A DESCENT SCHEME FOR THICK ELASTIC CURVES WITH SELF-CONTACT AND CONTAINER CONSTRAINTS

SHAWN W. WALKER

Abstract. We present a numerical method to simulate thick elastic curves that accounts for self-contact and container (obstacle) constraints under large deformations. The base model includes bending and torsion effects, as well as inextensibility. A minimizing movements, descent scheme is proposed for computing energy minimizers, under the non-convex inextensibility, self-contact, and container constraints (if the container is non-convex). At each pseudo time-step of the scheme, the constraints are linearized, which yields a convex minimization problem (at every time-step) with affine equality and inequality constraints. First order conditions are established for the descent scheme at each time-step, under reasonable assumptions on the admissible set. Furthermore, under a mild time-step restriction, we prove energy decrease for the descent scheme, and show that all constraints are satisfied to second order in the time-step, regardless of the total number of time-steps taken.

Moreover, we give a modification of the scheme that regularizes the inequality constraints, and establish convergence of the regularized solution. We then discretize the regularized problem with a finite element method using Hermite and Lagrange elements. Several numerical experiments are shown to illustrate the method, including an example that exhibits *massive* amounts of self-contact for a tightly packed curve inside a sphere.

MSC: 65K15, 74K10, 49M37

1. Introduction

The packaging of material into confined spaces is an ubiquitous process in both nature and in many engineering applications. Examples are: DNA packaging in viral and bacteriophage capsids, as well as cell nuclei [36, 1, 4, 38, 43, 45, 64], mechanical packing of filaments or slender rods [50, 55], self-deployable structures [40], drill strings in oil ducts [61], folding of tree leaves [39], and hindwing folding in ladybugs [51].

The physics at play in the various applications involves several pieces, such as elastic bending, biological diffusion, ion effects, etc. But the most salient feature in all of the examples is that the material cannot *self-penetrate*, which leads to the phenomena of *self-contact*, i.e. where the elastic body under consideration deforms so much that distinct regions of its boundary are in direct contact. Modeling the mechanics, and the numerical computation, of collisions, multi-body contact, and self-contact has a long history. In computer graphics [16, 54, 13, 22, 42], collisions and contact must be accounted for to yield realistic animations. Many practical problems exhibit contact effects, such as modeling aortic valves [8] and particulate suspensions in Stokes flow [44]. As a result, a host of methods and techniques have been developed to handle contact phenomena in a variety of contexts; for instance, see [37, 66, 23].

In the case of multi-body contact, a popular method is the so-called "master-slave" approach. It works by (arbitrarily) choosing one of the bodies as a reference (master) and the other body as the slave and treating the no-penetration condition as an (essentially) standard variational inequality or linear complimentarity problem [66, 23]. Unfortunately, when many bodies are present, this approach is quite complicated and may be unstable. Furthermore, the approach is not adequate for handling self-contact problems, at least not without ad hoc adjustments.

Focusing now on elastic curves with thickness, more detailed and mathematically sophisticated methods for modeling self-contact are available. In [60, 56], they develop analytical approaches to closed curves that exhibit self-contact at a finite number of distinct points. Along similar lines, [63] describe a Cosserat rod theory that accounts for isolated self-contact points to model various elastic effects, such as jump or pop-out of (open ended) elastic rods. Equilibrium equations for Kirchhoff rods with non-trivial self-contact (i.e. the

Key words and phrases. elastic curve, inequality constraint, self-contact, Newton method, curve packing.

SWW has been supported in part by NSF grant DMS-2111474.

contact regions are unknown a priori and not assumed to be distinct points) are developed in [21], as well as [31, 30]. Another concept, related to self-contact, is known as the *global radius of curvature* [26], which was used in characterizing ideal knots [25, 53, 18]. Lastly, the theory in [52], to the best of our knowledge, gives the most general treatment on the equilibrium conditions for nonlinearly elastic rods with self-contact.

The main concern of the present paper is the development of an algorithm that can model inextensible elastic curves, with finite thickness, subject to external container constraints as well as the no self-penetration condition. A variety of methods have been proposed for inextensible elastic curves, e.g. finite element based methods [5, 14, 3, 10, 19]. For ideal knots, which are closed curves of a given thickness and knot class, [7, 6] gave a gradient descent method based on polygonal curves that imposed the no self-penetration condition. A Monte Carlo method for computing ideal knots can be found in [41].

Of direct relevance to our work are the methods in [12, 10, 11]. A gradient descent method is proposed in [10] for a bending-torsion rod whose discretization satisfies a Γ-convergence result. In [12, 11], a singular "tangent-point" potential, based on [58, 59], is used to penalize self-intersections of knotted closed curves.

1.1. **Summary of Results.** In this paper, we model an inextensible bending-torsion rod with a given constant thickness, subject to two inequality constraints. The first is a constraint on self-contact, which maintains the no self-penetration condition, and is the central aspect of this paper. The second is an obstacle-type inequality constraint that represents a container (capsid) which confines the rod. The bending-torsion model is based on that in [10] with some minor simplification. We handle self-contact by formulating an inequality constraint on the distances between distinct points on the curve (similar to [21]). We then linearize the condition by a simple, but powerful, technique presented in [48, 49]. The capsid constraint is formulated as a non-linear inequality constraint, which is then linearized. This is all combined with a minimizing movements scheme (i.e. gradient descent) for finding a (constrained) energy minimizer. We establish well-posedness and KKT conditions, i.e. existence of Lagrange multipliers, at a single time-step, for our discrete in time and continuous in space minimizing movements scheme for inextensible, thick elastic curves inside a capsid that exhibit self-contact. A quasi energy decrease property of the descent scheme is proved under a mild time-step restriction. In addition, the self-contact constraint is exactly satisfied, and all other constraints are satisfied to *second order* accuracy in the time-step.

Furthermore, we give a regularization of the method that is simple to compute and converges to the unregularized problem as the parameter c goes to infinity. Convergence of a Newton scheme for solving the regularized problem is also verified for sufficiently good initial guesses. Our finite element discretization is easy to implement, and we demonstrate that the method can efficiently compute complex, tightly packed curve configurations. We also describe a damped Newton scheme that allows for taking the regularization parameter c to approximately 10^8 to 10^9 in our numerical experiments.

Despite the extensive literature on the analytical aspects of self-contact of thick elastic curves, few numerical schemes are available for their computation. The descent scheme that we present can handle self-contact and the capsid obstacle in a rigorous, stable, and efficient manner. Indeed, to the best of our knowledge, the fully discrete scheme appears to be the first to handle *massive* self-contact of elastic curves in an efficient manner. Moreover, no a priori condition on the contact set is needed by the algorithm.

1.2. **Outline.** Section 2 describes the base inextensible bending-torsion rod model and the descent scheme used to find an approximate energy minimizer. In Section 3, we present how the modeling and linearization of the inequality constraints, with respect to self-contact and the capsid container, is done. Section 4 discusses the optimality conditions for a single time-step of the descent scheme, and presents energy estimates under mild assumptions on the time-step. We then describe our regularization procedure for handling the inequality constraints, as well as a semi-smooth Newton method for solving the regularized problem, in Section 5. Section 6 describes our finite element discretization and discusses implementation issues. Numerical results are given in Section 7 which illustrate the robustness and efficacy of our method. We then make final comments in Section 8.

1.3. **Notation.** For denoting general (non-linear) functionals, we shall use brackets. For example, if $F : V \to \mathbb{R}$, where V is a function space, we write $F[v] \in \mathbb{R}$. In addition, F may depend on other fixed quantities, e.g. $F[v; u]$, where v is the argument and u is fixed, and we use this same convention for functions.

For the case of linear and bilinear forms, we shall use parenthesis. For example, if $f : V \to \mathbb{R}$ is a linear form (linear in its argument), then we write $f(v) \in \mathbb{R}$; the same goes for a *linear* functional, $F(v)$. Similarly, for a bilinear form $a : V \times V \to \mathbb{R}$ (linear in each argument separately), we write $a(v, w) \in \mathbb{R}$. In general, we use parenthesis to emphasize *linearity*, an inner product, or just collecting a tuple of arguments. If a linear or bilinear form depends on a "fixed" quantity, we use a bracket, e.g. $f[u](v)$ and $a[u](v, w)$ are linear and bilinear forms, respectively, that depend on the fixed function u.

Moreover, given a non-linear functional $J : U \to \mathbb{R}$, its variational derivative is denoted $\delta J[u](v)$, which is the derivative in the direction v, evaluated at the point u. If J is a non-linear functional on $U \times W$, then $\delta_u J[u, w](v)$ and $\delta_w J[u, w](q)$ specify which argument is being differentiated. Furthermore, the second variational derivative is denoted $\delta^2 F[u](v, q)$, where v and q are two different perturbations, and u is fixed.

2. An Inextensible Elastic Rod with Bending and Torsion

We model a three dimensional curve as an inextensible and unshearable elastic rod, and is inspired by [10].

2.1. **Model Problem.** We parameterize the rod, with length $L > 0$, by the map $\boldsymbol{x} : U \to \mathbb{R}^3$, where $U = (0, L)$. When denoting standard $L^p(U)$ spaces, and Sobolev spaces, e.g. $H^2(U)$ and $W^{1,p}(U)$, we simply write $L^p \equiv L^p(U)$, $H^2 \equiv H^2(U)$, and $W^{1,p} \equiv W^{1,p}(U)$ when the domain is clear.

The reference shape of the rod is straight and is captured by the map $\boldsymbol{x}(\xi) = (\xi, 0, 0)^T$, for $\xi \in U$. Let $' \equiv \partial_\xi$. Next, define the unit tangent vector of $\boldsymbol{x}$ by $\boldsymbol{t} = \partial_s \boldsymbol{x} \equiv |\boldsymbol{x}'|^{-1} \boldsymbol{x}'$, where ∂_s is differentiation with respect to arc-length. We also introduce the *director* $\boldsymbol{b} : U \to \mathbb{R}^3$ that represents the twisting of the rod. Furthermore, we demand that $\{\boldsymbol{t}, \boldsymbol{b}, \boldsymbol{d}\}$ form an *orthonormal frame*, where $\boldsymbol{d} = \boldsymbol{t} \times \boldsymbol{b}$.

The (dimensional) energy of the rod is given by

$$\widetilde{\mathcal{E}}_{\mathrm{r}}[\boldsymbol{x}, \boldsymbol{b}] = \frac{\tilde{c}_{\mathrm{b}}}{2} \int_0^L |\boldsymbol{x}''|^2 \, d\xi + \frac{\tilde{c}_{\mathrm{t}}}{2} \int_0^L (\boldsymbol{b}' \cdot \boldsymbol{d})^2 \, d\xi. \tag{1}$$

Further details on the derivation of this model can be found in [47, 9, 10]; see [2] for a general reference. Non-dimensionalizing, and including external effects, we have

$$\widehat{\mathcal{E}}_{\mathrm{r}}[\boldsymbol{x}, \boldsymbol{b}] = \frac{c_{\mathrm{b}}}{2} \int_0^L |\boldsymbol{x}''|^2 \, d\xi + \frac{c_{\mathrm{t}}}{2} \int_0^L (\boldsymbol{b}' \cdot \boldsymbol{d})^2 \, d\xi - \mathcal{F}(\boldsymbol{x}, \boldsymbol{b}), \tag{2}$$

where c_{b}, c_{t}, and L are non-dimensional parameters, and the external energy term $\mathcal{F}(\boldsymbol{x}, \boldsymbol{b})$ is due to external forces and boundary force effects, and we assume it to be a linear functional given by:

$$\mathcal{F}(\boldsymbol{x}, \boldsymbol{b}) = \mathcal{F}_{\mathrm{x}}(\boldsymbol{x}) + \mathcal{F}_{\mathrm{b}}(\boldsymbol{b}), \quad \mathcal{F}_{\mathrm{x}}(\boldsymbol{x}) := \langle \boldsymbol{x}, \boldsymbol{f} \rangle + \boldsymbol{h} \cdot \boldsymbol{x}\Big|_0^L + \boldsymbol{k} \cdot \boldsymbol{x}'\Big|_0^L, \quad \mathcal{F}_{\mathrm{b}}(\boldsymbol{b}) := \langle \boldsymbol{b}, \boldsymbol{g} \rangle + \boldsymbol{l} \cdot \boldsymbol{b}\Big|_0^L, \tag{3}$$

where $\boldsymbol{f} \in (H^2)^*$, $\boldsymbol{g} \in (H^1)^*$, and $\omega\Big|_0^L \equiv \omega(L) - \omega(0)$, for some functions $\boldsymbol{h}$, $\boldsymbol{k}$, and $\boldsymbol{l}$ defined on $\{0, L\}$. We minimize (2) over the admissible set A_{r} defined by

$$\begin{aligned} A_{\mathrm{r}} &:= \{(\boldsymbol{x}, \boldsymbol{b}) \in A_{\mathrm{r},\boldsymbol{x}} \times A_{\mathrm{r},\boldsymbol{b}} \mid \boldsymbol{x}' \cdot \boldsymbol{b} = 0, \text{ a.e.}\}, \\ A_{\mathrm{r},\boldsymbol{x}} &:= \{\boldsymbol{x} \in W(\boldsymbol{x}_{\mathrm{bc}}) \mid |\boldsymbol{x}'| = 1, \text{ a.e.}\}, \quad A_{\mathrm{r},\boldsymbol{b}} := \{\boldsymbol{b} \in V(\boldsymbol{b}_{\mathrm{bc}}) \mid |\boldsymbol{b}| = 1, \text{ a.e.}\}, \\ W(\boldsymbol{x}_{\mathrm{bc}}) &:= \{\boldsymbol{x} \in [H^2]^3 \mid \boldsymbol{X}_{\mathrm{bc}}(\boldsymbol{x}) = \boldsymbol{x}_{\mathrm{bc}}\}, \quad V(\boldsymbol{b}_{\mathrm{bc}}) := \{\boldsymbol{b} \in [H^1]^3 \mid \boldsymbol{B}_{\mathrm{bc}}(\boldsymbol{b}) = \boldsymbol{b}_{\mathrm{bc}}\}. \end{aligned} \tag{4}$$

Note that A_{r}, $A_{\mathrm{r},\boldsymbol{x}}$, and $A_{\mathrm{r},\boldsymbol{b}}$ are *not* convex, but are weakly closed (because of the compact embeddings $W(\boldsymbol{x}_{\mathrm{bc}}) \subset H^1$ and $V(\boldsymbol{b}_{\mathrm{bc}}) \subset L^2$). Dirichlet conditions are encoded in $\boldsymbol{X}_{\mathrm{bc}}(\boldsymbol{x}) = \boldsymbol{x}_{\mathrm{bc}} \in \mathbb{R}^{N_{\boldsymbol{x}}}$ and $\boldsymbol{B}_{\mathrm{bc}}(\boldsymbol{b}) = \boldsymbol{b}_{\mathrm{bc}} \in \mathbb{R}^{N_{\boldsymbol{b}}}$, which we assume to be linear; note: $0 \le N_{\boldsymbol{x}} \le 4$ and $0 \le N_{\boldsymbol{b}} \le 2$.

Assumption 2.1. *A_{r}, $A_{\mathrm{r},\boldsymbol{x}}$, and $A_{\mathrm{r},\boldsymbol{b}}$ are nonempty, which is guaranteed if the Dirichlet conditions $\boldsymbol{X}_{\mathrm{bc}}(\boldsymbol{x}) = \boldsymbol{x}_{\mathrm{bc}}$, $\boldsymbol{B}_{\mathrm{bc}}(\boldsymbol{b}) = \boldsymbol{b}_{\mathrm{bc}}$ are compatible with the curve parametrization and orthonormal frame condition. Note that one must also have $|\boldsymbol{x}(0) - \boldsymbol{x}(L)| < L$ as a necessary condition. Any imposed boundary data can be matched by a suitable adjustment; see [10] for a discussion.*

Moreover, we assume $\|\boldsymbol{x}''\|_{L^2} \approx \|\boldsymbol{x}\|_{H^2}$ over $A_{\mathrm{r},\boldsymbol{x}}$, and $\|\boldsymbol{b}'\|_{L^2} \approx \|\boldsymbol{b}\|_{H^1}$ over $A_{\mathrm{r},\boldsymbol{b}}$. This is fulfilled if suitable Dirichlet conditions are imposed, or other conditions are imposed, such as removing all rotations in $A_{\mathrm{r},\boldsymbol{x}}$ and imposing mean-value zero in $A_{\mathrm{r},\boldsymbol{b}}$.

Since $\{\boldsymbol{t}, \boldsymbol{b}, \boldsymbol{d}\}$ is an orthonormal frame, using the constraints in the admissible set we have that

$$|\boldsymbol{b}'|^2 = (\boldsymbol{b}' \cdot \boldsymbol{x}')^2 + (\boldsymbol{b}' \cdot \boldsymbol{d})^2, \tag{5}$$

because $\boldsymbol{b}' \cdot \boldsymbol{b} = 0$. Thus, the energy can be rewritten as

$$\widehat{\mathcal{E}}_{\mathrm{r}}[\boldsymbol{x}, \boldsymbol{b}] = \frac{c_{\mathrm{b}}}{2} \int_0^L |\boldsymbol{x}''|^2 \, d\xi + \frac{c_{\mathrm{t}}}{2} \int_0^L \left\{ |\boldsymbol{b}'|^2 |\boldsymbol{x}'|^2 - (\boldsymbol{b}' \cdot \boldsymbol{x}')^2 \right\} d\xi - \mathcal{F}(\boldsymbol{x}, \boldsymbol{b}). \tag{6}$$

Clearly, $\widehat{\mathcal{E}}_{\mathrm{r}}[\cdot, \cdot]$ over $H^2 \times H^1$ is bounded below, and is convex in each argument separately (when ignoring the orthonormal frame constraints), but not jointly convex. Note that this formulation is slightly different from the one in [10].

Proposition 2.2. *Adopt Assumption 2.1. Then there exists a minimizer in A_{r} for the energy in* (6).

Proof. Let $\{(\boldsymbol{x}_k, \boldsymbol{b}_k)\} \subset A_{\mathrm{r}}$ be a minimizing sequence. Clearly, $\|\boldsymbol{x}_k\|_{H^2}$ is uniformly bounded. Furthermore, because of the constraint $\boldsymbol{b} \cdot \boldsymbol{x}' = 0$, we have that $\boldsymbol{b}' \cdot \boldsymbol{x}' = -\boldsymbol{b} \cdot \boldsymbol{x}''$. Since, $|\boldsymbol{x}'| = |\boldsymbol{b}| = 1$, a.e., one can show that

$$\widehat{\mathcal{E}}_{\mathrm{r}}[\boldsymbol{x}, \boldsymbol{b}] \geq C_0 \|\boldsymbol{b}'\|_{L^2}^2 - C_1, \tag{7}$$

for some fixed positive constants C_0, C_1. Hence, $\|\boldsymbol{b}_k\|_{H^1}$ is uniformly bounded.

Therefore, $\boldsymbol{x}_k \rightharpoonup \boldsymbol{x}$ in H^2 and $\boldsymbol{b}_k \rightharpoonup \boldsymbol{b}$ in H^1. By compact embedding, $\boldsymbol{x}_k' \to \boldsymbol{x}'$ strongly in L^2, which allows to show that $\widehat{\mathcal{E}}_{\mathrm{r}}[\cdot, \cdot]$ is weakly lower semi-continuous. Thus, since A_{r} is weakly closed, by standard calculus of variations [35, 20], there exists a minimizer. □

In order to develop a practical numerical algorithm to find local minimizers, we follow the approach of [10] and relax the orthogonality constraint, $\boldsymbol{x}' \cdot \boldsymbol{b} = 0$, by penalizing it. Furthermore, in order to ensure a well-posed formulation and maintain simplicity, we include an additional regularization term $\int_0^L |\boldsymbol{b}'|^2 \, d\xi$. Thus, the energy we consider is as follows:

$$\begin{aligned} \mathcal{E}_{\mathrm{r}}[\boldsymbol{x}, \boldsymbol{b}] &= \frac{c_{\mathrm{b}}}{2} \|\boldsymbol{x}''\|_{L^2}^2 + \frac{c_{\mathrm{r}}}{2} \|\boldsymbol{b}'\|_{L^2}^2 + \frac{c_{\mathrm{t}}}{2} \int_0^L \left\{ |\boldsymbol{b}'|^2 |\boldsymbol{x}'|^2 - (\boldsymbol{b}' \cdot \boldsymbol{x}')^2 \right\} d\xi + \frac{c_{\mathrm{o}}}{2} \|\boldsymbol{x}' \cdot \boldsymbol{b}\|_{L^2}^2 - \mathcal{F}(\boldsymbol{x}, \boldsymbol{b}) \\ &= \frac{c_{\mathrm{b}}}{2} \|\boldsymbol{x}''\|_{L^2}^2 + \frac{c_{\mathrm{r}}}{2} \|\boldsymbol{b}'\|_{L^2}^2 + \frac{1}{2} m[\boldsymbol{b}](\boldsymbol{x}, \boldsymbol{x}) - \mathcal{F}(\boldsymbol{x}, \boldsymbol{b}), \end{aligned} \tag{8}$$

where $c_{\mathrm{o}} > 0$ is a (large) penalty parameter, $c_{\mathrm{b}}, c_{\mathrm{t}} > 0$, and $c_{\mathrm{r}} > 0$ is a (small) regularization parameter, and $m[\boldsymbol{b}](\boldsymbol{x}, \boldsymbol{x}) \equiv m[\boldsymbol{x}](\boldsymbol{b}, \boldsymbol{b})$ are continuous parameterized bilinear forms defined by

$$\begin{aligned} m[\tilde{\boldsymbol{b}}](\boldsymbol{y}, \boldsymbol{w}) &:= c_{\mathrm{t}} \int_0^L \left\{ |\tilde{\boldsymbol{b}}'|^2 (\boldsymbol{y}' \cdot \boldsymbol{w}') - \left(\tilde{\boldsymbol{b}}' \cdot \boldsymbol{y}'\right) \left(\tilde{\boldsymbol{b}}' \cdot \boldsymbol{w}'\right) \right\} d\xi + c_{\mathrm{o}} \int_0^L \left(\boldsymbol{y}' \cdot \tilde{\boldsymbol{b}}\right) \left(\boldsymbol{w}' \cdot \tilde{\boldsymbol{b}}\right) d\xi, \\ m[\tilde{\boldsymbol{x}}](\boldsymbol{z}, \boldsymbol{v}) &:= c_{\mathrm{t}} \int_0^L \left\{ (\boldsymbol{z}' \cdot \boldsymbol{v}') |\tilde{\boldsymbol{x}}'|^2 - (\boldsymbol{z}' \cdot \tilde{\boldsymbol{x}}') (\boldsymbol{v}' \cdot \tilde{\boldsymbol{x}}') \right\} d\xi + c_{\mathrm{o}} \int_0^L (\tilde{\boldsymbol{x}}' \cdot \boldsymbol{z}) (\tilde{\boldsymbol{x}}' \cdot \boldsymbol{v}) \, d\xi, \\ m[\tilde{\boldsymbol{b}}](\boldsymbol{w}, \boldsymbol{w}) &\geq c_{\mathrm{t}} \|(\boldsymbol{w}')^T \left[|\tilde{\boldsymbol{b}}'|^2 \boldsymbol{I} - \tilde{\boldsymbol{b}}' \otimes \tilde{\boldsymbol{b}}' \right] \boldsymbol{w}'\|_{L^2}^2 + c_{\mathrm{o}} \|\boldsymbol{w}' \cdot \tilde{\boldsymbol{b}}\|_{L^2}^2 \geq 0, \\ m[\tilde{\boldsymbol{x}}](\boldsymbol{v}, \boldsymbol{v}) &\geq c_{\mathrm{t}} \|(\boldsymbol{v}')^T \left[|\tilde{\boldsymbol{x}}'|^2 \boldsymbol{I} - \tilde{\boldsymbol{x}}' \otimes \tilde{\boldsymbol{x}}' \right] \boldsymbol{v}'\|_{L^2}^2 + c_{\mathrm{o}} \|\boldsymbol{v} \cdot \tilde{\boldsymbol{x}}'\|_{L^2}^2 \geq 0, \\ m[\tilde{\boldsymbol{b}}](\boldsymbol{y}, \boldsymbol{w}) &\leq c_{\mathrm{t}} \|\boldsymbol{y}'\|_{L^\infty} \|\boldsymbol{w}'\|_{L^\infty} \|\tilde{\boldsymbol{b}}'\|_{L^2}^2 + c_{\mathrm{o}} \|\boldsymbol{y}' \cdot \tilde{\boldsymbol{b}}\|_{L^2} \|\boldsymbol{w}' \cdot \tilde{\boldsymbol{b}}\|_{L^2} \\ &\leq C_1 \|\tilde{\boldsymbol{b}}'\|_{L^2}^2 \left(c_{\mathrm{t}} + c_{\mathrm{o}} C_2\right) \|\boldsymbol{y}''\|_{L^2} \|\boldsymbol{w}''\|_{L^2}, \\ m[\tilde{\boldsymbol{x}}](\boldsymbol{z}, \boldsymbol{v}) &\leq \|\tilde{\boldsymbol{x}}'\|_{L^\infty}^2 (c_{\mathrm{t}} + c_{\mathrm{o}} C_3) \|\boldsymbol{z}'\|_{L^2} \|\boldsymbol{v}'\|_{L^2}, \end{aligned} \tag{9}$$

where $C_1, C_2, C_3 > 0$ are uniform constants. We seek to find (local) minimizers of the following problem:

$$\inf_{(\boldsymbol{x}, \boldsymbol{b}) \in A_{\mathrm{r},\boldsymbol{x}} \times A_{\mathrm{r},\boldsymbol{b}}} \mathcal{E}_{\mathrm{r}}[\boldsymbol{x}, \boldsymbol{b}], \quad \text{subject to } \tau(\boldsymbol{x}) = 0, \ \ell(\boldsymbol{b}) = 0, \text{ a.e.}, \tag{10}$$

i.e., we minimize $\mathcal{E}_{\mathrm{r}}[\cdot,\cdot]$ over $W(\boldsymbol{x}_{\mathrm{bc}}) \times V(\boldsymbol{b}_{\mathrm{bc}})$, subject to the constraints $\tau(\boldsymbol{x}) = 0$ and $\ell(\boldsymbol{b}) = 0$, a.e., which capture the unit length constraints:

$$\begin{aligned} \tau &: [H^2]^3 \to H^1,\ \tau(\boldsymbol{w})(\xi) := \frac{1}{2}(|\boldsymbol{w}'(\xi)|^2 - 1),\ \forall\, \xi \in U, \\ \ell &: [H^1]^3 \to H^1,\ \ell(\boldsymbol{v})(\xi) := \frac{1}{2}(|\boldsymbol{v}(\xi)|^2 - 1),\ \forall\, \xi \in U, \end{aligned} \tag{11}$$

with variational derivatives

$$\begin{aligned} \tau'_{\boldsymbol{x}}(\boldsymbol{w})(\xi) = \boldsymbol{x}'(\xi)\cdot\boldsymbol{w}'(\xi), \quad [\tau''_{\boldsymbol{x}}(\boldsymbol{w},\boldsymbol{y})](\xi) = \boldsymbol{w}'(\xi)\cdot\boldsymbol{y}'(\xi),\ \forall\, \xi \in U, \\ \ell'_{\boldsymbol{b}}(\boldsymbol{v})(\xi) = \boldsymbol{b}(\xi)\cdot\boldsymbol{v}(\xi), \quad [\ell''_{\boldsymbol{b}}(\boldsymbol{v},\boldsymbol{z})](\xi) = \boldsymbol{v}(\xi)\cdot\boldsymbol{z}(\xi),\ \forall\, \xi \in U. \end{aligned}$$

A minimizer of (10) can be shown by the direct method, similar to the proof of Prop. 2.2.

Proposition 2.3. *Adopt Assumption 2.1. Then there exists a minimizer in $A_{\mathrm{r},\boldsymbol{x}} \times A_{\mathrm{r},\boldsymbol{b}}$ for the energy in (8).*

Proof. Let $\{(\boldsymbol{x}_k, \boldsymbol{b}_k)\} \subset A_{\mathrm{r},\boldsymbol{x}} \times A_{\mathrm{r},\boldsymbol{b}}$ be a minimizing sequence. Clearly, $\|\boldsymbol{x}_k\|_{H^2}$ is uniformly bounded. Furthermore, since $c_{\mathrm{r}} > 0$, $\|\boldsymbol{b}_k\|_{H^1}$ is uniformly bounded. Therefore, $\boldsymbol{x}_k \rightharpoonup \boldsymbol{x}$ in H^2 and $\boldsymbol{b}_k \rightharpoonup \boldsymbol{b}$ in H^1. By compact embedding, $\boldsymbol{x}'_k \to \boldsymbol{x}'$ strongly in L^2 and $\boldsymbol{b}_k \to \boldsymbol{b}$ strongly in L^2, which allows to show that $\mathcal{E}_{\mathrm{r}}[\cdot,\cdot]$ is weakly lower semi-continuous. Thus, since $A_{\mathrm{r},\boldsymbol{x}} \times A_{\mathrm{r},\boldsymbol{b}}$ is weakly closed, by standard calculus of variations [35, 20], there exists a minimizer. □

Standard arguments show that (8) and (11) are twice continuously Fréchet differentiable [62], i.e. taking advantage of Sobolev embeddings when the domain is one dimensional, one can show that

$$\begin{aligned} \delta_{(\boldsymbol{x},\boldsymbol{b})}\mathcal{E}_{\mathrm{r}}[\boldsymbol{x},\boldsymbol{b}](\cdot,\cdot) \in \mathcal{L}([H^2]^3 \times [H^1]^3, \mathbb{R}), \\ \delta^2_{(\boldsymbol{x},\boldsymbol{b})}\mathcal{E}_{\mathrm{r}}[\boldsymbol{x},\boldsymbol{b}]((\cdot,\cdot),(\cdot,\cdot)) \in \mathcal{L}([H^2]^3 \times [H^1]^3, \mathcal{L}([H^2]^3 \times [H^1]^3, \mathbb{R})), \end{aligned} \tag{12}$$

and

$$\begin{aligned} \tau'_{\boldsymbol{x}}(\cdot) \in \mathcal{L}([H^2]^3, H^1), \quad \tau''_{\boldsymbol{x}}(\cdot,\cdot) \in \mathcal{L}([H^2]^3, \mathcal{L}([H^2]^3, H^1)), \\ \ell'_{\boldsymbol{b}}(\cdot) \in \mathcal{L}([H^1]^3, H^1), \quad \ell''_{\boldsymbol{b}}(\cdot,\cdot) \in \mathcal{L}([H^1]^3, \mathcal{L}([H^1]^3, H^1)), \end{aligned} \tag{13}$$

are continuous in terms of $(\boldsymbol{x}, \boldsymbol{b})$. Moreover, we have surjectivity for the Fréchet derivatives of the unit length constraints $\tau(\boldsymbol{x})$, $\ell(\boldsymbol{b})$, under the following basic assumption.

Assumption 2.4. *Henceforth, we set $W(\mathbf{0}) := \{\boldsymbol{w} \in [H^2]^3 \mid \boldsymbol{w}' \in [H^1_0]^3, \boldsymbol{w}(0) = \mathbf{0}\}$ and $V(\mathbf{0}) \equiv [H^1_0]^3$.*

Other boundary conditions can also be used, provided $\boldsymbol{w} \in W(\mathbf{0})$ does not vanish at *both* end-points. For any other choice of boundary conditions, Propositions 2.5, 2.6 still hold with obvious modifications to the space for the dual variable.

Proposition 2.5. *Under Assumption 2.4, for all $\tilde{\boldsymbol{b}} \in [H^1]^3$, such that $|\tilde{\boldsymbol{b}}| \geq a_0 > 0$ a.e., $\ell'_{\tilde{\boldsymbol{b}}}(\cdot) : V(\mathbf{0}) \to H^1_0$, is surjective, i.e.*

$$\sup_{\boldsymbol{v} \in V(\mathbf{0})} \frac{\left\langle \mu, \ell'_{\tilde{\boldsymbol{b}}}(\boldsymbol{v}) \right\rangle_{H^{-1} \times H^1_0}}{\|\boldsymbol{v}'\|_{L^2}} \geq \beta \|\mu\|_{H^{-1}}, \quad \text{for all } \mu \in H^{-1}, \tag{14}$$

for some $\beta > 0$ that depends on $\tilde{\boldsymbol{b}}$ and a_0.

Proof. The proof is mainly standard (c.f. [32, Thm. 3.1]) and is similar to the proof of Prop. 2.6. □

Proposition 2.6. *Under Assumption 2.4, for all $\tilde{\boldsymbol{x}} \in [H^2]^3$, such that $|\tilde{\boldsymbol{x}}'| \geq a_0 > 0$ a.e., $\tau'_{\tilde{\boldsymbol{x}}}(\cdot) : W(\mathbf{0}) \to H^1_0$, is surjective, i.e.*

$$\sup_{\boldsymbol{w} \in W(\mathbf{0})} \frac{\langle \mu, \tau'_{\tilde{\boldsymbol{x}}}(\boldsymbol{w}) \rangle_{H^{-1} \times H^1_0}}{\|\boldsymbol{w}''\|_{L^2}} \geq \beta \|\mu\|_{H^{-1}}, \quad \text{for all } \mu \in H^{-1}, \tag{15}$$

for some $\beta > 0$ that depends on $\tilde{\boldsymbol{x}}$ and a_0.

