be triangulated[2]. Moreover, let $H \subset (0,\infty)$ with $\inf H = 0$ and let $Y_h \subset W^{1,\infty}(\Omega;\mathbb{R}^d)$, $h \in H$, such*

---

[2]i.e., $\Omega = \bigcup_k \overline{T_k}$ with finitely many simplices T_k with pairwise disjoint interior (triangles for $d = 2$, tetrahedra for $d = 3$)

that each Y_h is the set of all functions which are piecewise affine with respect to a simplicial mesh associated to Y_h, of mesh size at most h and triangulating Ω. Then (3.9) holds.

Proof. To simplify notation, we will write the proof only for the case $l = \tilde{l}$ and $L = \tilde{L}$. Let $y \in \mathcal{Y}(\varrho; l, L)$, fix $\tau > 1$ and abbreviate

$$m := \frac{l}{\tau} < l, \ M := L\tau > L.$$

We suffices to show that for every $\varepsilon > 0$ and all $h \in H$ sufficiently small (depending on ε and y), there exists $y_h \in Y_h \cap \mathcal{Y}(\varrho; m, M)$ such that

$$\|y - y_h\|_{W^{1,2}(\Omega;\mathbb{R}^d)} < 2\varepsilon. \tag{3.19}$$

Here, recall that $\mathcal{Y}(\varrho; m, M)$ is the set of all $y \in W^{1,2}(\Omega;\mathbb{R}^d)$ such that $\det \nabla y > 0$ a.e. and (3.6) holds with m, M instead of l, L. The latter requires that y is ϱ-locally Lipschitz with constant M and satisfies the angle condition

$$(y(x_1) - y(x_2)) \cdot (x_1 - x_2) \geq m |x_1 - x_2|^2$$
$$\text{for all } x_1, x_2 \in \Omega \text{ with } |x_1 - x_2| < \varrho. \tag{3.20}$$

The approximating maps are constructed in two consecutive steps, mollification and interpolation. The former is a bit more subtle than usual, because we lack extension results that would preserve (3.20). So we use domain shrinking instead to avoid troubles near the boundary.

Step 1: Mollification.
Choose a family of standard mollifying kernels φ_r, $r > 0$, i.e., $\varphi_r \in C_c^\infty(B_r(0); [0, \infty))$, $\varphi_r(rz) = r^{-d}\varphi_1(z)$ with $\int \varphi_1 = 1$, let $*$ the convolution operator. With the domain shrinking maps Ψ_j of Lemma 3.1 and Lemma 3.2, $(\varphi_r * y) \circ \Psi_j$ is well defined on all of Ω for r small enough, more precisely, $r < \text{Dist}(\Psi_j(\Omega), \partial\Omega)$.
We define
$$\hat{y} = \hat{y}_{r,j} := (\varphi_r * y) \circ \Psi_j \in C^2(\overline{\Omega}; \mathbb{R}^d),$$
We claim that for suitable r (small enough) and j (big enough), we have that

$$\|y - \hat{y}\|_{W^{1,2}(\Omega;\mathbb{R}^d)} < \varepsilon \quad \text{and} \quad \hat{y} \in \mathcal{Y}(\varrho; \hat{m}, \hat{M}), \tag{3.21}$$

where

$$\hat{\tau} := 1 + \frac{\tau - 1}{2}, \ \hat{m} := \frac{l}{\hat{\tau}} \text{ and } \hat{M} := L\hat{\tau}.$$

Notice that $l > \hat{m} > m$ and $L < \hat{M} < M$.

The bound for $\|y - \hat{y}\|_{W^{1,2}}$ in (3.21) follow for large enough j from the properties of Ψ_j obtained in Lemma 3.1 and Lemma 3.2. It remains to show that $\hat{y} \in \mathcal{Y}(\varrho; \hat{m}, \hat{M})$, i.e., that the upper and lower bounds of (3.6) hold for $\hat{y}$ with the constants $\hat{m}, \hat{M}$ instead of l, L, and that $\det \nabla \hat{y} > 0$ a.e..

Upper bound of (3.6): Since y is ϱ-locally L-Lipschitz and $\hat{M} > L$, $\hat{y}$ is ϱ-locally $\tilde{M}$-Lipschitz as long as r and j are small and big enough, respectively.

Lower bound of (3.6): We have to show that

$$(\hat{y}(x_1) - \hat{y}(x_2)) \cdot (x_1 - x_2) \geq \hat{m} \left| x_1 - x_2 \right|^2 \tag{3.22}$$
$$\text{for all } x_1, x_2 \in \Omega \text{ with } |x_1 - x_2| < \varrho.$$

In view of the definition of $\hat{y}$, this amounts to

$$[(\varphi_r * y)(\Psi_j(x_1)) - (\varphi_r * y)(\Psi_j(x_2))] \cdot (x_1 - x_2) \geq \hat{m} \left| \xi_1 - \xi_2 \right|^2 . \tag{3.23}$$

for $x_1, x_2 \in \Omega$ with $|x_1 - x_2| < \varrho$. For every $\xi_1, \xi_2 \in \Omega$ with $\operatorname{dist}(\xi_i; \partial\Omega) \geq r$ and $|\xi_1 - \xi_2| < \varrho$, (3.20) implies that

$$\begin{aligned}
&[(\varphi_r * y)(\xi_1) - (\varphi_r * y)(\xi_2)] \cdot (\xi_1 - \xi_2) \\
&= \int_{B_r(0)} \varphi_r(z) \big(y(z + \xi_1) - y(z + \xi_2) \big) \cdot (\xi_1 + z - (\xi_2 + z)) \, \mathrm{d}z \\
&\geq l \int_{B_r(0)} \varphi_r(z) \left| \xi_1 + z - (\xi_2 + z) \right|^2 \\
&= l \left| \xi_1 - \xi_2 \right|^2 \, \mathrm{d}z
\end{aligned} \tag{3.24}$$

We now set $\xi_i = \xi_i(j) := \Psi_j(x_i)$, $i = 1, 2$, in (3.24). In addition, we can replace all differences $\xi_1 - \xi_2$ occurring in (3.24) by $x_1 - x_2$ with small enough error to obtain (3.23). Here, the gap between l and the smaller $\hat{m}$ can be used to absorb the error for big enough j, since

$$\left| \xi_1(j) - \xi_2(j) - (x_1 - x_2) \right| \leq \operatorname{Lip}(\Psi_j - \operatorname{id}) \left| x_1 - x_2 \right| . \tag{3.25}$$

Above, $\operatorname{Lip}(\Psi_j - \operatorname{id})$ denotes the global Lipschitz constant of $\Psi_j - \operatorname{id}$ which converges to zero as $j \to \infty$.

det $\nabla\hat{y} > 0$ a.e.: We know that $\det \nabla\hat{y} \geq \hat{m}^d > 0$ in Ω. as a consequence of (3.22).

