SIAM J. CONTROL OPTIM.
Vol. 51, No. 4, pp. 3093–3126

ANALYSIS OF SHAPE OPTIMIZATION FOR MAGNETIC MICROSWIMMERS*

SHAWN W. WALKER† AND ERIC E. KEAVENY‡

Abstract. We analyze an infinite dimensional, geometrically constrained shape optimization problem for magnetically driven microswimmers (locomotors) in three-dimensional (3-D) Stokes flow and give a well-posed descent scheme for computing optimal shapes. The problem is inspired by recent experimental work in this area. We show the existence of a minimizer of the optimization problem using analytical tools for elastic rods that respect the excluded volume constraint. We derive a variational gradient descent method for computing optimal locomotor shapes using the tools of shape differential calculus. The descent direction is obtained by solving a saddle-point system, which we prove is well-posed. We also introduce a finite element approximation of the gradient descent method and prove its stability. We present numerical results illustrating our method and the effect that finite aspect ratio and external cargo can have on the optimal shape. The 3-D Stokes equations are solved by a boundary integral method.

Key words. shape optimization, microswimmer, 3-D Stokes flow, existence of minimizer, gradient descent method

AMS subject classifications. Primary, 49M25; Secondary, 76D55

DOI. 10.1137/110845823

1. Introduction.

Microorganisms swimming at low Reynolds number, where fluid inertia is negligible, must use time-irreversible motions of their control surfaces (e.g., their bodies or flagellum) in order to have a net translation after one stroke [42, 54]. This is achieved by sperm and small nematodes by propagating bending waves along the length of a flexible slender body, and bacteria that utilize rotating helices whose shape couples the rotation and translation. The flow fields generated by the motion of the control surfaces couple the motion of nearby bodies leading to swarms that induce vigorous fluid mixing [17, 56, 57, 60], which can have a large effect on the distribution of chemicals and nutrients within the fluid.

There has been a recent effort to develop microswimmers that mimic the low Reynolds number swimming strategies employed by microorganisms. There are numerous applications for these artificial systems, such as targeted drug delivery in the human body, microsurgery, automated transport of cargo/payloads in microfluidic chips, and filtering of toxic substances from polluted water streams. In addition, these systems can help illuminate the basic mechanisms of micron-scale swimming. Artificial swimmers based on flexible magnetic filaments [19, 37] that can be actuated using applied magnetic fields or bimetallic synthetic nanorods [9, 50, 51, 64] that propel themselves via a catalytic reaction have been the first of these devices. More recently, researchers have developed swimmers based on a rigid corkscrew design that contain magnetic material and can be rotated and directed in three dimensions us-

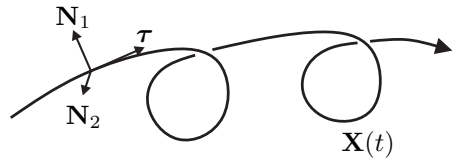
*Received by the editors August 25, 2011; accepted for publication (in revised form) May 23, 2013; published electronically August 1, 2013.
http://www.siam.org/journals/sicon/51-4/84582.html
†Department of Mathematics, Louisiana State University, Baton Rouge, LA 70803 (walker@math.lsu.edu). This author acknowledges funding support given by NSF-RTG grant DMS-0602235, NSF-FRG grant DMS-0652775, NSF grant DMS-1115636, and DOE grant DE-FG02-88ER25053.
‡Department of Mathematics, Imperial College, SW7 2AZ London, UK (e.keaveny@imperial.ac.uk). This author acknowledges funding support given by NSF-FRG grant DMS-0652775, NSF-Collaborative Research grant MSPA-ENG-0652775, and the NSF-MRSEC program DMR-0820341.

ing applied magnetic fields. In [23], the authors constructed submicron scale helical structures using a glass (SiO_2) deposition process, whereas in [70] they built their swimmer from a prestressed strip of a layered material attached to a magnetic body. In both cases, these swimmers have been shown to be highly controllable and able to execute complex swimming paths in all three spatial dimensions. To further the design of the swimmers constructed in these studies, we pose an infinite dimensional shape optimization problem to determine the swimmer shapes that deliver the highest speed for a given applied torque about the swimming direction, have the highest speed for a given power output, or travel the greatest distance per rotation.

Optimization techniques have been used extensively to examine the hydrodynamic efficiency and speed of locomotion gaits employed by microorganisms [2, 45, 53, 62, 65] leading to the establishment of fundamental results regarding the optimal shapes of the thin, high aspect ratio flagellar filaments utilized by sperm and bacteria. Using drag-based models, for example, Lighthill [44] showed that in two dimensions an infinitely long, infinitely thin body propagating a saw-tooth wave with a slope of 42° maximizes hydrodynamic efficiency. Establishing similar results for the swimmer shapes with finite length and attached payloads corresponding to those in [23, 70] requires a more precise treatment of the hydrodynamic forces as well as a general characterization of how variations in their geometry change cost functionals such as speed or efficiency. We incorporate both of these necessary requirements into our optimization approach by considering the exact Stokes flow on these bodies, posing the optimization problem at the continuous level, and evaluating the variations in the cost functionals using shape differential calculus. We analyze the infinite dimensional optimization problem and establish the existence of a minimizer for functionals associated with maximizing swimmer speed, stroke efficiency, and hydrodynamic efficiency. In addition, we present a variational descent method for this problem and show that it and a discretization based on cubic splines are well-posed. To compute the cost functionals and their variations, we employ a boundary integral formulation of Stokes flow and discretize it to second-order using a collocation method [38]. The usage of boundary integral formulations has been successfully employed in the optimization of axisymmetric swimming bodies [3]. Our computational results show this approach is equally successful for the optimization of more complex swimmer shapes directly related to those fabricated in [23, 70].

The paper is outlined as follows. Section 2 describes the physical model and shape optimization problem. Section 3 states the equivalent weak formulation of the model. In section 4, we prove the existence of a minimizer of the infinite dimensional locomotor shape optimization problem. Section 5 presents a variational descent method for obtaining optimal locomotor centerline shapes using the tools of shape differential calculus. This leads to a saddle-point system to solve for a descent direction at each step of the optimization, which we show is well-posed. Section 6 describes the details of a finite element approximation of the descent method, and we prove a stability result when solving for a discrete shape descent direction. In section 7, we show some numerical results to illustrate our method and give some discussion on the effects of aspect ratio and external cargo size. Last, we include appendices as a convenient reference for the reader which contain some basic analysis results, derivations of adjoint equations, and some facts from differential geometry of curves.

2. Problem description. We first introduce our representation of the swimmer geometry, which captures the shapes realized in experiments and includes a meaningful control over which to optimize. We state the functionals we intend on maximizing,



(a) Curve parametrization; $\boldsymbol{\tau}$ is the tangent vector of $\mathbf{X}$. The triplet $\{\boldsymbol{\tau}, \mathbf{N}_1, \mathbf{N}_2\}$ is an orthogonal frame.

(b) Example 3-D locomotor shape.

FIG. 2.1. *Part* (a) *describes the curve parametrization that is used in* (2.1). *Part* (b) *shows an example of a swimmer shape with circular cross section that varies in radius along the centerline.*

such as speed and efficiency, and form a PDE model for the swimmer motion (in a Stokesian fluid) induced by the application of time-dependent magnetic fields. In addition, we describe the constraints implemented to maintain certain characteristics (length, etc.) of the swimmer geometry throughout the optimization.