Proof. Given any $\mu \in H^{-1}$, there exists a $\varphi \in H^1_0$ such that $\langle \mu, \varphi \rangle / \|\varphi'\|_{L^2} = \|\mu\|_{H^{-1}}$. Let $\boldsymbol{y}' = \varphi \tilde{\boldsymbol{x}}' |\tilde{\boldsymbol{x}}'|^{-2}$, where $\tilde{\boldsymbol{x}}' \in [H^1]^3$ such that $|\tilde{\boldsymbol{x}}'| \geq a_0 > 0$ a.e. By Sobolev embedding and Poincaré, $\|\varphi\|_{L^\infty} \leq C_0 \|\varphi'\|_{L^2}$, whence $\boldsymbol{y}' \in [H^1_0]^3$ and $\boldsymbol{y} \in T$ (by an anti-derivative). Moreover, $\|\boldsymbol{y}''\|_{L^2} \leq a_0^{-1}(1 + a_0^{-1} C_0 \sqrt{3} \|\tilde{\boldsymbol{x}}''\|_{L^2}) \|\varphi'\|_{L^2}$. Thus, we obtain

$$\sup_{\boldsymbol{w} \in W(\boldsymbol{0})} \frac{\langle \mu, \tau'_{\tilde{\boldsymbol{x}}}(\boldsymbol{w}) \rangle}{\|\boldsymbol{w}''\|_{L^2}} \geq \frac{\langle \mu, \tau'_{\tilde{\boldsymbol{x}}}(\boldsymbol{y}) \rangle}{\|\boldsymbol{y}''\|_{L^2}} \geq \beta \frac{\langle \mu, \varphi \rangle}{\|\varphi'\|_{L^2}} = \beta \|\mu\|_{H^{-1}},$$

where $\beta = a_0(1 + a_0^{-1} C_0 \sqrt{3} \|\tilde{\boldsymbol{x}}''\|_{L^2})^{-1}$. □

With the above results, first order optimality conditions (i.e. KKT conditions) can be established for (10) (see [34]).

2.2. **Gradient Flow.** Even for the simple case of an inextensible elastic rod, finding an equilibrium solution is difficult. The energy is non-linear and non-convex. So we adopt a minimizing movements strategy. Given the current state $(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{b}}) \in A_{\mathrm{r}}$, i.e. a *feasible* point which will play the role of a linearization point, define the following modified energy motivated by minimizing movements:

$$\mathcal{J}_{\mathrm{r}}[\boldsymbol{x}, \boldsymbol{b}] \equiv \mathcal{J}_{\mathrm{r}}\left[\boldsymbol{x}, \boldsymbol{b}; \tilde{\boldsymbol{x}}, \tilde{\boldsymbol{b}}\right] = \mathcal{E}_{\mathrm{r}}[\boldsymbol{x}, \boldsymbol{b}] + \frac{1}{2\Delta t}(\boldsymbol{x} - \tilde{\boldsymbol{x}}, \boldsymbol{x} - \tilde{\boldsymbol{x}})_\star + \frac{1}{2\Delta t}\left(\boldsymbol{b} - \tilde{\boldsymbol{b}}, \boldsymbol{b} - \tilde{\boldsymbol{b}}\right)_\dagger, \tag{16}$$

where we have introduced a time-step $\Delta t > 0$. The continuous, symmetric bilinear forms $(\cdot,\cdot)_\star : H^2 \times H^2 \to \mathbb{R}$ and $(\cdot,\cdot)_\dagger : H^1 \times H^1 \to \mathbb{R}$ can be weighted inner products, but our typical choice is $(\boldsymbol{w}, \boldsymbol{y})_\star \equiv (\boldsymbol{w}, \boldsymbol{y})_{H^2}$ and $(\boldsymbol{v}, \boldsymbol{z})_\dagger \equiv (\boldsymbol{v}, \boldsymbol{z})_{H^1}$, as in [10]. Note that (16) is twice continuously Fréchet differentiable.

We now use a time-splitting method via alternating minimization to approximate a minimizer of (10), i.e. if $(\boldsymbol{x}^i, \boldsymbol{b}^i)$ is the solution at the current time-step, we perform a two-stage minimization to obtain the solution $(\boldsymbol{x}^{i+1}, \boldsymbol{b}^{i+1})$ at the next time-step (see Algorithm 1). Furthermore, note that the unit length constraints for $\boldsymbol{b}$ and $\boldsymbol{x}'$ can be written as

$$\tau(\boldsymbol{w}) = \tau_{\tilde{\boldsymbol{x}}}(\boldsymbol{w}) + \frac{1}{2}|\boldsymbol{w}' - \tilde{\boldsymbol{x}}'|^2, \qquad \ell(\boldsymbol{v}) = \ell_{\tilde{\boldsymbol{b}}}(\boldsymbol{v}) + \frac{1}{2}\left|\boldsymbol{v} - \tilde{\boldsymbol{b}}\right|^2, \tag{17}$$

where the linearizations are given by

$$\begin{aligned} &\tau_{\tilde{\boldsymbol{x}}}(\cdot) : [H^2]^3 \to H^1, \quad \tau_{\tilde{\boldsymbol{x}}}(\boldsymbol{w}) := \tau(\tilde{\boldsymbol{x}}) + \tau'_{\tilde{\boldsymbol{x}}}(\boldsymbol{w} - \tilde{\boldsymbol{x}}) = \frac{1}{2}\left(|\tilde{\boldsymbol{x}}'|^2 - 1\right) + \tilde{\boldsymbol{x}}' \cdot (\boldsymbol{w}' - \tilde{\boldsymbol{x}}'), \\ &\ell_{\tilde{\boldsymbol{b}}}(\cdot) : [H^1]^3 \to H^1, \quad \ell_{\tilde{\boldsymbol{b}}}(\boldsymbol{v}) := \ell(\tilde{\boldsymbol{b}}) + \ell'_{\tilde{\boldsymbol{b}}}\left(\boldsymbol{v} - \tilde{\boldsymbol{b}}\right) = \frac{1}{2}\left(|\tilde{\boldsymbol{b}}|^2 - 1\right) + \tilde{\boldsymbol{b}} \cdot \left(\boldsymbol{v} - \tilde{\boldsymbol{b}}\right). \end{aligned} \tag{18}$$

We use these *affine* constraints in Steps 1 and 2 of Algorithm 1. Additional constraints will be added later, but the alternating minimization scheme will be the same. Existence of minimizers for Steps 1 and 2 is established in Sections 4.1 and 4.2, respectively.

Algorithm 1 High level alternating minimization for (10).

Set a tolerance TOL > 0, time-step $\Delta t > 0$, and choose a feasible initial guess $(\boldsymbol{x}^0, \boldsymbol{b}^0) \in W(\boldsymbol{x}_{\mathrm{bc}}) \times V(\boldsymbol{b}_{\mathrm{bc}})$. Set $i := 0$ and do the following.

1. Find $\boldsymbol{x}^{i+1} \in W(\boldsymbol{x}_{\mathrm{bc}})$ that uniquely minimizes $\mathcal{J}_{\mathrm{r}}\left[\cdot, \boldsymbol{b}^i; \boldsymbol{x}^i, \boldsymbol{b}^i\right]$, subject to a convex constraint set; see (37).
2. Find $\boldsymbol{b}^{i+1} \in V(\boldsymbol{b}_{\mathrm{bc}})$ that uniquely minimizes $\mathcal{J}_{\mathrm{r}}\left[\boldsymbol{x}^{i+1}, \cdot\,; \boldsymbol{x}^{i+1}, \boldsymbol{b}^i\right]$, subject to $\ell_{\boldsymbol{b}^i}(\boldsymbol{b}) = 0$, on U; see (45).
3. If $\max\left(\|\boldsymbol{x}^{i+1} - \boldsymbol{x}^i\|_{W^{1,\infty}(U)}, \|\boldsymbol{b}^{i+1} - \boldsymbol{b}^i\|_{L^\infty(U)}\right) <$ TOL, then stop;
   else, replace $i \leftarrow i + 1$ and return to Step 1.

3. Inequality Constraints

We introduce two inequality constraints that are critical to modeling the packing of elastic rods, with finite thickness, inside rigid containers. The first accounts for no self-overlap of the rod, and the second accounts for the constraint of the container. In both cases, we introduce an effective linearization that is utilized in our minimizing movements scheme.

3.1. **Self-contact.** We define an elastic rod, with uniform thickness $\varepsilon > 0$, to be the set of points

$$\Sigma \equiv \Sigma(\boldsymbol{x}) := \{\boldsymbol{\eta} \in \mathbb{R}^3 \mid \operatorname{dist}(\boldsymbol{\eta}, \boldsymbol{x}(\cdot)) \leq \varepsilon/2\}, \tag{19}$$

using the curve parametrization $\boldsymbol{x}$. We seek to impose a *no self-overlap condition* on Σ, which is a central aspect of this paper, and has been investigated by various authors, e.g. [26, 25, 52, 59, 9, 11].

We approximate the no self-overlap condition through the following admissible set (c.f. [21, eqn. (8)] and [30, eqn. (2.2)]):

$$A_{\text{sc}} = \left\{\boldsymbol{x} \in [H^2]^3 \mid \varepsilon \leq |\boldsymbol{x}(\xi_1) - \boldsymbol{x}(\xi_2)|, \text{ for all } \xi_1, \xi_2 \in U, \text{ such that } |\xi_1 - \xi_2| \geq \pi\varepsilon/2\right\}, \tag{20}$$

which imposes a non-convex inequality constraint on *self-contact*. Note: the condition $|\xi_1 - \xi_2| \geq \pi\varepsilon/2$ ignores neighboring parametric points and ensures $A_{\text{sc}} \neq \emptyset$. An alternative method, that avoids the small "cut-off" of neighboring points, is the *global radius of curvature* [41, 26, 25], which is a powerful concept that simultaneously deals with "points of closest approach" and the local radius of curvature. Another related approach uses a non-local tangent-point repulsive potential [17, 59, 58, 57, 12, 9, 11].

In this paper, we focus only on (20) for the following reasons. It is simple, yet can be extended to use the *point-tangent radius*, instead of the simple distance function in (20), which would give a more exact definition of no self-overlap. Indeed, this is the crucial aspect of the global radius of curvature discussed in [41, 26, 25] and is an interesting point of future work. Moreover, the condition in (20) maintains the spatial locality of self-contact in the ambient space. The use of a non-local repulsive potential may not be physically relevant, depending on the phenomena being modeled.

3.1.1. *Linearization of Self-contact.* Of course, A_{sc} is a non-convex set. Therefore, we take advantage of the minimizing movements strategy and linearize the admissible set (20) about a given curve $\tilde{\boldsymbol{x}}$. This is inspired by the approach in [49] (also see [48] for more analytical discussion on this topic).

First note that, for any $\boldsymbol{y} \in [H^1]^3$, $|\boldsymbol{y}(\xi_1) - \boldsymbol{y}(\xi_2)| = -\boldsymbol{p}\,(\xi_1, \xi_2; \boldsymbol{y}) \cdot (\boldsymbol{y}(\xi_1) - \boldsymbol{y}(\xi_2))$, where

$$\boldsymbol{p}\,(\xi_1, \xi_2; \boldsymbol{y}) := \begin{cases} -\frac{\boldsymbol{y}(\xi_1) - \boldsymbol{y}(\xi_2)}{|\boldsymbol{y}(\xi_1) - \boldsymbol{y}(\xi_2)|}, & \text{if } \xi_1 \neq \xi_2, \\ \boldsymbol{0}, & \text{if } \xi_1 = \xi_2, \end{cases} \tag{21}$$

and is well-defined provided $\boldsymbol{y}$ is an injective map; the minus sign in the definition is for convenience. We call $\boldsymbol{p}$ the *proximity vector* and it satisfies $\boldsymbol{p}\,(\xi_1, \xi_2; \boldsymbol{y}) = -\boldsymbol{p}\,(\xi_2, \xi_1; \boldsymbol{y})$. With this, for a given $\tilde{\boldsymbol{x}} \in [H^1]^3$, we define the following *convex* set

$$\bar{A}_{\text{sc}}(\tilde{\boldsymbol{x}}) := \{\boldsymbol{y} \in [H^2]^3 \mid \varepsilon + \boldsymbol{p}\,(\xi_1, \xi_2; \tilde{\boldsymbol{x}}) \cdot (\boldsymbol{y}(\xi_1) - \boldsymbol{y}(\xi_2)) \leq 0, \ \forall\, \xi_1, \xi_2 \in U, \text{ such that } \chi^{\text{n}}(\xi_1, \xi_2) = 1\}, \tag{22}$$

where $\chi^{\text{n}} : U^2 \to \{0, 1\}$ is the characteristic function defined by

$$\chi^{\text{n}}(\xi_1, \xi_2) = \begin{cases} 1, & \text{if } |\xi_1 - \xi_2| \geq \pi\varepsilon/2, \\ 0, & \text{else.} \end{cases} \tag{23}$$

Because of χ^{n}, the value of $\boldsymbol{p}\,(\xi_1, \xi_2; \boldsymbol{y})$ when $\xi_1 = \xi_2$ is irrelevant.

3.1.2. *Properties.* Let $U^2 := U \times U$ be the Cartesian product domain and $\int_{U^2} f(\boldsymbol{\xi}) d\boldsymbol{\xi} \equiv \int_0^L \int_0^L f(\xi_1, \xi_2) d\xi_1 d\xi_2$, where $(\xi_1, \xi_2) \equiv \boldsymbol{\xi}$. The $L^2(U^2)$ inner product is denoted $((u, v))_{L^2(U^2)} = \int_{U^2} uv d\boldsymbol{\xi}$ and the duality pairing is denoted $\langle\langle \zeta, v \rangle\rangle_{D^* \times D}$, where D^* is the dual of D.

We introduce the point-to-point difference operator $\Theta : [H^2(U)]^3 \to [H^2(U^2)]^3$ defined by

$$[\Theta \boldsymbol{y}](\xi_1, \xi_2) := \boldsymbol{y}(\xi_1) - \boldsymbol{y}(\xi_2), \text{ for all } \xi_1, \xi_2 \in U, \tag{24}$$

where $[\Theta \boldsymbol{y}](\xi_1, \xi_2) = -[\Theta \boldsymbol{y}](\xi_2, \xi_1)$. Note that $\boldsymbol{p}\,(\xi_1, \xi_2; \boldsymbol{y}) = -[\Theta \boldsymbol{y} / |\Theta \boldsymbol{y}|](\xi_1, \xi_2)$ for $\xi_1 \neq \xi_2$.

The adjoint operator $\Theta^* : ([H^2(U^2)]^3)^* \to ([H^2(U)]^3)^*$ is defined formally through $\langle\langle \boldsymbol{\zeta}, \Theta \boldsymbol{y} \rangle\rangle = \langle \Theta^* \boldsymbol{\zeta}, \boldsymbol{y} \rangle$ for all $\boldsymbol{y} \in [H^2(U)]^3$ and $\boldsymbol{\zeta} \in ([H^2(U^2)]^3)^*$. In case $\boldsymbol{\zeta} \in [L^2(U^2)]^3$, we have a more explicit representation:

$$\begin{aligned} ((\boldsymbol{\zeta}, \Theta \boldsymbol{y}))_{L^2(U^2)} &= \int_0^L \int_0^L \boldsymbol{\zeta}(\xi_1, \xi_2) \cdot \boldsymbol{y}(\xi_1)\, d\xi_1 d\xi_2 - \int_0^L \int_0^L \boldsymbol{\zeta}(\xi_1, \xi_2) \cdot \boldsymbol{y}(\xi_2)\, d\xi_1 d\xi_2 \\ &= \int_0^L \boldsymbol{y}(\xi_1) \cdot \left(\int_0^L \boldsymbol{\zeta}(\xi_1, \xi_2)\, d\xi_2 \right) d\xi_1 - \int_0^L \boldsymbol{y}(\xi_1) \cdot \left(\int_0^L \boldsymbol{\zeta}(\xi_2, \xi_1)\, d\xi_2 \right) d\xi_1, \end{aligned}$$

which yields

$$(25)\quad ((\boldsymbol{\zeta},\Theta\boldsymbol{y}))_{L^2(U^2)}=\int_0^L[\Theta^*\boldsymbol{\zeta}](\xi)\cdot\boldsymbol{y}(\xi)d\xi=(\Theta^*\boldsymbol{\zeta},\boldsymbol{y})_{L^2(U)}\,,\quad [\Theta^*\boldsymbol{\zeta}](\xi_1)=\int_0^L\boldsymbol{\zeta}(\xi_1,\xi_2)-\boldsymbol{\zeta}(\xi_2,\xi_1)\,d\xi_2.$$

In the following, we need the space of symmetric functions in $H^2(U^2)$:

$$(26)\qquad T:=\{f\in H^2(U^2)\mid f(\xi_1,\xi_2)=f(\xi_2,\xi_1),\ \text{for all } \xi_1,\xi_2\in U\}.$$

For a given $\boldsymbol{w}\in[H^2]^3$, let $\psi(\boldsymbol{w}):U^2\to\mathbb{R}$ be defined as $\psi(\boldsymbol{w})(\xi_1,\xi_2):=\varepsilon-|\boldsymbol{w}(\xi_1)-\boldsymbol{w}(\xi_2)|$; we call ψ the *two-point self-contact function.* In addition, for a given curve $\tilde{\boldsymbol{x}}$, we have the mapping $\psi_{\tilde{\boldsymbol{x}}}(\cdot)\in\mathcal{L}([H^2(U)]^3,T)$, which is a linearization of $\psi(\cdot)$, given by

$$(27)\qquad \tilde{\psi}(\boldsymbol{x})\equiv\psi_{\tilde{\boldsymbol{x}}}(\boldsymbol{x})=\varepsilon+\widetilde{\mathcal{S}}\boldsymbol{x},\qquad \tilde{\psi}'(\boldsymbol{y})\equiv\psi'_{\tilde{\boldsymbol{x}}}(\boldsymbol{y})=\widetilde{\mathcal{S}}\boldsymbol{y},$$

where the operator $\widetilde{\mathcal{S}}\equiv\mathcal{S}_{\tilde{\boldsymbol{x}}}:[H^2(U)]^3\to T$ is defined by $\widetilde{\mathcal{S}}\boldsymbol{y}:=\Theta\boldsymbol{y}\cdot\tilde{\boldsymbol{p}}$, for all $\boldsymbol{y}\in H^2$, with $\tilde{\boldsymbol{p}}\equiv\boldsymbol{p}(\cdot,\cdot;\tilde{\boldsymbol{x}})$. The following monotone property is trivial, but extremely useful

$$(28)\qquad \psi_{\tilde{\boldsymbol{x}}}(\boldsymbol{x})\geq\psi(\boldsymbol{x}),\quad \text{for all } \tilde{\boldsymbol{x}},\boldsymbol{x}\in[H^2]^3.$$

Now define the closed convex cone $K_\psi\subset T$ and its dual cone:

$$(29)\qquad K_\psi=\{z\in T\mid z(\xi_1,\xi_2)\leq 0,\ \text{if } |\xi_1-\xi_2|\geq\pi\varepsilon/2\},\quad K_\psi^*=\{\zeta\in(T)^*\mid\langle\langle\zeta,z\rangle\rangle\geq 0,\ \forall z\in-K_\psi\},$$

which allows us to write the constraint $\boldsymbol{x}\in\bar{A}_{\rm sc}(\tilde{\boldsymbol{x}})$ as $\psi_{\tilde{\boldsymbol{x}}}(\boldsymbol{x})\in K_\psi$.

3.2. **Confinement Inside a Container.** Next, we impose that the curve lies strictly inside a container, which we call the capsid. For instance, Σ must be contained inside the capsid $\Omega_{\rm c}$, which we take to be convex for simplicity. Let $\phi:\mathbb{R}^3\to\mathbb{R}$ be the level set function for $\Omega_{\rm c}$ such that $\Omega_{\rm c}\equiv\{\boldsymbol{y}\in\mathbb{R}^3\mid\phi(\boldsymbol{y})<0\}$, where ϕ is assumed to satisfy

$$(30)\qquad \phi\in C^3(\mathbb{R}^3),\quad |\nabla\phi|=1\text{ in a neighborhood }\mathcal{N}\supset\{\phi=0\},\ \text{ and }\|\nabla\nabla\phi\|_{L^\infty}\leq H_0,$$

for some generic constant $H_0>0$, where $\mathcal{N}$ is an open set that contains all points $\boldsymbol{y}\in\mathbb{R}^3$ such that $\text{dist}\,(\boldsymbol{y},\{\phi=0\})<\varepsilon$. We then impose that the curve $\boldsymbol{x}(\cdot)$ stay within the following bounded, closed convex set:

$$(31)\qquad \left\{\boldsymbol{y}\in[H^2]^3\mid\varepsilon/2+\phi(\boldsymbol{y}(\xi))\leq 0,\ \text{for all }\xi\in U\right\}.$$

3.2.1. *Linearization of the Capsid Constraint.* We use an approximation of (31) at each time step of our minimizing movements strategy. Though the exact projection onto the convex set could be used, a linearization yields a simpler minimization problem at each time-step, even if the capsid is *not convex.* Thus, we linearize the capsid constraint about a given curve $\tilde{\boldsymbol{x}}$, as we did for self-contact (see Section 3.1.1), i.e. linearizing the constraint in (31) about $\tilde{\boldsymbol{x}}$ gives

$$(32)\qquad \varepsilon/2+\phi(\tilde{\boldsymbol{x}}(\xi))+\nabla\phi(\tilde{\boldsymbol{x}}(\xi))\cdot(\boldsymbol{y}(\xi)-\tilde{\boldsymbol{x}}(\xi))\leq 0,\ \text{for all }\xi\in U.$$

Note that $\nabla\phi=\boldsymbol{\nu}$, the unit normal vector of $\partial\Omega_{\rm c}$, in a neighborhood of the zero level set of ϕ.

In addition, to model the insertion of the curve into the container, we introduce the following characteristic function $\chi^{\rm a}:U\to\{0,1\}$. If $\chi^{\rm a}(\xi)=1$, then the point $\boldsymbol{x}(\xi)$ is considered to have entered the capsid; otherwise, $\chi^{\rm a}(\xi)=0$, and the point $\boldsymbol{x}(\xi)$ has not yet entered the capsid. In other words, the capsid constraint is only enforced at $\boldsymbol{x}(\xi)$ if $\chi^{\rm a}(\xi)=1$.

With the above, we rewrite (32) and define

$$(33)\quad \mathcal{C}(\tilde{\boldsymbol{x}}):=\left\{\boldsymbol{y}\in[H^2]^3\mid\varepsilon/2+\phi(\tilde{\boldsymbol{x}}(\xi))+\nabla\phi(\tilde{\boldsymbol{x}}(\xi))\cdot(\boldsymbol{y}(\xi)-\tilde{\boldsymbol{x}}(\xi))\leq 0,\ \forall\,\xi\in U,\ \text{such that }\chi^{\rm a}(\xi)=1\right\}.$$

3.2.2. *Properties.* For a given $\boldsymbol{y} \in [H^2]^3$, let $\gamma(\boldsymbol{y}) : U \to \mathbb{R}$ be the *capsid contact function* defined by $\gamma(\boldsymbol{y})(\xi) := \varepsilon/2 + \phi(\boldsymbol{y}(\xi))$. Next, define the operator $\widetilde{\Lambda} \equiv \Lambda^{\tilde{\boldsymbol{x}}} : [H^2]^3 \to H^2$ by $\widetilde{\Lambda}\boldsymbol{y} := \boldsymbol{y} \cdot \nabla\phi(\tilde{\boldsymbol{x}})$, for all $\boldsymbol{y} \in [H^2]^3$, with adjoint $\widetilde{\Lambda}^* : (H^2)^* \to ([H^2]^3)^*$ defined by $\widetilde{\Lambda}^*\zeta := \zeta\nabla\phi(\tilde{\boldsymbol{x}})$, for all $\zeta \in (H^2)^*$. Moreover, for a given curve $\tilde{\boldsymbol{x}}$, we have the mapping $\gamma_{\tilde{\boldsymbol{x}}}(\cdot) \in \mathcal{L}([H^2(U)]^3, H^2)$ given by

$$\tilde{\gamma}(\boldsymbol{x}) \equiv \gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}) = \varepsilon/2 + \phi(\tilde{\boldsymbol{x}}) + \widetilde{\Lambda}(\boldsymbol{x} - \tilde{\boldsymbol{x}}), \qquad \tilde{\gamma}'(\boldsymbol{y}) \equiv \gamma'_{\tilde{\boldsymbol{x}}}(\boldsymbol{y}) = \widetilde{\Lambda}\boldsymbol{y}, \tag{34}$$

which, by Taylor expansion, satisfies

$$\gamma(\boldsymbol{x}) = \gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}) + \frac{1}{2}(\boldsymbol{x} - \tilde{\boldsymbol{x}})^T\nabla\nabla\phi(\hat{\boldsymbol{x}})(\boldsymbol{x} - \tilde{\boldsymbol{x}}), \quad \text{for all } \tilde{\boldsymbol{x}}, \boldsymbol{x} \in [H^2]^3, \tag{35}$$

where $\hat{\boldsymbol{x}} = \eta\boldsymbol{x} + (1-\eta)\tilde{\boldsymbol{x}}$, for some $\eta \in (0,1)$. Now define the closed convex cone $K_\gamma \subset H^2$ and its dual cone:

$$K_\gamma = \{z \in H^2 \mid z(\xi) \le 0, \text{ if } \chi^{\mathrm{a}}(\xi) = 1\}, \quad K_\gamma^* = \{\zeta \in (H^2)^* \mid \langle \zeta, z\rangle \ge 0, \ \forall z \in -K_\gamma\}, \tag{36}$$

which allows us to write the constraint $\boldsymbol{x} \in \mathcal{C}(\tilde{\boldsymbol{x}})$ as $\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}) \in K_\gamma$.

4. Optimality Conditions

The well-posedness of the minimization problems in Algorithm 1 is now established, starting with existence of a minimizer. We then develop first order optimality, or Karush-Kuhn-Tucker (KKT), conditions involving Lagrange multipliers to enforce the equality and inequality constraints for both steps in Algorithm 1.

4.1. **Minimization for $\boldsymbol{x}$.** We consider the ith iteration of Step 1 of Algorithm 1 for fixed i. To simplify notation, we set $\tilde{\boldsymbol{x}} := \boldsymbol{x}^i$ and write $\boldsymbol{x} := \boldsymbol{x}^{i+1}$. Solving Step 1 is as follows. Define the closed convex cone $\boldsymbol{K} = \{0\} \times K_\gamma \times K_\psi \subset H_0^1 \times H^2 \times T$, so then $\boldsymbol{K}^* = H^{-1} \times K_\gamma^* \times K_\psi^*$. Thus, given $\tilde{\boldsymbol{x}} \in W(\boldsymbol{x}_{\mathrm{bc}})$ and $\tilde{\boldsymbol{b}} := \boldsymbol{b}^i \in V(\boldsymbol{b}_{\mathrm{bc}})$, we seek a solution $\boldsymbol{x}$ that satisfies

$$\inf_{\boldsymbol{x} \in W(\boldsymbol{x}_{\mathrm{bc}})} \mathcal{J}_{\mathrm{r}}\left[\boldsymbol{x}, \tilde{\boldsymbol{b}}; \tilde{\boldsymbol{x}}, \tilde{\boldsymbol{b}}\right], \text{ subject to } (\tau_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}), \gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}), \psi_{\tilde{\boldsymbol{x}}}(\boldsymbol{x})) \in \boldsymbol{K}. \tag{37}$$

It is clear that $\mathcal{J}_{\mathrm{r}}\left[\boldsymbol{x}, \tilde{\boldsymbol{b}}; \tilde{\boldsymbol{x}}, \tilde{\boldsymbol{b}}\right]$ is convex with respect to $\boldsymbol{x}$ and coercive. Moreover, the constraints $\tau_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}) = 0$, $\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}) \in K_\gamma$, $\psi_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}) \in K_\psi$, are all convex with respect to $\boldsymbol{x}$. To avoid issues at the boundary point $\xi = 0$, we assume that $\boldsymbol{x}_{\mathrm{bc}}(0)$ is strictly inside the capsid, i.e. $\gamma(\boldsymbol{x}_{\mathrm{bc}})(0) < 0$ or $\chi^{\mathrm{a}}(0) = 0$. So existence is trivial as long as the admissible set $G(\tilde{\boldsymbol{x}}) := \{\boldsymbol{x} \in W(\boldsymbol{x}_{\mathrm{bc}}) \mid \tau_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}) = 0, \text{ on } U, \ \gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}) \in K_\gamma, \ \psi_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}) \in K_\psi\}$ is non-empty. We have the following.

Proposition 4.1. *Suppose $G(\tilde{\boldsymbol{x}})$ is non-empty. Then there exists a unique minimizer of* (37).

Furthermore, by the Frechét differentiability of $\mathcal{J}_{\mathrm{r}}\left[\cdot, \tilde{\boldsymbol{b}}; \tilde{\boldsymbol{x}}, \tilde{\boldsymbol{b}}\right]$, the unique minimizer $\boldsymbol{x} \in G(\tilde{\boldsymbol{x}})$ of (37) is the unique solution of the variational inequality

$$c_{\mathrm{b}}(\boldsymbol{x}'', \boldsymbol{y}'' - \boldsymbol{x}'')_{L^2} + m[\tilde{\boldsymbol{b}}](\boldsymbol{x}, \boldsymbol{y} - \boldsymbol{x}) + \frac{1}{\Delta t}(\boldsymbol{x} - \tilde{\boldsymbol{x}}, \boldsymbol{y} - \boldsymbol{x})_\star \ge \mathcal{F}_{\mathrm{x}}(\boldsymbol{y} - \boldsymbol{x}), \ \forall \boldsymbol{y} \in G(\tilde{\boldsymbol{x}}). \tag{38}$$

4.1.1. *First and Second Order Conditions.* Deriving suitable KKT conditions for the minimizer of (37) requires more regularity than just non-emptiness of $G(\tilde{\boldsymbol{x}})$. To this end, we require the space

$$T_0 := \{f \in T \mid f(0,\xi) = f(\xi,0) = 0, \text{ for all } \xi \in U\}, \tag{39}$$

which is compatible with the boundary condition in $W(\boldsymbol{0})$. The following definition, taken from [34, Defn. 1.5], provides a sufficient condition to obtain KKT conditions.

Definition 1 (regular point)**.** *An element $\bar{\boldsymbol{x}} \in G(\tilde{\boldsymbol{x}})$ satisfies the* regular point condition *if*

$$(\tau'_{\tilde{\boldsymbol{x}}}(\cdot), \gamma'_{\tilde{\boldsymbol{x}}}(\cdot), \psi'_{\tilde{\boldsymbol{x}}}(\cdot))^T W(\boldsymbol{0}) - \boldsymbol{K}(\tau_{\tilde{\boldsymbol{x}}}(\bar{\boldsymbol{x}}), \gamma_{\tilde{\boldsymbol{x}}}(\bar{\boldsymbol{x}}), \psi_{\tilde{\boldsymbol{x}}}(\bar{\boldsymbol{x}})) = H_0^1 \times H^2 \times T_0, \tag{40}$$

where $\boldsymbol{K}(\tau_{\tilde{\boldsymbol{x}}}(\bar{\boldsymbol{x}}), \gamma_{\tilde{\boldsymbol{x}}}(\bar{\boldsymbol{x}}), \psi_{\tilde{\boldsymbol{x}}}(\bar{\boldsymbol{x}})) = \{a(0, y - \gamma_{\tilde{\boldsymbol{x}}}(\bar{\boldsymbol{x}}), w - \psi_{\tilde{\boldsymbol{x}}}(\bar{\boldsymbol{x}})) \mid y \in K_\gamma, \ w \in K_\psi, \ a \ge 0\}$. This condition is equivalent to the following set of conditions: $\tau'_{\tilde{\boldsymbol{x}}}(\cdot) : W(\boldsymbol{0}) \to H_0^1$ is surjective and there exists a $\boldsymbol{w} \in W(\boldsymbol{0})$ such that

(i) $\tau'_{\tilde{\boldsymbol{x}}}(\boldsymbol{w}) = 0$ *on* U*;*
(ii) $\gamma'_{\tilde{\boldsymbol{x}}}(\boldsymbol{w})(\xi) < 0$ *if* $\gamma_{\tilde{\boldsymbol{x}}}(\bar{\boldsymbol{x}})(\xi) = 0$*, for* $\xi \in U$ *such that* $\chi^{\mathrm{a}}(\xi) = 1$*;*

(iii) $\psi'_{\tilde{\boldsymbol{x}}}(\boldsymbol{w})(\xi_1,\xi_2) < 0$ *if* $\psi_{\tilde{\boldsymbol{x}}}(\bar{\boldsymbol{x}})(\xi_1,\xi_2) = 0$*, for* $(\xi_1,\xi_2) \in U^2$ *such that* $\chi^{\mathrm{n}}(\xi_1,\xi_2) = 1$.

Surjectivity was shown in Proposition 2.6; existence of a $\boldsymbol{w}$ such that (i)-(iii) are satisfied is guaranteed by the following assumption.

Assumption 4.2. *The set* $G(\tilde{\boldsymbol{x}})$ *is non-empty and contains a* strictly feasible *point, i.e. there exists a point* $\bar{\boldsymbol{x}} \in W(\boldsymbol{x}_{\mathrm{bc}})$ *such that* $\tau_{\tilde{\boldsymbol{x}}}(\bar{\boldsymbol{x}}) = 0$*, on* U*,* $\psi_{\tilde{\boldsymbol{x}}}(\bar{\boldsymbol{x}})(\xi_1,\xi_2) < 0$*, for all* (ξ_1,ξ_2) *in* U^2 *with* $|\xi_1 - \xi_2| \geq \pi\varepsilon/2$*, and* $\gamma_{\tilde{\boldsymbol{x}}}(\bar{\boldsymbol{x}})(\xi) < 0$*, for all* ξ *in* U *with* $\chi^{\mathrm{a}}(\xi) = 1$.

Deducing sufficient conditions to ensure a strictly feasible point is not at all obvious. This would require a detailed characterization of the properties of the self-contact set, as well as the capsid contact set, while simultaneously respecting the equality constraint $\tau_{\tilde{\boldsymbol{x}}}(\bar{\boldsymbol{x}}) = 0$. This appears to be out of reach for general, tightly packed curves, though special cases can be explicitly characterized. In [52, Thm. 1], they list certain transversality conditions to be satisfied by minimizers under the self-contact constraint, i.e. the self-contact set cannot be completely general (see [52, Remarks, pg. 51]). Indeed, for an extremely tightly packed curve, there may *not* be strict feasibility.