Step 2: Piecewise affine interpolation.
For a simplex $T = \operatorname{co}\{z_i \mid i = 0, \ldots, d\} \subset \mathbb{R}^d$ and a function $f : \overline{\Omega} \to \mathbb{R}^d$, we define the affine interpolation $I[f]$ of f on T as the unique affine function coinciding with f on all $(d+1)$ corners z_i of T. For any given triangulation of Ω into simplices, $I[f]$ is defined piecewise on each simplex of the triangulation, which gives a continuous, piecewise affine function on $\overline{\Omega}$.

Since the function $\hat{y}$ obtained in the previous step is of class C^2 up to the boundary and $\operatorname{diam} T \leq h$ for each simplex T, it is easy to see that

$$\| I_h[\hat{y}] - \hat{y} \|_{C^1} \leq h \| \hat{y} \|_{C^2} \xrightarrow[h \to 0]{} 0, \tag{3.26}$$

where I_h denotes the piecewise affine interpolation with respect to the triangulation associated to Y_h.

In addition, I_h preserves the local Lipschitz constant of $\hat{y}$, possibly up to a small change of the constant: For x_1, x_2 inf Ω with $|x_1 - x_2| < \varrho$, choose simplices $T_i \ni x_i$ of the triangulation of Ω associated to Y_h. If $\mathrm{diam}(T_1 \cup T_2) < \varrho$, the bound

$$|I_h[\hat{y}](x_1) - I_h[\hat{y}](x_2)| \leq (\hat{M} + h)\,|x_1 - x_2|$$

is inherited from $\hat{y}$ using (3.26). Otherwise, there exists at least two nodes z_1, z_2 of the grid (corners of some simplex belonging to the triangulation, so that $I_h[\hat{y}](z_i) = \hat{y}(z_i)$) such that $|z_1 - z_2| < \varrho$ and $|z_i - x_i| \leq 2h$, $i = 1, 2$. Assuming that $2h < \varrho$, by the triangle inequality we see that

$$|I_h[\hat{y}](x_1) - I_h[\hat{y}](x_1)|$$
$$\leq |I_h[\hat{y}](x_1) - I_h[\hat{y}](z_1)| + |I_h[\hat{y}](x_2) - I_h[\hat{y}](z_2)| + |\hat{y}(z_1) - \hat{y}(z_2)|$$
$$\leq 4\hat{M}h + \hat{M}\,|x_1 - x_2| \leq \left(\frac{4\hat{M}h}{\varrho - 2h} + \hat{M}\right)|x_1 - x_2|.$$

Here, for the last inequality, we used that this scenario can only occur if $|x_1 - x_2| \geq \varrho - 2h$.

Next, we show that (3.20) holds for $I_h[\hat{y}]$ instead of y. Let $x_1, x_2 \in \Omega$ with $|x_1 - x_2| < \varrho$, contained in simplices T_1, T_2 of the triangulation of Y_h. By (3.26), we in particular have that

$$\|\nabla I_h[\hat{y}] - \nabla \hat{y}\|_{L^\infty} \leq h\,\|D^2\hat{y}\|_{L^\infty}, \qquad (3.27)$$

and consequently, by the mean value theorem,

$$|I_h[\hat{y}](x_1) - I_h[\hat{y}](x_2) - (\hat{y}(x_1) - \hat{y}(x_2))| \leq 2h|x_1 - x_2|\,\|D^2\hat{y}\|_{L^\infty}. \tag{3.28}$$

With (3.28), (3.20) for $I_h[\hat{y}]$ thus follows from (3.22) if h is small enough. Finally, (3.20) for $I_h[\hat{y}]$ implies a positive sign of $\det \nabla I_h[\hat{y}]$ as in Step 1. $\qquad\square$

4. Numerical experiments

4.1. **A few useful explicit formulas.**

Example 4.1 (A test case to check implementations of $E_\varepsilon^{\partial\Omega}$)**.** We replace $\partial\Omega$ by two line segments in $\mathbb{R}^2$ (no longer the boundary of a domain, but this is irrelevant for the computation):

$$"\partial\Omega" := ([0,1] \times \{1\}) \cup ([0,1] \times \{0\}) \subset \mathbb{R}^2,$$

Two parameters $a, b \geq 0$ determine the deformation y we are interested in: a causes a horizontal shift of the upper line segment, b the distance of the lines after deformation. For $x = (x_1, x_2) \in "\partial\Omega"$, we set

$$y(x) = (y_1(x), y_2(x)) := (x_1, bx_2) + \begin{cases} (a, 0) & \text{if } x_2 = 1, \\ (0, 0) & \text{if } x_2 = 0. \end{cases}$$

For $\varepsilon < 1$, $E_\varepsilon^{\partial\Omega}(y)$ can now be calculated as follows:

$$\varepsilon^{\beta+d-1} E_\varepsilon^{\partial\Omega}(y) = \int_0^1 \int_0^1 H_\varepsilon(t, t-s) \, \mathrm{d}s \, \mathrm{d}t$$

$$= \int_0^1 \int_{\max\{t-1,\gamma-\delta\}}^{\min\{t,\gamma+\delta\}} H_\varepsilon(t, r) \, \mathrm{d}r \, \mathrm{d}t.$$

Here, we used the integrand

$$H_\varepsilon(t, r) := P\left(g\big(\sqrt{r^2+1}\big) - g\Big(\frac{1}{\varepsilon}\sqrt{b^2 + (r-a)^2}\Big) \right)$$

and the associated constants

$$\gamma := \frac{a}{\sqrt{1-\varepsilon^2}}, \quad \delta := \begin{cases} \sqrt{\dfrac{\varepsilon^2 - b^2}{1-\varepsilon^2}} & \text{if } \varepsilon \geq b, \\ 0 & \text{else.} \end{cases}$$

Notice that γ and δ were defined in such a way that

$$H_\varepsilon(t, r) > 0 \quad \text{if and only if} \quad \gamma - \delta < r < \gamma + \delta,$$

(provided that $P(\tau) > 0$ if and only if $\tau > 0$, in addition to (2.2) and (2.4)). For instance, in the special case $P(\cdot) = [\cdot]^+$ and $g = \mathrm{id}$, we get that

$$\varepsilon^{\beta+d-1} E_\varepsilon^{\partial\Omega}(y) = \int_0^1 \int_{\max\{t-1,\gamma-\delta\}}^{\min\{t,\gamma+\delta\}} \left(\sqrt{r^2+1} - \frac{1}{\varepsilon}\sqrt{b^2 + (r-a)^2} \right) \mathrm{d}r \, \mathrm{d}t,$$

which now an be evaluated using standard software like Mathematica.