2.1. Surface parametrization. The fundamental object to be optimized is a one-dimensional (1-D) curve in three dimensions denoted $\Sigma \subset \mathbb{R}^3$, which we parameterize by the vector function $\mathbf{X} : [-1, 1] \to \mathbb{R}^3$. The shape of Σ captures the basic form of the locomotor. The three-dimensional (3-D) solid form of the locomotor is defined through a surface parametrization attached to $\mathbf{X}$. Essentially, we restrict the shapes to long tube-like shapes very much in the spirit of the experimental results in [23, 70]; see Figure 2.1. This still allows for a lot of flexibility in the shape and is still practical for manufacturing purposes.

Let Γ be a closed manifold whose interior Ω_{B} represents the rigid body of the locomotor, i.e., $\Gamma \equiv \partial\Omega_{\mathrm{B}}$. We define Γ uniquely in terms of $\mathbf{X}$ by the parametrization $\boldsymbol{\Psi} : [-1, 1] \times [0, 2\pi] \to \mathbb{R}^3$,

$$\boldsymbol{\Psi}(t, \theta) = \mathbf{X}(t) + a_{\mathrm{c}}(t, \theta) \left[\cos\theta \mathbf{N}_1(t) + \sin\theta \mathbf{N}_2(t)\right], \quad -1 \le t \le 1, \quad 0 \le \theta \le 2\pi, \tag{2.1}$$

where $\mathbf{X}(t)$ parameterizes the 1-D curve Σ and θ is the azimuthal angle within a planar cross-section of Ω_{B}; a cross-section at $\mathbf{X}(t)$ is assumed to be orthogonal to $\boldsymbol{\tau}(t)$ (tangent vector). The radius of a cross-section located at $\mathbf{X}(t)$ is parameterized in terms of θ by the given smooth function $a_{\mathrm{c}}(t, \theta)$ (i.e., $a_{\mathrm{c}}(t, \theta)$ models a variable radius cross-section). The vectors $\{\mathbf{N}_1, \mathbf{N}_2\}$ form a basis of the cotangent space of $\mathbf{X}$ and are generated via parallel transport [5, 29].

The classic Frenet frame [16] is not satisfactory for our purposes. It can induce undesirable twisting of the surface because the normal and binormal are linked to the curvature and torsion. This is especially noticeable if the curve $\mathbf{X}$ has a small amplitude undulation. The Frenet frame also requires $\mathbf{X}$ to be in $C^3([-1, 1])$. For the parallel transport frame, $\mathbf{X}$ is only required to be $C^2([-1, 1])$, or even $C^{1,1}([-1, 1])$. Moreover, the effect of the torsion has been removed [5, 29], which reduces the twisting of the surface. This is especially useful when discretizing the surface in our boundary integral method [38, 40]. Removing the torsion prevents the grid from being distorted.

In the case of a circular cross-section, a_c is independent of θ. For our cases (section 7), we use

$$a_c(t) = A\sqrt{1-t^2}, \quad -1 \le t \le 1, \tag{2.2}$$

where $A > 0$ is the maximum cross-section radius that occurs at the midpoint (i.e., $t = 0$). Note that Γ depends on the parametrization $\mathbf{X}$ through (2.1), i.e., the shape of Γ will change if $\mathbf{X}$ is parameterized nonuniformly. So we require that $\mathbf{X}$ be an equal arc-length parametrization.

Remark 1. Enforcing an equal arc-length parametrization is easily accomplished by a Lagrange multiplier (see section 5.2.1). Moreover, making a_c depend on the arc-length function $s(t)$ would further complicate the shape perturbation formula (B.18) by introducing an additional term in the sensitivity analysis.

2.2. Magnetically driven objects in 3-D Stokes flow: PDE model. In the experiments [23, 70], a rotating magnetic field is used to both actuate the swimmers and control their swimming direction. Due to the presence of magnetic material in the swimmer, the field exerts a torque on the swimmer that rotates it about its helical axis while keeping this rotation aligned with a particular direction. We may capture these effects without considering the magnetic field explicitly by requiring that the total force on the swimmer be zero, the body rotate about the z-axis, and the z-component of the torque have unit value. These conditions on the rigid body lead to the Stokes flow problem [30, 40] in (2.3) that models the rigid swimmers in [23, 70].

Consider a *rigid* body (swimmer) $\Omega_B \subset \Omega_{ALL} \subset \mathbb{R}^3$ with surface denoted $\Gamma := \partial\Omega_B$, where Ω_{ALL} is a container (i.e., ball with large radius) with outer boundary denoted $\Gamma_O := \partial\Omega_{ALL}$. Assume Γ_O is far away from Ω_B. We call $\Omega := \Omega_{ALL} \setminus \Omega_B$ the fluid domain in the container but outside the rigid body (i.e., $\partial\Omega = \Gamma \cup \Gamma_O$). The governing equations for the flow field $(\mathbf{u}, p)$ induced by the rigid body are given by the Stokes equations (in strong form):

$$\begin{aligned}
-\nabla \cdot \boldsymbol{\sigma} &= \mathbf{0} \text{ in } \Omega, \\
\nabla \cdot \mathbf{u} &= 0 \text{ in } \Omega, \\
\mathbf{u} &= \mathbf{u}_B + \boldsymbol{\omega}_B \times (\mathbf{x} - \mathbf{x}_c) \text{ on } \Gamma \quad \text{(rigid motion)}, \\
\mathbf{u} &= \mathbf{0} \text{ on } \Gamma_O \\
-\int_{\Gamma_O} \boldsymbol{\sigma}\boldsymbol{\nu} = \int_{\Gamma} \boldsymbol{\sigma}\boldsymbol{\nu} &= \mathbf{f}_B := \mathbf{0} \quad \text{(net force)} \\
-\int_{\Gamma_O} \mathbf{x} \times (\boldsymbol{\sigma}\boldsymbol{\nu}) = \int_{\Gamma} \mathbf{x} \times (\boldsymbol{\sigma}\boldsymbol{\nu}) &= \boldsymbol{\tau}_B \quad \text{(net torque)},
\end{aligned} \tag{2.3}$$

where $\times$ denotes the cross-product of vectors in $\mathbb{R}^3$, $\boldsymbol{\nu}$ is the normal vector on Γ and points away from Ω, p is the pressure, $\mathbf{u}$ is the velocity, and $\mathbf{x}_c$ is the center of mass of Ω_B (see (3.7)). The constant vectors $\mathbf{u}_B$, $\boldsymbol{\omega}_B$, $\mathbf{f}_B$, $\boldsymbol{\tau}_B \in \mathbb{R}^3$ have the following form:

$$\begin{aligned}
\mathbf{u}_B &= \begin{pmatrix} u_{B,x} \text{ (unknown)} \\ u_{B,y} \text{ (unknown)} \\ u_{B,z} \text{ (unknown)} \end{pmatrix}, \ \boldsymbol{\omega}_B = \begin{pmatrix} 0 \\ 0 \\ \omega_{B,z} \text{ (unknown)} \end{pmatrix}, \\
\mathbf{f}_B &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \ \boldsymbol{\tau}_B = \begin{pmatrix} \tau_{B,x} \text{ (unknown)} \\ \tau_{B,y} \text{ (unknown)} \\ \tau_{B,z} \text{ (given)} \end{pmatrix},
\end{aligned} \tag{2.4}$$

where "unknown" indicates unknown quantities that are part of the PDE solution. Thus, the translation $\mathbf{u}_\mathrm{B}$ is an unknown to solve for, $\boldsymbol{\omega}_\mathrm{B}$ is constrained to only rotate about $\mathbf{e}_z$, the net force is zero, and $\boldsymbol{\tau}_\mathrm{B}$ has a prescribed z-component but x and y components are to be solved for. The Newtonian stress tensor $\boldsymbol{\sigma}$ is given by

$$\boldsymbol{\sigma} = -p\mathbf{I} + D(\mathbf{u}), \text{ where } D(\mathbf{u}) := \nabla\mathbf{u} + (\nabla\mathbf{u})^T, \tag{2.5}$$

where superscript ^{T} denotes the matrix transpose. To ensure uniqueness, p is taken to have mean value zero. A boundary integral version of this model is described in [38].