Remark 4.3. *In the case of the self-contact constraint only, strict feasibility is straightforward (see* [49]*). Suppose that* $\tilde{\boldsymbol{x}}$ *is feasible. Then* $\psi_{\tilde{\boldsymbol{x}}}(\tilde{\boldsymbol{x}}) \leq 0$ *and* $\psi'_{\tilde{\boldsymbol{x}}}(\tilde{\boldsymbol{x}}) \leq -\varepsilon < 0$ *on* U^2*. Next, assume* $\psi_{\tilde{\boldsymbol{x}}}(\tilde{\boldsymbol{x}})(0,\xi_2) < -2\delta_0$*, for* $\xi_2 \in [0,\delta_0)$*, for some* $\delta_0 > 0$*. Let* $\varphi \in C^\infty(U)$ *such that* $\varphi(0) = 0$ *and* $\varphi = 1$*, on* $[\delta_0, L]$.

Now set $\bar{\boldsymbol{x}} = \tilde{\boldsymbol{x}} + \delta_1\varphi\tilde{\boldsymbol{x}}$ *and note* $\bar{\boldsymbol{x}} \in W(\boldsymbol{x}_{\mathrm{bc}})$ *because* $\varphi\tilde{\boldsymbol{x}} \in W(\boldsymbol{0})$*. Thus, we obtain* $\psi_{\tilde{\boldsymbol{x}}}(\bar{\boldsymbol{x}}) = \psi_{\tilde{\boldsymbol{x}}}(\tilde{\boldsymbol{x}}) + \delta_1\psi'_{\tilde{\boldsymbol{x}}}(\varphi\tilde{\boldsymbol{x}}) < 0$ *on* U^2*, for some* $\delta_1 > 0$ *sufficiently small depending on* δ_0*. However,* $\tau_{\tilde{\boldsymbol{x}}}(\bar{\boldsymbol{x}}) \neq 0$.

We introduce a Lagrangian functional for the minimization problem (37). Note that $\mathcal{J}_{\mathrm{r}}\left[\cdot,\tilde{\boldsymbol{b}};\tilde{\boldsymbol{x}},\tilde{\boldsymbol{b}}\right]$ and $\tau_{\tilde{\boldsymbol{x}}}(\cdot)$, $\psi_{\tilde{\boldsymbol{x}}}(\cdot)$, and $\gamma_{\tilde{\boldsymbol{x}}}(\cdot)$ are twice continuously Fréchet differentiable. Introducing multipliers $\rho := \rho^{i+1} \in H^{-1}$, $\lambda := \lambda^{i+1} \in (H^2)^*$, $\sigma := \sigma^{i+1} \in T_0^*$, for any fixed $(\tilde{\boldsymbol{x}},\tilde{\boldsymbol{b}}) \in W(\boldsymbol{x}_{\mathrm{bc}}) \times V(\boldsymbol{b}_{\mathrm{bc}})$, we define the Lagrangian functional $\mathcal{L}_{\mathrm{x}} : [H^2]^3 \times H^{-1} \times (H^2)^* \times T^* \to \mathbb{R}$ by

$$\mathcal{L}_{\mathrm{x}}[\boldsymbol{x},\rho,\lambda,\sigma] := \mathcal{J}_{\mathrm{r}}\left[\boldsymbol{x},\tilde{\boldsymbol{b}};\tilde{\boldsymbol{x}},\tilde{\boldsymbol{b}}\right] + \langle\rho,\tau_{\tilde{\boldsymbol{x}}}(\boldsymbol{x})\rangle + \langle\lambda,\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x})\rangle + \langle\langle\sigma,\psi_{\tilde{\boldsymbol{x}}}(\boldsymbol{x})\rangle\rangle\,. \tag{41}$$

We have the following first order optimality conditions.

Proposition 4.4 (KKT conditions for (37))**.** *Adopt Assumption 4.2. Then, there exists an* $\boldsymbol{x} \in W(\boldsymbol{x}_{\mathrm{bc}})$ *that is the unique minimizer of* (37)*, with* $(\tau_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}), \gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}), \psi_{\tilde{\boldsymbol{x}}}(\boldsymbol{x})) \in \boldsymbol{K}$*, i.e.* $\tau_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}) = 0$*,* $\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}) \leq 0$ *on* U*, and* $\psi_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}) \leq 0$ *on* U^2*. Furthermore, there exist Lagrange multipliers* $(\rho,\lambda,\sigma) \in -\boldsymbol{K}^*$ *such that* $\delta_{\boldsymbol{x}}\mathcal{L}_{\mathrm{x}}[\boldsymbol{x},\rho,\lambda,\sigma](\boldsymbol{w}) = 0$*, for all* $\boldsymbol{w} \in W(\boldsymbol{0})$*, i.e. there exists* $\rho \in H^{-1}$*,* $\lambda \in (H^2)^*$ *with* $\langle\lambda, y\rangle \geq 0$ *for all* $y \in -K_\gamma$*, and* $\sigma \in T_0^*$ *with* $\langle\langle\sigma, z\rangle\rangle \geq 0$ *for all* $z \in -K_\psi$ *such that*

$$\begin{aligned} c_{\mathrm{b}}(\boldsymbol{x}'',\boldsymbol{w}'')_{L^2} + m[\tilde{\boldsymbol{b}}](\boldsymbol{x},\boldsymbol{w}) + \frac{1}{\Delta t}(\boldsymbol{x}-\tilde{\boldsymbol{x}},\boldsymbol{w})_\star \\ + \langle\rho,\tau'_{\tilde{\boldsymbol{x}}}(\boldsymbol{w})\rangle + \langle\lambda,\gamma'_{\tilde{\boldsymbol{x}}}(\boldsymbol{w})\rangle + \langle\langle\sigma,\psi'_{\tilde{\boldsymbol{x}}}(\boldsymbol{w})\rangle\rangle = \mathcal{F}_{\mathrm{x}}(\boldsymbol{w})\,,\ \forall \boldsymbol{w} \in W(\boldsymbol{0}), \end{aligned} \tag{42}$$

with $\tau_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}) = 0$ *on* U *and the following* complementarity conditions

$$\langle\lambda,\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x})\rangle = 0, \quad \langle\langle\sigma,\psi_{\tilde{\boldsymbol{x}}}(\boldsymbol{x})\rangle\rangle = 0, \tag{43}$$

where $\operatorname{supp}\lambda \subset \{\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}) = 0\}$*,* $\operatorname{supp}\sigma \subset \{\psi_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}) = 0\}$.

Proof. By Assumption 4.2, the regular point condition is satisfied. The rest follows by [34, Thm. 1.6]. □

Due to the linearization of the constraints, $\delta^2_{\boldsymbol{x}}\mathcal{L}_{\mathrm{x}}[\boldsymbol{x},\rho,\lambda,\sigma](\boldsymbol{y},\boldsymbol{w}) = c_{\mathrm{b}}(\boldsymbol{y}'',\boldsymbol{w}'')_{L^2} + m[\tilde{\boldsymbol{b}}](\boldsymbol{y},\boldsymbol{w}) + \Delta t^{-1}(\boldsymbol{y},\boldsymbol{w})_\star$, for all $\boldsymbol{y},\boldsymbol{w} \in W(\boldsymbol{0})$, so the second order condition is automatic:

$$\delta^2_{\boldsymbol{x}}\mathcal{L}_{\mathrm{x}}[\boldsymbol{x},\rho,\lambda,\sigma](\boldsymbol{w},\boldsymbol{w}) \geq c_{\mathrm{b}}\|\boldsymbol{w}''\|^2_{L^2},\ \forall \boldsymbol{w} \in W(\boldsymbol{0}). \tag{44}$$

4.2. **Minimization for $\boldsymbol{b}$.** We consider the ith iteration of Step 2 of Algorithm 1 for fixed i. To simplify notation, we set $\tilde{\boldsymbol{b}} := \boldsymbol{b}^i$, $\boldsymbol{x} := \boldsymbol{x}^{i+1}$, and write $\boldsymbol{b} := \boldsymbol{b}^{i+1}$. Solving Step 2 is as follows. Given $\boldsymbol{x} \in W(\boldsymbol{x}_{\rm bc})$ and $\tilde{\boldsymbol{b}} \in V(\boldsymbol{b}_{\rm bc})$, we seek a solution $\boldsymbol{b}$ that solves

$$\inf_{\boldsymbol{b}\in V(\boldsymbol{b}_{\rm bc})} \mathcal{J}_{\rm r}\left[\boldsymbol{x},\boldsymbol{b};\boldsymbol{x},\tilde{\boldsymbol{b}}\right], \text{ subject to } \ell_{\tilde{\boldsymbol{b}}}(\boldsymbol{b}) = 0, \text{ on } U. \tag{45}$$

Since $\mathcal{J}_{\rm r}\left[\boldsymbol{x},\boldsymbol{b};\boldsymbol{x},\tilde{\boldsymbol{b}}\right]$ is convex with respect to $\boldsymbol{b}$ and coercive, and the constraint $\ell_{\tilde{\boldsymbol{b}}}(\boldsymbol{b}) = 0$ is affine, existence is trivial as long as the admissible set $Q(\tilde{\boldsymbol{b}}) := \{\boldsymbol{b} \in V(\boldsymbol{b}_{\rm bc}) \mid \ell_{\tilde{\boldsymbol{b}}}(\boldsymbol{b}) = 0, \text{ on } U\}$ is non-empty.

The regular point condition associated with (45) simply reduces to the surjectivity of $\ell'_{\tilde{\boldsymbol{b}}}(\cdot) : V(\boldsymbol{0}) \to H^1_0$, which was established in Proposition 2.5. Next, we introduce a Lagrangian functional for the minimization problem (45). Note that $\mathcal{J}_{\rm r}\left[\boldsymbol{x},\cdot\,;\boldsymbol{x},\tilde{\boldsymbol{b}}\right]$ and $\ell_{\tilde{\boldsymbol{b}}}(\cdot)$ are twice continuously Fréchet differentiable. Introducing a multiplier $\omega := \omega^{i+1} \in H^{-1}$, for any fixed $(\boldsymbol{x},\tilde{\boldsymbol{b}}) \in W(\boldsymbol{x}_{\rm bc}) \times V(\boldsymbol{b}_{\rm bc})$, we define the Lagrangian functional $\mathcal{L}_{\rm b} : [H^1]^3 \times H^{-1} \to \mathbb{R}$ by

$$\mathcal{L}_{\rm b}[\boldsymbol{b},\omega] := \mathcal{J}_{\rm r}\left[\boldsymbol{x},\boldsymbol{b};\boldsymbol{x},\tilde{\boldsymbol{b}}\right] + \left\langle \omega, \ell_{\tilde{\boldsymbol{b}}}(\boldsymbol{b})\right\rangle. \tag{46}$$

We now obtain the following well-posedness result.

Proposition 4.5. *Suppose $Q(\tilde{\boldsymbol{b}})$ is non-empty. Then there exists a unique minimizer $\boldsymbol{b} \in V(\boldsymbol{b}_{\rm bc})$ of (45). Furthermore, there exists a unique Lagrange multiplier $\omega \in H^{-1}$ such that $\delta_{\boldsymbol{b}}\mathcal{L}_{\rm b}[\boldsymbol{b},\omega](\boldsymbol{v}) = 0$, for all $\boldsymbol{v} \in V(\boldsymbol{0})$, i.e. there exists a unique $\omega \in H^{-1}$ such that*

$$c_{\rm r}(\boldsymbol{b}',\boldsymbol{v}')_{L^2} + m[\boldsymbol{x}](\boldsymbol{b},\boldsymbol{v}) + \frac{1}{\Delta t}\left(\boldsymbol{b}-\tilde{\boldsymbol{b}},\boldsymbol{v}\right)_\dagger + \left\langle \omega, \ell'_{\tilde{\boldsymbol{b}}}(\boldsymbol{v})\right\rangle = \mathcal{F}_{\rm b}(\boldsymbol{v}), \ \forall \boldsymbol{v} \in V(\boldsymbol{0}), \tag{47}$$

with $\ell_{\tilde{\boldsymbol{b}}}(\boldsymbol{b}) = 0$ on U.

Proof. Existence of a unique minimizer follows by the direct method and coercivity. Since the regular point condition is satisfied, the rest follows by [34, Thm. 1.6]. Uniqueness of ω follows by surjectivity of $\ell'_{\tilde{\boldsymbol{b}}}(\cdot)$. □

Because the constraint is affine, $\delta^2_{\boldsymbol{b}}\mathcal{L}_{\rm b}[\boldsymbol{b},\omega](\boldsymbol{z},\boldsymbol{v}) = c_{\rm r}(\boldsymbol{z}',\boldsymbol{v}')_{L^2} + m[\boldsymbol{x}](\boldsymbol{z},\boldsymbol{v}) + \Delta t^{-1}(\boldsymbol{z},\boldsymbol{v})_\dagger$, for all $\boldsymbol{z},\boldsymbol{v} \in V(\boldsymbol{0})$, so the second order condition is automatic:

$$\delta^2_{\boldsymbol{b}}\mathcal{L}_{\rm b}[\boldsymbol{b},\omega](\boldsymbol{v},\boldsymbol{v}) \geq c_{\rm r}\|\boldsymbol{v}'\|^2_{L^2}, \ \forall \boldsymbol{v} \in V(\boldsymbol{0}). \tag{48}$$

4.3. **Energy Estimates.** Energy estimates for the descent scheme essentially follow by the convexity of the alternating minimization steps. To ensure robustness of the scheme, with respect to the parameters, we make some additional assumptions on the data and the descent metric used.

Assumption 4.6. *We take $(\boldsymbol{w},\boldsymbol{y})_\star := (\boldsymbol{w},\boldsymbol{y})_{H^2}$ and $(\boldsymbol{v},\boldsymbol{z})_\dagger := (\boldsymbol{v},\boldsymbol{z})_{H^1}$. Moreover, we assume that $\mathcal{F}_{\rm x}(\cdot) \in (H^2)^*$ and $\mathcal{F}_{\rm b} \equiv 0$.*

Since we derive energy estimates for the sequence generated by Algorithm 1, we make the following notational replacements in Proposition 4.4 (recall Section 4.1): $\boldsymbol{x} \equiv \boldsymbol{x}^{i+1}$, $\tilde{\boldsymbol{x}} \equiv \boldsymbol{x}^i$, $\rho \equiv \rho^{i+1}$, $\lambda \equiv \lambda^{i+1}$, and $\sigma \equiv \sigma^{i+1}$. In addition, We make the following notational replacements in Proposition 4.5 (recall Section 4.2): $\boldsymbol{b} \equiv \boldsymbol{b}^{i+1}$, $\tilde{\boldsymbol{b}} \equiv \boldsymbol{b}^i$, $\omega \equiv \omega^{i+1}$.

In order to obtain energy estimates, we need the following uniform bound on ρ^{i+1} and λ^{i+1} from Proposition 4.4, and on ω^{i+1} from Proposition 4.5.

Assumption 4.7. *For all $i \geq 0$, there holds $\|\rho^{i+1}\|_{H^{-1}} + \left|\left\langle \lambda^{i+1}, 1\right\rangle\right| \leq M_1$, for some fixed constant $M_1 > 0$, and $\|\omega^{i+1}\|_{H^{-1}} \leq M_2$, for some fixed constant $M_2 > 0$.*

We first derive estimates on the violation of the exact constraints. Considering the solution of (42) at time-step $i+1$, we have $\tau_{\boldsymbol{x}^i}\left(\boldsymbol{x}^{i+1}\right) = 0$, on U, which, from (17), gives

$$\begin{aligned} \tau(\boldsymbol{x}^{i+1}) = \frac{1}{2}\left(|(\boldsymbol{x}^{i+1})'|^2 - 1\right) = \tau_{\boldsymbol{x}^i}\left(\boldsymbol{x}^{i+1}\right) + \frac{1}{2}\left|(\boldsymbol{x}^{i+1})' - (\boldsymbol{x}^i)'\right|^2 = \frac{1}{2}\left|(\boldsymbol{x}^{i+1})' - (\boldsymbol{x}^i)'\right|^2 \geq 0, \\ \tau'_{\boldsymbol{x}^i}\left(\boldsymbol{x}^{i+1} - \boldsymbol{x}^i\right) = \tau_{\boldsymbol{x}^i}\left(\boldsymbol{x}^{i+1}\right) - \tau(\boldsymbol{x}^i) = -\frac{1}{2}\left|(\boldsymbol{x}^i)' - (\boldsymbol{x}^{i-1})'\right|^2, \end{aligned} \tag{49}$$

and also implies $|(\boldsymbol{x}^{i+1})'| \geq 1$. Moreover, $\gamma_{\boldsymbol{x}^i}\left(\boldsymbol{x}^{i+1}\right) \leq 0$ on U and $\left\langle \lambda^{i+1}, \gamma_{\boldsymbol{x}^i}\left(\boldsymbol{x}^{i+1}\right)\right\rangle = 0$, so (34), (35) yield

$$
\begin{aligned}
\gamma(\boldsymbol{x}^{i+1}) &\leq \frac{1}{2}(\boldsymbol{x}^{i+1}-\boldsymbol{x}^i)^T \nabla\nabla\phi(\hat{\boldsymbol{x}})(\boldsymbol{x}^{i+1}-\boldsymbol{x}^i), \\
\left\langle \lambda^{i+1}, \gamma'_{\boldsymbol{x}^i}\left(\boldsymbol{x}^{i+1}-\boldsymbol{x}^i\right)\right\rangle &= \left\langle \lambda^{i+1}, \gamma_{\boldsymbol{x}^i}\left(\boldsymbol{x}^{i+1}\right)\right\rangle - \left\langle \lambda^{i+1}, \gamma(\boldsymbol{x}^i)\right\rangle \\
&= \left\langle \lambda^{i+1}, \|[\gamma(\boldsymbol{x}^i)]^+\|_{L^\infty} - \gamma(\boldsymbol{x}^i)\right\rangle - \left\langle \lambda^{i+1}, \|[\gamma(\boldsymbol{x}^i)]^+\|_{L^\infty}\right\rangle \\
&\geq -\|[\gamma(\boldsymbol{x}^i)]^+\|_{L^\infty}\left\langle \lambda^{i+1}, 1\right\rangle.
\end{aligned} \tag{50}
$$

In addition, $\psi_{\boldsymbol{x}^i}\left(\boldsymbol{x}^{i+1}\right) \leq 0$ on U^2 and $\left\langle\left\langle \sigma^{i+1}, \psi_{\boldsymbol{x}^i}\left(\boldsymbol{x}^{i+1}\right)\right\rangle\right\rangle = 0$, which, from (28), gives

$$
\begin{aligned}
\psi(\boldsymbol{x}^{i+1}) &\leq \psi_{\boldsymbol{x}^i}\left(\boldsymbol{x}^{i+1}\right) \leq 0, \\
\left\langle\left\langle \sigma^{i+1}, \psi'_{\boldsymbol{x}^i}\left(\boldsymbol{x}^{i+1}-\boldsymbol{x}^i\right)\right\rangle\right\rangle &= \left\langle\left\langle \sigma^{i+1}, \psi_{\boldsymbol{x}^i}\left(\boldsymbol{x}^{i+1}\right)\right\rangle\right\rangle - \left\langle\left\langle \sigma^{i+1}, \psi(\boldsymbol{x}^i)\right\rangle\right\rangle \geq 0.
\end{aligned} \tag{51}
$$

Furthermore, the solution of (47) at time-step $i+1$ satisfies $\ell_{\boldsymbol{b}^i}\left(\boldsymbol{b}^{i+1}\right) = 0$ on U, which gives

$$
\begin{aligned}
\ell(\boldsymbol{b}^{i+1}) = \frac{1}{2}\left(|\boldsymbol{b}^{i+1}|^2 - 1\right) = \ell_{\boldsymbol{b}^i}\left(\boldsymbol{b}^{i+1}\right) + \frac{1}{2}\left|\boldsymbol{b}^{i+1}-\boldsymbol{b}^i\right|^2 = \frac{1}{2}\left|\boldsymbol{b}^{i+1}-\boldsymbol{b}^i\right|^2 \geq 0, \\
\ell'_{\boldsymbol{b}^i}\left(\boldsymbol{b}^{i+1}-\boldsymbol{b}^i\right) = \ell_{\boldsymbol{b}^i}\left(\boldsymbol{b}^{i+1}\right) - \ell(\boldsymbol{b}^i) = -\frac{1}{2}\left|\boldsymbol{b}^i-\boldsymbol{b}^{i-1}\right|^2,
\end{aligned} \tag{52}
$$

and also implies $|\boldsymbol{b}^{i+1}| \geq 1$.

From the above, there is a uniform constant $v_0 > 0$, that depends on L, embedding constants, and $\|\nabla\nabla\phi\|_{L^\infty}$, such that the violation of the constraints can be estimated by

$$
\|\tau(\boldsymbol{x}^i)\|_{H^1} \leq v_0\|\boldsymbol{x}^i-\boldsymbol{x}^{i-1}\|_{H^2}^2, \quad \|[\gamma(\boldsymbol{x}^i)]^+\|_{L^\infty} \leq v_0\|\boldsymbol{x}^i-\boldsymbol{x}^{i-1}\|_{H^2}^2, \quad \|\ell(\boldsymbol{b}^i)\|_{H^1} \leq v_0\|\boldsymbol{b}^i-\boldsymbol{b}^{i-1}\|_{H^1}^2. \tag{53}
$$

Note that $\psi(\boldsymbol{x}^i) \leq 0$ for all $i \geq 0$ by the monotone property (28). We have the following theorem.

Theorem 4.8. *Suppose Assumptions 4.6 and 4.7 hold, and define* $\Xi_1^i := M_1(\|\tau(\boldsymbol{x}^i)\|_{H^1} + \|[\gamma(\boldsymbol{x}^i)]^+\|_{L^\infty})$, $\Xi_2^i := M_2\|\ell(\boldsymbol{b}^i)\|_{H^1}$. *If the time-step* Δt *is sufficiently small, i.e. if* $\Delta t \leq (2v_0M_1)^{-1}$ *and* $\Delta t \leq (v_0M_2)^{-1}$, *then we have the following (quasi)-energy decrease property*

$$
\mathcal{E}_{\mathrm{r}}\left[\boldsymbol{x}^{i+1}, \boldsymbol{b}^{i+1}\right] + \Xi_1^{i+1} + \Xi_2^{i+1} \leq \mathcal{E}_{\mathrm{r}}\left[\boldsymbol{x}^i, \boldsymbol{b}^i\right] + \Xi_1^i + \Xi_2^i, \quad \textit{for all } i \geq 0, \tag{54}
$$

i.e. $E^i + \Xi_1^i + \Xi_2^i \leq E^0 + \Xi_1^0 + \Xi_2^0 =: F^0$, *for all* $i \geq 0$, *where* $E^i := \mathcal{E}_{\mathrm{r}}\left[\boldsymbol{x}^i, \boldsymbol{b}^i\right]$. *Moreover, for all* $i \geq 0$, *we have the a priori estimate*

$$
\begin{aligned}
(c_{\mathrm{b}}/4)\|(\boldsymbol{x}^{i+1})''\|_{L^2}^2 + (c_{\mathrm{r}}/2)\|(\boldsymbol{b}^{i+1})'\|_{L^2}^2 + \Delta t^{-1}\|\boldsymbol{x}^{i+1}-\boldsymbol{x}^i\|_{H^2}^2 + \Delta t^{-1}\|\boldsymbol{b}^{i+1}-\boldsymbol{b}^i\|_{H^1}^2 \\
\leq E^i + \Xi_1^i + \Xi_2^i + Q_0 \leq F^0 + Q_0,
\end{aligned} \tag{55}
$$

where $Q_0 := (C_P^2/c_{\mathrm{b}})\|\mathcal{F}_{\mathrm{x}}\|_{(H^2)^*}^2 + \|\mathcal{F}_{\mathrm{x}}\|_{(H^2)^*}[C_P\|\boldsymbol{x}''_{\mathrm{bc}}\|_{L^2} + \|\boldsymbol{x}_{\mathrm{bc}}\|_{H^2}]$, *and* $C_P > 0$ *is the Poincaré constant in the estimate* $\|\boldsymbol{y}\|_{H^2} \leq C_P\|\boldsymbol{y}''\|_{L^2}$, *for all* $\boldsymbol{y} \in W(\boldsymbol{0})$. *In addition, we have*

$$
\|\boldsymbol{x}^{i+1}-\boldsymbol{x}^i\|_{H^2} \leq R_1\Delta t, \quad \|\boldsymbol{b}^{i+1}-\boldsymbol{b}^i\|_{H^1} \leq R_2\Delta t, \quad \textit{for all } i \geq 0, \tag{56}
$$

where R_1, R_2 *are uniform constants that do not depend on the final time. Lastly, the violation of the unit length constraints, and the capsid constraint, are estimated by*

$$
\|\tau(\boldsymbol{x}^i)\|_{H^1} \leq v_0R_1^2\Delta t^2, \quad \|[\gamma(\boldsymbol{x}^i)]^+\|_{L^\infty} \leq v_0R_1^2\Delta t^2, \quad \|\ell(\boldsymbol{b}^i)\|_{H^1} \leq v_0R_2^2\Delta t^2, \quad \textit{for all } i \geq 0. \tag{57}
$$

The self-contact constraint is always satisfied, $\psi(\boldsymbol{x}^i) \leq 0$ *for all* $i \geq 0$.

The proof is given in Appendix A. Note that the violation of the constraint estimates hold for any $i \geq 0$ and do not depend on the final time.

5. Regularization

If the multipliers $\lambda \in (H^2)^*$, $\sigma \in T_0^*$ happen to be more regular, say $\lambda \in L^2(U)$, $\sigma \in L^2(U^2)$, then λ, σ satisfy

$$\lambda \geq 0,\ \lambda = \max\left(0, \lambda + c\,\gamma_{\tilde{\boldsymbol{x}}}\left(\boldsymbol{x}\right)\right)\chi^{\mathrm{a}},\ \text{a.e. in } U,\ \text{and } \sigma \geq 0,\ \sigma = \max\left(0, \sigma + c\,\psi_{\tilde{\boldsymbol{x}}}\left(\boldsymbol{x}\right)\right)\chi^{\mathrm{n}},\ \text{a.e. in } U^2, \tag{58}$$

where $c > 0$ is a constant. In this case, one could use a primal-dual active set strategy, or semi-smooth Newton scheme, to solve the first order conditions [33, 27, 28, 29, 34].

5.1. Regularization of the Inequality Constraints.

Therefore, we propose a regularization of the first order conditions (see [34]). Given $\tilde{\boldsymbol{x}} \in W(\boldsymbol{x}_{\mathrm{bc}})$, $\tilde{\boldsymbol{b}} \in V(\boldsymbol{b}_{\mathrm{bc}})$, we seek to find $\boldsymbol{x}_c \in W(\boldsymbol{x}_{\mathrm{bc}})$, $\rho_c \in H^{-1}$, $\lambda_c \in L^2$, and $\sigma_c \in L^2(U^2)$ such that

$$\begin{aligned}
c_{\mathrm{b}}\left(\boldsymbol{x}_c'', \boldsymbol{w}''\right)_{L^2} + m[\tilde{\boldsymbol{b}}]\left(\boldsymbol{x}_c, \boldsymbol{w}\right) + \frac{1}{\Delta t}\left(\boldsymbol{x}_c - \tilde{\boldsymbol{x}}, \boldsymbol{w}\right)_\star & \\
+ \left\langle \rho_c, \tau'_{\tilde{\boldsymbol{x}}}\left(\boldsymbol{w}\right)\right\rangle + \left(\lambda_c, \gamma'_{\tilde{\boldsymbol{x}}}\left(\boldsymbol{w}\right)\right) + \left(\left(\sigma_c, \psi'_{\tilde{\boldsymbol{x}}}\left(\boldsymbol{w}\right)\right)\right) &= \mathcal{F}_{\mathrm{x}}\left(\boldsymbol{w}\right),\ \forall \boldsymbol{w} \in W(\boldsymbol{0}), \\
\left\langle \eta, \tau_{\tilde{\boldsymbol{x}}}\left(\boldsymbol{x}_c\right)\right\rangle &= 0,\ \forall \eta \in H^{-1}, \\
\left(\varphi, \max\left(0, \bar{\lambda} + c\,\gamma_{\tilde{\boldsymbol{x}}}\left(\boldsymbol{x}_c\right)\right)\chi^{\mathrm{a}}\right) - \left(\varphi, \lambda_c\right) &= 0,\ \forall \varphi \in L^2, \\
\left(\left(\mu, \max\left(0, \bar{\sigma} + c\,\psi_{\tilde{\boldsymbol{x}}}\left(\boldsymbol{x}_c\right)\right)\chi^{\mathrm{n}}\right)\right) - \left(\left(\mu, \sigma_c\right)\right) &= 0,\ \forall \mu \in L^2(U^2),
\end{aligned} \tag{59}$$

where $c > 0$ is a regularization parameter and $\bar{\lambda} \in L^2(U)$, $\bar{\sigma} \in L^2(U^2)$ are fixed, given non-negative functions that may be zero. A Newton scheme is described in Section 5.2 for solving this system.

Lemma 5.1. *The solution of* (59) *exists and is unique.*

Proof. By eliminating λ_c and σ_c, (59) is equivalent to

$$\begin{aligned}
c_{\mathrm{b}}\left(\boldsymbol{x}_c'', \boldsymbol{w}''\right)_{L^2} + m[\tilde{\boldsymbol{b}}]\left(\boldsymbol{x}_c, \boldsymbol{w}\right) + \frac{1}{\Delta t}\left(\boldsymbol{x}_c - \tilde{\boldsymbol{x}}, \boldsymbol{w}\right)_\star + \left(\max\left(0, \bar{\lambda} + c\,\gamma_{\tilde{\boldsymbol{x}}}\left(\boldsymbol{x}_c\right)\right)\chi^{\mathrm{a}}, \gamma'_{\tilde{\boldsymbol{x}}}\left(\boldsymbol{w}\right)\right) & \\
+ \left(\left(\max\left(0, \bar{\sigma} + c\,\psi_{\tilde{\boldsymbol{x}}}\left(\boldsymbol{x}_c\right)\right)\chi^{\mathrm{n}}, \psi'_{\tilde{\boldsymbol{x}}}\left(\boldsymbol{w}\right)\right)\right) + \left\langle \rho_c, \tau'_{\tilde{\boldsymbol{x}}}\left(\boldsymbol{w}\right)\right\rangle &= \mathcal{F}_{\mathrm{x}}\left(\boldsymbol{w}\right),\ \forall \boldsymbol{w} \in W(\boldsymbol{0}), \\
\left\langle \eta, \tau_{\tilde{\boldsymbol{x}}}\left(\boldsymbol{x}_c\right)\right\rangle &= 0,\ \forall \eta \in H^{-1},
\end{aligned} \tag{60}$$

which is the first order condition corresponding to the following convex minimization problem

$$\begin{aligned}
\mathcal{J}_{\mathrm{r},c}\left[\boldsymbol{x}_c\right] := \mathcal{J}_{\mathrm{r}}\left[\boldsymbol{x}_c, \tilde{\boldsymbol{b}}; \tilde{\boldsymbol{x}}, \tilde{\boldsymbol{b}}\right] + \frac{1}{2c}\left\|\max\left(0, \bar{\lambda} + c\,\gamma_{\tilde{\boldsymbol{x}}}\left(\boldsymbol{x}_c\right)\right)\chi^{\mathrm{a}}\right\|^2_{L^2(U)} + \frac{1}{2c}\left\|\max\left(0, \bar{\sigma} + c\,\psi_{\tilde{\boldsymbol{x}}}\left(\boldsymbol{x}_c\right)\right)\chi^{\mathrm{n}}\right\|^2_{L^2(U^2)}, & \\
\min_{\boldsymbol{x}_c \in W(\boldsymbol{x}_{\mathrm{bc}})} \mathcal{J}_{\mathrm{r},c}\left[\boldsymbol{x}_c\right],\ \text{such that } \tau_{\tilde{\boldsymbol{x}}}\left(\boldsymbol{x}_c\right) = 0,\ \text{on } U. &
\end{aligned} \tag{61}$$

Since $\mathcal{J}_{\mathrm{r},c}$ is continuously Frechét differentiable and coercive, there is a unique minimizer. Moreover, by Proposition 2.6, the Lagrange multiplier $\rho_c \in H^{-1}$ is also unique. Lastly, λ_c and σ_c are determined uniquely through the last two equations of (59). □

Theorem 5.2. *The solution* $(\boldsymbol{x}_c, \rho_c, \lambda_c, \sigma_c) \in W(\boldsymbol{x}_{\mathrm{bc}}) \times H^{-1} \times L^2 \times L^2(U^2)$ *of* (59) *converges to a solution* $(\boldsymbol{x}^*, \rho^*, \lambda^*, \sigma^*) \in W(\boldsymbol{x}_{\mathrm{bc}}) \times H^{-1} \times (H^2)^* \times T_0^*$ *of* (42), (43)*, in the sense that* $\boldsymbol{x}_c \to \boldsymbol{x}^*$ *strongly in* H^2 *(where* $\boldsymbol{x}^*$ *is unique), and*

$$\lim_{c \to \infty} \left\langle \rho_c - \rho^*, \tau'_{\tilde{\boldsymbol{x}}}\left(\boldsymbol{w}\right)\right\rangle + \left\langle \lambda_c - \lambda^*, \gamma'_{\tilde{\boldsymbol{x}}}\left(\boldsymbol{w}\right)\right\rangle + \left\langle\left\langle \sigma_c - \sigma^*, \psi'_{\tilde{\boldsymbol{x}}}\left(\boldsymbol{w}\right)\right\rangle\right\rangle = 0,\ \forall \boldsymbol{w} \in W(\boldsymbol{0}). \tag{62}$$

Proof. Note that $(\boldsymbol{x}_c, \rho_c, \lambda_c, \sigma_c)$ satisfies (59) and

$$\max\left(0, \bar{\lambda} + c\,\gamma_{\tilde{\boldsymbol{x}}}\left(\boldsymbol{x}_c\right)\right)\chi^{\mathrm{a}} = \lambda_c, \qquad \max\left(0, \bar{\sigma} + c\,\psi_{\tilde{\boldsymbol{x}}}\left(\boldsymbol{x}_c\right)\right)\chi^{\mathrm{n}} = \sigma_c. \tag{63}$$