4.2. **Simulations in 3D for linear elasticity.** We are motivated by 2D simulations of [34] assuming the bulk version of penalization (2.1) and perform 3D energy minimization evaluations with surface penalty (2.3). Therefore, we consider $d = 3$ and the approximate deformation

$$y = (y_1, y_2, y_3) \in W^{1,s}(\Omega; \mathbb{R}^3)$$

is searched for as the (ideally global) minimizer of the

$$E_{\varepsilon,\mu}(y) = E_{\mathrm{lin}}^{\mathrm{el}}(y) - E^{body}(y) + \mu_{\partial\Omega} E_\varepsilon^{\partial\Omega}(y) \tag{4.1}$$

over a finite-dimensional space. We stay within the linear elastic model of Subsection 3.2, with the elastic energy (3.3). The local constraint (3.6) will not be enforced during the computation. In fact, with the forces and boundary conditions we use, there is no real incentive to violate (3.6) for, say, $l = \frac{1}{2}$ and $L = 2$ (and some small $\varrho > 0$). In any case, (3.6) could still be checked a posteriori.

The three components of the deformation y_1, y_2, y_3 are discretized using the finite element method in the space of P_1 tetrahedral elements, i.e., linear and globally continuous functions.

4.2.1. *Implementation details.* We assume the following sets of nodes of the tetrahedral mesh:

$\mathcal{N}$ - the set of all nodes,

$\mathcal{N}_D$ - the set of nodes corresponding to the Dirichlet boundary condition,

$\mathcal{N}_{NP}$ - the subset of surface nodes expected to contribute to the nonpenetration penalty term $E_\varepsilon^{\partial\Omega}(y)$.

The set $\mathcal{N}_{NP}$ is defined a priori in computations and it holds

$$|\mathcal{N}_{NP}| \ll |\mathcal{N}|.$$

The assumption of linear elasticity density (3.4) allows an efficient evaluation of energies $E_{\text{lin}}^{\text{el}}(y) - E^{body}(y)$. Instead of displacement y, we work with displacement $u(x) = y(x) - x, x \in \Omega$ approximated by the finite element method as

$$u(x) \approx u_h(x) = \sum_{j=1}^{3|\mathcal{N}|} (\mathbf{u})_j\, \phi_j(x)$$

with a vector $\mathbf{u} \in \mathbb{R}^{3|\mathcal{N}|}$ of degrees of freedom and finite element basis functions $\phi_j, j = 1, \ldots, 3|\mathcal{N}|$. Then we have approximated

$$E_{\text{lin}}^{\text{el}}(y) - E^{body}(y) \approx \frac{1}{2}\mathbf{u}^T K \mathbf{u} - \mathbf{b}^T \mathbf{u}, \tag{4.2}$$

where $K \in \mathbb{R}^{3|\mathcal{N}| \times 3|\mathcal{N}|}$ is a stiffness matrix and $\mathbf{b} \in \mathbb{R}^{3|\mathcal{N}|}$ is a loading vector. The vector $\mathbf{u}$ is further decomposed into three disjoint parts $\mathbf{u} = (\mathbf{u}_D, \mathbf{u}_{NP}, \mathbf{u}_R)$, where

$\mathbf{u}_D$ is the vector of (prescribed) displacements in nodes $\mathcal{N}_D$,

$\mathbf{u}_{NP}$ is the vector of displacements in nodes $\mathcal{N}_{NP}$,

$\mathbf{u}_R$ is the vector of displacements in the remaining nodes.

If $\mathbf{u}_D = 0$, the optimality conditions for $\frac{1}{2}\mathbf{u}^T K \mathbf{u} - \mathbf{b}^T \mathbf{u} \to \min$ yield a relation between the vectors $\mathbf{u}_{NP}$ and $\mathbf{u}_R$ in the form of the linear system of equations

$$\begin{pmatrix} K_{NP,NP} & K_{NP,R} \\ K_{R,NP} & K_{R,R} \end{pmatrix} \begin{pmatrix} \mathbf{u}_{NP} \\ \mathbf{u}_R \end{pmatrix} = \begin{pmatrix} \mathbf{b}_{NP} \\ \mathbf{b}_R \end{pmatrix} \tag{4.3}$$

with K and $\mathbf{b}$ decomposed to its parts given above. Then the well-known Schur complement technique of linear algebra provides

$$\mathbf{u}_R = K_{R,R}^{-1}(\mathbf{b}_R - K_{R,NP}\,\mathbf{u}_{NP}). \tag{4.4}$$

The relation (4.4), together with the Dirichlet condition $\mathbf{u}_D = 0$, allows us to express the quadratic energy (4.2) in reduced form featuring the vector $\mathbf{u}_{NP}$ as the only argument:

$$\hat{E}(\mathbf{u}_{NP}) = \frac{1}{2}\mathbf{u}_{NP}^T S \mathbf{u}_{NP} - \hat{\mathbf{b}}^T \mathbf{u}_{NP} - \hat{c}, \tag{4.5}$$

with the quadratic, linear and constant contributions given by

$$\begin{aligned}
S &:= K_{NP,NP} - K_{R,NP}^T K_{R,R}^{-1} K_{R,NP}, \\
\hat{\mathbf{b}}^T &:= \mathbf{b}_{NP}^T - \mathbf{b}_R^T K_{R,R}^{-1} K_{R,NP} + \mathbf{b}_R^T K_{R,NP}, \\
\hat{c} &:= \frac{1}{2} \mathbf{b}_R^T K_{R,R}^{-1} \mathbf{b}_R.
\end{aligned}$$

Here, we used that $K_{R,NP} = K_{NP,R}^T$ and $K_{R,R}^{-1} = K_{R,R}^{-T}$ due to the symmetry of K. When looking for minimizers or critical points, the constant contribution $\hat{c}$ can of course be ignored.