2.3. Cost functionals. With the body shape described by (2.1), we seek to determine the shape of the centerline of the swimmer's tail that minimizes a notion of cost. One relevant cost functional is $\mathbf{u}_\mathrm{B} \cdot \mathbf{e}_z$, i.e., maximize $\mathbf{u}_\mathrm{B} \cdot \mathbf{e}_z$ for a given $\boldsymbol{\tau}_\mathrm{B} \cdot \mathbf{e}_z$. This corresponds to finding the swimmer shape that maximizes the speed of the swimmer in the direction of the fixed component of the applied torque. It can also be thought of as determining the swimmer shape that provides the highest value of the entry of the low Reynolds number mobility matrix [30, 40] that couples the torque and translational velocity in the direction of the swimmer's axis. Another functional to consider is a measure of the hydrodynamic efficiency [13, 44], which is the viscous dissipation required to pull the swimmer at its swimming speed relative to the viscous dissipation associated with locomotion (i.e., the ratio of the "dead" power to the swimming power). Maximizing efficiency corresponds to determining the swimmer shape that gives the highest speed for a fixed input power.

We state these functionals more concretely in the following definitions.

DEFINITION 2.1. *Let $\mathbf{M} \in \mathbb{R}^{3\times 3}$ be a positive semidefinite matrix. Define the coupling of torque to speed as*

$$J_\mathrm{ts}(\mathbf{X}) \equiv J_\mathrm{ts}(\mathbf{u}_\mathrm{B}(\mathbf{X}), \boldsymbol{\tau}_\mathrm{B}(\mathbf{X})) = \mathbf{u}_\mathrm{B} \cdot \mathbf{M}\boldsymbol{\tau}_\mathrm{B} = \frac{\int_{\Omega_\mathrm{B}} \mathbf{u} \cdot \mathbf{M}\boldsymbol{\tau}_\mathrm{B}}{|\Omega_\mathrm{B}|} \quad \text{for all } (\mathbf{u}_\mathrm{B}, \boldsymbol{\tau}_\mathrm{B}) \text{ satisfying } (2.3). \tag{2.6}$$

Recall that $\mathbf{u}_\mathrm{B}$ and $\boldsymbol{\tau}_\mathrm{B}$ are constant over Ω_B. When $\mathbf{M}$ has all zero entries, except the last diagonal entry is 1, then J_ts reduces to $J_\mathrm{ts} = (\mathbf{u}_\mathrm{B} \cdot \mathbf{e}_z)(\boldsymbol{\tau}_\mathrm{B} \cdot \mathbf{e}_z)$.

DEFINITION 2.2. *The total rate of viscous dissipation for a swimmer satisfying* (2.3) *is given by*

$$\begin{aligned} J_\mathrm{diss}(\mathbf{X}) = J_\mathrm{diss}(\mathbf{X}, \mathbf{u}(\mathbf{X})) &:= \int_\Gamma \mathbf{u} \cdot \boldsymbol{\sigma}\boldsymbol{\nu} \\ &= \frac{1}{2}\int_\Omega D(\mathbf{u}) : D(\mathbf{u}) = (\boldsymbol{\omega}_\mathrm{B}(\mathbf{X}) \cdot \mathbf{e}_z)(\boldsymbol{\tau}_\mathrm{B}(\mathbf{X}) \cdot \mathbf{e}_z), \end{aligned} \tag{2.7}$$

where we used (3.6) *and $\mathbf{u}$, $\boldsymbol{\omega}_\mathrm{B}$ satisfy* (2.3). *Note that ":" is the inner product of two tensors.*

DEFINITION 2.3. *We consider the notion of* stroke efficiency *for a swimmer, which is a measure of how far one can go during one period of motion. Define*

$$J_\mathrm{eff}(\mathbf{X}) = \frac{J_\mathrm{ts}(\mathbf{X})}{J_\mathrm{diss}(\mathbf{X})} = \frac{(\mathbf{u}_\mathrm{B}(\mathbf{X}) \cdot \mathbf{e}_z)(\boldsymbol{\tau}_\mathrm{B}(\mathbf{X}) \cdot \mathbf{e}_z)}{(\boldsymbol{\omega}_\mathrm{B}(\mathbf{X}) \cdot \mathbf{e}_z)(\boldsymbol{\tau}_\mathrm{B}(\mathbf{X}) \cdot \mathbf{e}_z)} = \frac{u_\mathrm{B,z}}{\omega_\mathrm{B,z}}, \tag{2.8}$$

which is the ratio of net translating swimming velocity to angular velocity. (Note that $u_\mathrm{B,z}$ is not the tangential velocity of the swimmer due to rotation.) Therefore, a more efficient swimmer can move farther with fewer rotations.

Remark 2. Another notion of efficiency is the hydrodynamic efficiency. Let $\hat{\mathbf{u}}$ solve the Stokes problem (2.3), except with the following Dirichlet condition for velocity $\hat{\mathbf{u}} = \mathbf{e}_z$ on Γ and $\hat{\mathbf{u}} = \mathbf{0}$ on Γ_{O}. Then the viscous dissipation for this case is

$$\widehat{J_{\mathrm{diss}}}(\mathbf{X}) = \widehat{J_{\mathrm{diss}}}(\mathbf{X}, \hat{\mathbf{u}}(\mathbf{X})) := \int_{\Gamma} \hat{\mathbf{u}} \cdot \hat{\boldsymbol{\sigma}} \boldsymbol{\nu} = \frac{1}{2} \int_{\Omega} D(\hat{\mathbf{u}}) : D(\hat{\mathbf{u}}).$$

Then one can consider

$$\widehat{J_{\mathrm{eff}}}(\mathbf{X}) = \frac{u_{\mathrm{B,z}}^2 \widehat{J_{\mathrm{diss}}}(\mathbf{X})}{J_{\mathrm{diss}}(\mathbf{X})}, \tag{2.9}$$

which is essentially the ratio of the "dead" power to the swimming power [44, 13]. The factor $u_{\mathrm{B,z}}^2$ is for appropriate scaling. It is also common to approximate (2.9) by [44]

$$\widehat{J_{\mathrm{eff}}}(\mathbf{X}) \approx \frac{(J_{\mathrm{ts}}(\mathbf{X}))^2}{J_{\mathrm{diss}}(\mathbf{X})}.$$

Remark 3. For analyzing the infinite dimensional optimization problem (section 4), we will consider the functionals (2.6) and (2.8). For the optimization algorithm and results, we just consider (2.6). (Results on (2.8) will be reported in future work.)

To fit with the language of minimizing, we let $J_{\mathrm{mag}} = -J_{\mathrm{ts}}$ or $J_{\mathrm{mag}} = -J_{\mathrm{eff}}$. Hence, the statement of the optimization problem is to find the shape of the centerline $\mathbf{X}$ such that J_{mag} is *minimized* subject to appropriate constraints (see section 2.4).