For any $\boldsymbol{y} \in G(\tilde{\boldsymbol{x}})$, we have

$$
\begin{aligned}
&(\lambda_c, \gamma'_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c - \boldsymbol{y})) + ((\sigma_c, \psi'_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c - \boldsymbol{y}))) = \\
&\left(\lambda_c, \varepsilon/2 + \phi(\tilde{\boldsymbol{x}})|\nabla\phi(\tilde{\boldsymbol{x}})|^{-1} + \gamma'_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c - \tilde{\boldsymbol{x}}) - \left[\varepsilon/2 + \phi(\tilde{\boldsymbol{x}})|\nabla\phi(\tilde{\boldsymbol{x}})|^{-1} + \gamma'_{\tilde{\boldsymbol{x}}}(\boldsymbol{y} - \tilde{\boldsymbol{x}})\right]\right) \\
&\qquad + ((\sigma_c, \varepsilon + \psi'_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c) - [\varepsilon + \psi'_{\tilde{\boldsymbol{x}}}(\boldsymbol{y})])) = \\
&(\lambda_c, \gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)) + ((\sigma_c, \psi_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c))) - (\lambda_c, \gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{y})) - ((\sigma_c, \psi_{\tilde{\boldsymbol{x}}}(\boldsymbol{y}))) \geq (\lambda_c, \gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)) + ((\sigma_c, \psi_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c))) \\
&\qquad = (\lambda_c, \gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)) + ((\sigma_c, \psi_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)))
\end{aligned} \tag{64}
$$

since σ_c, λ_c are non-negative. Next, we estimate $(\lambda_c, \gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c))$, $((\sigma_c, \psi_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)))$ separately. First, writing $[f]^+ \equiv \max(0, f)$ and $[f]^- \equiv \max(0, -f)$, we have $f = [f]^+ - [f]^-$ and $|f| = [f]^+ + [f]^-$. Let $\xi \in U$ such that $\chi^{\mathrm{a}}(\xi) = 1$. Since $\lambda_c \geq 0$, $\lambda_c \geq \bar{\lambda} + c\,\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)$, and $\bar{\lambda} \geq 0$, we have

$$
\begin{aligned}
\lambda_c(\xi)\,\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)(\xi) &= \lambda_c(\xi)[\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)]^+(\xi) - \lambda_c(\xi)[\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)]^-(\xi) \\
&\geq \left(\bar{\lambda}(\xi) + c\,\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)(\xi)\right)[\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)]^+(\xi) - \lambda_c(\xi)[\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)]^-(\xi) \\
&= \bar{\lambda}(\xi)[\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)]^+(\xi) + c\left([\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)]^+(\xi)\right)^2 - \lambda_c(\xi)[\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)]^-(\xi).
\end{aligned} \tag{65}
$$

Either $\lambda_c(\xi)[\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)]^-(\xi) = 0$ or > 0. Suppose ξ is such that $\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)(\xi) < 0$. If $\lambda_c(\xi) = \bar{\lambda}(\xi) + c\,\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)(\xi) > 0$, then $\bar{\lambda}(\xi) > -c\,\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)(\xi) = c[\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)]^-(\xi) > 0$. This means that

$$
\begin{aligned}
\lambda_c(\xi)[\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)]^-(\xi) &= \left(\bar{\lambda}(\xi) + c\,\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)(\xi)\right)[\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)]^-(\xi) \\
&= \frac{\bar{\lambda}(\xi)}{c} c[\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)]^-(\xi) + c[\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)]^+(\xi)[\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)]^-(\xi) - c\left([\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)]^-(\xi)\right)^2 \\
&< \frac{(\bar{\lambda}(\xi))^2}{c}.
\end{aligned} \tag{66}
$$

Thus, combining with (65), we obtain

$$
\lambda_c\,\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c) \geq \bar{\lambda}[\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)]^+ + c\left([\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)]^+\right)^2 - \frac{\bar{\lambda}^2}{c} \geq \frac{c}{2}\left([\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)]^+\right)^2 - \frac{3\bar{\lambda}^2}{2c}, \text{ on } U. \tag{67}
$$

Similarly, we arrive at

$$
\sigma_c\psi_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c) \geq \bar{\sigma}[\psi_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)]^+ + c\left([\psi_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)]^+\right)^2 - \frac{\bar{\sigma}^2}{c} \geq \frac{c}{2}\left([\psi_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)]^+\right)^2 - \frac{3\bar{\sigma}^2}{2c}, \text{ on } U^2. \tag{68}
$$

Returning to (64), we get

$$
\begin{aligned}
&(\lambda_c, \gamma'_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c - \boldsymbol{y})) + ((\sigma_c, \psi'_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c - \boldsymbol{y}))) \\
&\quad \geq \frac{c}{2}\|[\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)]^+\|_{L^2}^2 + \frac{c}{2}\|[\psi_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)]^+\|_{L^2(U^2)}^2 - \frac{3}{2c}\left(\|\bar{\lambda}\|_{L^2}^2 + \|\bar{\sigma}\|_{L^2(U^2)}^2\right).
\end{aligned} \tag{69}
$$

Moreover, it holds that $\tau'_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c - \boldsymbol{y}) = 0$. Thus, the first equation in (63) yields

(70)

$$
\begin{aligned}
&c_{\mathrm{b}}(\boldsymbol{x}_c'', \boldsymbol{x}_c'' - \boldsymbol{y}'')_{L^2} + m[\tilde{\boldsymbol{b}}](\boldsymbol{x}_c, \boldsymbol{x}_c - \boldsymbol{y}) + \frac{1}{\Delta t}(\boldsymbol{x}_c - \tilde{\boldsymbol{x}}, \boldsymbol{x}_c - \boldsymbol{y})_\star + \frac{c}{2}\|[\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)]^+\|_{L^2}^2 + \frac{c}{2}\|[\psi_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)]^+\|_{L^2(U^2)}^2 \\
&\qquad \leq \mathcal{F}_{\mathrm{x}}(\boldsymbol{x}_c - \boldsymbol{y}) + \frac{3}{2c}\left(\|\bar{\lambda}\|_{L^2}^2 + \|\bar{\sigma}\|_{L^2(U^2)}^2\right), \ \forall \boldsymbol{y} \in G(\tilde{\boldsymbol{x}}).
\end{aligned}
$$

This implies that there is a $k_0 > 0$ such that

$$
k_0\|\boldsymbol{x}_c\|_{H^2}^2 + \frac{c}{2}\|[\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)]^+\|_{L^2}^2 + \frac{c}{2}\|[\psi_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)]^+\|_{L^2(U^2)}^2
$$

is bounded uniformly in c. So there exists a subsequence $\{\boldsymbol{x}_c\}$, not relabeled, such that $\boldsymbol{x}_c \to \boldsymbol{x}^*$ weakly in H^2. Since $[\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)]^+ \to 0$ strongly in L^2 and $[\psi_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)]^+ \to 0$ strongly in $L^2(U^2)$, and $\boldsymbol{x}_c \to \boldsymbol{x}^*$ strongly in L^2, we have

$$
0 = \lim_{c\to\infty}\left(\eta, [\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)]^+\right) + \left(\left(\zeta, [\psi_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)]^+\right)\right) \geq \lim_{c\to\infty}(\eta, \gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c)) + ((\zeta, \psi_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c))) = (\eta, \gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}^*)) + ((\zeta, \psi_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}^*))),
$$

for all $\eta \in L^2$ and $\zeta \in L^2(U^2)$ with $\eta, \zeta \geq 0$.

Furthermore, since $\tau_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c) = 0$ for all c, by compact Sobolev embedding, $0 = \lim_{c\to\infty} \tau_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}_c) = \tau_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}^*)$. Hence, $\boldsymbol{x}^* \in G(\tilde{\boldsymbol{x}})$. So, by weak lower semicontinuity, (70) yields that $\boldsymbol{x}^* \in G(\tilde{\boldsymbol{x}})$ satisfies

$$c_{\mathrm{b}}\left((\boldsymbol{x}^*)'', (\boldsymbol{x}^* - \boldsymbol{y})''\right)_{L^2} + m[\tilde{\boldsymbol{b}}](\boldsymbol{x}^*, \boldsymbol{x}^* - \boldsymbol{y}) + \frac{1}{\Delta t}(\boldsymbol{x}^* - \tilde{\boldsymbol{x}}, \boldsymbol{x}^* - \boldsymbol{y})_\star \le \mathcal{F}_{\mathrm{x}}(\boldsymbol{x}^* - \boldsymbol{y}), \ \forall \boldsymbol{y} \in G(\tilde{\boldsymbol{x}}), \tag{71}$$

which, by (38), implies that $\boldsymbol{x}^*$ is the unique minimizer of (37). Returning to (70), setting $\boldsymbol{y} = \boldsymbol{x}^*$, we get

(72)

$$\begin{aligned} c_{\mathrm{b}}\|(\boldsymbol{x}_c - \boldsymbol{x}^*)''\|_{L^2}^2 &\le c_{\mathrm{b}}\left((\boldsymbol{x}_c - \boldsymbol{x}^*)'', (\boldsymbol{x}_c - \boldsymbol{x}^*)''\right)_{L^2} + m[\tilde{\boldsymbol{b}}](\boldsymbol{x}_c - \boldsymbol{x}^*, \boldsymbol{x}_c - \boldsymbol{x}^*) + \frac{1}{\Delta t}(\boldsymbol{x}_c - \boldsymbol{x}^*, \boldsymbol{x}_c - \boldsymbol{x}^*)_\star \\ &\le \mathcal{F}_{\mathrm{x}}(\boldsymbol{x}_c - \boldsymbol{x}^*) - c_{\mathrm{b}}\left((\boldsymbol{x}^*)'', (\boldsymbol{x}_c - \boldsymbol{x}^*)''\right)_{L^2} - m[\tilde{\boldsymbol{b}}](\boldsymbol{x}^*, \boldsymbol{x}_c - \boldsymbol{x}^*) \\ &\quad + \frac{1}{\Delta t}(\tilde{\boldsymbol{x}} - \boldsymbol{x}^*, \boldsymbol{x}_c - \boldsymbol{x}^*)_\star + \frac{3}{2c}\left(\|\bar{\lambda}\|_{L^2}^2 + \|\bar{\sigma}\|_{L^2(U^2)}^2\right), \end{aligned}$$

and utilizing the weak convergence of $\boldsymbol{x}_c$, we find that $\boldsymbol{x}_c \to \boldsymbol{x}^*$ strongly in H^2.

Furthermore, by Proposition 4.4, there exist Lagrange multipliers $(\rho^*, \lambda^*, \sigma^*) \in -\boldsymbol{K}^*$ such that (42), (43) are satisfied. From this, one can show that (62) holds. □

5.2. **Newton Scheme for the Regularized Problem.** We begin by recalling some basic facts from [34, Sec. 8.3]. Let $\mathbb{X}$ and $\mathbb{Z}$ be real Banach spaces. Moreover, let $\mathbb{X}$ be a subspace of a larger space $\widehat{\mathbb{X}}$ and note that $\mathbb{X} + \hat{x}$ is an affine space for any $\hat{x} \in \widehat{\mathbb{X}}$. For a fixed $\hat{x} \in \widehat{\mathbb{X}}$, we say $F : \mathcal{D} \subset \mathbb{X} + \hat{x} \to \mathbb{Z}$, where $\mathcal{D}$ is an open set, is *Newton differentiable* at $x \in \mathbb{X} + \hat{x}$, for any direction h in $\mathbb{X}$, if there exists an open neighborhood $\mathcal{N}(x) \subset \mathcal{D}$ and a mapping $G : \mathcal{N}(x) \to \mathcal{L}(\mathbb{X}, \mathbb{Z})$ such that

$$\lim_{\|h\|_{\mathbb{X}}\to 0} \frac{\|F(x+h) - F(x) - G(x+h)h\|_{\mathbb{Z}}}{\|h\|_{\mathbb{X}}} = 0. \tag{73}$$

The family $\{G(s) \mid s \in \mathcal{N}(x)\}$ is called an *N-derivative* of F at x.

Upon setting $\widehat{\mathbb{X}} = [H^2]^3 \times H^{-1} \times L^2 \times L^2(U^2)$, $\mathbb{X} = W(\mathbf{0}) \times H^{-1} \times L^2 \times L^2(U^2)$, and $\mathbb{Z} = \mathbb{X}^* \equiv (W(\mathbf{0}))^* \times H_0^1 \times L^2 \times L^2(U^2)$, define $\Phi[\cdot] : \mathbb{X} + (\boldsymbol{x}_{\mathrm{bc}}, 0, 0, 0) \to \mathbb{X}^*$ to be the residual functional associated with (59) by

$$\begin{aligned} \langle (\boldsymbol{w}, \eta, \varphi, \mu), \Phi[\boldsymbol{x}, \rho, \lambda, \sigma] \rangle &:= c_{\mathrm{b}}(\boldsymbol{w}'', \boldsymbol{x}'')_{L^2} + m[\tilde{\boldsymbol{b}}](\boldsymbol{w}, \boldsymbol{x}) + \frac{1}{\Delta t}(\boldsymbol{w}, \boldsymbol{x} - \tilde{\boldsymbol{x}})_\star + \langle \rho, \tau'_{\tilde{\boldsymbol{x}}}(\boldsymbol{w}) \rangle \\ &\quad + (\lambda, \gamma'_{\tilde{\boldsymbol{x}}}(\boldsymbol{w})) + ((\sigma, \psi'_{\tilde{\boldsymbol{x}}}(\boldsymbol{w}))) - \mathcal{F}_{\mathrm{x}}(\boldsymbol{w}) + \langle \eta, \tau'_{\tilde{\boldsymbol{x}}}(\boldsymbol{x} - \tilde{\boldsymbol{x}}) \rangle + \langle \eta, \tau(\tilde{\boldsymbol{x}}) \rangle \\ &\quad - (\varphi, \lambda) + \left(\varphi, \max\left(0, \bar{\lambda} + c\,\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x})\right)\chi^{\mathrm{a}}\right) - ((\mu, \sigma)) + ((\mu, \max\left(0, \bar{\sigma} + c\,\psi_{\tilde{\boldsymbol{x}}}(\boldsymbol{x})\right)\chi^{\mathrm{n}})), \end{aligned} \tag{74}$$

for all $(\boldsymbol{x}, \rho, \lambda, \sigma) \in \mathbb{X} + (\boldsymbol{x}_{\mathrm{bc}}, 0, 0, 0)$, and $(\boldsymbol{w}, \eta, \varphi, \mu) \in \mathbb{X}$, i.e. if $\langle (\boldsymbol{w}, \eta, \varphi, \mu), \Phi[\boldsymbol{x}, \rho, \lambda, \sigma] \rangle = 0$, for all $(\boldsymbol{w}, \eta, \varphi, \mu) \in \mathbb{X}$, then $(\boldsymbol{x}, \rho, \lambda, \sigma)$ yields the solution of (59).

Next, for any $(\boldsymbol{x}, \rho, \lambda, \sigma) \in \mathbb{X} + (\boldsymbol{x}_{\mathrm{bc}}, 0, 0, 0)$, we define $\Upsilon[\boldsymbol{x}, \rho, \lambda, \sigma](\cdot) \in \mathcal{L}(\mathbb{X}, \mathbb{X}^*)$ by

$$\begin{aligned} \langle (\boldsymbol{w}, \eta, \varphi, \mu), \Upsilon[\boldsymbol{x}, \rho, \lambda, \sigma](\hat{\boldsymbol{w}}, \hat{\eta}, \hat{\varphi}, \hat{\mu}) \rangle &:= c_{\mathrm{b}}(\boldsymbol{w}'', \hat{\boldsymbol{w}}'')_{L^2} + m[\tilde{\boldsymbol{b}}](\boldsymbol{w}, \hat{\boldsymbol{w}}) + \frac{1}{\Delta t}(\boldsymbol{w}, \hat{\boldsymbol{w}})_\star \\ &\quad + \langle \hat{\eta}, \tau'_{\tilde{\boldsymbol{x}}}(\boldsymbol{w}) \rangle + \langle \eta, \tau'_{\tilde{\boldsymbol{x}}}(\hat{\boldsymbol{w}}) \rangle + (\hat{\varphi}, \gamma'_{\tilde{\boldsymbol{x}}}(\boldsymbol{w})) + ((\hat{\mu}, \psi'_{\tilde{\boldsymbol{x}}}(\boldsymbol{w}))) \\ &\quad - (\varphi, \hat{\varphi}) + (\varphi, c\,\tilde{\chi}^{\gamma}(\boldsymbol{x})\gamma'_{\tilde{\boldsymbol{x}}}(\hat{\boldsymbol{w}})\chi^{\mathrm{a}}) - ((\mu, \hat{\mu})) + \left((\mu, c\,\tilde{\chi}^{\psi}(\boldsymbol{x})\psi'_{\tilde{\boldsymbol{x}}}(\hat{\boldsymbol{w}})\chi^{\mathrm{n}})\right), \end{aligned} \tag{75}$$

for all $(\boldsymbol{w}, \eta, \varphi, \mu), (\hat{\boldsymbol{w}}, \hat{\eta}, \hat{\varphi}, \hat{\mu}) \in \mathbb{X}$, where $\tilde{\chi}^{\gamma}(\boldsymbol{x})$ and $\tilde{\chi}^{\psi}(\boldsymbol{x})$ are characteristic functions of the active sets

$$\left\{\bar{\lambda} + c\,\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x}) > 0\right\} \subset U, \quad \left\{\bar{\sigma} + c\,\tilde{\psi}(\boldsymbol{x}) > 0\right\} \subset U^2, \ \text{ respectively.} \tag{76}$$

We recall that the max function $\max(0, \cdot)$, viewed as a mapping from $L^p(U) \to L^2(U)$, or from $L^p(U^2) \to L^2(U^2)$, is Newton differentiable if $p > 2$ (see [34, Sec. 8.4]). Moreover, $\gamma_{\tilde{\boldsymbol{x}}}(\boldsymbol{x})\chi^{\mathrm{a}} \in L^\infty$ and $\psi_{\tilde{\boldsymbol{x}}}(\boldsymbol{x})\chi^{\mathrm{n}} \in L^\infty(U^2)$ for all $\boldsymbol{x} \in [H^2]^3$. Thus, standard arguments yield the following result [34].

Proposition 5.3. *Let $(\boldsymbol{x} - \boldsymbol{x}_{\mathrm{bc}}, \rho, \lambda, \sigma) \in \mathbb{X}$. Then $\Phi[\cdot]$ is Newton differentiable at $(\boldsymbol{x}, \rho, \lambda, \sigma)$, with $\Upsilon : \mathcal{N}(\boldsymbol{x}, \rho, \lambda, \sigma) \to \mathcal{L}(\mathbb{X}, \mathbb{X}^*)$ the N-derivative, where $\mathcal{N}(\boldsymbol{x}, \rho, \lambda, \sigma) = \mathcal{N}' + (\boldsymbol{x}_{\mathrm{bc}}, 0, 0, 0)$, and $\mathcal{N}'$ is any bounded open neighborhood of $(\boldsymbol{x} - \boldsymbol{x}_{\mathrm{bc}}, \rho, \lambda, \sigma)$ in $\mathbb{X}$.*

Next, we have the lemma.

Lemma 5.4. *Suppose* $(\boldsymbol{x}^* - \boldsymbol{x}_{\rm bc}, \rho^*, \lambda^*, \sigma^*) \in \mathbb{X}$ *satisfies* $\langle \Phi[\boldsymbol{x}^*, \rho^*, \lambda^*, \sigma^*], (\boldsymbol{w}, \eta, \varphi, \mu)\rangle = 0$, *for all* $(\boldsymbol{w}, \eta, \varphi, \mu) \in \mathbb{X}$, *i.e.* $(\boldsymbol{x}^*, \rho^*, \lambda^*, \sigma^*)$ *solves* (59). *Let* $\mathcal{N}(\boldsymbol{x}^*, \rho^*, \lambda^*, \sigma^*) = \mathcal{N}' + (\boldsymbol{x}_{\rm bc}, 0, 0, 0)$, *where* $\mathcal{N}'$ *is any bounded open neighborhood of* $(\boldsymbol{x}^* - \boldsymbol{x}_{\rm bc}, \rho^*, \lambda^*, \sigma^*)$ *in* $\mathbb{X}$. *Then,* $\Upsilon[\boldsymbol{x}, \rho, \lambda, \sigma]$ *is non-singular for all* $(\boldsymbol{x}, \rho, \lambda, \sigma) \in \mathcal{N}(\boldsymbol{x}^*, \rho^*, \lambda^*, \sigma^*)$ *and*

$$\{\|[\Upsilon[\boldsymbol{x}, \rho, \lambda, \sigma]]^{-1}\| \mid \text{for } (\boldsymbol{x}, \rho, \lambda, \sigma) \in \mathcal{N}(\boldsymbol{x}^*, \rho^*, \lambda^*, \sigma^*)\} \text{ is bounded.} \tag{77}$$

Proof. Let $\mathcal{G} = (\boldsymbol{g}_1, g_2, g_3, g_4) \in \mathbb{X}^*$ and $(\boldsymbol{x}, \rho, \lambda, \sigma) \in \mathbb{X} + (\boldsymbol{x}_{\rm bc}, 0, 0, 0)$ be arbitrary, and consider the weak formulation: find $(\hat{\boldsymbol{w}}, \hat{\eta}, \hat{\varphi}, \hat{\mu}) \in \mathbb{X}$ such that

$$\langle (\boldsymbol{w}, \eta, \varphi, \mu), \Upsilon[\boldsymbol{x}, \rho, \lambda, \sigma](\hat{\boldsymbol{w}}, \hat{\eta}, \hat{\varphi}, \hat{\mu})\rangle := \mathcal{G}(\boldsymbol{w}, \eta, \varphi, \mu), \quad \text{for all } (\boldsymbol{w}, \eta, \varphi, \mu) \in \mathbb{X}. \tag{78}$$

First, we show existence and uniqueness of a solution, which will establish the non-singularity of Υ. Rewriting (78), we have

$$\begin{aligned} c_{\rm b}(\boldsymbol{w}'', \hat{\boldsymbol{w}}'')_{L^2} + m[\tilde{\boldsymbol{b}}](\boldsymbol{w}, \hat{\boldsymbol{w}}) + \frac{1}{\Delta t}(\boldsymbol{w}, \hat{\boldsymbol{w}})_\star + \langle \hat{\eta}, \tau'_{\tilde{\boldsymbol{x}}}(\boldsymbol{w})\rangle & \\ + (\hat{\varphi}, \gamma'_{\tilde{\boldsymbol{x}}}(\boldsymbol{w})) + ((\hat{\mu}, \psi'_{\tilde{\boldsymbol{x}}}(\boldsymbol{w}))) &= \langle \boldsymbol{g}_1, \boldsymbol{w}\rangle, \ \forall \boldsymbol{w} \in W(\boldsymbol{0}), \\ \langle \eta, \tau'_{\tilde{\boldsymbol{x}}}(\hat{\boldsymbol{w}})\rangle &= \langle \eta, g_2\rangle, \ \forall \eta \in H^{-1}, \\ (\varphi, c\,\tilde{\chi}^\gamma(\boldsymbol{x})\gamma'_{\tilde{\boldsymbol{x}}}(\hat{\boldsymbol{w}})\chi^{\rm a}) - (\varphi, \hat{\varphi}) &= (\varphi, g_3)_{L^2}, \ \forall \varphi \in L^2, \\ \big((\mu, c\,\tilde{\chi}^\psi(\boldsymbol{x})\psi'_{\tilde{\boldsymbol{x}}}(\hat{\boldsymbol{w}})\chi^{\rm n})\big) - ((\mu, \hat{\mu})) &= ((\mu, g_4))_{L^2(U^2)}, \ \forall \mu \in L^2(U^2). \end{aligned} \tag{79}$$

Eliminating $\hat{\varphi}$ and $\hat{\mu}$, we obtain the following saddle-point problem: find $\hat{\boldsymbol{w}} \in W(\boldsymbol{0})$, $\hat{\eta} \in H^{-1}$ such that

$$\begin{aligned} c_{\rm b}(\boldsymbol{w}'', \hat{\boldsymbol{w}}'')_{L^2} + m[\tilde{\boldsymbol{b}}](\boldsymbol{w}, \hat{\boldsymbol{w}}) + \frac{1}{\Delta t}(\boldsymbol{w}, \hat{\boldsymbol{w}})_\star + \langle \hat{\eta}, \tau'_{\tilde{\boldsymbol{x}}}(\boldsymbol{w})\rangle & \\ + c\,(\tilde{\chi}^\gamma(\boldsymbol{x})\gamma'_{\tilde{\boldsymbol{x}}}(\hat{\boldsymbol{w}})\chi^{\rm a}, \gamma'_{\tilde{\boldsymbol{x}}}(\boldsymbol{w})) + c\,\big((\tilde{\chi}^\psi(\boldsymbol{x})\psi'_{\tilde{\boldsymbol{x}}}(\hat{\boldsymbol{w}})\chi^{\rm n}, \psi'_{\tilde{\boldsymbol{x}}}(\boldsymbol{w}))\big) & \\ = \langle \boldsymbol{g}_1, \boldsymbol{w}\rangle + \Big(\widetilde{\Lambda}^* g_3, \boldsymbol{w}\Big)_{L^2} + \Big(\widetilde{\mathcal{S}}^* g_4, \boldsymbol{w}\Big)_{L^2}, \ \forall \boldsymbol{w} \in W(\boldsymbol{0}), & \\ \langle \eta, \tau'_{\tilde{\boldsymbol{x}}}(\hat{\boldsymbol{w}})\rangle = \langle \eta, g_2\rangle, \ \forall \eta \in H^{-1}. & \end{aligned} \tag{80}$$

By Proposition 2.6, and standard saddle point theory [15], there exists a unique solution to (80), for all finite $c > 0$. In addition, $\hat{\varphi}$ and $\hat{\mu}$ are uniquely determined. Thus, Υ is non-singular. Again by standard saddle point theory, we get a priori estimates for the solution in terms of $\mathcal{G}$, which yields the boundedness of the inverse, i.e. we obtain (77). □

Solving the system of equations in (59) is equivalent to finding $\boldsymbol{x}_c \in W(\boldsymbol{x}_{\rm bc})$, $\rho_c \in H^{-1}$, $\lambda_c \in L^2$, and $\sigma_c \in L^2(U^2)$ such that $\langle (\boldsymbol{w}, \eta, \varphi, \mu), \Phi[\boldsymbol{x}_c, \rho_c, \lambda_c, \sigma_c]\rangle = 0$, for all $\boldsymbol{w} \in W(\boldsymbol{0})$, $\eta \in L^2$, $\varphi \in L^2$, $\mu \in L^2(U^2)$. The Newton scheme for this is described in Algorithm 2.

Algorithm 2 Newton's method for (59).

Set a tolerance TOL > 0 and regularization parameter $c > 0$.
Assume $(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{b}}) \in W(\boldsymbol{x}_{\rm bc}) \times V(\boldsymbol{b}_{\rm bc})$ and $\bar{\lambda} \in L^2(U)$, $\bar{\sigma} \in L^2(U^2)$ given. Choose initial guess $\hat{\boldsymbol{x}}^0 \in W(\boldsymbol{x}_{\rm bc})$ and choose $\hat{\rho}^0 \in H^{-1}$, $\hat{\lambda}^0 \in L^2$, $\hat{\sigma}^0 \in L^2(U^2)$, with $\hat{\lambda}^0$, $\hat{\sigma}^0 \geq 0$ a.e.
Set $i := 0$ and do the following.

1. Find $\delta\hat{\boldsymbol{x}}^i \in W(\boldsymbol{0})$, $\delta\hat{\rho}^i \in H^{-1}$, $\delta\hat{\lambda}^i \in L^2$, $\delta\hat{\sigma}^i \in L^2(U^2)$, such that
$$\Big\langle (\boldsymbol{w}, \eta, \varphi, \mu), \Upsilon\Big[\hat{\boldsymbol{x}}^i, \hat{\rho}^i, \hat{\lambda}^i, \hat{\sigma}^i\Big]\Big(\delta\hat{\boldsymbol{x}}^i, \delta\hat{\rho}^i, \delta\hat{\lambda}^i, \delta\hat{\sigma}^i\Big)\Big\rangle = -\Big\langle (\boldsymbol{w}, \eta, \varphi, \mu), \Phi\Big[\hat{\boldsymbol{x}}^i, \hat{\rho}^i, \hat{\lambda}^i, \hat{\sigma}^i\Big]\Big\rangle, \tag{81}$$
for all $\boldsymbol{w} \in W(\boldsymbol{0})$, $\eta \in H^{-1}$, $\varphi \in L^2$, $\mu \in L^2(U^2)$.
2. Set $\hat{\boldsymbol{x}}^{i+1} := \hat{\boldsymbol{x}}^i + \delta\hat{\boldsymbol{x}}^i$, $\hat{\rho}^{i+1} := \hat{\rho}^i + \delta\hat{\rho}^i$, $\hat{\lambda}^{i+1} := \hat{\lambda}^i + \delta\hat{\lambda}^i$, $\hat{\sigma}^{i+1} := \hat{\sigma}^i + \delta\hat{\sigma}^i$.
3. If $\|\delta\hat{\boldsymbol{x}}^i\|_{W^{1,\infty}(U)} <$ TOL, then goto Step 4;
else, replace $i \leftarrow i + 1$ and return to Step 1.
4. Set $\boldsymbol{x}_c := \hat{\boldsymbol{x}}^{i+1}$, $\rho_c := \hat{\rho}^{i+1}$, $\lambda_c := \hat{\lambda}^{i+1}$, $\sigma_c := \hat{\sigma}^{i+1}$; this is the solution.

In lieu of Lemma 5.4 and [34, Thm 8.16], we have the following.

Theorem 5.5. *Suppose that* $(\boldsymbol{x}_c, \rho_c, \lambda_c, \sigma_c)$ *is the unique solution of* (59). *Then the iterates* $\left(\hat{\boldsymbol{x}}^i, \hat{\rho}^i, \hat{\lambda}^i, \hat{\sigma}^i\right)$ *generated by Algorithm 2 converge superlinearly to* $(\boldsymbol{x}_c, \rho_c, \lambda_c, \sigma_c)$ *provided*

$$\| \left(\hat{\boldsymbol{x}}^0, \hat{\rho}^0, \hat{\lambda}^0, \hat{\sigma}^0\right) - (\boldsymbol{x}_c, \rho_c, \lambda_c, \sigma_c) \|_{\mathbb{X}} \quad \textit{is sufficiently small.}$$

6. Discretization

We begin with a triangulation (set of sub-intervals) of the 1-dimensional U. Let $\mathcal{T}_h$ consist of N sub-intervals such that $\overline{U} = \cup_{I \in \mathcal{T}_h} \overline{I}$; the vertices are denoted $\mathcal{V}_h = \{\xi_i\}_{i=1}^{N+1}$.

6.1. Finite Element Spaces.

We introduce the standard Lagrange and Hermite spaces

$$\begin{aligned} M_h^0 &:= \{v \in C^0([0,L]) \mid v|_I \in \mathcal{P}_1(I), \quad \text{for every } I \in \mathcal{T}_h\}, \quad \text{(Lagrange)} \\ M_h^1 &:= \{w \in C^1([0,L]) \mid w|_I \in \mathcal{P}_3(I), \quad \text{for every } I \in \mathcal{T}_h\}, \quad \text{(Hermite)}, \end{aligned} \tag{82}$$

with interpolation operators $\mathcal{I}_h^0 : C^0([0,L]) \to M_h^0$, $\mathcal{I}_h^1 : C^1([0,L]) \to M_h^1$ defined by

$$\mathcal{I}_h^0 v = \sum_{\xi_i \in \mathcal{V}_h} v(\xi_i)\phi_i, \qquad \mathcal{I}_h^1 w = \sum_{\xi_i \in \mathcal{V}_h} \left(w(\xi_i)\psi_i^0 + w'(\xi_i)\psi_i^1\right), \tag{83}$$

where $\{\phi_i\}$ are basis functions spanning M_h^0 and $\left\{\{\psi_i^0\}, \{\psi_i^1\}\right\}$ are basis functions spanning M_h^1.