As long as $\mathcal{N}_{NP}$ has been chosen well, all surface nodes outside of $\mathcal{N}_{NP}$ remain far from any self-contact and thus contribute nothing to the nonpenetration penalty term, so that $E_\varepsilon^{\partial\Omega}(y)$ also only depends $\mathbf{u}_{NP}$, say,

$$E_\varepsilon^{\partial\Omega}(y) = \hat{E}_\varepsilon^{\partial\Omega}(\mathbf{u}_{NP}).$$

The total energy (4.1) then can be likewise expressed as a function of $\mathbf{u}_{NP} \in \mathbb{R}^{3|\mathcal{N}_{NP}|}$ only. This reduces the number of degrees of freedom from $3(|\mathcal{N}| - |\mathcal{N}_D|)$ to $3|\mathcal{N}_{NP}|$ at the cost of precomputations of the terms in (4.5).

Besides this reduction, the matrix S appearing in the leading part of the reduced elastic and potential energy $\hat{E}$ offers us an easy way of preconditioning the problem. To this end, we use the following transformation of the reduced total energy in the computations:

$$E_{\text{precond}}(\mathbf{v}_{NP}) := \hat{E}\big(S^{-\frac{1}{2}}\mathbf{v}_{NP}\big) + \hat{E}_\varepsilon^{\partial\Omega}\big(S^{-\frac{1}{2}}\mathbf{v}_{NP}\big).$$

Thus we have replaced the original variable $\mathbf{u}_{NP}$ by $\mathbf{v}_{NP} = S^{\frac{1}{2}}\mathbf{u}_{NP}$. Here, recall that as a symmetric positive definite matrix (corresponding to the coercivity of the linear elastic energy), S has a well-defined square root that for instance can be computed by Cholesky decomposition.

Remark 4.2. Heuristically, our preconditioning means that we choose L^2 as the natural energy space for the piecewise affine boundary elements represented by the variable $\mathbf{v}_{NP}$ in our computations. By contrast, the natural space for (piecewise affine interpolations of) $\mathbf{u}_{NP}$ would be the trace space $H^{\frac{1}{2}}$ on the boundary piece of Ω corresponding to the nodes in $\mathcal{N}_{NP}$. Here, notice that formally, S is a discretization of a symmetric and invertible nonlocal pseudo-differential operator of order 1 defining a continuous and coercive bilinear form on $H^{\frac{1}{2}}$. Its condition number and largest eigenvalue are expected to scale like h^{-1} with respect to the grid size h.

Remark 4.3. The interaction of the preconditioning with the penalty energy $\hat{E}_\varepsilon^{\partial\Omega}$ is not so obvious. In that regard, our choice is vaguely inspired by the tests conducted in [50, Fig. 10] (for a quite different nonlocal interaction requiring higher regularity, though), where the use

of H^s-gradient flows turned out to be favorable. In our notation, this corresponds to preconditioning by a (discretized) fractional Laplacian of order $\frac{s}{2}$. However, unlike in [50], our nonlocal penalty term is well-defined even in L^2 and so does not suggest a natural choice for s. For this reason, we proceeded with the precondition best matching the "leading" term with respect to local regularity, i.e., the linear elasticity energy. This suggests the use of the square root of the matrix S. As its inverse has a smoothening effect, it is reasonable to expect that it stabilizes the contributions of $\hat{E}_\varepsilon^{\partial\Omega}$ as well.

4.2.2. *Computational benchmark.* We assume a pincer-shaped domain $\Omega = \Omega_1 \cup \Omega_2 \cup \Omega_3$ which consists of block subdomains

$$\Omega_1 = (0, 6) \times (0, 0.5) \times (2.5, 3), \qquad \text{(upper part)}$$
$$\Omega_2 = (0, 0.5) \times (0, 0.5) \times (0.5, 2.5), \qquad \text{(middle part)}$$
$$\Omega_3 = (0, 6) \times (0, 0.5) \times (0, 0.5), \qquad \text{(lower part)}$$
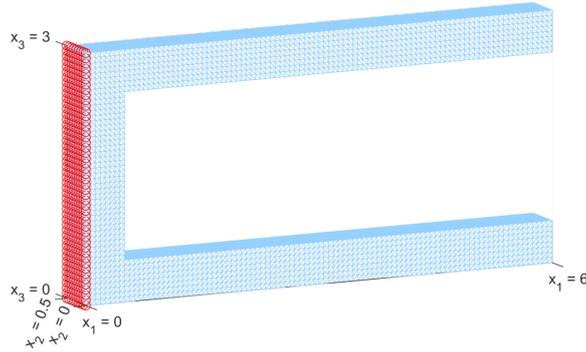
(A) A nondeformed mesh.
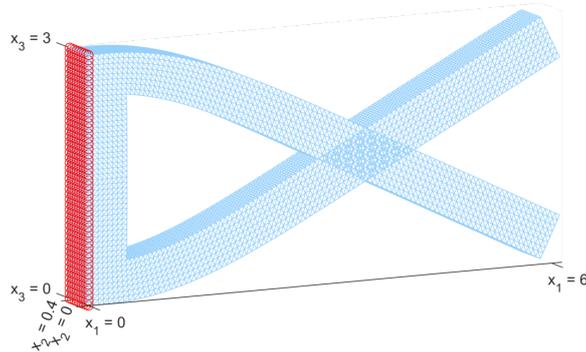
(B) An elastically deformed mesh ($\mu = 0$).

FIGURE 3. Nondeformed and elastically deformed meshes (level 3) are discretized by tetrahedral elements. The red nodes indicate Dirichlet nodes $\mathcal{N}_D$ at which zero displacements are required.