2.4. Constraints. Some constraints must be placed on $\mathbf{X}$ to ensure a meaningful optimization problem. All of the constraints are geometric and are listed as follows:

- Non-self-intersecting: $\mathbf{X}$ must have bounded global radius of curvature (sections 4.2.2 and 4.2.4). Basically, we demand that the solid volume defined by (2.1) does not intersect itself. This is the so-called excluded volume constraint.
- Local inextensibility: $\|\mathbf{X}'(t)\| = L/2$, where L is the total length of the centerline $|\mathbf{X}([-1, 1])|$ (section 5.2.1).

Note that a constant volume constraint is effectively imposed because we enforce inextensibility of $\mathbf{X}$, assume a known surface parametrization (2.1) (recall a_{c}), and require the locomotor shape to avoid self-intersections. Of course, we could consider additional constraints (see section 5.2.2), but the above constraints are the most important for the mathematical analysis in section 4.

3. Weak formulation. We recall some standard notation, give the weak formulation of (2.3), and note some basic estimates.

3.1. Notation. We adopt the following Sobolev space notation:

$$\begin{aligned} L^p(D) &= \left\{ f : D \to \mathbb{R} \ : \ \int_D |f|^p < \infty \right\}, \\ H^k(D) &= \left\{ f : D \to \mathbb{R} \ : \ \int_D |\partial_\alpha f|^2 < \infty, |\alpha| \le k \right\}, \end{aligned} \tag{3.1}$$

where α is a multiindex and D is an open set in $\mathbb{R}^n$ for $n = 1, 2$, or 3. If E is a function space, then E^* denotes the dual space of E (i.e., the set of functionals defined on E).

3.2. Function spaces. To facilitate proving that there is a minimizer, we rewrite (2.3) into a weak formulation. The velocity and pressure spaces are defined as

$$\mathbb{V} \equiv \mathbb{V}(\mathbf{X}) \tag{3.2}$$
$$= \{\mathbf{v} \in H^1(\Omega) : \mathbf{v} = \mathbf{v}_{\mathrm{B}} + \boldsymbol{\vartheta}_{\mathrm{B}} \times (\mathbf{x} - \mathbf{x}_c) \text{ on } \Gamma, \mathbf{v}_{\mathrm{B}}, \boldsymbol{\vartheta}_{\mathrm{B}} \in \mathbb{R}^3, \mathbf{v} = \mathbf{0} \text{ on } \Gamma_{\mathrm{O}}\}$$
$$\mathbb{V}_0 \equiv \mathbb{V}_0(\mathbf{X}) \tag{3.3}$$
$$= \{\mathbf{v} \in \mathbb{V} : \mathbf{v} = \mathbf{v}_{\mathrm{B}} + \boldsymbol{\vartheta}_{\mathrm{B}} \times (\mathbf{x} - \mathbf{x}_c) \text{ on } \Gamma, \mathbf{v}_{\mathrm{B}} \in \mathbb{R}^3, \boldsymbol{\vartheta}_{\mathrm{B}} = (0, 0, a), a \in \mathbb{R}\},$$
$$\mathbb{Q} = \left\{ q \in L^2(\Omega) : \int_\Omega q = 0 \right\}, \tag{3.4}$$

all of which clearly depend on $\mathbf{X}$ (the parametrization of Σ).

3.3. Weak form. The weak formulation is derived by multiplying the first equation in (2.3) by a test function $\mathbf{v}$ in $\mathbb{V}_0$ and integrating by parts, i.e.,

$$\begin{aligned} \int_\Omega \boldsymbol{\sigma} : \nabla \mathbf{v} &= \int_\Gamma \boldsymbol{\sigma}\boldsymbol{\nu} \cdot \mathbf{v} = \mathbf{v}_{\mathrm{B}} \cdot \int_\Gamma \boldsymbol{\sigma}\boldsymbol{\nu} + \boldsymbol{\vartheta}_{\mathrm{B}} \times \int_\Gamma (\mathbf{x} - \mathbf{x}_c) \cdot \boldsymbol{\sigma}\boldsymbol{\nu} \\ &= \boldsymbol{\vartheta}_{\mathrm{B}} \cdot \int_\Gamma (\mathbf{x} - \mathbf{x}_c) \times \boldsymbol{\sigma}\boldsymbol{\nu} = \boldsymbol{\vartheta}_{\mathrm{B}} \cdot \int_\Gamma \mathbf{x} \times \boldsymbol{\sigma}\boldsymbol{\nu} - \boldsymbol{\vartheta}_{\mathrm{B}} \cdot \left(\mathbf{x}_c \times \int_\Gamma \boldsymbol{\sigma}\boldsymbol{\nu} \right) \\ &= \boldsymbol{\vartheta}_{\mathrm{B}} \cdot \boldsymbol{\tau}_{\mathrm{B}}, \end{aligned} \tag{3.5}$$

where we used the net force and torque conditions. Hence, we obtain the following.

VARIATIONAL FORMULATION 3.1. *Let Ω be a Lipschitz domain. Then there exists a unique solution* [21, 59, 66] $\mathbf{u}$ *in* $\mathbb{V}_0$ *and* p *in* $\mathbb{Q}$ *such that*

$$\begin{aligned} \int_\Omega \boldsymbol{\sigma} : \nabla \mathbf{v} \equiv \int_\Omega D(\mathbf{u}) : \nabla \mathbf{v} - \int_\Omega p \nabla \cdot \mathbf{v} &= (\boldsymbol{\vartheta}_{\mathrm{B}} \cdot \mathbf{e}_z) \tau_{\mathrm{B,z}}, \\ \int_\Omega q \nabla \cdot \mathbf{u} &= 0 \end{aligned} \tag{3.6}$$

for all $\mathbf{v}$ *in* $\mathbb{V}_0$ *and* q *in* $\mathbb{Q}$*, where* $\tau_{\mathrm{B,z}}$ *is a given number. Moreover, if* $\partial\Omega$ *is* C^2*, then* $\mathbf{u}$ *is* $H^2(\Omega) \cap \mathbb{V}_0$ *and* p *is* $H^1(\Omega) \cap \mathbb{Q}$ *(see* [59, 66]*).*

3.4. Basic estimates.

3.4.1. Domain geometry. We note some facts related to the geometry of Ω_{B}, which are useful in the analysis in section 4. Let $\mathbf{x}_c$ and $\mathbf{x}_g$ denote the center of mass and geometric center, respectively, i.e.,

$$\mathbf{x}_c = \frac{\int_{\Omega_{\mathrm{B}}} \mathbf{x}}{|\Omega_{\mathrm{B}}|}, \qquad \mathbf{x}_g = \frac{\int_\Gamma \mathbf{x}}{|\Gamma|}, \quad \text{where} \quad |\Omega_{\mathrm{B}}| = \int_{\Omega_{\mathrm{B}}} 1, \qquad |\Gamma| = \int_\Gamma 1, \tag{3.7}$$

where we have dropped the integral measure $d\mathbf{x}$, $dS(\mathbf{x})$ for convenience. Furthermore, we will sometimes denote the dependence of Ω_{B} on $\mathbf{X}$ by writing $\Omega_{\mathrm{B}}(\mathbf{X})$.