Next, define the following conforming finite element spaces

$$V_h(\boldsymbol{b}_{\rm bc}) := V(\boldsymbol{b}_{\rm bc}) \cap [M_h^0]^3, \qquad W_h(\boldsymbol{x}_{\rm bc}) := W(\boldsymbol{x}_{\rm bc}) \cap [M_h^1]^3. \tag{84}$$

Now introduce the nodal sets at which boundary conditions are imposed in (84):

$$\begin{aligned} \mathcal{N}_{\boldsymbol{x}} &:= \{\xi \in \mathcal{V}_h \mid \boldsymbol{X}_{\rm bc}(\boldsymbol{w}) \text{ enforces a Dirichlet condition on } \boldsymbol{w}(\xi)\}, \\ \mathcal{N}_{\boldsymbol{x}'} &:= \{\xi \in \mathcal{V}_h \mid \boldsymbol{X}_{\rm bc}(\boldsymbol{w}) \text{ enforces a Dirichlet condition on } \boldsymbol{w}'(\xi)\}, \\ \mathcal{N}_{\boldsymbol{b}} &:= \{\xi \in \mathcal{V}_h \mid \boldsymbol{B}_{\rm bc}(\boldsymbol{v}) \text{ enforces a Dirichlet condition on } \boldsymbol{v}(\xi)\}. \end{aligned} \tag{85}$$

For the unit length constraints, we have the discrete spaces

$$\begin{aligned} R_h &:= \left\{\rho_h \in M_h^0 \mid \rho_h(\xi) = 0, \text{ for all } \xi \in \mathcal{N}_{\boldsymbol{x}'}\right\}, \\ O_h &:= \left\{\omega_h \in M_h^0 \mid \omega_h(\xi) = 0, \text{ for all } \xi \in \mathcal{N}_{\boldsymbol{b}}\right\}, \end{aligned} \tag{86}$$

and for the capsid constraint, we have

$$L_h := \left\{\lambda_h \in M_h^0 \mid \lambda_h(\xi) = \boldsymbol{0}, \text{ for all } \xi \in \mathcal{N}_{\boldsymbol{x}}\right\}. \tag{87}$$

Discretizing the self-contact constraint must account for the product domain U^2, thus we make use of bilinear interpolation. First, define the space of shape functions

$$\begin{aligned} \mathcal{Q}_1(I \times J) := \Big\{ & q \in C^0(I \times J) \mid v(x,y) = \sum_{i,j} a_{i,j}\phi_i^I(x)\phi_j^J(y), \\ & \text{where } \{\phi_i^I\} \text{ are basis functions of } \mathcal{P}_1(I), \text{ and } \{\phi_i^J\} \text{ are basis functions of } \mathcal{P}_1(J)\Big\}, \end{aligned} \tag{88}$$

where $I, J \in \mathcal{T}_h$. Now, define

$$M_h^{00} := \{v \in C^0(U^2) \mid v|_{I\times J} \in \mathcal{Q}_1(I \times J), \quad \text{for every } I, J \in \mathcal{T}_h\}, \quad \text{(bilinear)}, \tag{89}$$

with interpolation operator $\mathcal{I}_h^{00} : C^0(U^2) \to M_h^{00}$ defined by

$$\mathcal{I}_h^{00} v = \sum_{\xi_i \in \mathcal{V}_h} \sum_{\xi_j \in \mathcal{V}_h} v(\xi_i, \xi_j)\vartheta_{i,j}, \tag{90}$$

where $\{\vartheta_{i,j}\}$ are basis functions spanning M_h^{00}. The discrete space for the self-contact multiplier is

$$S_h := \left\{\sigma_h \in M_h^{00} \mid \sigma_h(\xi_1, \xi_2) = 0, \text{ for all } \xi_1, \xi_2 \in \mathcal{N}_{\boldsymbol{x}}\right\}. \tag{91}$$

6.2. **Discrete Energy and Decent Steps.** We introduce the following discrete inner products:

$$(u,v)^h_{L^2} = \sum_{I\in\mathcal{T}_h}\int_I \mathcal{I}^0_h\{uv\}\,d\xi, \quad ((\sigma,\mu))^h_{L^2} = \int_0^L\int_0^L \mathcal{I}^{00}_h\{\sigma\mu\}d\xi_1 d\xi_2, \tag{92}$$

for all u, v in $L^2(U)$, such that $u|_I$ and $v|_I$ are continuous for each $I\in\mathcal{T}_h$, and for all $\sigma,\mu\in C^0(U^2)$. These definitions generalize to vector-valued functions in the obvious way. For all $\boldsymbol{x}_h\in[M^1_h]^3$ and $\boldsymbol{b}_h\in[M^0_h]^3$, the discrete version of the energy (8) is given by

$$\begin{aligned}\mathcal{E}^h_{\mathrm{r}}\left[\boldsymbol{x}_h,\boldsymbol{b}_h\right] = {}& \frac{c_{\mathrm{b}}}{2}\|\boldsymbol{x}_h''\|^2_{L^2} + \frac{c_{\mathrm{r}}}{2}\|\boldsymbol{b}_h'\|^2_{L^2} + \frac{c_{\mathrm{t}}}{2}\left[(|\boldsymbol{b}_h'||\boldsymbol{x}_h'|,|\boldsymbol{b}_h'||\boldsymbol{x}_h'|)^h_{L^2} - (\boldsymbol{b}_h'\cdot\boldsymbol{x}_h',\boldsymbol{b}_h'\cdot\boldsymbol{x}_h')^h_{L^2}\right]\\ &+\frac{c_{\mathrm{o}}}{2}\|\mathcal{I}^0_h\{\boldsymbol{x}_h'\cdot\boldsymbol{b}_h\}\|^2_{L^2} - \mathcal{F}\left(\boldsymbol{x}_h,\boldsymbol{b}_h\right),\end{aligned} \tag{93}$$

and the discrete minimizing movements energy (c.f. (16)) is simply $\mathcal{J}_h\left[\boldsymbol{x}_h,\boldsymbol{b}_h\right]\equiv\mathcal{J}_h\left[\boldsymbol{x}_h,\boldsymbol{b}_h;\tilde{\boldsymbol{x}}_h,\tilde{\boldsymbol{b}}_h\right]$:

$$\mathcal{J}_h\left[\boldsymbol{x}_h,\boldsymbol{b}_h\right] = \mathcal{E}^h_{\mathrm{r}}\left[\boldsymbol{x}_h,\boldsymbol{b}_h\right] + \frac{1}{2\Delta t}\left(\boldsymbol{x}_h-\tilde{\boldsymbol{x}}_h,\boldsymbol{x}_h-\tilde{\boldsymbol{x}}_h\right)_\star + \frac{1}{2\Delta t}\left(\boldsymbol{b}_h-\tilde{\boldsymbol{b}}_h,\boldsymbol{b}_h-\tilde{\boldsymbol{b}}_h\right)_\dagger. \tag{94}$$

Throughout the remainder, we treat the regularization parameter c as fixed and we no longer write the c subscript. Step 1 of Algorithm 1 consists in solving a discrete version of (59), which is as follows. Given $\tilde{\boldsymbol{x}}_h\in W_h(\boldsymbol{x}_{\mathrm{bc}})$, $\tilde{\boldsymbol{b}}_h\in V_h(\boldsymbol{b}_{\mathrm{bc}})$ (solution at the previous time-step), find $\boldsymbol{x}_h\in W_h(\boldsymbol{x}_{\mathrm{bc}})$, $\rho_h\in R_h$, $\lambda_h\in L_h$, and $\sigma_h\in S_h$ such that

$$\begin{aligned} c_{\mathrm{b}}\left(\boldsymbol{w}_h'',\boldsymbol{x}_h''\right)_{L^2} + c_{\mathrm{t}}\left[\left(|\tilde{\boldsymbol{b}}_h'|\boldsymbol{w}_h',|\tilde{\boldsymbol{b}}_h'|\boldsymbol{x}_h'\right)^h_{L^2} - \left(\tilde{\boldsymbol{b}}_h'\cdot\boldsymbol{w}_h',\tilde{\boldsymbol{b}}_h'\cdot\boldsymbol{x}_h'\right)^h_{L^2}\right] + c_{\mathrm{o}}\left(\boldsymbol{w}_h'\cdot\tilde{\boldsymbol{b}}_h,\boldsymbol{x}_h'\cdot\tilde{\boldsymbol{b}}_h\right)^h \\ +\frac{1}{\Delta t}\left(\boldsymbol{w}_h,\boldsymbol{x}_h-\tilde{\boldsymbol{x}}_h\right)_\star + \left(\lambda_h,\gamma'_{\tilde{\boldsymbol{x}}_h}\left(\boldsymbol{w}_h\right)\right)^h + \left(\left(\sigma_h,\psi'_{\tilde{\boldsymbol{x}}_h}\left(\boldsymbol{w}_h\right)\right)\right)^h \\ + \left(\rho_h,\tau'_{\tilde{\boldsymbol{x}}_h}\left(\boldsymbol{w}_h\right)\right)^h = \mathcal{F}_{\mathrm{x}}\left(\boldsymbol{w}_h\right),\ \forall\boldsymbol{w}_h\in W_h(\boldsymbol{0}), \\ \left(\eta_h,\tau_{\tilde{\boldsymbol{x}}_h}\left(\boldsymbol{x}_h\right)\right)^h = 0,\ \forall\eta_h\in R_h, \\ \left(\varphi_h,\max\left(0,\bar{\lambda}_h + c\,\gamma_{\tilde{\boldsymbol{x}}_h}\left(\boldsymbol{x}_h\right)\right)\chi^{\mathrm{a}}\right)^h - \left(\varphi_h,\lambda_h\right)^h = 0,\ \forall\varphi_h\in L_h, \\ \left(\left(\mu_h,\max\left(0,\bar{\sigma}_h + c\,\psi_{\tilde{\boldsymbol{x}}_h}\left(\boldsymbol{x}_h\right)\right)\chi^{\mathrm{n}}\right)\right)^h - \left(\left(\mu_h,\sigma_h\right)\right)^h = 0,\ \forall\mu_h\in S_h, \end{aligned} \tag{95}$$

where $\bar{\lambda}_h\in L_h$, $\bar{\sigma}_h\in S_h$ are fixed given functions that may be zero. Step 2 of Algorithm 1 consists in solving a discrete version of (47), which is as follows. Given $\boldsymbol{x}_h\in W_h(\boldsymbol{x}_{\mathrm{bc}})$, $\tilde{\boldsymbol{b}}_h\in V_h(\boldsymbol{b}_{\mathrm{bc}})$, find $\boldsymbol{b}_h\in V_h(\boldsymbol{b}_{\mathrm{bc}})$, $\omega_h\in O_h$ such that

$$\begin{aligned} c_{\mathrm{r}}\left(\boldsymbol{v}_h',\boldsymbol{b}_h'\right)_{L^2} + c_{\mathrm{t}}\left[\left(|\boldsymbol{x}_h'|\boldsymbol{v}_h',|\boldsymbol{x}_h'|\boldsymbol{b}_h'\right)^h_{L^2} - \left(\boldsymbol{x}_h'\cdot\boldsymbol{v}_h',\boldsymbol{x}_h'\cdot\boldsymbol{b}_h'\right)^h_{L^2}\right] + c_{\mathrm{o}}\left(\boldsymbol{x}_h'\cdot\boldsymbol{v}_h,\boldsymbol{x}_h'\cdot\boldsymbol{b}_h\right)^h \\ +\frac{1}{\Delta t}\left(\boldsymbol{v}_h,\boldsymbol{b}_h-\tilde{\boldsymbol{b}}_h\right)_\dagger + \left(\omega_h,\ell'_{\tilde{\boldsymbol{b}}_h}\left(\boldsymbol{v}_h\right)\right)^h = \mathcal{F}_{\mathrm{b}}\left(\boldsymbol{v}_h\right),\ \forall\,\boldsymbol{v}_h\in V_h(\boldsymbol{0}), \\ \left(\theta_h,\ell_{\tilde{\boldsymbol{b}}_h}\left(\boldsymbol{b}_h\right)\right)^h = 0,\ \forall\,\theta_h\in O_h. \end{aligned} \tag{96}$$

The well-posedness of (95) and (96) follows the same arguments as in the continuous case.

6.3. **Implementation Issues.** The discrete finite element spaces are conforming and completely standard. Furthermore, the use of discrete inner products facilitates fast matrix assembly. The main issue for this method is in handling the product domain $U^2\equiv U\times U$ in the inner product $((\cdot,\cdot))$ that appears in (95). A straightforward implementation can be very slow, especially when L is very large and h is very small.

The first step is to eliminate λ_h and σ_h from (95) and obtain a reduced system analogous to (60). It is now a matter of (efficiently) assembling the linear form

$$\Phi_{\mathrm{sc}}(\boldsymbol{w}_h) = \left(\left(\sigma_h,\chi^{\mathrm{n}}\left[\Theta\boldsymbol{w}_h\cdot\tilde{\boldsymbol{p}}_h\right]\right)\right)^h,\ \text{where } \sigma_h = \max\left[0,\bar{\sigma}_h + c\left(\varepsilon+\Theta\boldsymbol{x}_h\cdot\tilde{\boldsymbol{p}}_h\right)\right]\chi^{\mathrm{n}},\quad \tilde{\boldsymbol{p}}_h = -\frac{\Theta\tilde{\boldsymbol{x}}_h}{|\Theta\tilde{\boldsymbol{x}}_h|}, \tag{97}$$

for all $\boldsymbol{w}_h\in W_h(\boldsymbol{0})$, and the bilinear form

$$\Upsilon_{\mathrm{sc}}(\boldsymbol{w}_h,\hat{\boldsymbol{w}}_h) = \left(\left(\chi^{\mathrm{n}}\left[\Theta\boldsymbol{w}_h\cdot\tilde{\boldsymbol{p}}_h\right],\tilde{\chi}^{\psi}(\boldsymbol{x}_h)\left[\Theta\hat{\boldsymbol{w}}_h\cdot\tilde{\boldsymbol{p}}_h\right]\right)\right)^h, \tag{98}$$

for all $\boldsymbol{w}_h, \hat{\boldsymbol{w}}_h \in W_h(\mathbf{0})$, both of which appear in the inner iteration of Algorithm 2. All other forms only involve assembly over U, which is standard.

Next, we note that a sparse matrix representation of Θ can be computed once and for all. Moreover, $\tilde{\boldsymbol{p}}_h$ need only be computed once per time-step. Thus, the column vector realization of Φ_{sc} is only computed once per time-step. Of course, σ_h must be evaluated at each inner iteration, but this is straightforward as a sparse vector-vector dot product.

As for Υ_{sc}, this can be *partially* assembled once per time-step. Only the term $\tilde{\chi}^{\psi}(\boldsymbol{x}_h)$ changes in the inner iteration, which can be accounted for by multiplying with a sparse diagonal matrix; here, we take advantage of the discrete inner product. The final matrix realization of Υ_{sc} is merely $N \times N$, where N is the number of basis functions of $W_h(\mathbf{0})$. Then the formation of the saddle-point system to solve for updating $\boldsymbol{x}_h$ in the inner iteration is completely standard.

Remark 6.1 (nearest neighbors)**.** *The main trick in speeding up the computations is to take advantage of nearest neighbors in the ambient space* $\mathbb{R}^3$*. Even for situations of massive self-contact, most pairwise interactions of the curve* do not *take place.*

Thus, an octree [7, 24] *is used to determine which parts of the curve may possibly interact, e.g. find all pairwise distances which are less than* $1.5 \cdot \varepsilon$*; all other pairwise interactions are ignored. This has the practical advantage of making the matrix realization of* Θ more sparse*, which speeds up all related sparse matrix-vector products. If the time-step is sufficiently small, or alternatively the cutoff distance sufficiently large, this technique will not miss any pairwise interactions that would violate no self-penetration.*

Updating the octree and finding K *nearest neighbors, where* K *is sufficiently large to find all points within* $1.5 \cdot \varepsilon$*, has a negligible computational cost, so we do this every time-step. The choice of* K *depends on the mesh size (which is stated in the numerical experiments below). For our typical computations,* $K \approx 20$ *to* 60*.*

Remark 6.2 (mesh size vs. thickness)**.** *The self-contact constraint for the discrete problem only considers point-to-point distances between* nodes *of the mesh (instead of all points along the curve). Thus, if the mesh size* h *is too large, e.g.* $h > 2\varepsilon$*, then the curve could pass through itself* without detecting any violation *of the self-contact inequality constraint. Therefore, one must choose* $h < 2\varepsilon$*. In our numerical experiments, we always choose* $h < \varepsilon/2$ *as an extra safety margin.*

7. Numerical Results

The algorithm was implemented using the Matlab/C++ finite element toolbox FELICITY [65]. The computations were carried out on a Dell XPS 8700 desktop (circa 2014), 4 core Intel i7-4770 3.40GHz with 32GB RAM. The linear system solves were done using Matlab's "backslash" command. For the most expensive simulations, most of the computational time ($> 70\%$) is spent on "backslash" for solving the inner iteration of Newton's method when updating the curve $\boldsymbol{x}$.

To ensure robustness, we use a damped Newton scheme for solving Step 1 of the minimizing movements scheme, using a convergence tolerance of 10^{-9} to 10^{-8}, an initial step-size of $\alpha = 1.0$, a minimum step-size of 10^{-4}, and a back-tracking line search with the following step-size acceptance criterion

$$\mathcal{J}_{\text{r},c}\left[\hat{\boldsymbol{x}}^i + \alpha\delta\hat{\boldsymbol{x}}^i\right] < \mathcal{J}_{\text{r},c}\left[\hat{\boldsymbol{x}}^i\right] \quad \Rightarrow \quad \hat{\boldsymbol{x}}^{i+1} := \hat{\boldsymbol{x}}^i + \alpha\delta\hat{\boldsymbol{x}}^i, \tag{99}$$

where $\mathcal{J}_{\text{r},c}[\cdot]$ is defined in (61), $\delta\hat{\boldsymbol{x}}^i$ is the current Newton search direction, and i is the inner Newton iteration index. The initial guess $\hat{\boldsymbol{x}}^0 := \boldsymbol{x}^k$ is taken to be the solution from the previous time index k, with a slight modification to ensure $\tau_{\boldsymbol{x}^k}\left(\hat{\boldsymbol{x}}^0\right) = 0$, i.e. a feasible point.

The capsid, when present, is taken to be a sphere of unit radius with level set (distance) function $\phi(\boldsymbol{y}) := |\boldsymbol{y}| - 1$, which is a smooth function for all $|\boldsymbol{y}| > 1/2$. The smoothness of $\phi(\boldsymbol{y})$ for $|\boldsymbol{y}| \leq 1/2$ is irrelevant.

7.1. A Curve in a Capsid.

We simulate an open curve inside a capsid with thickness $\varepsilon = 0.05$ that is clamped at both ends. The initial $\boldsymbol{x}$ is given by the parametrization

$$\boldsymbol{x}(v) = (v - a_0, c_0 v(v - a_1)\sin(\omega v), 0)\,, \quad a_0 = 0.9,\ a_1 = 1.8,\ c_0 = 1.12,\ \omega = 5\pi/a_1, \tag{100}$$

on the interval $[0, M]$, where M is found so that an equal arc-length parametrization is on the interval $U = [0, L]$ with $L = 6.76$. The initial $\boldsymbol{b}$ is given by

$$\boldsymbol{b}(\xi) = \cos(\theta_0\xi)\boldsymbol{n}_1(\xi) + \sin(\theta_0\xi)\boldsymbol{n}_2(\xi), \quad \theta_0 = 1.0, \tag{101}$$

where ξ is the arc-length coordinate, and $\{\boldsymbol{t}, \boldsymbol{n}_1, \boldsymbol{n}_2\}$ is an orthonormal frame with $\boldsymbol{n}_1$ generated through parallel transport along the initial curve (100) with initial condition $\boldsymbol{n}_1(0) = (0, 0, 1)$ and $\boldsymbol{n}_2 := \boldsymbol{t} \times \boldsymbol{n}_1$. The clamped conditions for $\boldsymbol{x}$, and the Dirichlet conditions for $\boldsymbol{b}$, are taken from (100) and (101).

The coefficients in the energy (8) are

$$c_{\mathrm{r}} = 0.001,\ c_{\mathrm{b}} = 4,\ c_{\mathrm{t}} = 1,\ c_{\mathrm{o}} = 10^5, \tag{102}$$

and there are no body forces: $\mathcal{F}(\boldsymbol{x}, \boldsymbol{b}) \equiv 0$. The capsid activation function is $\chi^{\mathrm{a}} \equiv 1$. The regularization parameter is $c = 10^8$. The time-step is $\Delta t = 0.05$. The triangulation $\mathcal{T}_h$ of U consists of $N = 676$ sub-intervals of uniform length $h = 0.01$. We use $K = 30$ for the nearest neighbor search.

Figure 1 shows snapshots of the evolution of the curve driven by the minimization scheme in Algorithm 1. The minimizing movements scheme was run for 6400 iterations, for a total wall time of 8.75 minutes. A small amount of self-contact occurs, while a large amount of contact with the capsid occurs toward the end of the evolution. Figure 2 shows the energy decrease for the evolution.

In Figure 3, the max norm of the violation of the equality constraints is plotted versus time. Moreover, Figure 4 verifies the $O(\Delta t^2)$ error control of the unit length constraints by running several simulations for different values of Δt. The violation of the inequality constraints is plotted versus time in Figure 5.

7.2. **Closed and Open Knots.**

7.2.1. *The Trefoil Knot.* We simulate a closed curve with thickness $\varepsilon = 0.2$ that forms the trefoil knot; no capsid is present in this example. The initial $\boldsymbol{x}$ is given by the parametrization

$$\boldsymbol{x}(v) = (1/2)\left(\sin(v) + 2\sin(2v), \cos(v) - 2\cos(2v), -\sin(3v)\right), \tag{103}$$

on the interval $[0, 2\pi]$, and then reparameterized to yield an equal arc-length parametrization on the interval $U = [0, L]$ with $L = 6$. The initial $\boldsymbol{b}$ is given by

$$\boldsymbol{b}(\xi) = \cos(\theta_0 \xi)\boldsymbol{n}_1(\xi) + \sin(\theta_0 \xi)\boldsymbol{n}_2(\xi), \quad \theta_0 = 2\pi/L, \tag{104}$$

where ξ is the arc-length coordinate, and $\{\boldsymbol{t}, \boldsymbol{n}_1, \boldsymbol{n}_2\}$ is an orthonormal frame with $\boldsymbol{n}_1$ generated through parallel transport along the initial curve (103) with initial condition $\boldsymbol{n}_1(0) \parallel \boldsymbol{t}(0) \times (0, 0, 1)$ and $\boldsymbol{n}_2 := \boldsymbol{t} \times \boldsymbol{n}_1$.

The coefficients in the energy (8) are

$$c_{\mathrm{r}} = 0.001,\ c_{\mathrm{b}} = 1,\ c_{\mathrm{t}} = 2,\ c_{\mathrm{o}} = 10^5, \tag{105}$$

and there are no body forces: $\mathcal{F}(\boldsymbol{x}, \boldsymbol{b}) \equiv 0$. The regularization parameter is $c = 10^8$. The time-step is $\Delta t = 0.05$. The triangulation $\mathcal{T}_h$ of U consists of $N = 300$ sub-intervals of uniform length $h = 0.02$. We use $K = 40$ for the nearest neighbor search.

Figure 6 shows snapshots of the evolution of the curve driven by the minimization scheme in Algorithm 1. The minimizing movements scheme was run for 12000 iterations, for a total wall time of 9.003 minutes. A large amount of self-contact occurs toward the end of the evolution. Figure 7 shows the energy decrease for the evolution.

In Figure 8, the max norm of the violation of the equality constraints is plotted versus time. The violation of the self-contact inequality constraint is plotted versus time in Figure 9.

7.2.2. *A Practical Knot.* We simulate an open curve with thickness $\varepsilon = 0.2$ that forms a figure eight "knot"; no capsid is present in this example. The initial $\boldsymbol{x}$ is given by the parametrization

$$\boldsymbol{x}(v) = \begin{pmatrix} (2((v-73/20)^2 - 7)((v-73/20)^2 - 11)(v-73/20))/5 \\ (v-73/20)^4 - (67(v-73/20)^2)/5 \\ (((v-73/20)^2 - 3)((v-73/20)^2 - 9)((v-73/20)^2 - 13)(v-73/20))/10 \end{pmatrix}, \tag{106}$$

on the interval $[0, 7.3]$, and then reparameterized to yield an equal arc-length parametrization on the interval $U = [0, L]$ with $L = 7.3$. The initial $\boldsymbol{b}$ is given by

$$\boldsymbol{b}(\xi) = \boldsymbol{n}_1(\xi), \tag{107}$$

where ξ is the arc-length coordinate, and $\{\boldsymbol{t}, \boldsymbol{n}_1, \boldsymbol{n}_2\}$ is an orthonormal frame with $\boldsymbol{n}_1$ generated through parallel transport along the initial curve (106) with initial condition $\boldsymbol{n}_1(0) \parallel \boldsymbol{t}(0) \times (0, 0, 1)$.

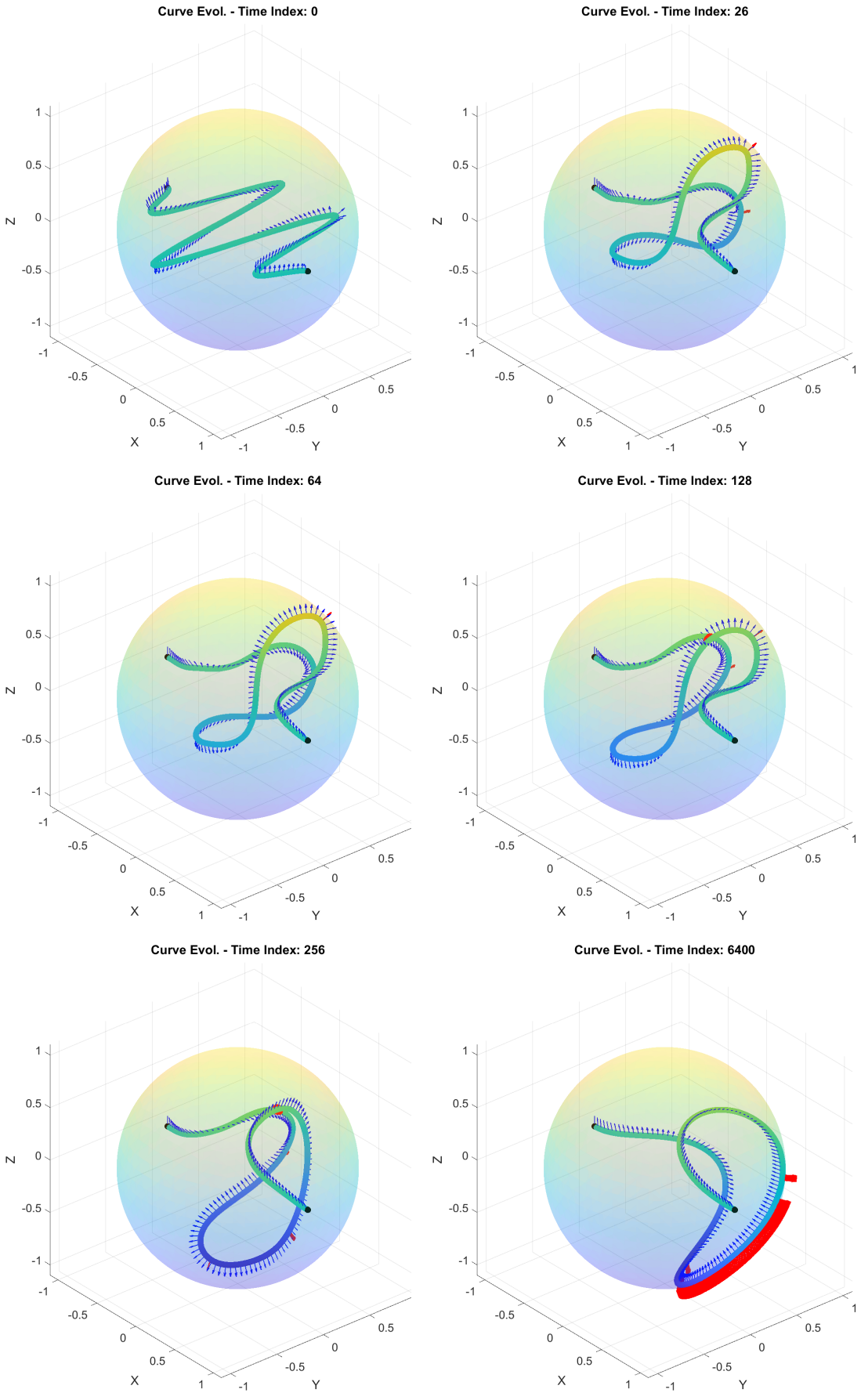

FIGURE 1. Evolution of a curve in a capsid (Section 7.1). A tube of thickness $\varepsilon = 0.05$ is plotted around the centerline curve $\boldsymbol{x}$. The blue arrows show $\boldsymbol{b}$, and red arrows represent $\nabla\phi$ at points that touch the capsid. At time index 128 and 256, a red disk is shown at a point of self-contact.

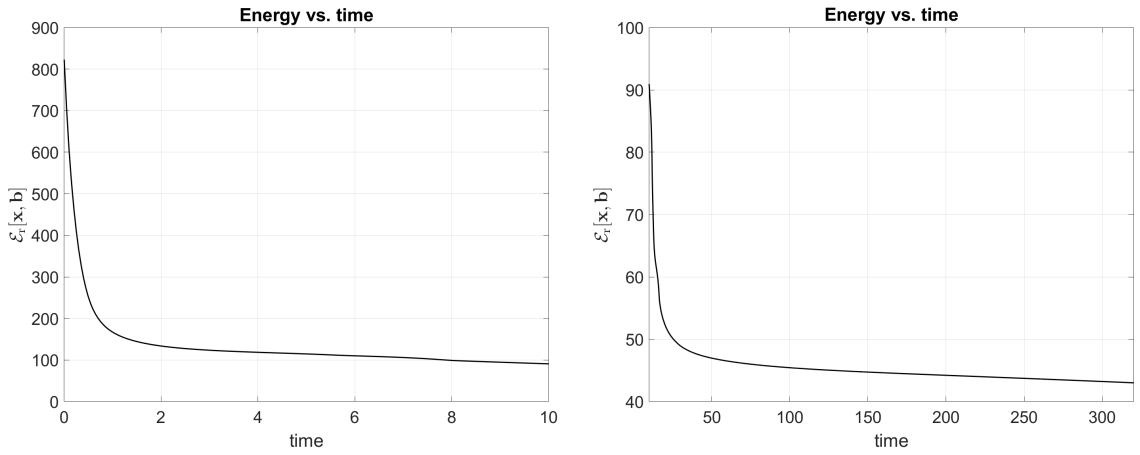

FIGURE 2. Energy decay for a curve in a capsid (Section 7.1). Left figure goes to time index 200.

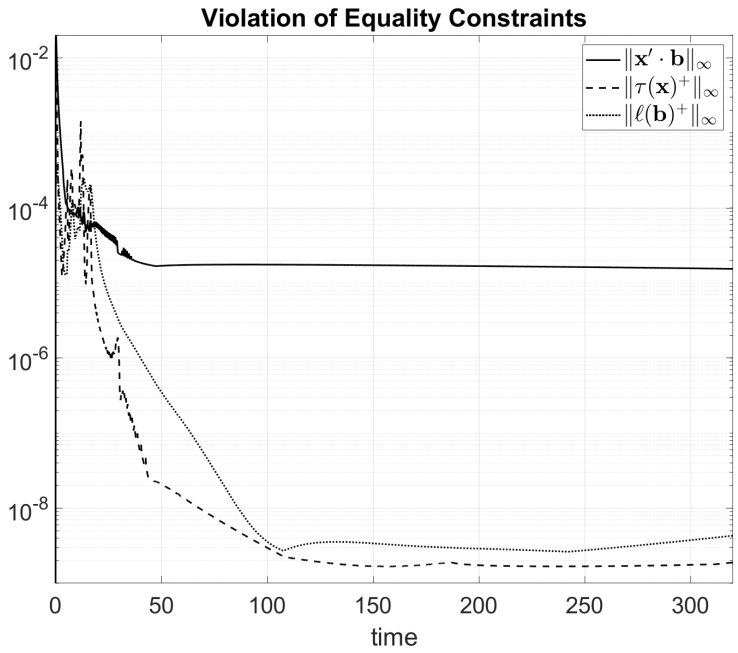

FIGURE 3. Equality constraint violation for a curve in a capsid (Section 7.1).

The coefficients in the energy (8) are

$$c_{\mathrm{r}} = 0.001,\ c_{\mathrm{b}} = 1,\ c_{\mathrm{t}} = 0.1,\ c_{\mathrm{o}} = 10^5. \tag{108}$$

No Dirichlet boundary conditions are imposed, and the only forces present are end point forces that pull the knot tight:

$$\mathcal{F}(\boldsymbol{x}, \boldsymbol{b}) \equiv \mathcal{F}_{\mathrm{x}}(\boldsymbol{x}) = \boldsymbol{h} \cdot \boldsymbol{x}\Big|_0^L, \quad \boldsymbol{h}(t) = (10 + t)(1, 0, 1). \tag{109}$$

The regularization parameter is $c = 10^8$. The time-step is $\Delta t = 0.02$. The triangulation $\mathcal{T}_h$ of U consists of $N = 365$ sub-intervals of uniform length $h = 0.02$. We use $K = 50$ for the nearest neighbor search.

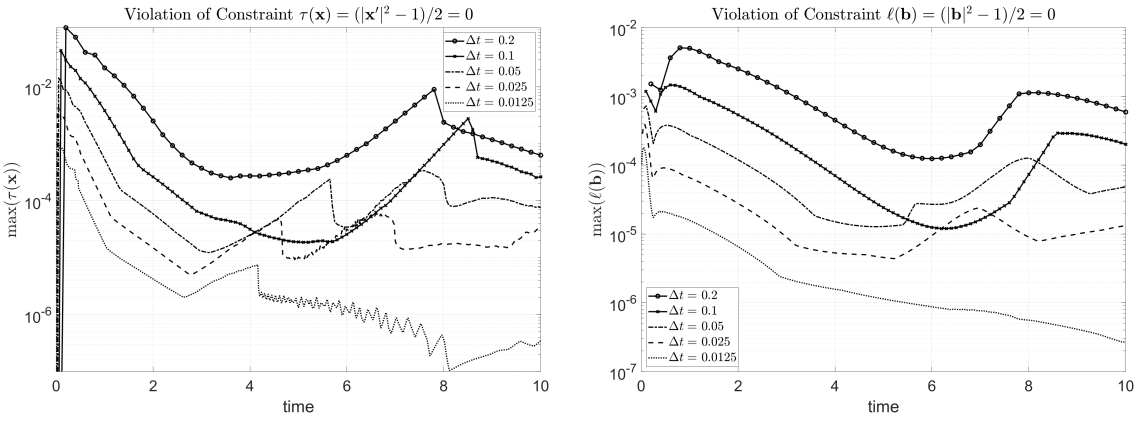

FIGURE 4. $O(\Delta t^2)$ check for $\tau(\boldsymbol{x})$ and $\ell(\boldsymbol{b})$ for a curve in a capsid (Section 7.1). Recall (57).

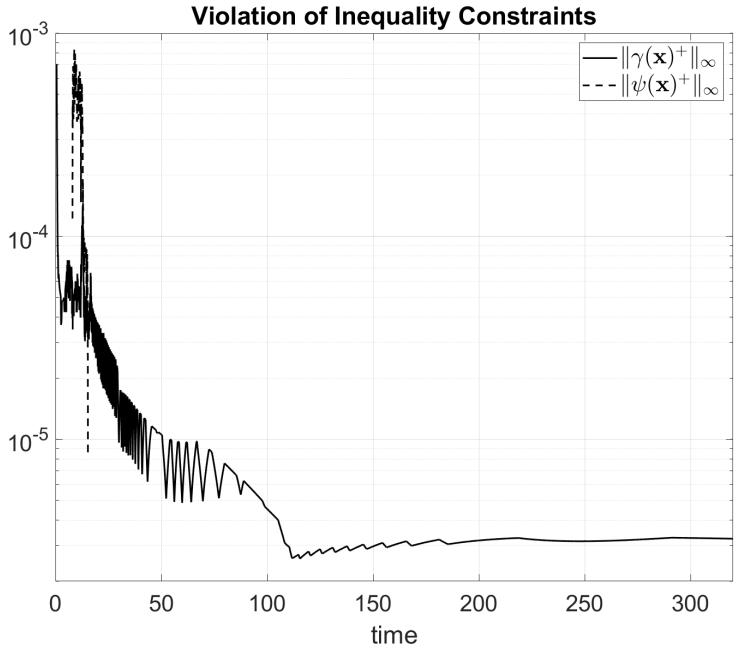

FIGURE 5. Inequality constraint violation for a curve in a capsid (Section 7.1). Note that $\psi(\boldsymbol{x})^+$ is only violated during a brief period of self-contact early in the simulation.