The domain Ω is depicted in Figure 3a along with its tetrahedral triangulation and nodes corresponding to the homogeneous Dirichlet boundary conditions defined on a part of the domain boundary.

$$y = 0 \quad \text{for } x \in \Lambda_D = \{1\} \times (0, 0.5) \times (0.5, 2.5).$$

There is a linear body force term $E^{body}(y)$ with the energy contribution

$$E^{body}(y) := \int_\Omega g_{\text{body}}(x) \cdot (y(x) - x)\, \mathrm{d}x$$

and a linear body force density

$$g_{\text{body}}(x_1, x_2, x_3) = g_{\text{load}}(0, 0, -H(x_1 - 2)H(x_1 - 4)\,\text{sign}(x_3 - 1.5))$$

on $(x_1, x_2, x_3) \in \Omega$, where $H(\cdot)$ denotes the Heaviside step function, $\text{sign}(\cdot)$ the signum function and $g_{\text{load}} > 0$ is a given loading parameter. This form of the linear body force density presses the tips of both pincer parts Ω_1 and Ω_3 against each other.

The surface penalty term $E_\varepsilon^{\partial\Omega}(y)$ is taken according to (2.3) and assumes the choice

$$\beta = d - 1 + 0.1 = 2.1, \qquad \varepsilon = s\,h/r,$$

where $s = 3, r = 2$ and h is the grid size. As to P and g, we use C^2 functions with a fixed parameter $a := 0.01$ given by

$$P(t) = g(t) := \begin{cases} 0 & \text{if } t < 0, \\ \frac{1}{a^2}t^3 - \frac{1}{2a^3}t^4 & \text{if } 0 \leq t \leq a, \\ t - \frac{a}{2} & \text{if } t > a. \end{cases} \qquad (4.6)$$

For computational simulations, we consider the elastic material parameters - Young's modulus and Poisson's ratio

$$E = 2e8, \quad \nu = 0.3$$

corresponding to the Lamé's parameters

$$\lambda = E\nu/(1+\nu)(1-2\nu) \approx 1.15e08, \quad \mu = E/(2(1+\nu)) \approx 7.69e07$$

and the linear body force defined by $g_{\text{load}} = 4e05$. The solution of the purely linear elastic system with the nonpenetration penalty switched off (i.e., $\mu = 0$) leads to the body interpenetration, see Figure 3b. In order to prevent the interpenetration, the penalty term $\mu E_\varepsilon^{\partial\Omega}(y)$ must be switched on and we set $\mu_{\partial\Omega} = 0.001\, E$.

Remark 4.4. In our example, $r = 2$ is precisely the distance of Ω_1 and Ω_3, the distance in reference configuration of the two inner pincer surfaces on opposing sides which will approach self-contact under the influence of the body force. As long as local deformations remain close to the identity, the above choice for ε corresponds to an effective range of the penalty term of $s = 3$ grid boxes on these contact surfaces. Here, recall that the penalty term is active at any pair of points x_1, x_2

iff $|y(x_1)-y(x_2)| < \varepsilon|x_1-x_2|$, and $|x_1-x_2| \approx r$ for the relevant material points where we expect self-contact in our example.

Remark 4.5. We do not enforce the local constraint (3.6) numerically, and this also means that the local bi-Lischitz property (3.7) and the local invertibility of deformations as required by our theoretical results are *not* guaranteed. However, we can check (3.7) a posteriori. In fact, it suffices to check the global behavior on the boundary and local invertibility on each element: For any $y \in W^{1,\infty}$ such that y is invertible on $\partial\Omega$ and $\det \nabla y > 0$ a.e., y is already a homeomorphism on $\overline{\Omega}$, as long as Ω is a Lipschitz domain without holes (i.e., $\mathbb{R}^d \setminus \partial\Omega$ has only two connected components) [32]. If, in addition, y is also continuous and piecewise affine (e.g.), (3.7) holds (with suitable constants) if and only if $|\nabla y|, |(\nabla y)^{-1}| \in L^\infty$, where $|\cdot|$ denotes the operator norm (modulus of the largest singular value, which never exceeded 1.2 in our computations). Here, in the interior, one could actually choose $l^{-1} = \sup |\nabla y^{-1}|$, but smaller l can still occur due to boundary effects if the deformation reduces angles from the outside between boundary elements. Notice that if y is piecewise affine and invertible on the boundary, the formation of outer cusps in the deformed configuration (which would break (3.7)) is impossible.

Depending on different initial deformations in the minimization of the total energy (4.1) we discuss two numerical solutions. The first is symmetric and probably corresponds to a local minimum. The second is asymmetric, and likely approximates one of two global minima related to each other by reflection.

4.2.3. *Symmetric initial deformation.* The elastic deformation y_{elast} is evaluated by the solution of the purely linear elastic system and the initial deformation y_{init} is generated as