Next, we assume that Ω_{ALL} is an open ball with large radius r_0 centered at the origin that strictly contains the locomotor $\Omega_{\mathrm{B}}(\mathbf{X})$, i.e., $\Omega_{\mathrm{B}} \subset\subset \Omega_{\mathrm{ALL}}$. The fluid domain is denoted Ω and defined by $\Omega := \Omega_{\mathrm{ALL}} \backslash \overline{\Omega_{\mathrm{B}}}$. Moreover, we assume throughout that Ω_{B}, Ω_{ALL} satisfy

$$\frac{1}{2}|\Omega_{\mathrm{ALL}}| \leq |\Omega| \leq |\Omega_{\mathrm{ALL}}|, \tag{3.8}$$

which is trivial to guarantee by taking r_0 sufficiently large. Furthermore, by the parametrization (2.1), there exists a constant $C_{a_c} > 1$ depending only on a_c such that

$$\frac{1}{C_{a_c}}|\Gamma| \leq |\Omega_B| \leq C_{a_c}|\Gamma|, \qquad \frac{1}{C_{a_c}}|\Sigma| \leq |\Omega_B| \leq C_{a_c}|\Sigma|, \tag{3.9}$$

provided that $\Gamma = \partial\Omega_B$ does not intersect itself (section 4.2 and Lemma 4.3). We further note that Γ is at least uniformly Lipschitz continuous (i.e., $C^{0,1}([-1,1])$), provided $\mathbf{X}$ is at least $C^{1,1}([-1,1])$.

3.4.2. A priori estimates. In order to show the existence of a minimizer of the optimization problem (section 4), we need the following uniform a priori estimates.

LEMMA 3.1 (Korn's inequality). *Let $\Omega = \Omega_{ALL} \setminus \Omega_B$ be a Lipschitz continuous bounded domain. Then there is a constant $C > 0$ such that*

$$C\|\mathbf{v}\|^2_{H^1(\Omega)} \leq \frac{1}{2}\int_\Omega D(\mathbf{v}) : D(\mathbf{v}) \quad \textit{for all } \mathbf{v} \in \mathbb{V}. \tag{3.10}$$

Proof. Recall that $\mathbf{v} = \mathbf{0}$ on Γ_O. See [20] for the rest. ∎

LEMMA 3.2. *Let $(\mathbf{u}, p)$ be a solution of (3.6). Then there are constants C_1, C_2, and C_3 that only depend on $|\Sigma|$, $|\Omega_{ALL}|$, and a_c such that*

$$\frac{1}{C_1}|\tau_{B,z}| \leq \|\mathbf{u}\|_{H^1(\Omega)} \leq C_1|\tau_{B,z}|, \qquad \|p\|_{L^2(\Omega)} \leq C_2|\tau_{B,z}|, \tag{3.11}$$

and moreover

$$|\mathbf{u}_B| + |\omega_{B,z}| + |\tau_{B,x}| + |\tau_{B,y}| \leq C_3|\tau_{B,z}|. \tag{3.12}$$

Proof. See Appendix A. ∎

4. Existence of a minimizer. We show that a minimizer of the infinite dimensional (PDE-constrained) shape optimization problem does exist (with suitable constraints). The most critical part is in defining the admissible set of shapes, which takes advantage of some tools coming from the study of self-contact of curves [26]. Some related examples of optimization in fluids can be found in [28, 35, 53].

4.1. Uniform boundedness. For the existence proof, it is important that J_{ts} and J_{eff} are uniformly bounded in some sense, which is the purpose of the following proposition.

PROPOSITION 4.1. *The functionals defined by (2.6) and (2.8) satisfy*

$$-C_1|\tau_{B,z}|^2 \leq J_{ts} \leq C_1|\tau_{B,z}|^2, \quad -C_2 \leq J_{eff} \leq C_2 \tag{4.1}$$

for suitable constants C_1 and C_2, depending only on $|\Sigma|$, Ω_{ALL}, and a_c (cross-section radius of locomotor).

Proof. Clearly, by (3.12), we have

$$|J_{ts}| = |\mathbf{u}_B \cdot \mathbf{M}\boldsymbol{\tau}_B| \leq \|\mathbf{M}\|\,|\mathbf{u}_B|\,|\boldsymbol{\tau}_B| \leq c_1\|\mathbf{M}\||\tau_{B,z}|^2 = C_1|\tau_{B,z}|^2, \tag{4.2}$$

where $\|\mathbf{M}\|$ is the max norm of the matrix $\mathbf{M}$ and c_1 is a constant depending only on $|\Sigma|$, Ω_{ALL}, and a_c. Next, by (3.10) and (3.11), we get

$$|\tau_{B,z}|^2 \leq c_2\|\mathbf{u}\|^2_{H^1(\Omega)} \leq c_3|J_{diss}|. \tag{4.3}$$

Thus,

$$|J_{eff}| = \left|\frac{u_{B,z}\tau_{B,z}}{J_{diss}}\right| \leq c_3c_6\frac{\tau^2_{B,z}}{\tau^2_{B,z}} = C_2, \tag{4.4}$$

where we again used (3.12); this gives the assertion. ∎

4.2. Admissible shapes. In order to have a well-posed optimization problem, a suitable admissible set must be defined for the set of controls. Hence, we must define an admissible set $\mathcal{X}$ for the base parametrization $\mathbf{X}$. The main virtue of this set will be to ensure that (2.1) gives a well-defined surface Γ, i.e., a surface that is non-self-intersecting.

4.2.1. Curve parametrization. To do this, we use a few concepts from [26] that relate to self-contact of elastic curves and rods. Let $\mathbb{G}$ be the set of continuous maps $\gamma : [-1,1] \to \mathbb{R}^3$ (3-D curves) that have a Lipschitz continuous equal arc-length parametrization $\mathbf{X}_\gamma : [-1,1] \to \mathbb{R}^3$, such that $\|\mathbf{X}'_\gamma(t)\| = L/2$ for almost all t in $[-1,1]$, where L is the total length of the parameterized curve, i.e., $L = |\mathbf{X}([-1,1])|$. Note that we sometimes drop the γ subscript notation and identify γ with $\mathbf{X}_\gamma$.

Remark 4. We point out that maps γ in $W^{1,q}([-1,1],\mathbb{R}^3)$ for $1 \le q \le \infty$ are also in $\mathbb{G}$. This is because $W^{1,q}([-1,1],\mathbb{R}^3)$ is a subset of the functions of bounded variation, and one can always find a Lipschitz continuous arc-length parametrization in this case [24, p. 255], [26].

For convenience of the reader, we list some results from [26], but we deviate slightly in that our curves are not closed loops. However, the results we use are still true with a suitable modification [26]. The concept of tubular neighborhood will be useful. Define

$$B_r(\Sigma) = \{\mathbf{x} \in \mathbb{R}^3 : \operatorname{dist}(\mathbf{x}, \Sigma) < r\}, \tag{4.5}$$

where Σ is any set in $\mathbb{R}^3$, $r > 0$, and $B_r(\Sigma)$ is an open set containing Σ. The solid region $B_r(\Sigma)$ is said to be non-self-intersecting if the closest-point projection map $\Pi_\Sigma : B_r(\Sigma) \to \Sigma$ is single-valued and continuous.

4.2.2. Global radius of curvature. Next, let $R(\mathbf{x},\mathbf{y},\mathbf{z}) \ge 0$ be the radius of the smallest circle containing $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$. When $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ are noncollinear (and distinct) we have $R(\mathbf{x},\mathbf{y},\mathbf{z}) = \frac{|\mathbf{x}-\mathbf{y}|}{|2\sin[\angle(\mathbf{x}-\mathbf{z},\mathbf{y}-\mathbf{z})]|}$, where $\angle(\mathbf{a},\mathbf{b})$ is the positive measure of the angle made by the vectors $\mathbf{a}$, $\mathbf{b}$ (see [26] for more details). Now we define the global radius of curvature functions.