Figure 10 shows snapshots of the evolution of the curve driven by the minimization scheme in Algorithm 1. The minimizing movements scheme was run for 5000 iterations, for a total wall time of 5.533 minutes. A large amount of self-contact occurs toward the end of the evolution when the knot is pulled tight. Figure 11 shows the energy decrease for the evolution.

In Figure 12, the max norm of the violation of the equality constraints is plotted versus time. The violation of the self-contact inequality constraint is plotted versus time in Figure 13. Note that a small time-step is needed to ensure energy decrease and adequate control of the constraints (because of the large end point forces).

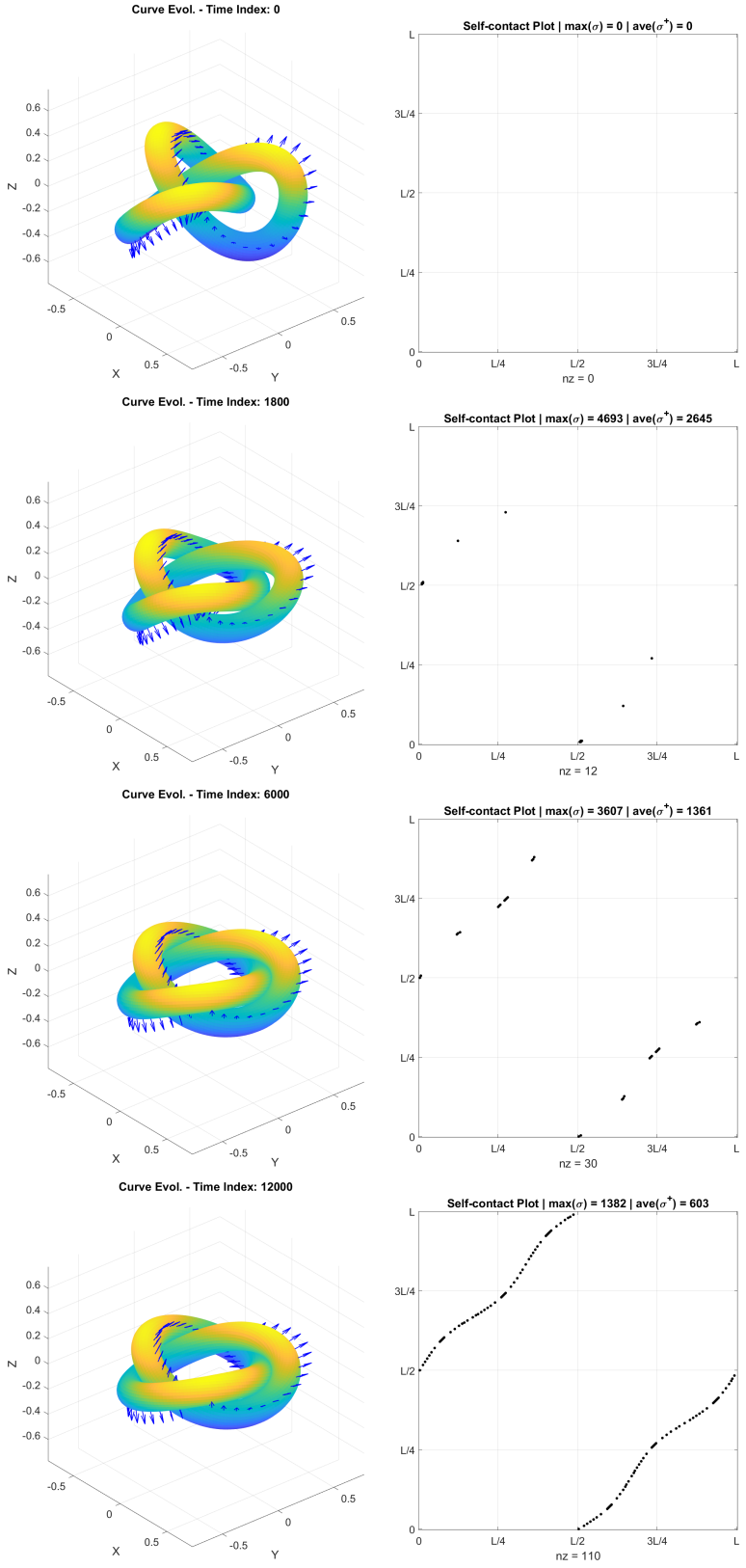

FIGURE 6. Evolution of a trefoil knot (Section 7.2.1). A tube of thickness $\varepsilon = 0.2$ is plotted around the centerline curve $\boldsymbol{x}$. The blue arrows show $\boldsymbol{b}$. The right side shows a plot of the characteristic function of $\{\sigma > 0\}$ over the square U^2. A significant amount of self-contact is present toward the end of the simulation.

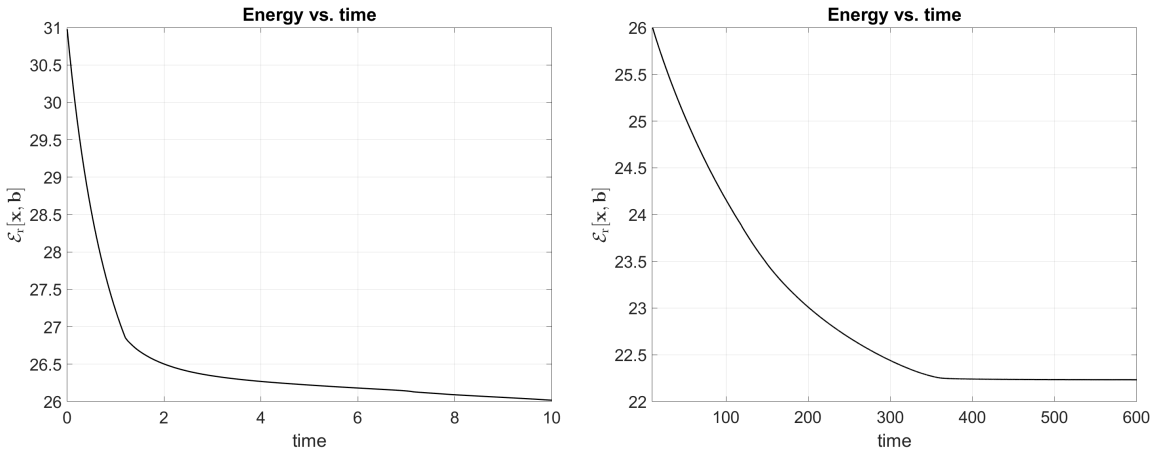

FIGURE 7. Energy decay for a trefoil knot (Section 7.2.1). Left figure goes to time index 200.

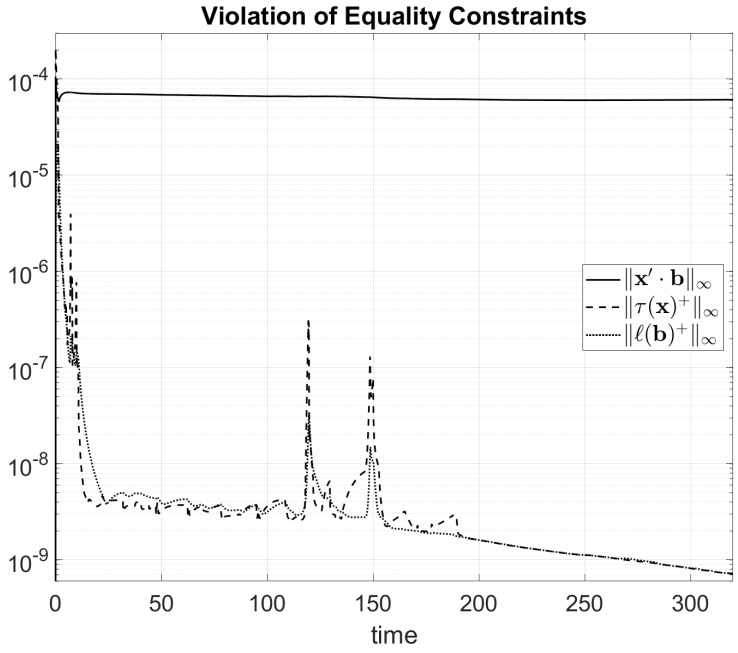

FIGURE 8. Equality constraint violation for a trefoil knot (Section 7.2.1).

7.3. **Packing in Two Dimensions.** We simulate an open curve, with thickness $\varepsilon = 0.05$, being packed inside a two dimensional capsid. The initial $\boldsymbol{x}$ is given by the parametrization

$$\boldsymbol{x}(v) = \left(-a_0((v-c_0)^+)^2, v - c_1\right), \quad a_0 = 3.0,\ c_0 = L - 0.8,\ c_1 = L - 0.9, \tag{110}$$

on the interval $[0, M]$, where M is found so that an equal arc-length parametrization is on the interval $U = [0, L]$ with $L = 40$. Since this example is in two dimensions, $\boldsymbol{b} \equiv (0, 0, 1)$ and Step 2 of Algorithm 1 is omitted. No Dirichlet conditions are imposed at either end point of the curve.

In addition, we model a simple "motor" at the bottom of the capsid that inserts the elastic curve into the capsid. The details are as follows. Let $U_{\text{Cap}} \subset U$ be the subset of points ξ such that $\boldsymbol{x}(\xi)$ has entered the capsid. We initialize the set U^0_{Cap} to be the set of ξ values such that $\phi(\boldsymbol{x}^0(\xi)) < 0$ and $\boldsymbol{x}^0(\xi) \cdot \boldsymbol{e}_2 > -1 + \varepsilon/2$.

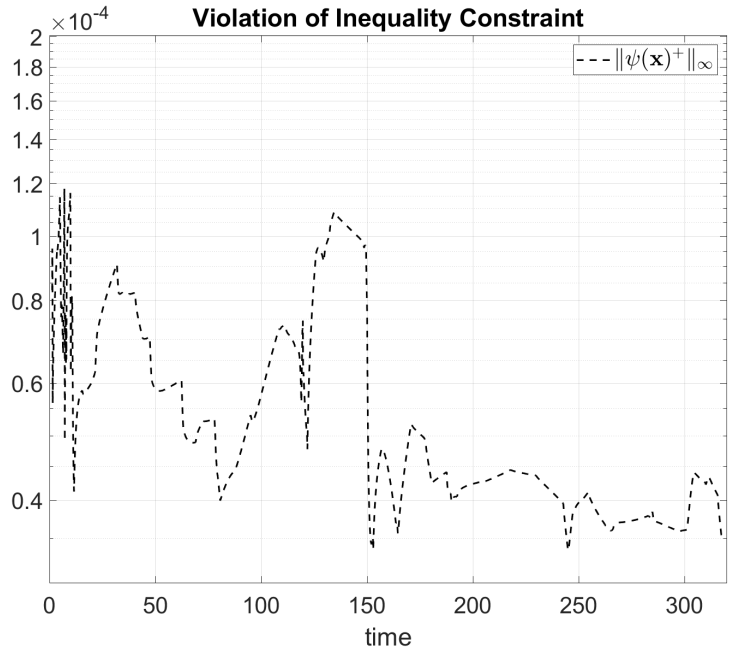

FIGURE 9. Inequality constraint violation for a trefoil knot (Section 7.2.1). Note that the capsid constraint is not present in this example.

We also assume that the initial curve is oriented so there exists a ξ^0 such that $U^0_{\text{Cap}} = (\xi^0, L]$. The point ξ^0 is where the motor is located.

Thus, the coordinates of the motor are $(0, -1 + \varepsilon/2)$ and the following property is satisfied:

$$\xi^j := \underset{\xi \in U \setminus U^j_{\text{Cap}}}{\arg\max}\, \boldsymbol{x}^j(\xi) \cdot \boldsymbol{e}_2 \quad \Rightarrow \quad \boldsymbol{x}^j(\xi^j) = (0, -1 + \varepsilon/2), \tag{111}$$

for $j = 0$ (i.e. the initial curve). Suppose that (111) is satisfied for all $j = 0, 1, ..., k$, for a sequence $\{\xi^j\}_{j=0}^k$, such that $\xi^{j-1} > \xi^j$, with $U^j_{\text{Cap}} = (\xi^j, L]$. During the kth step, when solving for $\boldsymbol{x}^{k+1}$, we set the following Dirichlet condition at the point ξ^k:

$$\boldsymbol{x}^{k+1}(\xi^k) = \big(0, \boldsymbol{x}^k(\xi^k) \cdot \boldsymbol{e}_2\big) + \Delta t\,(0, 0.2)\,, \quad (\boldsymbol{x}^{k+1})'(\xi^k) = (0, 1)\,. \tag{112}$$

Since the Dirichlet condition decouples the curve into two disjoint parts, we solve for the curve corresponding to U^k_{Cap} using the method outlined earlier in the paper. The other part of the curve that is outside the capsid, is solved for "analytically," i.e. $\boldsymbol{x}^{k+1}(\xi) \cdot \boldsymbol{e}_1 = 0$ and $\boldsymbol{x}^{k+1}(\xi)$ is parallel to the $\boldsymbol{e}_2$ axis for all $\xi \notin U^k_{\text{Cap}}$. Hence, we do not have to account for self-contact outside of the capsid. Therefore, we set $U^{k+1}_{\text{Cap}} = (\xi^{k+1}, L]$ where $\xi^{k+1} = \xi^k - 0.2\Delta t$. The capsid activation function, $(\chi^{\text{a}})^k$, is simply taken to be the characteristic function of U^k_{Cap}. The motor causes the curve to be "pumped" into the capsid, with the capsid inequality constraint only applied to the points in U^k_{Cap} at iteration k.

The coefficients in the energy (8) are

$$c_{\text{r}} = 0,\ c_{\text{b}} = 1,\ c_{\text{t}} = 0,\ c_{\text{o}} = 0, \tag{113}$$

and there are no body forces: $\mathcal{F}(\boldsymbol{x}, \boldsymbol{b}) \equiv 0$. The regularization parameter is $c = 10^8$. The time-step is $\Delta t = 0.05$. The triangulation $\mathcal{T}_h$ of U consists of $N = 2400$ sub-intervals of uniform length $h = 0.016667$. We use $K = 30$ for the nearest neighbor search.

Figure 14 shows snapshots of the evolution of the curve driven by the scheme in Algorithm 1 combined with the "motor" described earlier. The scheme was run for 6000 iterations, for a total wall time of 22.13 minutes. The overall capsid contact and self-contact increases as the curve is pumped into the capsid. Figure

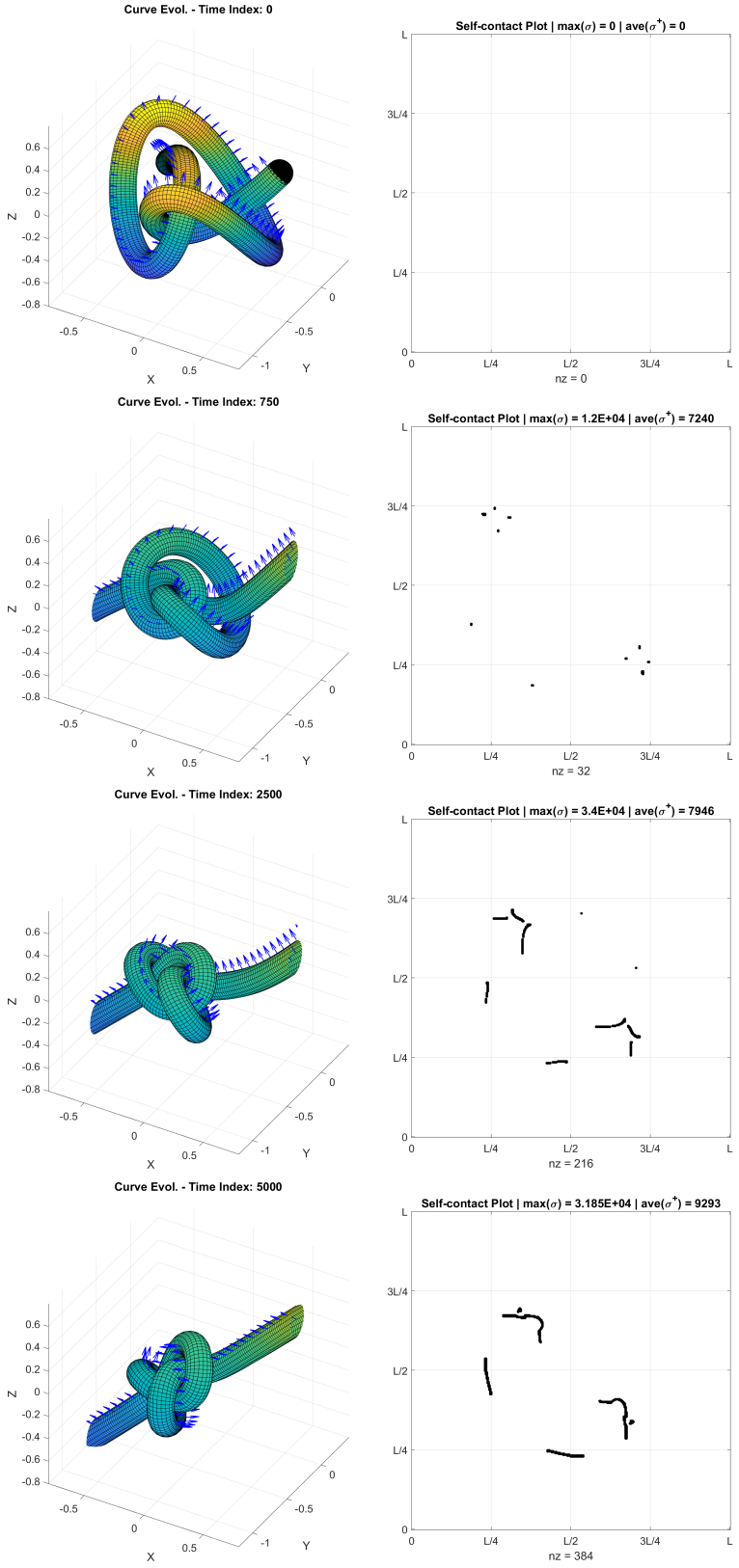

FIGURE 10. Evolution of a figure eight "knot" (Section 7.2.2). A tube of thickness $\varepsilon = 0.2$ is plotted around the centerline curve $\boldsymbol{x}$. The blue arrows show $\boldsymbol{b}$. The right side shows a plot of the characteristic function of $\{\sigma > 0\}$ over the square U^2. A significant amount of self-contact is present toward the end of the simulation, because the knot is pulled tight.

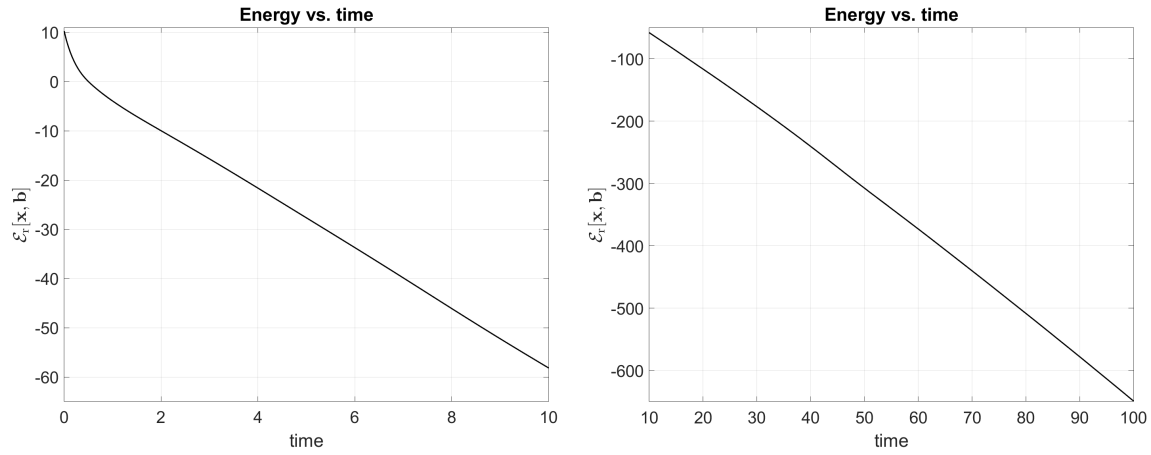

FIGURE 11. Energy decay for a figure eight "knot" (Section 7.2.2). Left figure goes to time index 500. The energy decreases linearly because of the applied forces at the end points.

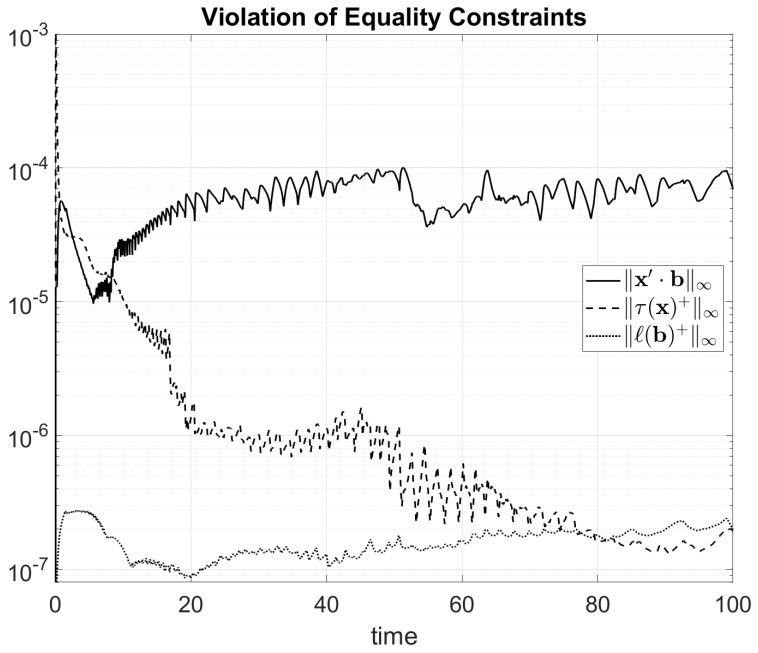

FIGURE 12. Equality constraint violation for a figure eight "knot" (Section 7.2.2).

15 shows the energy plot for the evolution, which increases (because energy is being pumped into the system) until about $t = 190$, when the motor releases the tail of the curve.

In Figure 16, the max norm of the violation of the equality constraints is plotted versus time. The violation of the inequality constraints is plotted versus time in Figure 17.

7.4. **Packing in Three Dimensions.** We simulate an open curve, with thickness $\varepsilon = 0.15$, being packed inside a three dimensional capsid. The initial $\boldsymbol{x}$ is given by the parametrization

$$\boldsymbol{x}(v) = \left(-a_0((v - c_0)^+)^2, 0, v - c_1\right), \quad a_0 = 3.0,\ c_0 = L - 0.8,\ c_1 = L - 0.95, \tag{114}$$

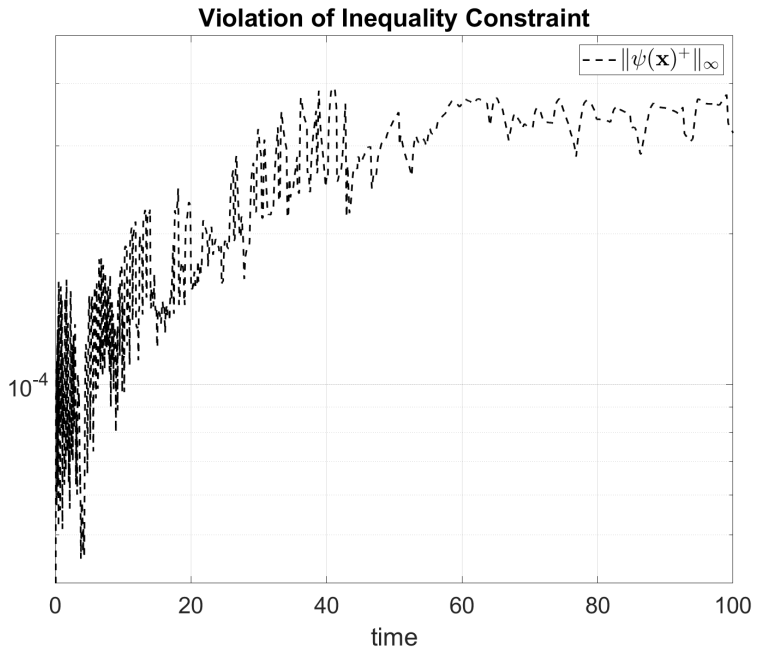

FIGURE 13. Inequality constraint violation for a figure eight "knot" (Section 7.2.2). Note that the capsid constraint is not present in this example.

on the interval $[0, M]$, where M is found so that an equal arc-length parametrization is on the interval $U = [0, L]$ with $L = 200$. The initial $\boldsymbol{b}$ is given by

$$\boldsymbol{b}(\xi) = (0, 1, 0)\,, \quad \text{for all } \xi \in U. \tag{115}$$

No Dirichlet conditions are imposed at either end point of the curve.

In addition, we model a simple "motor" at the south pole of the capsid that inserts the elastic curve into the capsid. The model is similar to Section 7.3 with the following modifications. We initialize the set U^0_{Cap} to be the set of ξ values such that $\phi(\boldsymbol{x}^0(\xi)) < 0$ and $\boldsymbol{x}^0(\xi) \cdot \boldsymbol{e}_3 > -1 + \varepsilon/2$. We also assume that the initial curve is oriented so that $U^0_{\text{Cap}} = (\xi^0, L]$ for some ξ^0 (i.e. the motor point).

Thus, the coordinates of the motor are $(0, 0, -1 + \varepsilon/2)$ and the following property is satisfied:

$$\xi^j := \underset{\xi \in U \setminus U^j_{\text{Cap}}}{\arg\max}\, \boldsymbol{x}^j(\xi) \cdot \boldsymbol{e}_3 \quad \Rightarrow \quad \boldsymbol{x}^j(\xi^j) = (0, 0, -1 + \varepsilon/2), \quad \text{and } \boldsymbol{b}^j(\xi^j) \cdot \boldsymbol{e}_3 = 0, \tag{116}$$

for $j = 0$ (i.e. the initial curve). Suppose that (116) is satisfied for all $j = 0, 1, ..., k$, for a sequence $\{\xi^j\}_{j=0}^k$, such that $\xi^{j-1} > \xi^j$, with $U^j_{\text{Cap}} = (\xi^j, L]$. During the kth step, when solving for $\boldsymbol{x}^{k+1}$, we set the following Dirichlet condition at the point ξ^k:

$$\begin{aligned} &\boldsymbol{x}^{k+1}(\xi^k) = \big(0, 0, \boldsymbol{x}^k(\xi^k) \cdot \boldsymbol{e}_3\big) + \Delta t\,(0, 0, 0.2)\,, \quad (\boldsymbol{x}^{k+1})'(\xi^k) = (0, 0, 1)\,, \\ &\boldsymbol{b}^{k+1}(\xi^k) = R(\Delta\theta)\boldsymbol{b}^k(\xi^k), \quad \Delta\theta = 5\Delta t(\pi/180), \end{aligned} \tag{117}$$

where $R(\Delta\theta)$ is a 3×3 rotation matrix about $\boldsymbol{e}_3$ by the (small) angle $\Delta\theta$, i.e. $\boldsymbol{b}^{k+1}(\xi^k)$ is simply $\boldsymbol{b}^k(\xi^k)$ rotated in the x-y plane by 5 degrees per unit time. This models a twisting motion as the curve is inserted into the capsid.

Since the Dirichlet condition decouples the curve into two disjoint parts, we solve for the two parts of the curve separately as described in Section 7.3. The part outside of the capsid is solved for "analytically," i.e. $\boldsymbol{x}^{k+1}(\xi) \cdot \boldsymbol{e}_1 = 0$, $\boldsymbol{x}^{k+1}(\xi) \cdot \boldsymbol{e}_2 = 0$, $\boldsymbol{x}^{k+1}(\xi)$ is parallel to the $\boldsymbol{e}_3$ axis, and $\boldsymbol{b}^{k+1}(\xi)$ is constant for all $\xi \notin U^k_{\text{Cap}}$.

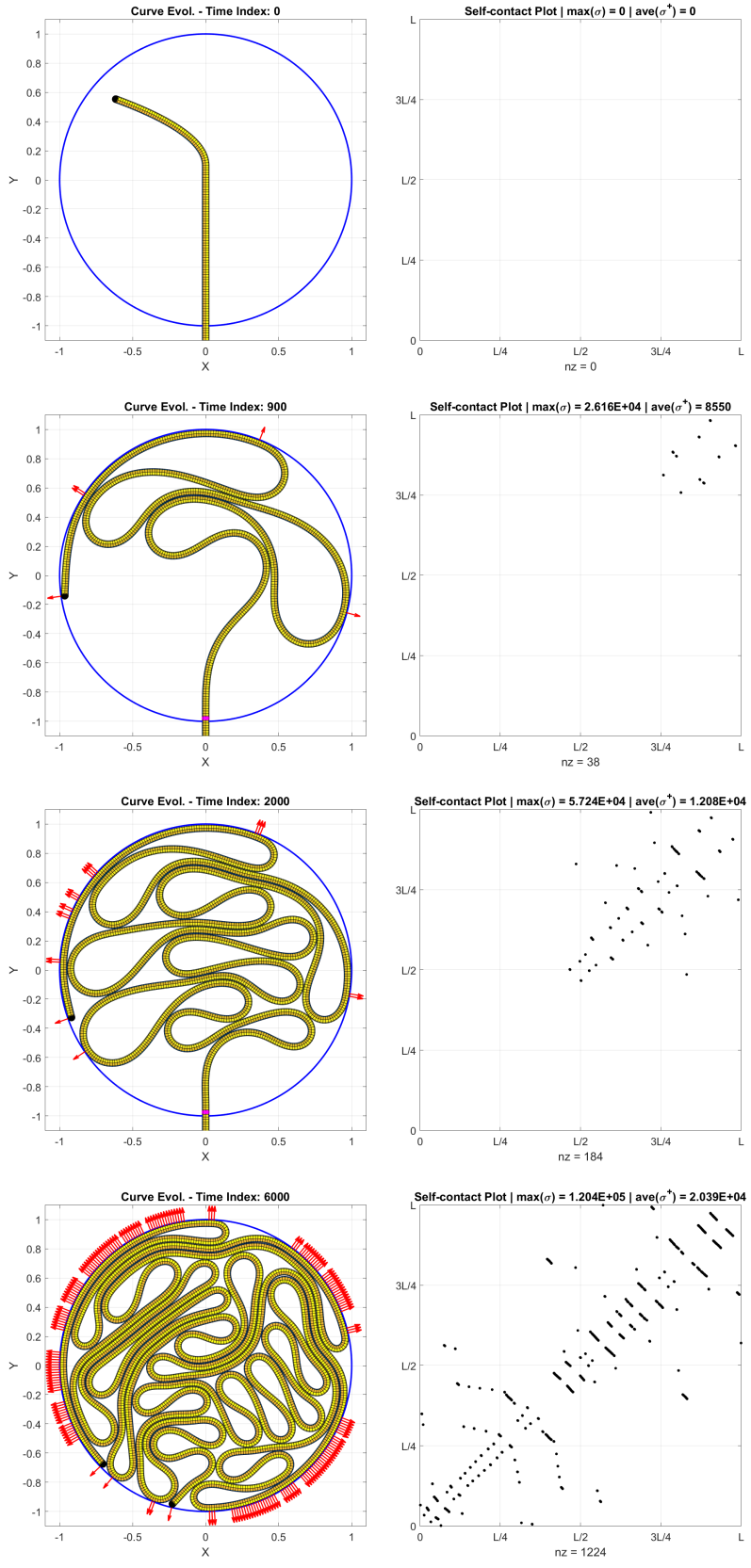

FIGURE 14. Packing of a two dimensional curve in a capsid (Section 7.3). A tube of thickness $\varepsilon = 0.05$ is plotted around the centerline curve $\boldsymbol{x}$. The "motor" (at the bottom) is shown in magenta. The red arrows represent $\nabla\phi$ at points that touch the capsid. The right side shows a plot of the characteristic function of $\{\sigma > 0\}$ over the square U^2. A significant amount of self-contact is present toward the end of the simulation.

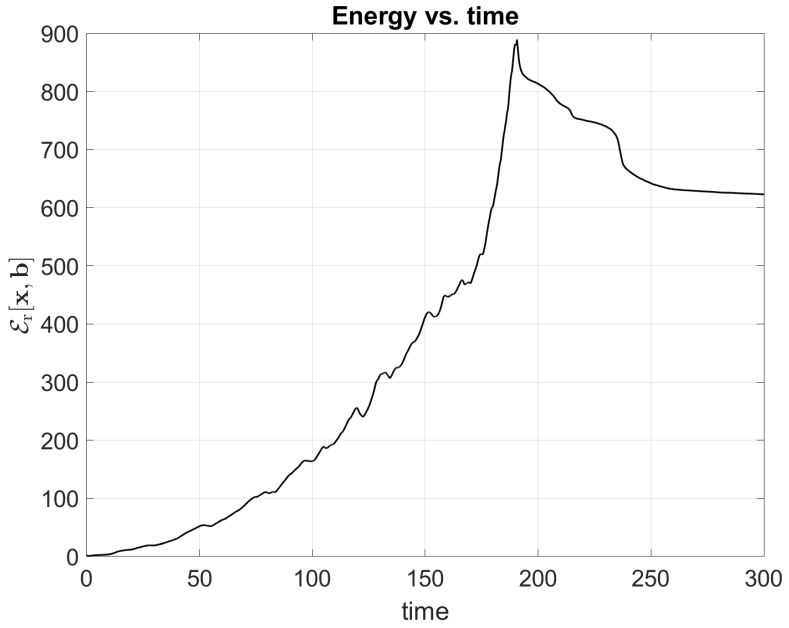

FIGURE 15. Energy plot for packing a two dimensional curve in a capsid (Section 7.3). Note: the energy increases because energy is put into the system by the "motor." At approximately $t = 190$, the motor releases the tail of the curve, which causes the curve to relax and the energy to decrease at later times.