$$y_{\text{init}}(x) = x + 0.05(y_{\text{elast}}(x) - x), \qquad x \in \Omega.$$

The constant 0.05 is chosen so that the initial deformation is out of self-contact, see the left part of Figure 4a. The full minimization of (4.1) respecting the nonpenetration term $\mu =$ and taking $y_{\text{init}}(x)$ is the minimization procedure converges to the deformation displayed with linear elasticity and non-penetration density and its zoomed view in the left column of Figure 4b, 4c, 4d. The detailed evaluation data is given in Table 2.
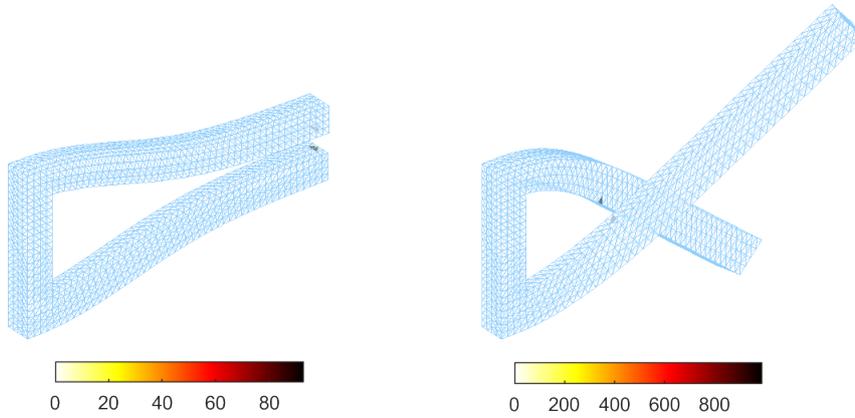
4.2.4. *Asymmetric initial deformation.* Starting from a symmetric starting condition, the solver consistently terminated at a (almost) symmetric finale state where the two pincer ends are flatly pressed together. While this probably always is a local minimum, with sufficiently strong forces we expect to find a another candidate for the global minimum with less energy, a deformation where the pincer arms slide past each
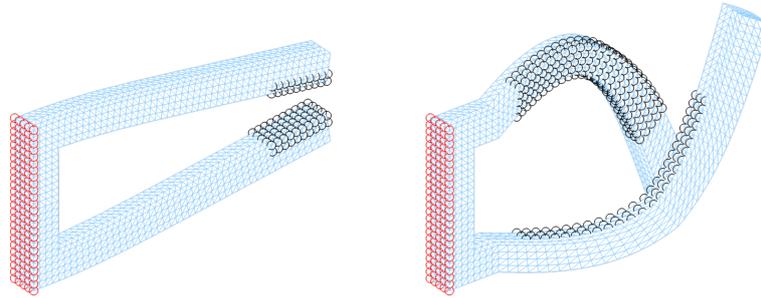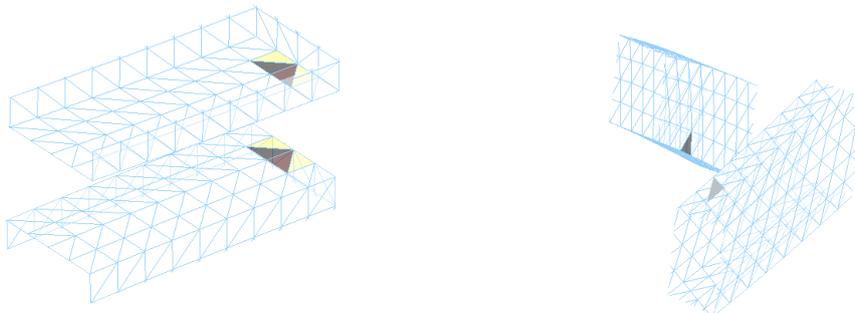
(A) Initially deformed meshes.
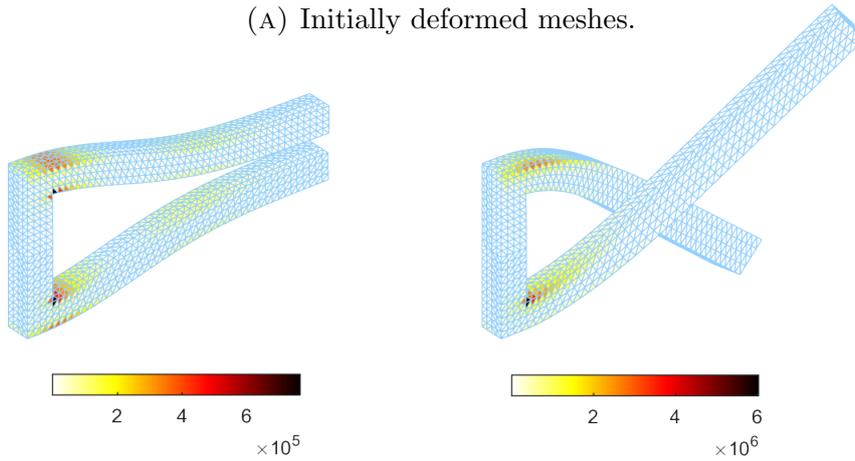
(B) Deformed meshes with the underlying elastic densities.

(C) Deformed meshes with the underlying nonpenetration densities.

(D) The above pictures zoomed in nonpenetration nodes.

FIGURE 4. Solutions for level 2 mesh: the symmetric initial deformation (the left column) and the asymmetric initial deformation (the right column).

other. If this happens, there are obviously two such solutions that can be mapped into each other by reflection across the plane $\{x_2 = 0.25\}$, with the upper pincer (occupying $\Omega \cap \{x_3 > 1.5\}$ in its undeformed state) passing in front or in the back of the lower pincer. To find such a deformation numerically, we artificially specify an explicitly defined starting deformation y_0 with the pincer arms passing each other. More specifically, we fix $X = (X_1, X_2, X_3) := (3, 0.25, 1.5)$, a point in the center between the two "arms" of Ω, and introduce the (reflected) planar polar coordinates with respect to the center point X given by

$$r = r(x_1, x_3) := \sqrt{(x_1 - X_1)^2 + (x_3 - X_3)^2} > 0,$$
$$\theta = \theta(x_1, x_3) := -\sphericalangle\big((x_1 - X_1, x_3 - X_3), (-1, 0)\big) \in (-\pi, \pi).$$

Here, notice that the half-plane $X + \{(t, 0, z) \mid t \geq 0, z \in \mathbb{R}\}$ does not intersect Ω. Consequently, the angle θ is a well defined continuous extension of the function $-\arctan((x_3 - X_3)/(x_1 - X_1))$ from $\{x_1 < X_1\}$ to Ω. It satisfies $\theta(x_1, X_3) = 0$ for all $x_1 < X_1$, $\theta(x_1, x_3) < 0$ for $x_3 < X_3$ and $\theta(x_1, x_3) > 0$ for $x_3 > X_3$. With the auxiliary "twist parameter" function

$$T = T(x_1) := \frac{1}{2} \min\{[x_1 - X_1 - 0.5]^+, 2\}$$

we now define the starting deformation as

$$y_0(x) := \begin{pmatrix} -r \cos\big((1 + 0.2T)\theta\big) \\ x_2 + 0.3T(x_3 - X_3) \\ r \sin\big((1 + 0.2T)\theta\big) \end{pmatrix}$$

mesh			energy				performance	
lev.	nodes:		total	elastic	nonpenet.	body	iters.	time
	$|\mathcal{N}|$	$|\mathcal{N}_{NP}|$	$E_{\varepsilon,\mu}(y)$	$E_{\text{lin}}^{\text{el}}(y)$	$\mu E_{\varepsilon}^{\partial\Omega}(y)$	$E^{body}(y)$		(sec)
1	513	0	-3.88e+05	3.88e+05	0	-7.76e+05	1	3.03e-03
2	2825	0	-5.38e+05	5.38e+05	0	-1.08e+06	1	2.48e-02
3	18225	0	-6.03e+05	6.03e+05	0	-1.21e+06	1	3.62e-01
4	129761	0	-6.19e+05	6.19e+05	0	-1.24e+06	1	7.67e+00

TABLE 1. Purely elastic material.