DEFINITION 4.1. *Let γ be in $\mathbb{G}$ with an equal arc-length parametrization $\mathbf{X}$ defined on $[-1,1]$, and assume that $|\mathbf{X}([-1,1])| > 0$. Then the global radius of curvature of γ at the point $\mathbf{X}(t_0)$ (for t_0 in $[-1,1]$) is given by*

$$\rho_{\text{global}}(\gamma, t_0) := \inf\{R(\mathbf{X}(t_0), \mathbf{X}(t), \mathbf{X}(t')) : t, t' \in [-1,1], \text{ and } t \ne t_0, t' \ne t_0, t \ne t'\}, \tag{4.6}$$

and denote its infimum by

$$\mathcal{R}_{\text{global}}(\gamma) := \inf_{-1 \le t_0 \le 1} \rho_{\text{global}}(\gamma, t_0). \tag{4.7}$$

From [26], we have the following.

LEMMA 4.1. *Suppose γ in $\mathbb{G}$ with equal arc-length parametrization $\mathbf{X}$. Assume γ has a double point, i.e., t, t' in $[-1,1]$ such that $t \ne t'$ but $\mathbf{X}(t) = \mathbf{X}(t')$. Then, $\rho_{\text{global}}(\gamma, t) = \rho_{\text{global}}(\gamma, t') = 0$. If $\mathcal{R}_{\text{global}}(\gamma) > 0$, then γ is simple (i.e., has no self-intersection).*

4.2.3. Some analysis results. The following regularity result [26] is also useful.

LEMMA 4.2. *Let γ be in $\mathbb{G}$ with equal arc-length parametrization $\mathbf{X}$. Assume $\mathcal{R}_{\text{global}}(\gamma) \ge d$ for some constant $d > 0$. Then $\mathbf{X}'$ is Lipschitz continuous, i.e., $\mathbf{X}$ is*

$C^{1,1}([-1,1])$ *(or* $W^{2,\infty}([-1,1])$*) and*

$$\|\mathbf{X}'(t_1) - \mathbf{X}'(t_2)\| \leq d^{-1}|t_1 - t_2| \quad \textit{for all } t_1, t_2 \in [-1,1]. \tag{4.8}$$

We will also need a slight modification of two lemmas from [26], which are essential for defining the admissible set of shapes for Ω_{B}.

LEMMA 4.3 (non-self-intersection). *Let* γ *be in* $\mathbb{G}$ *and assume* $\mathcal{R}_{\mathrm{global}}(\gamma) \geq d$ *for some given constant* $d > 0$*. Take* $\mathbf{X} : [-1,1] \to \mathbb{R}^3$ *to be the equal arc-length parametrization of* γ *and assume that* $\|\mathbf{X}(-1) - \mathbf{X}(1)\| \geq 2d$*, which implies that* $L \geq 2d$*. Then*

- $\mathrm{diam}(\mathbf{X}([-1,1])) \geq 2d$*,*
- $B_d(\mathbf{X}([-1,1]))$ *does not self-intersect.*

The second item is the so-called excluded volume constraint.

LEMMA 4.4 (weak closure). *Let* $\{\gamma_n\} \subset W^{1,q}([-1,1],\mathbb{R}^3)$*,* $q \in (1,\infty]$*, be a sequence of maps with equal arc-length parametrizations* $\mathbf{X}_{\gamma_n} : [-1,1] \to \mathbb{R}^3$*. Suppose* $\gamma_n \rightharpoonup \gamma \in W^{1,q}([-1,1],\mathbb{R}^3)$ *and*

$$\mathcal{R}_{\mathrm{global}}(\gamma_n) \geq d, \quad \|\mathbf{X}_{\gamma_n}(-1) - \mathbf{X}_{\gamma_n}(1)\| \geq 2d, \quad \textit{for all } n \geq 1$$

for some constant $d > 0$*. Then*

$$\mathcal{R}_{\mathrm{global}}(\gamma) \geq d, \quad \|\mathbf{X}_{\gamma}(-1) - \mathbf{X}_{\gamma}(1)\| \geq 2d. \tag{4.9}$$

Remark 5. The modification $\|\mathbf{X}(-1) - \mathbf{X}(1)\| \geq 2d$ is to prevent a pathological case and ensure that $B_d(\mathbf{X}([-1,1]))$ does not self-intersect. Consider a perfect circular arc defined by

$$\mathbf{X}(t) = (\sin((1-\epsilon)\pi t), \cos((1-\epsilon)\pi t), 0) \text{ for } t \in [-1,1], \tag{4.10}$$

where $\epsilon > 0$ is small. Computing the global radius of curvature gives $\mathcal{R}_{\mathrm{global}}(\gamma) = 1$ for all $\epsilon > 0$, yet it is clear that $B_1(\mathbf{X}([-1,1]))$ *does* intersect itself. Moreover, if $\epsilon = 0$, then $\mathbf{X}(-1) = \mathbf{X}(1)$, and it follows from Definition 4.1 that $\mathcal{R}_{\mathrm{global}}(\gamma) = 0$. This issue was avoided in [26] by considering only *closed* curves. The proofs of Lemmas 4.3 and 4.4 are a straightforward modification of that in [26].

4.2.4. Admissible parametrizations and locomotor shapes. Define the set of admissible parametrizations $\mathcal{X}$

$$\begin{aligned}\mathcal{X}(d,L) = \{\mathbf{X} : [-1,1] \to \mathbb{R}^3 \in \mathbb{G} : \|\mathbf{X}(t)\| \leq r_0/2,\ \|\mathbf{X}'(t)\| = L/2,\ \text{for all } t \in [-1,1],\\ \mathcal{R}_{\mathrm{global}}(\mathbf{X}) \geq d,\ \|\mathbf{X}(-1) - \mathbf{X}(1)\| \geq 2d\}\end{aligned} \tag{4.11}$$

for any fixed "thickness" constant $d > 0$ and specified length $L \geq 2d > 0$. (Note that $\mathcal{X}$ is not a convex set.) For compatibility reasons, we choose r_0 such that $r_0 \gg d$, $r_0 \gg A = \max_{t,\theta} |a_{\mathrm{c}}(t,\theta)|$, and $d > A$. Note that because we impose an equal arc-length parametrization, we have by Lemma 4.2 that $\mathcal{X}(d,L) \subset C^{1,1}([-1,1])$. Recall that $\mathbf{\Psi}$ is uniquely determined given any $\mathbf{X}$ in $\mathcal{X}$. Ergo, $\mathcal{X}$ is equivalent to the admissible shape set $\mathcal{O}$

$$\mathcal{O} = \{\Omega_{\mathrm{B}} \subset \Omega_{\mathrm{ALL}} :\ \Gamma \equiv \partial\Omega_{\mathrm{B}} \text{ is parameterized by } \mathbf{\Psi}(\mathbf{X}),\ \text{where } \mathbf{X} \in \mathcal{X}(d,L)\}. \tag{4.12}$$

Thus, since $d > A$, any Ω_{B} in $\mathcal{O}$ is non-self-intersecting (by Lemma 4.3), so all shapes in $\mathcal{O}$ are well defined. Also note that all Ω_{B} in $\mathcal{O}$ are at least Lipschitz continuous.

The condition $\|\mathbf{X}(-1)-\mathbf{X}(1)\| \geq 2d$ is important here, because it ensures that shapes in $\mathcal{O}$ do not have cusp like regions (e.g., if the endpoints of Σ were touching; see Remark 5). The constraint $\|\mathbf{X}(t)\| \leq r_0/2$ prevents any contact of the swimmer with the outer boundary Γ_{O}.