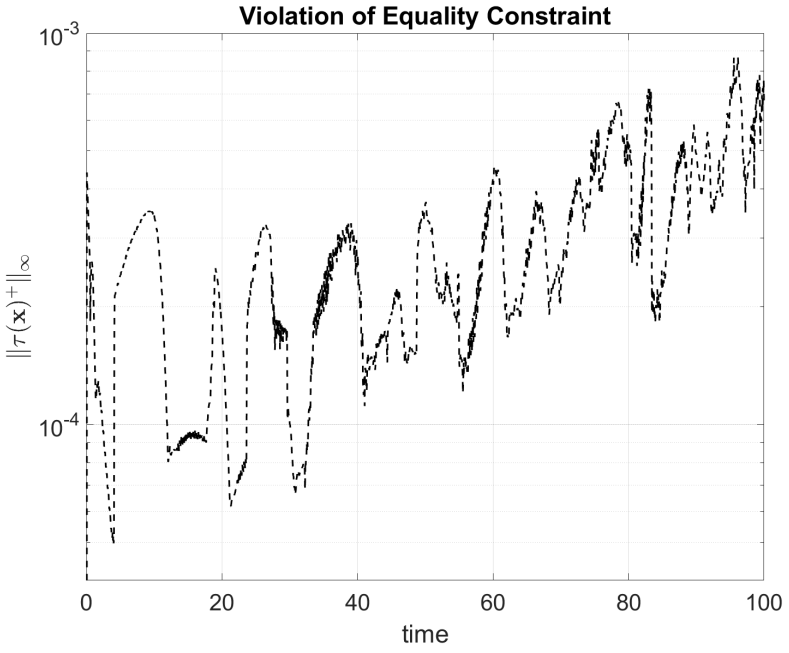

FIGURE 16. Violation of the unit length constraint $\tau(\boldsymbol{x}) = 0$ for packing a two dimensional curve in a capsid (Section 7.3).

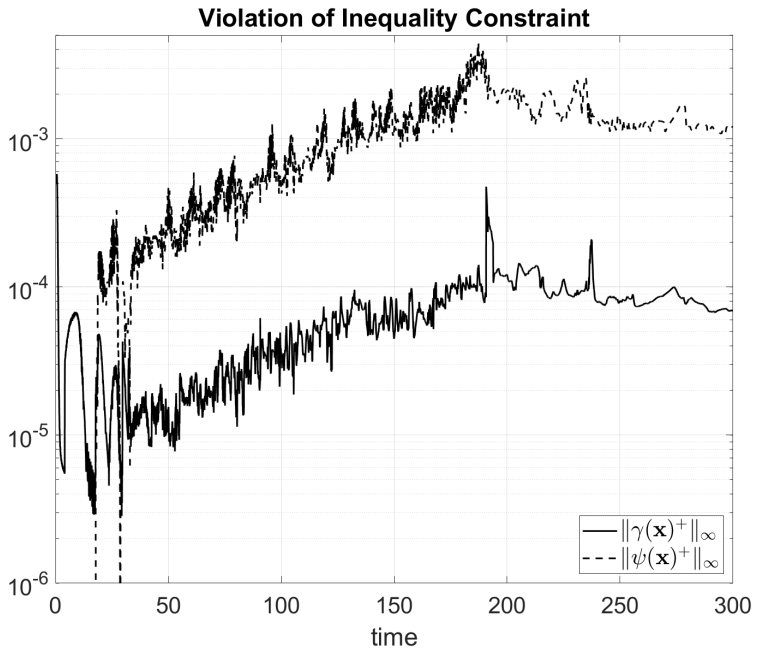

FIGURE 17. Inequality constraint violation for packing a two dimensional curve in a capsid (Section 7.3).

The coefficients in the energy (8) are

$$c_{\mathrm{r}} = 0.001,\ c_{\mathrm{b}} = 4,\ c_{\mathrm{t}} = 1,\ c_{\mathrm{o}} = 10^5, \tag{118}$$

and there are no body forces: $\mathcal{F}(\boldsymbol{x}, \boldsymbol{b}) \equiv 0$. For the first 14000 iterations, the regularization parameter is $c = 10^8$ and the time-step is $\Delta t = 0.05$; afterward, $c = 10^9$ and $\Delta t = 0.02$. The triangulation $\mathcal{T}_h$ of U consists of $N = 6000$ sub-intervals of uniform length $h = 1/30$. We use $K = 25$ for the nearest neighbor search.

Figures 18 and 19 show snapshots of the evolution of the curve driven by the scheme in Algorithm 1 combined with the "motor" described earlier. The scheme was run for 14000 iterations (at $\Delta t = 0.05$), followed by another 15000 iterations (at $\Delta t = 0.02$), for a total wall time of 13.256 hours. The overall capsid contact and self-contact increases as the curve is pumped into the capsid. Figure 20 shows the energy plot for the evolution, which increases (because energy is being pumped into the system) until about $t = 980$, when the motor releases the tail of the curve.

In Figure 21, the max norm of the violation of the equality constraints is plotted versus time. The violation of the inequality constraints is plotted versus time in Figure 22.

8. Conclusions

We have presented a robust and efficient method for simulating elastic curves with self-contact and capsid obstacle constraints. The minimizing movements scheme we use can find complex energy minimizers under non-trivial inequality constraints (our numerical results demonstrate this). Each step of the scheme involves solving a convex problem, which is achieved through linearization of the inequality constraints. Indeed, the linearization of the self-contact constraint enjoys a powerful monotone property. Violation of the exact constraints is $O(\Delta t^2)$, regardless of the number of time-steps that have run (provided a mild time-step restriction is satisfied). Furthermore, our regularization of the constraints yields a system that is effectively solved by a damped Newton scheme that appears to be robust with respect to c.

There are many questions and points of future work, however, that remain. Some simple extensions would be the following. Using the point-tangent radius (as in [41, 26, 25]) in the definition of the proximity

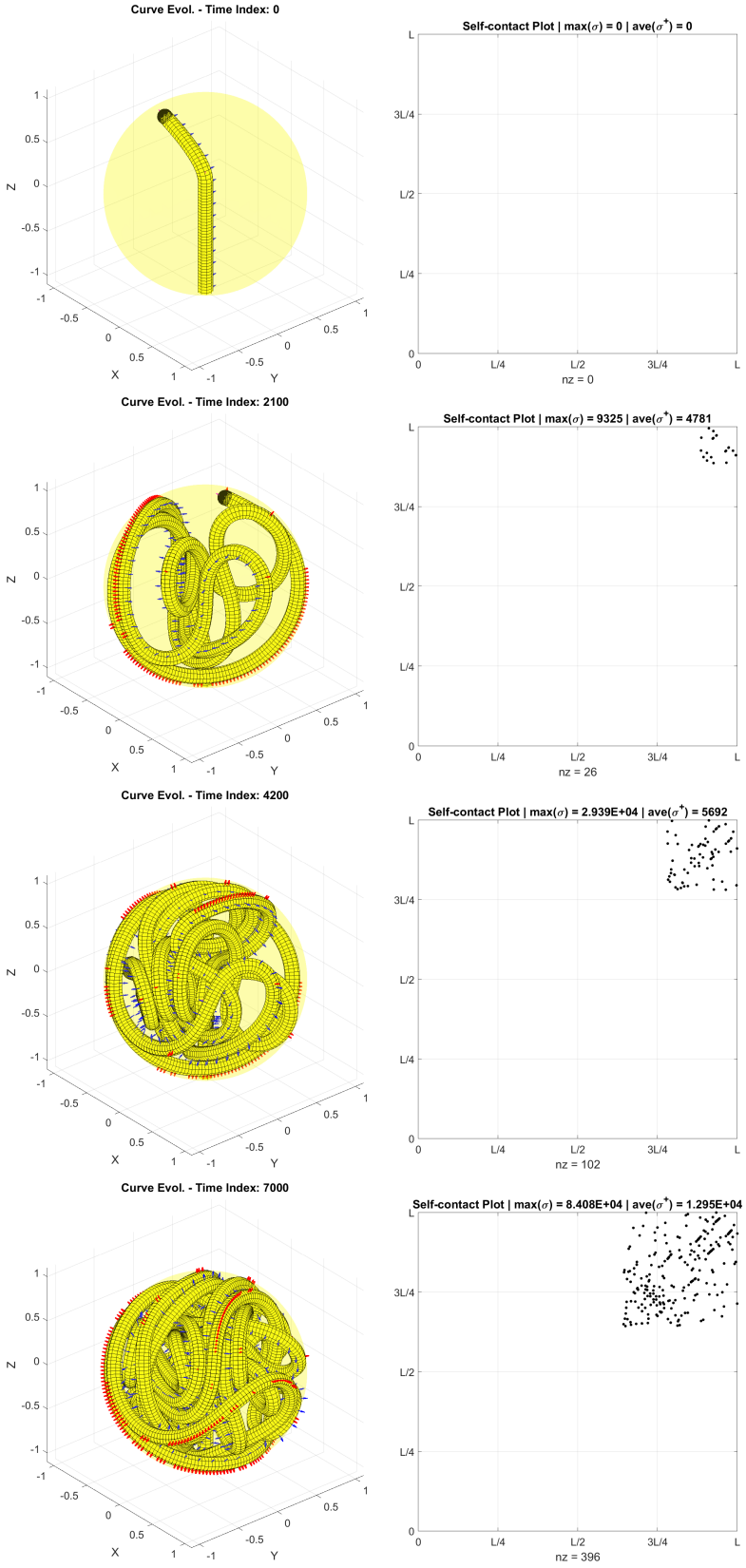

FIGURE 18. Packing of a three dimensional curve in a capsid (Section 7.4); continued in Figure 19. A tube of thickness $\varepsilon = 0.15$ is plotted around the centerline curve $\boldsymbol{x}$. The "motor" is at the bottom (not shown). Blue arrows show $\boldsymbol{b}$, and red arrows represent $\nabla\phi$ at points that touch the capsid. The right side shows a plot of the characteristic function of $\{\sigma > 0\}$ over the square U^2. The self-contact square gets filled in as more of the curve is packed into the capsid.

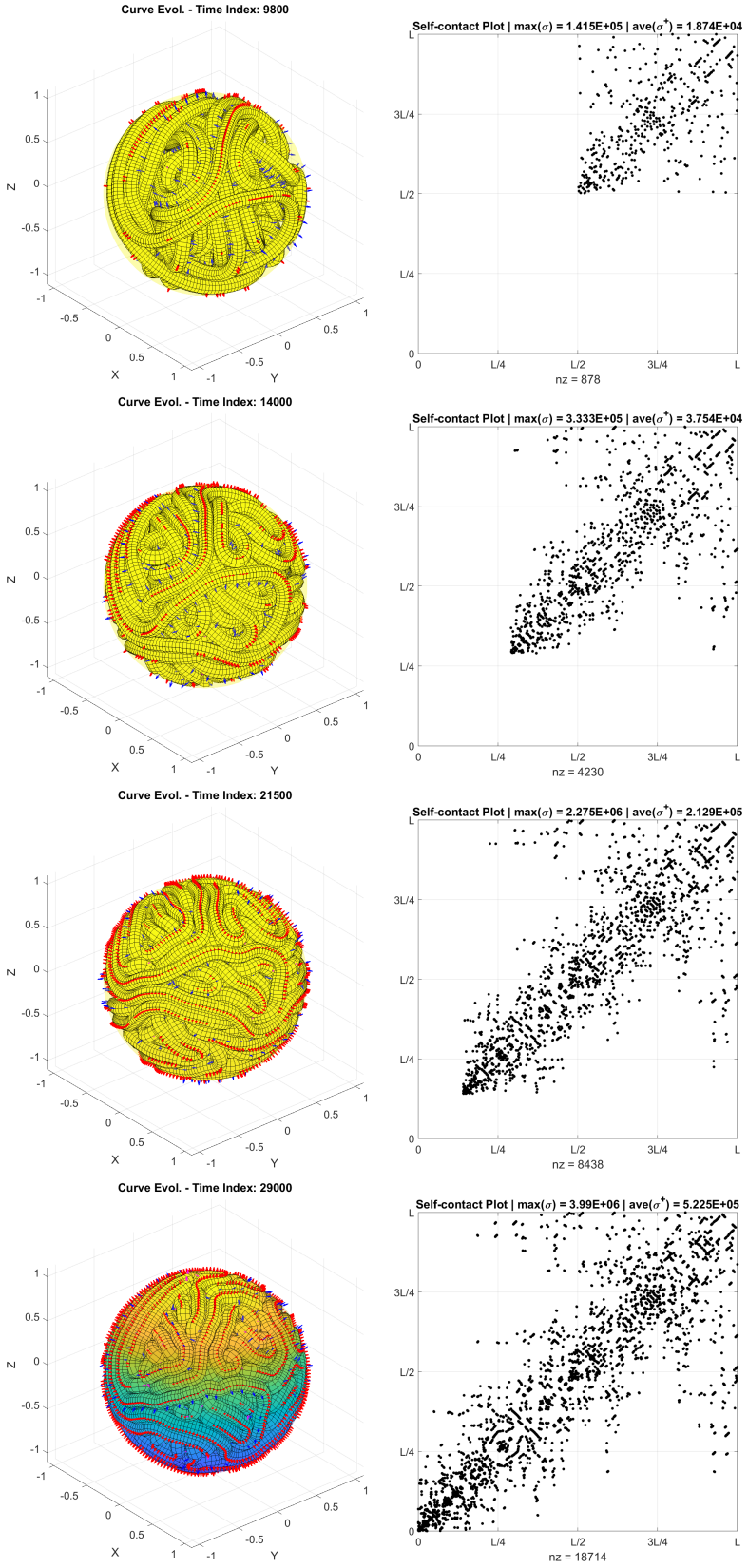

FIGURE 19. Packing of a three dimensional curve in a capsid (Section 7.4); continued from Figure 18. A tube of thickness $\varepsilon = 0.15$ is plotted around the centerline curve $\boldsymbol{x}$. The "motor" is at the bottom (not shown). Blue arrows show $\boldsymbol{b}$, and red arrows represent $\nabla\phi$ at points that touch the capsid. The right side shows a plot of the characteristic function of $\{\sigma > 0\}$ over the square U^2. A significant amount of self-contact is present at the end of the simulation.

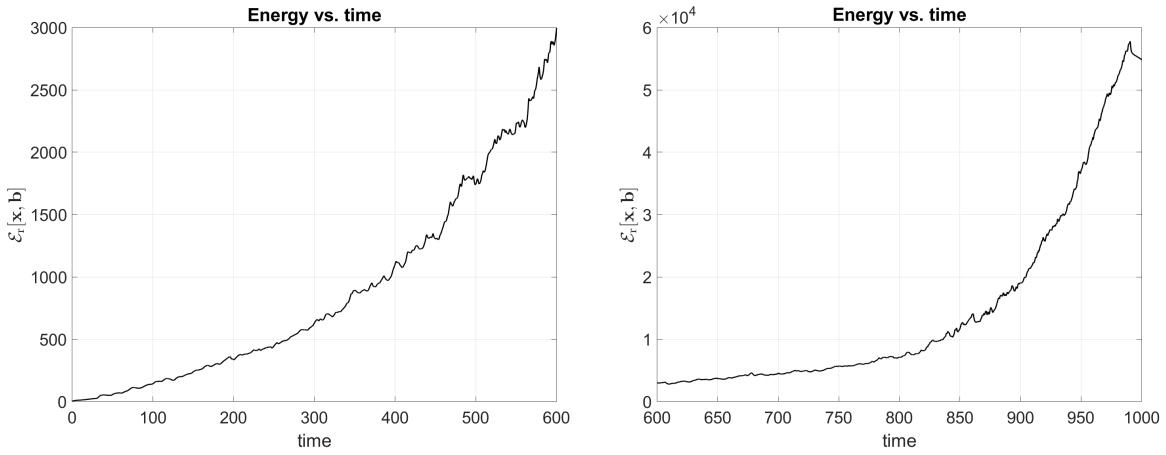

FIGURE 20. Energy plot for packing a three dimensional curve in a capsid (Section 7.4). The left plot goes to time index 12000. Note: the energy increases because energy is put into the system by the "motor." At approximately $t = 980$, the motor releases the tail of the curve, which causes the curve to relax and the energy to decrease at later times.

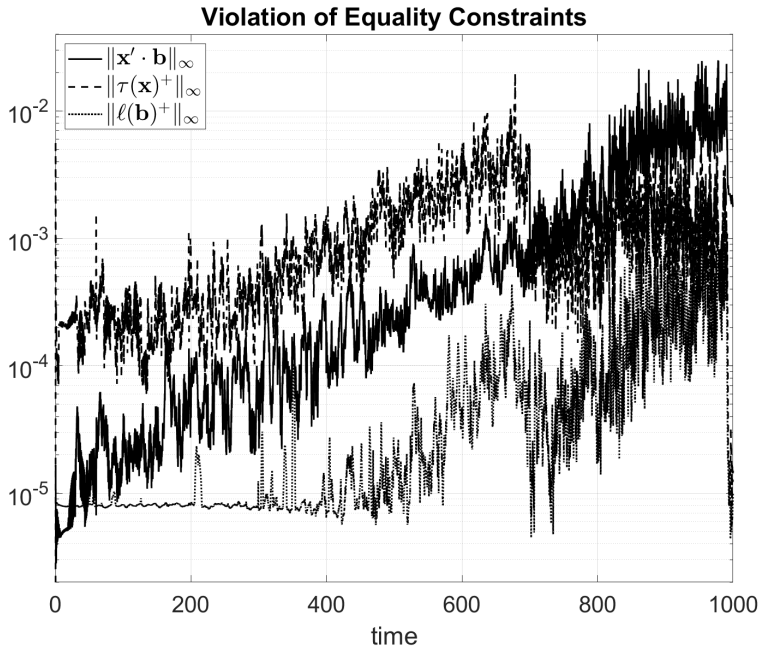

FIGURE 21. Equality constraint violation for packing a three dimensional curve in a capsid (Section 7.4).

vector (21) would have the advantage of avoiding the mesh size restriction noted in Remark 6.2. Moreover, deriving a convergent time-stepping scheme for realistic dynamics that includes frictional effects is necessary from the application perspective. In addition, developing a multi-grid or parallel method to solve the linear systems resulting from applying Newton's method to (60) could greatly speed up the method, as well as handle parameter regimes where c is chosen extremely large. And, of course, extending the method to two dimensional surfaces with thickness would be very interesting.

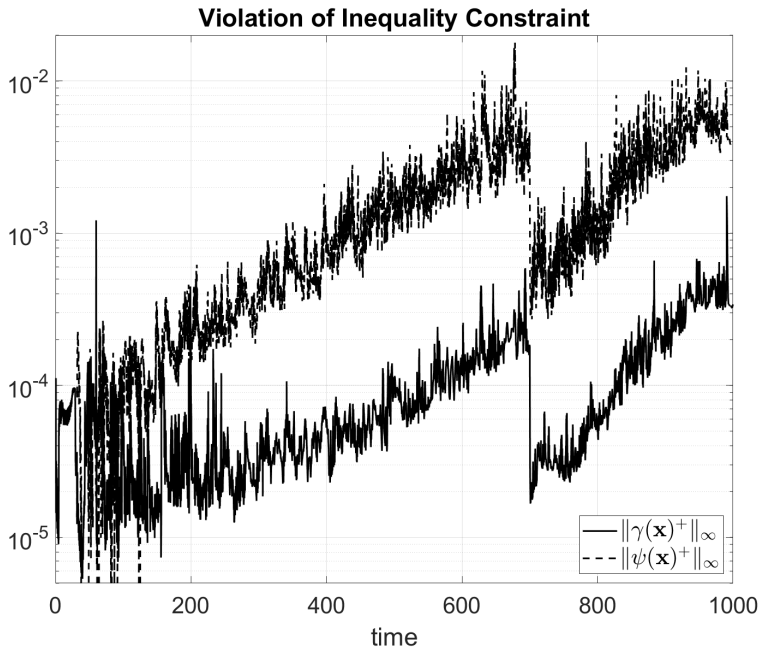

FIGURE 22. Inequality constraint violation for packing a three dimensional curve in a capsid (Section 7.4).

Furthermore, it is not clear at the moment how to derive an efficient scheme for solving the exact problem (without regularization). Solvers for convex, inequality constrained problems exist for finite dimensional problems, but it is not clear how to extend these, in an efficient way, to problems posed in infinite dimensional function space. Alternatively, how can the constraints be better satisfied in extreme situations, such as in Figure 19? Increasing c cannot be done arbitrarily.

Lastly, can the method be extended to compute ropelength [7, 18, 46], i.e. compute the shortest elastic closed curve with a given thickness and knot class. Indeed, are knot classes preserved with our algorithm for the regularized problem? For example, what conditions are needed on c, mesh size h, and time step Δt to ensure that a given knot class is preserved? Moreover, an interesting extension would be to handle networks of curves, i.e. curves connected at multi-junctions.

APPENDIX A. ENERGY ESTIMATES

Proof of Theorem 4.8. Recall the following notational replacements $\boldsymbol{x} \equiv \boldsymbol{x}^{i+1}$, $\tilde{\boldsymbol{x}} \equiv \boldsymbol{x}^i$, $\rho \equiv \rho^{i+1}$, $\lambda \equiv \lambda^{i+1}$, and $\sigma \equiv \sigma^{i+1}$, as well as $\boldsymbol{b} \equiv \boldsymbol{b}^{i+1}$, $\tilde{\boldsymbol{b}} \equiv \boldsymbol{b}^i$, $\omega \equiv \omega^{i+1}$. In addition, for notational convenience, we set $a^i(\boldsymbol{w}, \boldsymbol{y}) := c_{\mathrm{b}}(\boldsymbol{w}'', \boldsymbol{y}'')_{L^2} + m[\boldsymbol{b}^i](\boldsymbol{w}, \boldsymbol{y})$, noting that $\mathcal{E}_{\mathrm{r}}[\boldsymbol{y}, \boldsymbol{b}^i] = (1/2)a^i(\boldsymbol{y}, \boldsymbol{y}) + (c_{\mathrm{r}}/2)\|(\boldsymbol{b}^i)'\|_{L^2}^2 - \mathcal{F}(\boldsymbol{y}, \boldsymbol{b}^i)$.

Setting $\boldsymbol{w} = \boldsymbol{x}^{i+1} - \boldsymbol{x}^i$ in (42), we get

$$
\begin{aligned}
a^i\left(\boldsymbol{x}^{i+1}, \boldsymbol{x}^{i+1} - \boldsymbol{x}^i\right) + \frac{1}{\Delta t}\|\boldsymbol{x}^{i+1} - \boldsymbol{x}^i\|_{H^2}^2 + \left\langle \rho^{i+1}, \tau'_{\boldsymbol{x}^i}\left(\boldsymbol{x}^{i+1} - \boldsymbol{x}^i\right)\right\rangle \\
+ \left\langle \lambda^{i+1}, \gamma'_{\boldsymbol{x}^i}\left(\boldsymbol{x}^{i+1} - \boldsymbol{x}^i\right)\right\rangle + \left\langle\left\langle \sigma^{i+1}, \psi'_{\boldsymbol{x}^i}\left(\boldsymbol{x}^{i+1} - \boldsymbol{x}^i\right)\right\rangle\right\rangle = \mathcal{F}_{\mathrm{x}}\left(\boldsymbol{x}^{i+1} - \boldsymbol{x}^i\right),
\end{aligned} \tag{119}
$$

Combining with (49), (50), and (51), we have

$$
a^i\left(\boldsymbol{x}^{i+1}, \boldsymbol{x}^{i+1} - \boldsymbol{x}^i\right) + \frac{1}{\Delta t}\|\boldsymbol{x}^{i+1} - \boldsymbol{x}^i\|_{H^2}^2 \leq \mathcal{F}_{\mathrm{x}}\left(\boldsymbol{x}^{i+1} - \boldsymbol{x}^i\right) + \|\tau(\boldsymbol{x}^i)\|_{H^1}\|\rho^{i+1}\|_{H^{-1}} + \|[\gamma(\boldsymbol{x}^i)]^+\|_{L^\infty}\left\langle \lambda^{i+1}, 1\right\rangle. \tag{120}
$$

Utilizing the classic identity $2a(a-b)=a^2-b^2+(a-b)^2$, we get

$$\begin{aligned}\frac{1}{2}a^i\left(\boldsymbol{x}^{i+1}-\boldsymbol{x}^i,\boldsymbol{x}^{i+1}-\boldsymbol{x}^i\right)+\frac{1}{2}a^i\left(\boldsymbol{x}^{i+1},\boldsymbol{x}^{i+1}\right)-\frac{1}{2}a^i\left(\boldsymbol{x}^i,\boldsymbol{x}^i\right)\\ +\frac{1}{\Delta t}\|\boldsymbol{x}^{i+1}-\boldsymbol{x}^i\|_{H^2}^2\leq\mathcal{F}_{\mathrm{x}}\left(\boldsymbol{x}^{i+1}-\boldsymbol{x}^i\right)+\|\tau(\boldsymbol{x}^i)\|_{H^1}\|\rho^{i+1}\|_{H^{-1}}+\|[\gamma(\boldsymbol{x}^i)]^+\|_{L^\infty}\left\langle\lambda^{i+1},1\right\rangle.\end{aligned}\tag{121}$$

From (121), we can obtain a (quasi) energy decrease property, i.e. we first have

$$\begin{aligned}&\frac{1}{2}a^i\left(\boldsymbol{x}^{i+1},\boldsymbol{x}^{i+1}\right)+\frac{c_{\mathrm{r}}}{2}\|(\boldsymbol{b}^i)'\|_{L^2}^2-\mathcal{F}_{\mathrm{x}}\left(\boldsymbol{x}^{i+1}\right)+\frac{1}{\Delta t}\|\boldsymbol{x}^{i+1}-\boldsymbol{x}^i\|_{H^2}^2\\ &\quad\leq\frac{1}{2}a^i\left(\boldsymbol{x}^i,\boldsymbol{x}^i\right)+\frac{c_{\mathrm{r}}}{2}\|(\boldsymbol{b}^i)'\|_{L^2}^2-\mathcal{F}_{\mathrm{x}}\left(\boldsymbol{x}^i\right)+\|\tau(\boldsymbol{x}^i)\|_{H^1}\|\rho^{i+1}\|_{H^{-1}}+\|[\gamma(\boldsymbol{x}^i)]^+\|_{L^\infty}\left\langle\lambda^{i+1},1\right\rangle,\end{aligned}\tag{122}$$

which gives

$$\mathcal{E}_{\mathrm{r}}\left[\boldsymbol{x}^{i+1},\boldsymbol{b}^i\right]+\frac{1}{\Delta t}\|\boldsymbol{x}^{i+1}-\boldsymbol{x}^i\|_{H^2}^2\leq\mathcal{E}_{\mathrm{r}}\left[\boldsymbol{x}^i,\boldsymbol{b}^i\right]+\|\tau(\boldsymbol{x}^i)\|_{H^1}\|\rho^{i+1}\|_{H^{-1}}+\|[\gamma(\boldsymbol{x}^i)]^+\|_{L^\infty}\left\langle\lambda^{i+1},1\right\rangle.\tag{123}$$

Next, for notational convenience, we set $e^i(\boldsymbol{v},\boldsymbol{z}):=c_{\mathrm{r}}(\boldsymbol{v}',\boldsymbol{z}')_{L^2}+m[\boldsymbol{x}^{i+1}](\boldsymbol{v},\boldsymbol{z})$, noting that $\mathcal{E}_{\mathrm{r}}\left[\boldsymbol{x}^{i+1},\boldsymbol{v}\right]=(1/2)e^i(\boldsymbol{v},\boldsymbol{v})+(c_{\mathrm{b}}/2)\|(\boldsymbol{x}^{i+1})''\|_{L^2}^2-\mathcal{F}\left(\boldsymbol{x}^{i+1},\boldsymbol{v}\right)$. Setting $\boldsymbol{v}=\boldsymbol{b}^{i+1}-\boldsymbol{b}^i$ in (47), we have

$$e^i\left(\boldsymbol{b}^{i+1},\boldsymbol{b}^{i+1}-\boldsymbol{b}^i\right)+\frac{1}{\Delta t}\|\boldsymbol{b}^{i+1}-\boldsymbol{b}^i\|_{H^1}^2=-\left\langle\omega^{i+1},\ell'_{\boldsymbol{b}^i}\left(\boldsymbol{b}^{i+1}-\boldsymbol{b}^i\right)\right\rangle,\tag{124}$$

Again, using the classic identity $2a(a-b)=a^2-b^2+(a-b)^2$, and (52), we get

$$\begin{aligned}\frac{1}{2}e^i\left(\boldsymbol{b}^{i+1}-\boldsymbol{b}^i,\boldsymbol{b}^{i+1}-\boldsymbol{b}^i\right)+\frac{1}{2}e^i\left(\boldsymbol{b}^{i+1},\boldsymbol{b}^{i+1}\right)-\frac{1}{2}e^i\left(\boldsymbol{b}^i,\boldsymbol{b}^i\right)\\ +\frac{1}{\Delta t}\|\boldsymbol{b}^{i+1}-\boldsymbol{b}^i\|_{H^1}^2\leq\|\omega^{i+1}\|_{H^{-1}}\|\ell(\boldsymbol{b}^i)\|_{H^1}.\end{aligned}\tag{125}$$

From (125), we can obtain a (quasi) energy decrease property, i.e. we first have

$$\begin{aligned}&\frac{c_{\mathrm{b}}}{2}\|(\boldsymbol{x}^{i+1})''\|_{L^2}^2+\frac{1}{2}e^i\left(\boldsymbol{b}^{i+1},\boldsymbol{b}^{i+1}\right)-\mathcal{F}_{\mathrm{x}}\left(\boldsymbol{x}^{i+1}\right)+\frac{1}{\Delta t}\|\boldsymbol{b}^{i+1}-\boldsymbol{b}^i\|_{H^1}^2\\ &\quad\leq\frac{c_{\mathrm{b}}}{2}\|(\boldsymbol{x}^{i+1})''\|_{L^2}^2+\frac{1}{2}e^i\left(\boldsymbol{b}^i,\boldsymbol{b}^i\right)-\mathcal{F}_{\mathrm{x}}\left(\boldsymbol{x}^{i+1}\right)+\|\ell(\boldsymbol{b}^i)\|_{H^1}\|\omega^{i+1}\|_{H^{-1}}.\end{aligned}\tag{126}$$

which gives

$$\mathcal{E}_{\mathrm{r}}\left[\boldsymbol{x}^{i+1},\boldsymbol{b}^{i+1}\right]+\frac{1}{\Delta t}\|\boldsymbol{b}^{i+1}-\boldsymbol{b}^i\|_{H^1}^2\leq\mathcal{E}_{\mathrm{r}}\left[\boldsymbol{x}^{i+1},\boldsymbol{b}^i\right]+\|\ell(\boldsymbol{b}^i)\|_{H^1}\|\omega^{i+1}\|_{H^{-1}}.\tag{127}$$

Combining (123) and (127), for all $i\geq 0$, we get

$$\mathcal{E}_{\mathrm{r}}\left[\boldsymbol{x}^{i+1},\boldsymbol{b}^{i+1}\right]+\frac{1}{\Delta t}\|\boldsymbol{x}^{i+1}-\boldsymbol{x}^i\|_{H^2}^2+\frac{1}{\Delta t}\|\boldsymbol{b}^{i+1}-\boldsymbol{b}^i\|_{H^1}^2\leq\mathcal{E}_{\mathrm{r}}\left[\boldsymbol{x}^i,\boldsymbol{b}^i\right]+\Xi_1^i+\Xi_2^i,\tag{128}$$

where, upon recalling Assumption 4.7, $\Xi_1^i:=M_1(\|\tau(\boldsymbol{x}^i)\|_{H^1}+\|[\gamma(\boldsymbol{x}^i)]^+\|_{L^\infty})$, $\Xi_2^i:=M_2\|\ell(\boldsymbol{b}^i)\|_{H^1}$. Upon recalling (53), if $\Delta t\leq(2v_0M_1)^{-1}$ and $\Delta t\leq(v_0M_2)^{-1}$, then the previous estimates give the following (quasi)-energy decrease result

$$\mathcal{E}_{\mathrm{r}}\left[\boldsymbol{x}^{i+1},\boldsymbol{b}^{i+1}\right]+\Xi_1^{i+1}+\Xi_2^{i+1}\leq\mathcal{E}_{\mathrm{r}}\left[\boldsymbol{x}^i,\boldsymbol{b}^i\right]+\Xi_1^i+\Xi_2^i,\quad\text{for all } i\geq 0,\tag{129}$$

i.e. $E^i+\Xi_1^i+\Xi_2^i\leq E^0+\Xi_1^0+\Xi_2^0=:F^0$, for all $i\geq 0$, where $E^i:=\mathcal{E}_{\mathrm{r}}\left[\boldsymbol{x}^i,\boldsymbol{b}^i\right]$.