mesh			energy				performance	
lev.	nodes:		total	elastic	nonpenet.	body	iters.	time
	$|\mathcal{N}|$	$|\mathcal{N}_{NP}|$	$E_{\varepsilon,\mu}(y)$	$E_{\text{lin}}^{\text{el}}(y)$	$\mu E_{\varepsilon}^{\partial\Omega}(y)$	$E^{body}(y)$		(sec)
1	513	52	-1.94e+05	9.16e+04	2.80e+01	-2.86e+05	10	1.12e+00
2	2825	146	-2.52e+05	1.18e+05	7.42e+00	-3.71e+05	18	1.98e+01
3	18225	454	-2.78e+05	1.28e+05	1.63e+00	-4.06e+05	25	4.49e+02

TABLE 2. The initial symmetric deformation.

mesh			energy				performance	
lev.	nodes:		total	elastic	nonpenet.	body	iters.	time
	$|\mathcal{N}|$	$|\mathcal{N}_{NP}|$	$E_{\varepsilon,\mu}(y)$	$E_{\text{lin}}^{\text{el}}(y)$	$\mu E_{\varepsilon}^{\partial\Omega}(y)$	$E^{body}(y)$		(sec)
1	513	130	-2.70e+05	3.22e+05	3.79e+02	-5.93e+05	14	6.99e+00
2	2825	450	-4.67e+05	4.86e+05	3.09e+01	-9.52e+05	23	2.66e+02

TABLE 3. The initial asymmetric deformation.

A visualization of $y_0(\Omega)$ can be seen on the right in Figure 4a. The optimal deformation displayed with linear elasticity and non-penetration density and its zoomed view on the right of Figure 4b, 4c, 4d. The detailed evaluation data is given in Table 3.

Details on implementation and running times: Our MATLAB code is based on the FEM vectorization ideas of [42] combined with fast assembly routines of [48] for linear elasticity. Practical energy minimization is based on the first-order quasi-Newton method applied to E_{precond}, with an explicit differentiation behind and the construction of an approximate Hessian by the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm. The code is available at

$$\texttt{https://www.mathworks.com/matlabcentral/fileexchange/124290}$$

for download and testing. It requires running the Optimization Toolbox (function "fminunc") and Statistics and Machine Learning Toolbox (function "pdist2"). Assembly times were obtained on a MacBook Air (M1 processor, 2020) with 16 GB memory running MATLAB R2022a.

Acknowledgments

References

[1] Virginia Agostiniani, Gianni Dal Maso, and Antonio DeSimone. Linear elasticity obtained from finite elasticity by Γ-convergence under weak coerciveness conditions. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 29(5):715–735, 2012.

[2] Noam Aigerman and Yaron Lipman. Injective and bounded distortion mappings in 3d. *ACM Trans. Graph.*, 32(4):106:1–106:14, July 2013.

[3] Stefano Almi, Elisa Davoli, and Manuel Friedrich. Non-interpenetration conditions in the passage from nonlinear to linearized Griffith fracture. Preprint arXiv:2204.10622, 2022.

[4] Stuart S. Antman. *Nonlinear problems of elasticity*, volume 107 of *Applied Mathematical Sciences*. Springer, New York, second edition, 2005.

[5] Ivo Babuška and Manil Suri. Locking effects in the finite element approximation of elasticity problems. *Numer. Math.*, 62(4):439–463, 1992.

[6] John M. Ball. Convexity conditions and existence theorems in nonlinear elasticity. *Arch. Rational Mech. Anal.*, 63(4):337–403, 1977.

[7] John M. Ball. Global invertibility of Sobolev functions and the interpenetration of matter. *Proc. R. Soc. Edinb., Sect. A, Math.*, 88:315–328, 1981.

[8] John M. Ball. Some open problems in elasticity. In *Geometry, mechanics, and dynamics*, pages 3–59. Springer, New York, 2002.

[9] Sören Bartels and Philipp Reiter. Numerical solution of a bending-torsion model for elastic rods. *Numer. Math.*, 146(4):661–697, 2020.

[10] Sören Bartels and Philipp Reiter. Stability of a simple scheme for the approximation of elastic knots and self-avoiding inextensible curves. *Math. Comp.*, 90(330):1499–1526, 2021.

[11] Sören Bartels, Philipp Reiter, and Johannes Riege. A simple scheme for the approximation of self-avoiding inextensible curves. *IMA J. Numer. Anal.*, 38(2):543–565, 2018.

[12] Barbora Benešová, Martin Kružík, and Anja Schlömerkemper. A note on locking materials and gradient polyconvexity. *Math. Models Methods Appl. Sci.*, 28(12):2367–2401, 2018.

[13] Simon Blatt, Philipp Reiter, Armin Schikorra, and Nicole Vorderobermeier. Scale-invariant tangent-point energies for knots. Preprint arXiv:2104.10238, 2022.

[14] Tino Bog, Nils Zander, Stefan Kollmannsberger, and Ernst Rank. Normal contact with high order finite elements and a fictitious contact material. *Computers & Mathematics with Applications*, 70(7):1370–1390, 2015.

[15] Daniel Campbell, Stanislav Hencl, and Ville Tengvall. Approximation of $W^{1,p}$ Sobolev homeomorphism by diffeomorphisms and the signs of the Jacobian. *Adv. Math.*, 331:748–829, 2018.

[16] Philippe G. Ciarlet. *Mathematical elasticity. Vol. I*, volume 20 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam, 1988. Three-dimensional elasticity.

[17] Philippe G. Ciarlet and Jindřich Nečas. Injectivity and self-contact in nonlinear elasticity. *Arch. Ration. Mech. Anal.*, 97:173–188, 1987.

[18] Bernard Dacorogna. *Direct methods in the calculus of variations*, volume 78 of *Applied Mathematical Sciences*. Springer, New York, second edition, 2008.

[19] G. Dal Maso, M. Negri, and D. Percivale. Linearized elasticity as Γ-limit of finite elasticity. *Set-Valued Anal.*, 10(2-3):165–183, 2002. Calculus of variations, nonsmooth analysis and related topics.

[20] Gianni Dal Maso. *An introduction to Γ-convergence.* Number 8 in Progress in Nonlinear Differential Equations and their Applications. Birkhäuser, Basel, 1993.

[21] Sara Daneri and Aldo Pratelli. Smooth approximation of bi-Lipschitz orientation-preserving homeomorphisms. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 31(3):567–589, 2014.