4.3. Formal optimization statement. The formal statement of the optimization problem is that we want to find the shape of the centerline $\Sigma \equiv \mathbf{X}([-1,1])$ such that J_{mag} is minimized over the admissible set $\mathcal{O}$. More precisely, the minimization problem is the following: find an *optimal pair* $(\mathbf{X}^*, \mathbf{u}^*(\mathbf{X}^*))$ such that

$$J_{\mathrm{mag}}\Big(\mathbf{X}^*, \mathbf{u}^*(\mathbf{X}^*)\Big) = \min_{\Omega_{\mathrm{B}}(\mathbf{X})\in\mathcal{O}} J_{\mathrm{mag}}\Big(\mathbf{X}, \mathbf{u}(\mathbf{X})\Big), \tag{4.13}$$

where $J_{\mathrm{mag}} = -J_{\mathrm{ts}}$ or $J_{\mathrm{mag}} = -J_{\mathrm{eff}}$ and $(\mathbf{X}, \mathbf{u}(\mathbf{X}))$ solves (2.3) for a particular shape $\Omega_{\mathrm{B}}(\mathbf{X})$. Again, the set $\mathcal{O}$ is general in that we are not restricting the optimization to a small set of parameters.

We introduce the admissibility set of controls and velocities

$$\begin{aligned}\mathcal{V} = \{(\mathbf{X}, \mathbf{u}(\mathbf{X})) \in \mathcal{X}(d,L)\times \mathbb{V}_0(\mathbf{X}) : J_{\mathrm{mag}}(\mathbf{X}, \mathbf{u}(\mathbf{X})) < \infty,\ \text{where there is a } p\in\mathbb{Q}\\ \text{such that } (\mathbf{u}, p) \text{ is a solution of (3.6)}\}.\end{aligned} \tag{4.14}$$

Then the extremal problem (4.13) can be restated as

$$J_{\mathrm{mag}}\Big(\mathbf{X}^*, \mathbf{u}^*(\mathbf{X}^*)\Big) = \min_{(\mathbf{X},\mathbf{u}(\mathbf{X}))\in\mathcal{V}} J_{\mathrm{mag}}\Big(\mathbf{X}, \mathbf{u}(\mathbf{X})\Big). \tag{4.15}$$

4.4. Convergence of domains. We now clarify notions of convergence of a sequence of functions $\{u_n\}$ when the domain itself (Ω_n) is also changing [28]. First, we use the fact that the domain shape Ω_{B} (and also Ω) is directly parameterized in terms of $\mathbf{X}$ to *define* domain convergence.

DEFINITION 4.2. *Let $\{\mathbf{X}_n\}$ be a sequence in $\mathcal{X}(d,L)$ for some $d > A = \max_{t,\theta} |a_{\mathrm{c}}(t,\theta)|$ and $L > 0$ satisfying $L \geq 2d$. For each $\mathbf{X}_n$ in $\mathcal{X}$, let $\Omega_{\mathrm{B},n}$ be in $\mathcal{O}$ such that $\Omega_{\mathrm{B},n} = \Omega_{\mathrm{B}}(\mathbf{X}_n)$. Let $\Omega_n := \Omega(\mathbf{X}_n) = \Omega_{\mathrm{ALL}} \setminus \overline{\Omega_{\mathrm{B},n}}$. Then we define the convergence of $\Omega_{\mathrm{B},n}$ to $\Omega_{\mathrm{B}}(\mathbf{X})$ by*

$$\Omega_{\mathrm{B},n} \to \Omega_{\mathrm{B}}(\mathbf{X}) \Longleftrightarrow \|\mathbf{X}_n - \mathbf{X}\|_{L^\infty([-1,1])} = \max_{-1\leq t\leq 1} |\mathbf{X}_n(t) - \mathbf{X}(t)| \to 0. \tag{4.16}$$

Note that $\Omega_{\mathrm{B}}(\mathbf{X})$ in $\mathcal{O}$ is well defined by Lemmas 4.3 and 4.4. Convergence of Ω_n is similarly defined.

Next, we must extend functions defined on Ω_n to Ω_{ALL} in order to make clear the statement "$u_n \to u$." The following theorem is adapted from the Calderón extension theorem (see [1]).

THEOREM 4.1. *Let Ω be a uniformly Lipschitz domain in $\mathbb{R}^n$. Then there is a linear continuous extension operator $P : H^1(\Omega) \to H^1(\mathbb{R}^n)$, such that for u in $H^1(\Omega)$ we have that $\|Pu\|_{H^1(\mathbb{R}^n)} \leq C\|u\|_{H^1(\Omega)}$, where $C > 0$ depends on the Lipschitz constant of Ω. Similarly, we can define an extension operator $P_{\Omega_{\mathrm{ALL}}} : H^1(\Omega) \to H^1(\Omega_{\mathrm{ALL}})$ such that*

$$\hat{u} = P_{\Omega_{\mathrm{ALL}}} u, \quad \|\hat{u}\|_{H^1(\Omega_{\mathrm{ALL}})} \leq \widehat{C}\|u\|_{H^1(\Omega)},\ \text{for } \widehat{C} > 0. \tag{4.17}$$

Hence, convergence of a sequence of functions $\{u_n\}$ (each u_n defined on Ω_n) will be understood in terms of convergence of their extensions $\{\hat{u}_n\}$ to the fixed domain Ω_{ALL}.

4.4.1. Lower semicontinuity. Clearly, the functional J_{ts} is strongly continuous. As for J_{diss}, the following lemma can be obtained from [22, 28].

LEMMA 4.5. *Let* $\mathbf{X}$ *and* $\{\mathbf{X}_n\}$ *be in* $\mathcal{X}(d,L)$ *for some compatible* d *and* L *(see Definition* 4.2*). Note that* $\Omega = \Omega(\mathbf{X})$ *and* $\Omega_n = \Omega_n(\mathbf{X}_n)$ *are uniformly Lipschitz. Let* $\mathbf{u} \in \mathbb{V}_0(\mathbf{X})$ *and* $\mathbf{u}_n \in \mathbb{V}_0(\mathbf{X}_n)$ *and define the extensions* $\hat{\mathbf{u}} := P_{\Omega_{\text{ALL}}}\mathbf{u}$, $\hat{\mathbf{u}}_n := P_{\Omega_{\text{ALL}}}\mathbf{u}_n$. *Assume that* $\hat{\mathbf{u}}_n \rightharpoonup \hat{\mathbf{u}}$ *in* $H^1(\Omega_{\text{ALL}})$. *Then the functionals defined by* (2.6) *and* (2.7) *satisfy*

$$\begin{aligned} J_{\text{ts}}(\mathbf{u}_{\text{B}}(\mathbf{X}), \boldsymbol{\tau}_{\text{B}}(\mathbf{X})) &= \liminf_{n\to\infty} J_{\text{ts}}(\mathbf{u}_{\text{B},n}(\mathbf{X}_n), \boldsymbol{\tau}_{\text{B},n}(\mathbf{X}_n)), \\ J_{\text{diss}}(\mathbf{X}, \mathbf{u}(\mathbf{X})) &\le \liminf_{n\to\infty} J_{\text{diss}}(\mathbf{X}_n, \mathbf{u}_n(\mathbf{X}_n)). \end{aligned} \tag{4.18}$$