A simple bound on $\mathcal{F}_{\mathrm{x}}\left(\boldsymbol{x}^{i+1}\right)=\mathcal{F}_{\mathrm{x}}\left(\boldsymbol{x}^{i+1}-\boldsymbol{x}_{\mathrm{bc}}\right)+\mathcal{F}_{\mathrm{x}}\left(\boldsymbol{x}_{\mathrm{bc}}\right)$ gives

$$\begin{aligned}|\mathcal{F}_{\mathrm{x}}\left(\boldsymbol{x}^{i+1}\right)|&\leq\|\mathcal{F}_{\mathrm{x}}\|_{(H^2)^*}C_P\|(\boldsymbol{x}^{i+1}-\boldsymbol{x}_{\mathrm{bc}})''\|_{L^2}+|\mathcal{F}_{\mathrm{x}}\left(\boldsymbol{x}_{\mathrm{bc}}\right)|\\ &\leq C_P\|\mathcal{F}_{\mathrm{x}}\|_{(H^2)^*}\|(\boldsymbol{x}^{i+1})''\|_{L^2}+\|\mathcal{F}_{\mathrm{x}}\|_{(H^2)^*}C_P\|\boldsymbol{x}''_{\mathrm{bc}}\|_{L^2}+|\mathcal{F}_{\mathrm{x}}\left(\boldsymbol{x}_{\mathrm{bc}}\right)|\\ &\leq(C_P^2/c_{\mathrm{b}})\|\mathcal{F}_{\mathrm{x}}\|_{(H^2)^*}^2+(c_{\mathrm{b}}/4)\|(\boldsymbol{x}^{i+1})''\|_{L^2}^2+\|\mathcal{F}_{\mathrm{x}}\|_{(H^2)^*}\left[C_P\|\boldsymbol{x}''_{\mathrm{bc}}\|_{L^2}+\|\boldsymbol{x}_{\mathrm{bc}}\|_{H^2}\right]\end{aligned}\tag{130}$$

where $\|\boldsymbol{y}\|_{H^2}\leq C_P\|\boldsymbol{y}''\|_{L^2}$, which then yields, for all $i\geq 0$,

$$\begin{aligned}(c_{\mathrm{b}}/4)\|(\boldsymbol{x}^{i+1})''\|_{L^2}^2+(c_{\mathrm{r}}/2)\|(\boldsymbol{b}^{i+1})'\|_{L^2}^2+\Delta t^{-1}\|\boldsymbol{x}^{i+1}-\boldsymbol{x}^i\|_{H^2}^2+\Delta t^{-1}\|\boldsymbol{b}^{i+1}-\boldsymbol{b}^i\|_{H^1}^2\\ \leq E^i+\Xi_1^i+\Xi_2^i+Q_0\leq F^0+Q_0,\end{aligned}\tag{131}$$

where $Q_0:=(C_P^2/c_{\mathrm{b}})\|\mathcal{F}_{\mathrm{x}}\|_{(H^2)^*}^2+\|\mathcal{F}_{\mathrm{x}}\|_{(H^2)^*}\left[C_P\|\boldsymbol{x}''_{\mathrm{bc}}\|_{L^2}+\|\boldsymbol{x}_{\mathrm{bc}}\|_{H^2}\right]$.

From (121), we have

$$\begin{aligned}\frac{1}{2}a^i\left(\boldsymbol{x}^{i+1}-\boldsymbol{x}^i,\boldsymbol{x}^{i+1}-\boldsymbol{x}^i\right)+\frac{1}{\Delta t}\|\boldsymbol{x}^{i+1}-\boldsymbol{x}^i\|_{H^2}^2\\ \leq\frac{1}{2}a^i\left(\boldsymbol{x}^i,\boldsymbol{x}^i\right)-\frac{1}{2}a^i\left(\boldsymbol{x}^{i+1},\boldsymbol{x}^{i+1}\right)+\mathcal{F}_{\mathrm{x}}\left(\boldsymbol{x}^{i+1}-\boldsymbol{x}^i\right)+\Xi_1^i.\end{aligned}\tag{132}$$

Again, using $a^2-b^2=2a(a-b)-(a-b)^2$, with $a^2=a^i\left(\boldsymbol{x}^i,\boldsymbol{x}^i\right)$ and $b^2=a^i\left(\boldsymbol{x}^{i+1},\boldsymbol{x}^{i+1}\right)$, we get

$$a^i\left(\boldsymbol{x}^{i+1}-\boldsymbol{x}^i,\boldsymbol{x}^{i+1}-\boldsymbol{x}^i\right)+\frac{1}{\Delta t}\|\boldsymbol{x}^{i+1}-\boldsymbol{x}^i\|_{H^2}^2\leq a^i\left(\boldsymbol{x}^i,\boldsymbol{x}^i-\boldsymbol{x}^{i+1}\right)+\|\mathcal{F}_{\mathrm{x}}\|_{(H^2)^*}\|\boldsymbol{x}^{i+1}-\boldsymbol{x}^i\|_{H^2}+\Xi_1^i.\tag{133}$$

Focusing on $a^i\left(\boldsymbol{x}^i,\boldsymbol{x}^i-\boldsymbol{x}^{i+1}\right)$, and defining the 3×3 matrix $B^i=\left(|(\boldsymbol{b}^i)'|^2\boldsymbol{I}-(\boldsymbol{b}^i)'\otimes(\boldsymbol{b}^i)'\right)^{1/2}$, with $|B^i|\leq|(\boldsymbol{b}^i)'|$, we find that

$$\begin{aligned}a^i\left(\boldsymbol{x}^i,\boldsymbol{x}^i-\boldsymbol{x}^{i+1}\right)&=c_{\mathrm{b}}\left((\boldsymbol{x}^i)'',(\boldsymbol{x}^i-\boldsymbol{x}^{i+1})''\right)_{L^2}+c_{\mathrm{t}}\left(B^i(\boldsymbol{x}^i)',B^i(\boldsymbol{x}^i-\boldsymbol{x}^{i+1})'\right)_{L^2}\\&\quad+c_{\mathrm{o}}\left(\boldsymbol{b}^i\cdot(\boldsymbol{x}^i)',\boldsymbol{b}^i\cdot(\boldsymbol{x}^i-\boldsymbol{x}^{i+1})'\right)_{L^2}\\&\leq c_{\mathrm{b}}\|(\boldsymbol{x}^{i+1}-\boldsymbol{x}^i)''\|_{L^2}\|(\boldsymbol{x}^i)''\|_{L^2}+(c_{\mathrm{t}}/c_{\mathrm{r}})^{1/2}\|(\boldsymbol{x}^{i+1}-\boldsymbol{x}^i)'\|_{L^\infty}c_{\mathrm{t}}^{1/2}\|B^i(\boldsymbol{x}^i)'\|_{L^2}c_{\mathrm{r}}^{1/2}\|B^i\|_{L^2}\\&\quad+c_{\mathrm{o}}^{1/2}\|(\boldsymbol{x}^{i+1}-\boldsymbol{x}^i)'\|_{L^\infty}c_{\mathrm{o}}^{1/2}\|\boldsymbol{b}^i\cdot(\boldsymbol{x}^i)'\|_{L^2}\|\boldsymbol{b}^i\|_{L^2}\\&\leq\alpha_1\|\boldsymbol{x}^{i+1}-\boldsymbol{x}^i\|_{H^2}\left[c_{\mathrm{b}}^{1/2}\|(\boldsymbol{x}^i)''\|_{L^2}+c_{\mathrm{t}}^{1/2}\|B^i(\boldsymbol{x}^i)'\|_{L^2}c_{\mathrm{r}}^{1/2}\|B^i\|_{L^2}+c_{\mathrm{o}}^{1/2}\|\boldsymbol{b}^i\cdot(\boldsymbol{x}^i)'\|_{L^2}\|\boldsymbol{b}^i\|_{L^2}\right],\end{aligned}\tag{134}$$

where $\alpha_1=\max\left(c_{\mathrm{b}}^{1/2},C_1(c_{\mathrm{t}}/c_{\mathrm{r}})^{1/2},C_1c_{\mathrm{o}}^{1/2}\right)$ and $C_1>0$ is a uniform embedding constant. The term in brackets $[*]$ is bounded by

$$\begin{aligned}[*]&\leq 1+\frac{c_{\mathrm{b}}}{4}\|(\boldsymbol{x}^i)''\|_{L^2}^2+\frac{c_{\mathrm{t}}}{2}\|B^i(\boldsymbol{x}^i)'\|_{L^2}^2+\frac{c_{\mathrm{r}}}{2}\|(\boldsymbol{x}^i)'\|_{L^2}^2+\frac{c_{\mathrm{o}}}{2}\|\boldsymbol{b}^i\cdot(\boldsymbol{x}^i)'\|_{L^2}^2-\mathcal{F}_{\mathrm{x}}\left(\boldsymbol{x}^i\right)\\&\quad+C_P\|\mathcal{F}_{\mathrm{x}}\|_{(H^2)^*}\|(\boldsymbol{x}^i)''\|_{L^2}+\frac{1}{2}\|\boldsymbol{b}^i\|_{L^2}^2+\|\mathcal{F}_{\mathrm{x}}\|_{(H^2)^*}\\&\leq 1+\frac{c_{\mathrm{b}}}{4}\|(\boldsymbol{x}^i)''\|_{L^2}^2+\frac{c_{\mathrm{t}}}{2}\|B^i(\boldsymbol{x}^i)'\|_{L^2}^2+\frac{c_{\mathrm{r}}}{2}\|(\boldsymbol{x}^i)'\|_{L^2}^2+\frac{c_{\mathrm{o}}}{2}\|\boldsymbol{b}^i\cdot(\boldsymbol{x}^i)'\|_{L^2}^2-\mathcal{F}_{\mathrm{x}}\left(\boldsymbol{x}^i\right)\\&\quad+\frac{C_P^2}{c_{\mathrm{b}}}\|\mathcal{F}_{\mathrm{x}}\|_{(H^2)^*}^2+\frac{c_{\mathrm{b}}}{4}\|(\boldsymbol{x}^i)''\|_{L^2}^2+\frac{1}{2}\|\boldsymbol{b}^i\|_{L^2}^2+\|\mathcal{F}_{\mathrm{x}}\|_{(H^2)^*}\\&=1+\frac{1}{2}\|\boldsymbol{b}^i\|_{L^2}^2+\|\mathcal{F}_{\mathrm{x}}\|_{(H^2)^*}+\frac{C_P^2}{c_{\mathrm{b}}}\|\mathcal{F}_{\mathrm{x}}\|_{(H^2)^*}^2+E^i.\end{aligned}\tag{135}$$

Now, note that

$$\|\boldsymbol{b}^i\|_{L^2}^2=\||\boldsymbol{b}^i|^2-1\|_{L^1}+L\leq C_3\|\ell(\boldsymbol{b}^i)\|_{H^1}+L=\frac{C_3}{M_2}\Xi_2^i+L,\tag{136}$$

for some embedding constant $C_3>0$. So, (135) simplifies to

$$\begin{aligned}[*]&\leq 1+\frac{L}{2}+\|\mathcal{F}_{\mathrm{x}}\|_{(H^2)^*}+\frac{C_P^2}{c_{\mathrm{b}}}\|\mathcal{F}_{\mathrm{x}}\|_{(H^2)^*}^2+E^i+\frac{C_3}{2M_2}\Xi_2^i\\&\leq 1+\frac{L}{2}+\|\mathcal{F}_{\mathrm{x}}\|_{(H^2)^*}+\frac{C_P^2}{c_{\mathrm{b}}}\|\mathcal{F}_{\mathrm{x}}\|_{(H^2)^*}^2+C_4F^0=:\widetilde{R}_1,\end{aligned}\tag{137}$$

where $C_4:=\max(C_3/(2M_2),1)$. Hence, combining (133), (134), and (137), we obtain

$$\begin{aligned}\frac{1}{\Delta t}\|\boldsymbol{x}^{i+1}-\boldsymbol{x}^i\|_{H^2}^2&\leq\left(\alpha_1\widetilde{R}_1+\|\mathcal{F}_{\mathrm{x}}\|_{(H^2)^*}\right)\|\boldsymbol{x}^{i+1}-\boldsymbol{x}^i\|_{H^2}+\Xi_1^i\\&\leq\left(\alpha_1\widetilde{R}_1+\|\mathcal{F}_{\mathrm{x}}\|_{(H^2)^*}\right)\|\boldsymbol{x}^{i+1}-\boldsymbol{x}^i\|_{H^2}+2v_0M_1\Delta t\left(F^0+Q_0\right),\end{aligned}\tag{138}$$

where we used (131). Hence, there is a uniform constant R_1, such that

$$\|\boldsymbol{x}^{i+1}-\boldsymbol{x}^i\|_{H^2}\leq R_1\Delta t,\quad\text{for all } i\geq 0.\tag{139}$$

From (125), we have

$$\frac{1}{2}e^i\left(\boldsymbol{b}^{i+1}-\boldsymbol{b}^i,\boldsymbol{b}^{i+1}-\boldsymbol{b}^i\right)+\frac{1}{\Delta t}\|\boldsymbol{b}^{i+1}-\boldsymbol{b}^i\|_{H^1}^2\leq\frac{1}{2}e^i\left(\boldsymbol{b}^i,\boldsymbol{b}^i\right)-\frac{1}{2}e^i\left(\boldsymbol{b}^{i+1},\boldsymbol{b}^{i+1}\right)+\Xi_2^i.\tag{140}$$

Again, using $a^2-b^2=2a(a-b)-(a-b)^2$, with $a^2=e^i\left(\boldsymbol{b}^i,\boldsymbol{b}^i\right)$ and $b^2=e^i\left(\boldsymbol{b}^{i+1},\boldsymbol{b}^{i+1}\right)$, we get

$$e^i\left(\boldsymbol{b}^{i+1}-\boldsymbol{b}^i,\boldsymbol{b}^{i+1}-\boldsymbol{b}^i\right)+\frac{1}{\Delta t}\|\boldsymbol{b}^{i+1}-\boldsymbol{b}^i\|_{H^1}^2\leq e^i\left(\boldsymbol{b}^i,\boldsymbol{b}^i-\boldsymbol{b}^{i+1}\right)+\Xi_2^i. \tag{141}$$

Focusing on $e^i\left(\boldsymbol{b}^i,\boldsymbol{b}^i-\boldsymbol{b}^{i+1}\right)$, and defining the 3×3 matrix $X^{i+1}=\left(|(\boldsymbol{x}^{i+1})'|^2\boldsymbol{I}-(\boldsymbol{x}^{i+1})'\otimes(\boldsymbol{x}^{i+1})'\right)^{1/2}$, with $|X^{i+1}|\leq|(\boldsymbol{x}^{i+1})'|$, we find that

$$\begin{aligned}
e^i\left(\boldsymbol{b}^i,\boldsymbol{b}^i-\boldsymbol{b}^{i+1}\right)&=c_{\mathrm{r}}\left((\boldsymbol{b}^i)',(\boldsymbol{b}^i-\boldsymbol{b}^{i+1})'\right)_{L^2}+c_{\mathrm{t}}\left(X^{i+1}(\boldsymbol{b}^i)',X^{i+1}(\boldsymbol{b}^i-\boldsymbol{b}^{i+1})'\right)_{L^2}\\
&\quad+c_{\mathrm{o}}\left((\boldsymbol{x}^{i+1})'\cdot\boldsymbol{b}^i,(\boldsymbol{x}^{i+1})'\cdot(\boldsymbol{b}^i-\boldsymbol{b}^{i+1})\right)_{L^2}\\
&\leq c_{\mathrm{r}}^{1/2}\|(\boldsymbol{b}^{i+1}-\boldsymbol{b}^i)'\|_{L^2}c_{\mathrm{r}}^{1/2}\|(\boldsymbol{b}^i)'\|_{L^2}\\
&\quad+(c_{\mathrm{t}}/c_{\mathrm{b}})^{1/2}\|(\boldsymbol{b}^{i+1}-\boldsymbol{b}^i)'\|_{L^2}c_{\mathrm{t}}^{1/2}\|X^{i+1}(\boldsymbol{b}^i)'\|_{L^2}c_{\mathrm{b}}^{1/2}\|X^{i+1}\|_{L^\infty}\\
&\quad+(c_{\mathrm{o}}/c_{\mathrm{b}})^{1/2}\|\boldsymbol{b}^{i+1}-\boldsymbol{b}^i\|_{L^2}c_{\mathrm{o}}^{1/2}\|(\boldsymbol{x}^{i+1})'\cdot\boldsymbol{b}^i\|_{L^2}c_{\mathrm{b}}^{1/2}\|(\boldsymbol{x}^{i+1})'\|_{L^\infty}\\
&\leq\alpha_2\|\boldsymbol{b}^{i+1}-\boldsymbol{b}^i\|_{H^1}\Big[c_{\mathrm{r}}^{1/2}\|(\boldsymbol{b}^i)'\|_{L^2}+c_{\mathrm{t}}^{1/2}\|X^{i+1}(\boldsymbol{b}^i)'\|_{L^2}(c_{\mathrm{b}}/(2C_2))^{1/2}\|X^{i+1}\|_{L^\infty}\\
&\qquad+c_{\mathrm{o}}^{1/2}\|(\boldsymbol{x}^{i+1})'\cdot\boldsymbol{b}^i\|_{L^2}(c_{\mathrm{b}}/(2C_2))^{1/2}\|(\boldsymbol{x}^{i+1})'\|_{L^\infty}\Big],
\end{aligned} \tag{142}$$

where $\alpha_2^2=\max\left(c_{\mathrm{r}},2C_2(c_{\mathrm{t}}/c_{\mathrm{b}}),2C_2(c_{\mathrm{o}}/c_{\mathrm{b}})\right)$, with $C_2\geq 1$ the uniform embedding constant for $\|\boldsymbol{x}'\|_{L^\infty}^2\leq C_2\|\boldsymbol{x}''\|_{L^2}^2$. The term in brackets $[*]$ is bounded by

$$\begin{aligned}
[*]&\leq\frac{1}{2}+\frac{c_{\mathrm{r}}}{2}\|(\boldsymbol{b}^i)'\|_{L^2}^2+\frac{c_{\mathrm{t}}}{2}\|X^{i+1}(\boldsymbol{b}^i)'\|_{L^2}^2+\frac{c_{\mathrm{o}}}{2}\|(\boldsymbol{x}^{i+1})'\cdot\boldsymbol{b}^i\|_{L^2}^2+\frac{c_{\mathrm{b}}}{2}\|(\boldsymbol{x}^{i+1})''\|_{L^2}^2\\
&=\frac{1}{2}+\mathcal{E}_{\mathrm{r}}\left[\boldsymbol{x}^{i+1},\boldsymbol{b}^i\right]+\mathcal{F}_{\mathrm{x}}\left(\boldsymbol{x}^{i+1}\right)\leq\frac{1}{2}+\mathcal{E}_{\mathrm{r}}\left[\boldsymbol{x}^{i+1},\boldsymbol{b}^i\right]+\frac{c_{\mathrm{b}}}{4}\|(\boldsymbol{x}^{i+1})''\|_{L^2}^2+Q_0\\
&\leq\frac{1}{2}+2E^i+2\Xi_1^i+\Xi_2^i+2Q_0\leq\frac{1}{2}+2F^0+2Q_0=:\widetilde{R}_2,
\end{aligned} \tag{143}$$

where $\mathcal{E}_{\mathrm{r}}\left[\boldsymbol{x}^{i+1},\boldsymbol{b}^i\right]\leq E^i+\Xi_1^i$ and we used (131). Combining with (141), (142) yields

$$\begin{aligned}
\frac{1}{\Delta t}\|\boldsymbol{b}^{i+1}-\boldsymbol{b}^i\|_{H^1}^2&\leq\alpha_2\|\boldsymbol{b}^{i+1}-\boldsymbol{b}^i\|_{H^1}\widetilde{R}_2+\Xi_2^i\\
&\leq\Delta t\frac{(\alpha_2\widetilde{R}_2)^2}{2}+\frac{1}{2\Delta t}\|\boldsymbol{b}^{i+1}-\boldsymbol{b}^i\|_{H^1}^2+v_0M_2\Delta t\left(F^0+Q_0\right),
\end{aligned} \tag{144}$$

where we used (131). Hence, there is a uniform constant R_2, such that

$$\|\boldsymbol{b}^{i+1}-\boldsymbol{b}^i\|_{H^1}\leq R_2\Delta t,\quad\text{for all } i\geq 0. \tag{145}$$

Furthermore, by (53), we obtain (57).

References

[1] I. Ali, D. Marenduzzo, and J. M. Yeomans. Polymer packaging and ejection in viral capsids: Shape matters. *Phys. Rev. Lett.*, 96:208102, May 2006.

[2] Stuart S. Antman. *Nonlinear Problems of Elasticity*. Springer, 2nd edition, 2005.

[3] F. Armero and J. Valverde. Invariant hermitian finite elements for thin kirchhoff rods. ii: The linear three-dimensional case. *Computer Methods in Applied Mechanics and Engineering*, 213-216:458–485, 2012.

[4] Javier Arsuaga, Robert K.-Z Tan, Mariel Vázquez, De Witt Sumners, and Stephen C. Harvey. Investigation of viral dna packaging using molecular mechanics models. *Biophysical Chemistry*, 101-102:475 – 484, 2002. Special issue in honour of John A Schellman.

[5] K. Arunakirinathar and B. D. Reddy. Mixed finite element methods for elastic rods of arbitrary geometry. *Numerische Mathematik*, 64(1):13–43, Dec 1993.

[6] T. Ashton, J. Cantarella, M. Piatek, and E. J. Rawdon. Knot tightening by constrained gradient descent. *Experimental Mathematics*, 20(1):57–90, 2011.

[7] Ted Ashton and Jason Cantarella. *Physical and Numerical Models in Knot Theory and their Application to the Life Sciences,*, chapter A fast octree-based algorithm for computing ropelength, pages 323 – 341. World Scientific Press, 2005.

[8] Matteo Astorino, Jean-Frédéric Gerbeau, Olivier Pantz, and Karim-Frédéric Traoré. Fluid–structure interaction and multi-body contact: Application to aortic valves. *Computer Methods in Applied Mechanics and Engineering*, 198(45):3603 – 3612, 2009. Models and Methods in Computational Vascular and Cardiovascular Mechanics.

[9] Sören Bartels. Chapter 3 - finite element simulation of nonlinear bending models for thin elastic rods and plates. In Andrea Bonito and Ricardo H. Nochetto, editors, *Geometric Partial Differential Equations - Part I*, volume 21 of *Handbook of Numerical Analysis*, pages 221–273. Elsevier, 2020.
[10] Sören Bartels and Philipp Reiter. Numerical solution of a bending-torsion model for elastic rods. *Numerische Mathematik*, Oct 2020.
[11] Sören Bartels and Philipp Reiter. Stability of a simple scheme for the approximation of elastic knots and self-avoiding inextensible curves. *Mathematics of Computation*, 90:1499–1526, 2021.
[12] Sören Bartels, Philipp Reiter, and Johannes Riege. A simple scheme for the approximation of self-avoiding inextensible curves. *IMA Journal of Numerical Analysis*, 38(2):543–565, 05 2017.
[13] Jan Bender, Kenny Erleben, Jeff Trinkle, and Erwin Coumans. Interactive Simulation of Rigid Body Dynamics in Computer Graphics. In Marie-Paule Cani and Fabio Ganovelli, editors, *Eurographics 2012 - State of the Art Reports*. The Eurographics Association, 2012.
[14] Miklós Bergou, Max Wardetzky, Stephen Robinson, Basile Audoly, and Eitan Grinspun. Discrete elastic rods. In *ACM SIGGRAPH 2008 Papers*, SIGGRAPH '08, New York, NY, USA, 2008. Association for Computing Machinery.
[15] Daniele Boffi, Franco Brezzi, and Michel Fortin. *Mixed Finite Element Methods and Applications*, volume 44 of *Springer Series in Computational Mathematics*. Springer-Verlag, New York, NY, 2013.
[16] Robert Bridson, Ronald Fedkiw, and John Anderson. Robust treatment of collisions, contact and friction for cloth animation. *ACM Trans. Graph.*, 21(3):594–603, July 2002.
[17] Gregory Buck and Jeremey Orloff. A simple energy function for knots. *Topology and its Applications*, 61(3):205–214, 1995.
[18] Jason Cantarella, Joseph H G Fu, Robert B Kusner, and John M Sullivan. Ropelength criticality. *Geometry & Topology 18 (2014) 2595–2665*, 18(4):1973–2043, 2014.
[19] Raphaël Charrondière, Florence Bertails-Descoubes, Sébastien Neukirch, and Victor Romero. Numerical modeling of inextensible elastic ribbons with curvature-based elements. *Computer Methods in Applied Mechanics and Engineering*, 364:112922, 2020.
[20] Michel Chipot. *Elements of Nonlinear Analysis*. Birkhäuser, 2000.
[21] N. Clauvelin, B. Audoly, and S. Neukirch. Matched asymptotic expansions for twisted elastic knots: A self-contact problem with non-trivial contact topology. *Journal of the Mechanics and Physics of Solids*, 57(9):1623–1656, 2009.
[22] Murilo G. Coutinho. *Guide to Dynamic Simulations of Rigid Bodies and Particle Systems*. Simulation Foundations, Methods and Applications. Springer-Verlag, 1st edition.
[23] Zdeněk Dostál, Tomáš Kozubek, Marie Sadowská, and Vít Vondrák. *Scalable Algorithms for Contact Problems*. Advances in Mechanics and Mathematics. Springer, 1st edition.
[24] Sarah F. Frisken and Ronald N. Perry. Simple and efficient traversal methods for quadtrees and octrees. *Journal of Graphics Tools*, 7(3):1–11, 2002.
[25] O. Gonzalez, J. H. Maddocks, F. Schuricht, and H. von der Mosel. Global curvature and self-contact of nonlinearly elastic curves and rods. *Calculus of Variations*, 14:29–68, 2002.
[26] Oscar Gonzalez and John H. Maddocks. Global curvature, thickness and the ideal shapes of knots. In *Proceedings of the National Academy of Sciences, USA 96*, pages 4769–4773, 1999.
[27] Michael Hintermüller, Kazufumi Ito, and Karl Kunisch. The primal-dual active set strategy as a semismooth newton method. *SIAM Journal on Optimization*, 13(3):865–888, 2003.
[28] Michael Hintermüller and Karl Kunisch. Feasible and noninterior path-following in constrained minimization with low multiplier regularity. *SIAM Journal on Control and Optimization*, 45(4):1198–1221, 2006.
[29] Michael Hintermüller and Karl Kunisch. Path-following methods for a class of constrained minimization problems in function space. *SIAM Journal on Optimization*, 17(1):159–187, 2006.
[30] Kathleen A. Hoffman and Thomas I. Seidman. A variational characterization of a hyperelastic rod with hard self-contact. *Nonlinear Analysis: Theory, Methods & Applications*, 74(16):5388–5401, 2011.
[31] Kathleen A. Hoffman and Thomas I. Seidman. A variational rod model with a singular nonlocal potential. *Archive for Rational Mechanics and Analysis*, 200(1):255–284, Apr 2011.
[32] Qiya Hu, Xue-Cheng Tai, and Ragnar Winther. A saddle point approach to the computation of harmonic maps. *SIAM Journal on Numerical Analysis*, 47(2):1500–1523, 2009.
[33] Kazufumi Ito and Karl Kunisch. Semi-smooth newton methods for variational inequalities of the first kind. *ESAIM: Mathematical Modelling and Numerical Analysis*, 37(1):41–62, 2003.
[34] Kazufumi Ito and Karl Kunisch. *Lagrange Multiplier Approach to Variational Problems and Applications*. Advances in Design and Control. SIAM, 2008.
[35] Jürgen Jost and Xianqing Li-Jost. *Calculus of Variations*. Cambridge, 1998.
[36] James Kindt, Shelly Tzlil, Avinoam Ben-Shaul, and William M. Gelbart. Dna packaging and ejection forces in bacteriophage. *Proceedings of the National Academy of Sciences*, 98(24):13671–13674, 2001.
[37] G. Kloosterman, R. M. J. van Damme, A. H. van den Boogaard, and J. Huétink. A geometrical-based contact algorithm using a barrier method. *International Journal for Numerical Methods in Engineering*, 51(7):865–882, 2001.
[38] W.S. Klug and M. Ortiz. A director-field model of dna packaging in viral capsids. *Journal of the Mechanics and Physics of Solids*, 51(10):1815 – 1847, 2003.
[39] H. Kobayashi, B. Kresling, and J. F. V. Vincent. The geometry of unfolding tree leaves. *Proceedings: Biological Sciences*, 265(1391):147–154, 1998.

[40] Carlye Lauff, Timothy W. Simpson, Mary Frecker, Zoubeida Ounaies, Saad Ahmed, Paris von Lockette, Rebecca Strzelec, Robert Sheridan, and Jyh-Ming Lien. Differentiating bending from folding in origami engineering using active materials. In *ASME 2014 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, volume Volume 5B: 38th Mechanisms and Robotics Conference, 2014.

[41] B. Laurie, J. Smutny, M. Carlen, and J. H. Maddocks. Biarcs, global radius of curvature, and the computation of ideal knot shapes. In Jorge A. Calvo, Kenneth C. Millett, Eric J. Rawdon, and Andrzej Stasiak, editors, *Physical And Numerical Models In Knot Theory*, volume 36 of *Series on Knots and Everything*, chapter 5, pages 75–108. World Scientific Publishing, Sept 2005.

[42] Minchen Li, Zachary Ferguson, Teseo Schneider, Timothy Langlois, Denis Zorin, Daniele Panozzo, Chenfanfu Jiang, and Danny M. Kaufman. Incremental potential contact: Intersection-and inversion-free, large-deformation dynamics. *ACM Trans. Graph.*, 39(4), jul 2020.

[43] Pei Liu, Javier Arsuaga, M. Carme Calderer, Dmitry Golovaty, Mariel Vazquez, and Shawn Walker. Ion-dependent dna configuration in bacteriophage capsids. *Biophysical Journal*, 120(16):3292–3302, 2021.

[44] Libin Lu, Abtin Rahimian, and Denis Zorin. Contact-aware simulations of particulate stokesian suspensions. *Journal of Computational Physics*, 347:160–182, 2017.

[45] Davide Marenduzzo, Enzo Orlandini, Andrzej Stasiak, De Witt Sumners, Luca Tubiana, and Cristian Micheletti. Dna-dna interactions in bacteriophage capsids are responsible for the observed dna knotting. *Proceedings of the National Academy of Sciences*, 106(52):22269–22274, 2009.

[46] Kenneth C. Millett and Eric J. Rawdon. Energy, ropelength, and other physical aspects of equilateral knots. *J. Comput. Phys.*, 186(2):426–456, April 2003.

[47] Maria Giovanna Mora and Stefan Müller. Derivation of the nonlinear bending-torsion theory for inextensible rods by γ-convergence. *Calculus of Variations and Partial Differential Equations*, 18(3):287–305, Nov 2003.

[48] Olivier Pantz. The modeling of deformable bodies with frictionless (self-)contacts. *Archive for Rational Mechanics and Analysis*, 188(2):183–212, 2008.

[49] Olivier Pantz. A frictionless contact algorithm for deformable bodies. *ESAIM: Mathematical Modelling and Numerical Analysis*, 45(2):235–254, 2011.

[50] Piñeirua, M., Adda-Bedia, M., and Moulinet, S. Spooling and disordered packing of elastic rods in cylindrical cavities. *EPL*, 104(1):14005, 2013.

[51] Kazuya Saito, Shuhei Nomura, Shuhei Yamamoto, Ryuma Niiyama, and Yoji Okabe. Investigation of hindwing folding in ladybird beetles by artificial elytron transplantation and microcomputed tomography. *Proceedings of the National Academy of Sciences*, 114(22):5624–5628, 2017.

[52] Friedemann Schuricht and Heiko von der Mosel. Euler-lagrange equations for nonlinearly elastic rods with self-contact. *Archive for Rational Mechanics and Analysis*, 168(1):35–82, 2003.

[53] Friedemann Schuricht and Heiko von der Mosel. Characterization of ideal knots. *Calculus of Variations and Partial Differential Equations*, 19(3):281–305, Apr 2004.

[54] Andrew Selle, Jonathan Su, Geoffrey Irving, and Ronald Fedkiw. Robust high-resolution cloth using parallelism, history-based collisions, and accurate friction. *IEEE Transactions on Visualization and Computer Graphics*, 15(2):339–350, Mar 2009.

[55] M. Reza Shaebani, Javad Najafi, Ali Farnudi, Daniel Bonn, and Mehdi Habibi. Compaction of quasi-one-dimensional elastoplastic materials. *Nature Communications*, 8:15568–, 2017.

[56] E. L. Starostin. Symmetric equilibria of a thin elastic rod with self-contacts. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 362(1820):1317–1334, 2004.

[57] Paweł Strzelecki and Heiko von der Mosel. Global curvature for surfaces and area minimization under a thickness constraint. *Calculus of Variations and Partial Differential Equations*, 25(4):431–467, 2006.

[58] Paweł Strzelecki and Heiko von der Mosel. On rectifiable curves with l^p-bounds on global curvature: self-avoidance, regularity, and minimizing knots. *Mathematische Zeitschrift*, 257(1):107–130, 2007.

[59] Paweł Strzelecki and Heiko von der Mosel. Tangent-point repulsive potentials for a class of non-smooth m-dimensional sets in $\mathbb{R}^n$. part I: Smoothing and self-avoidance effects. *Journal of Geometric Analysis*, 23(3):1085–1139, 2013.

[60] J. M. T. Thompson, Bernard D. Coleman, and David Swigon. Theory of self-contact in kirchhoff rods with applications to supercoiling of knotted and unknotted dna plasmids. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 362(1820):1281–1299, 2004.

[61] J. M. T. Thompson, M. Silveira, G. H. M. van der Heijden, and M. Wiercigroch. Helical post-buckling of a rod in a cylinder: with applications to drill-strings. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 468(2142):1591–1614, 2012.

[62] Fredi Tröltzsch. *Optimal Control of Partial Differential Equations*. Graduate Studies in Mathematics. American Mathematical Society, April 2010.

[63] G. H. M. van der Heijden, S. Neukirch, V. G. A. Goss, and J. M. T. Thompson. Instability and self-contact phenomena in the writhing of clamped rods. *International Journal of Mechanical Sciences*, 45(1):161–196, 2003.

[64] Shawn Walker, Javier Arsuaga, Lindsey Hiltner, M. Carme Calderer, and Mariel Vázquez. Fine structure of viral dsdna encapsidation. *Phys. Rev. E*, 101:022703, Feb 2020.

[65] Shawn W. Walker. FELICITY: A Matlab/C++ toolbox for developing finite element methods and simulation modeling. *SIAM Journal on Scientific Computing*, 40(2):C234–C257, 2018.

[66] Peter Wriggers. *Computational Contact Mechanics*. Springer, 2nd edition.

Statements and Declarations

Funding: SWW has been supported in part by NSF grant DMS-2111474.

Financial interests: SWW declares they have no financial or competing interests.

Data availability: The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

(Shawn W. Walker) Department of Mathematics and Center for Computation and Technology (CCT) Louisiana State University, Baton Rouge, LA 70803

Email address: walker@math.lsu.edu