[22] G. De Philippis and A. Pratelli. The closure of planar diffeomorphisms in Sobolev spaces. *Ann. Inst. H. Poincaré C Anal. Non Linéaire*, 37(1):181–224, 2020.

[23] Mathieu Fabre, Jérôme Pousin, and Yves Renard. A fictitious domain method for frictionless contact problems in elasticity using Nitsche's method. *SMAI J. Comput. Math.*, 2:19–50, 2016.

[24] Roger Fosdick, Francesco Freddi, and Gianni Royer-Carfagni. Bifurcation instability in linear elasticity with the constraint of local injectivity. *J. Elasticity*, 90(1):99–126, 2008.

[25] Roger Fosdick and Gianni Royer-Carfagni. The constraint of local injectivity in linear elasticity theory. *R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci.*, 457(2013):2167–2187, 2001.

[26] M. Foss, W. J. Hrusa, and V. J. Mizel. The Lavrentiev gap phenomenon in nonlinear elasticity. *Arch. Ration. Mech. Anal.*, 167(4):337–365, 2003.

[27] Alessandro Giacomini and Marcello Ponsiglione. Non-interpenetration of matter for SBV deformations of hyperelastic brittle materials. *Proc. Roy. Soc. Edinburgh Sect. A*, 138(5):1019–1041, 2008.

[28] Timothy J. Healey and Stefan Krömer. Injective weak solutions in second-gradient nonlinear elasticity. *ESAIM, Control Optim. Calc. Var.*, 15(4):863–871, 2009.

[29] Duvan Henao, Carlos Mora-Corral, and Marcos Oliva. Global invertibility of Sobolev maps. *Adv. Calc. Var.*, 14(2):207–230, 2021.

[30] Stanislav Hencl and Pekka Koskela. *Lectures on mappings of finite distortion*, volume 2096 of *Lecture Notes in Mathematics*. Springer, Cham, 2014.

[31] Tadeusz Iwaniec, Jani Onninen, Pekka Pankka, and Teresa Radice. A neo-hookean model of plates. *SIAM J. Math. Anal.*, 53(1):509–529, 2021.

[32] Stefan Krömer. Global invertibility for orientation-preserving Sobolev maps via invertibility on or near the boundary. *Arch. Ration. Mech. Anal.*, 238(3):1113–1155, 2020.

[33] Stefan Krömer and Philipp Reiter. Nonlinear elasticity with vanishing nonlocal self-repulsion. Preprint arXiv:2206.09594, 2022.

[34] Stefan Krömer and Jan Valdman. Global injectivity in second-gradient non-linear elasticity and its approximation with penalty terms. *Mathematics and Mechanics of Solids*, 24(8):2619–2635, 2019.

[35] Francesco Maddalena, Danilo Percivale, and Franco Tomarelli. Adhesive flexible material structures. *Discrete Contin. Dyn. Syst. Ser. B*, 17(2):553–574, 2012.

[36] Francesco Maddalena, Danilo Percivale, and Franco Tomarelli. The gap between linear elasticity and the variational limit of finite elasticity in pure traction problems. *Arch. Ration. Mech. Anal.*, 234(3):1091–1120, 2019.

[37] Francesco Maddalena, Danilo Percivale, and Franco Tomarelli. A new variational approach to linearization of traction problems in elasticity. *J. Optim. Theory Appl.*, 182(1):383–403, 2019.

[38] Edoardo Mainini and Danilo Percivale. Sharp conditions for the linearization of finite elasticity. *Calc. Var. Partial Differential Equations*, 60(5):Paper No. 164, 31, 2021.

[39] Cy Maor and Maria Giovanna Mora. Reference configurations versus optimal rotations: a derivation of linear elasticity from finite elasticity for all traction forces. *J. Nonlinear Sci.*, 31(3):Paper No. 62, 28, 2021.

[40] Alexander Mielke and Tomáš Roubíček. Rate-independent elastoplasticity at finite strains and its numerical approximation. *Math. Models Methods Appl. Sci.*, 26(12):2203–2236, 2016.

[41] Umberto Mosco. Convergence of convex sets and of solutions of variational inequalities. *Advances in Math.*, 3:510–585, 1969.

[42] Alexej Moskovka and Jan Valdman. Fast MATLAB evaluation of nonlinear energies using FEM in 2D and 3D: nodal elements. *Appl. Math. Comput.*, 424:Paper No. 127048, 18, 2022.

[43] Pablo V. Negrón Marrero. A numerical method for detecting singular minimizers of multidimensional problems in nonlinear elasticity. *Numer. Math.*, 58(2):135–144, 1990.

[44] J. A. Nitsche. On Korn's second inequality. *RAIRO Anal. Numér.*, 15(3):237–248, 1981.

[45] Aaron Z. Palmer. Variations of deformations with self-contact on Lipschitz domains. *Set-Valued Var. Anal (online first)*, pages 1–11, 2018.

[46] Aaron Z. Palmer and Timothy J. Healey. Injectivity and self-contact in second-gradient nonlinear elasticity. *Calc. Var. Partial Differ. Equ.*, 56(4):11, 2017.

[47] Olivier Pantz. A frictionless contact algorithm for deformable bodies. *ESAIM Math. Model. Numer. Anal.*, 45(2):235–254, 2011.

[48] Talal Rahman and Jan Valdman. Fast MATLAB assembly of FEM matrices in 2D and 3D: nodal elements. *Appl. Math. Comput.*, 219(13):7151–7158, 2013.

[49] Miroslav Šilhavý. *The mechanics and thermodynamics of continuous media.* Texts and Monographs in Physics. Springer, Berlin, 1997.

[50] Chris Yu, Caleb Brakensiek, Henrik Schumacher, and Keenan Crane. Repulsive surfaces. *ACM Trans. Graph.*, 40(6), 2021.

Stefan Krömer, Institute of Information Theory and Automation, Czech Academy of Sciences, Pod vodárenskou věží 4, 182 08 Praha 8, Czech Republic, *Email address*: skroemer@utia.cas.cz

Jan Valdman, Institute of Information Theory and Automation, Czech Academy of Sciences, Pod vodárenskou věží 4, 182 08 Praha 8, Czech Republic and Faculty of Information Technology, Czech Technical University in Prague, Thákurova 9, 160 00 Praha 6, Czech Republic, *Email address*: jan.valdman@utia.cas.cz