COROLLARY 4.1. *Recall* (2.6), (2.8) *and let* $J_{\text{mag}} = -J_{\text{ts}}$ *or* $J_{\text{mag}} = -J_{\text{eff}}$. *Then* J_{mag} *satisfies*

$$J_{\text{mag}}(\mathbf{X}, \mathbf{u}(\mathbf{X})) \le \liminf_{n\to\infty} J_{\text{mag}}(\mathbf{X}_n, \mathbf{u}_n(\mathbf{X}_n)). \tag{4.19}$$

Proof. The proof follows directly from Lemma 4.5. ∎

4.4.2. Compactness of admissible set.

LEMMA 4.6 (compactness of $\mathcal{X}$). *Let* $\{\mathbf{X}_n\}$ *be a sequence in* $\mathcal{X}(d,L)$ *with* $d, L > 0$ *and* $L \ge 2d$. *Then there is a subsequence converging uniformly in* $W^{1,\infty}([-1,1])$ *to an* $\mathbf{X}$ *in* $\mathcal{X}(d,L)$, *i.e.,*

$$\|\mathbf{X}_{n_k} - \mathbf{X}\|_{W^{1,\infty}([-1,1])} \to 0, \ \text{as } k \to \infty. \tag{4.20}$$

Proof. Let $\mathbf{X}_n$ be in $\mathcal{X}(d,L)$ for all $n \ge 1$. Then the family of functions $\{\mathbf{X}_n\}$ is defined on a compact set $[-1,1]$ and is uniformly bounded, and both $\{\mathbf{X}_n\}$ and $\{\mathbf{X}_n'\}$ are equicontinuous families of functions; recall $\mathbf{X}_n$ is Lipschitz continuous as well as $\mathbf{X}_n'$ by Lemma 4.2. Thus, by the Arzela–Ascoli theorem [41, 43], there is a uniformly convergent subsequence $\{\mathbf{X}_{n_k}\}$:

$$\|\mathbf{X}_{n_k} - \mathbf{X}\|_{L^\infty([-1,1])} \to 0, \qquad \|\mathbf{X}_{n_k}' - \mathbf{X}'\|_{L^\infty([-1,1])} \to 0. \tag{4.21}$$

Hence, $\mathbf{X}$ is in $W^{1,\infty}([-1,1])$. By the definition of $\mathcal{X}(d,L)$, the subsequence $\{\mathbf{X}_{n_k}\}$ satisfies the hypothesis of Lemma 4.4. So $\mathbf{X}$ satisfies

$$\mathcal{R}_{\text{global}}(\mathbf{X}) \ge d, \quad \|\mathbf{X}(-1) - \mathbf{X}(1)\| \ge 2d. \tag{4.22}$$

Likewise, uniform convergence implies $\|\mathbf{X}(t)\| \le r_0/2$ and $\|\mathbf{X}'(t)\| = L/2$. Thus, $\mathbf{X}$ is in $\mathcal{X}(d,L)$. ∎

4.5. A minimizer exists. We apply the direct method in the calculus of variations to prove the following theorem.

THEOREM 4.2. *There exists at least one minimizer* $(\mathbf{X}^*, \mathbf{u}^*(\mathbf{X}^*))$ *in* $\mathcal{V}$ *for the problem* (4.15).

Proof. Clearly, $\mathcal{V}$ is nonempty because there exists a unique solution of (3.6). Assume d is sufficiently large, i.e., $d > A = \max_{t,\theta} |a_{\text{c}}(t,\theta)|$, and $L \ge 2d$. So all shapes in $\mathcal{O}$ are well defined.

For any sequence $\{\mathbf{X}_n\}$ in $\mathcal{X}(d,L)$, let $\Omega_{\text{B},n} = \Omega_{\text{B}}(\mathbf{X}_n)$, $p_n = p(\mathbf{X}_n)$, and $\mathbf{u}_n = \mathbf{u}(\mathbf{X}_n)$ denote the dependence on $\mathbf{X}_n$, where $(\mathbf{u}_n, p_n)$ is the solution of (3.6) for the locomotor shape given by $\Omega_{\text{B},n}$. From Proposition 4.1, we know that J_{mag} is bounded

below for all $\mathbf{X}$ in $\mathcal{X}$. Thus, one can find a minimizing sequence $\{(\mathbf{X}_n, \mathbf{u}_n)\}$ in $\mathcal{V}$ such that

$$\lim_{n\to\infty} J_{\text{mag}}(\mathbf{X}_n, \mathbf{u}_n) = \inf_{(\mathbf{X},\mathbf{u}(\mathbf{X}))\in\mathcal{V}} J_{\text{mag}}(\mathbf{X}, \mathbf{u}(\mathbf{X})).$$

Therefore, by Lemma 4.6, there exists a subsequence of $\{\mathbf{X}_n\}$, again denoted $\{\mathbf{X}_n\}$, and an $\mathbf{X}^*$ in $\mathcal{X}$ such that $\mathbf{X}_n \to \mathbf{X}^*$ uniformly in $W^{1,\infty}([-1,1])$.

By the definition of the admissible set $\mathcal{X}$ and Lemma 3.2, we have that $\|\mathbf{u}_n\|_{H^1(\Omega_n)}$ is uniformly bounded, i.e., $\|\mathbf{u}_n\|_{H^1(\Omega_n)} \le K_0$ for some independent constant $K_0 > 0$. Now we extend $\mathbf{u}_n$ to Ω_{ALL} by setting $\hat{\mathbf{u}}_n(\mathbf{x}) = \mathbf{u}_{\text{B},n} + \boldsymbol{\omega}_{\text{B},n} \times (\mathbf{x} - \mathbf{x}_c)$ for all $\mathbf{x}$ in $\Omega_{\text{B},n}$, and $\hat{\mathbf{u}}_n(\mathbf{x}) = \mathbf{u}_n(\mathbf{x})$ for all $\mathbf{x}$ in Ω_n. One can show that $\|\hat{\mathbf{u}}_n\|_{H^1(\Omega_{\text{ALL}})} \le C\|\mathbf{u}_n\|_{H^1(\Omega_n)} \le K_1$ uniformly in n. Using the "inf-sup" condition [8, 6, 25], one can also show that $\|p_n\|_{L^2(\Omega_n)}$ is uniformly bounded (p_n being the pressure associated with $\mathbf{u}_n$). Next, we define an extension of p_n by setting $\hat{p}_n(\mathbf{x}) = 0$ for all $\mathbf{x}$ in $\Omega_{\text{B},n}$, and $\hat{p}_n(\mathbf{x}) = p_n(\mathbf{x})$ for all $\mathbf{x}$ in Ω_n. Clearly, $\|\hat{p}_n\|_{L^2(\Omega_{\text{ALL}})}$ is uniformly bounded.

Hence, we can take weakly convergent subsequences, and using the Rellich lemma [21], we get

$$\hat{\mathbf{u}}_n \rightharpoonup \hat{\mathbf{u}}, \text{ in } H^1(\Omega_{\text{ALL}}), \quad \hat{\mathbf{u}}_n \to \hat{\mathbf{u}}, \text{ in } L^2(\Omega_{\text{ALL}}), \quad \hat{p}_n \rightharpoonup \hat{p}, \text{ in } L^2(\Omega_{\text{ALL}}) \tag{4.23}$$

for some $\hat{\mathbf{u}}$ in $H^1(\Omega_{\text{ALL}})$ and $\hat{p}$ in $L^2(\Omega_{\text{ALL}})$. Now define $\mathbf{u}(\mathbf{X}^*) = \hat{\mathbf{u}}|_{\Omega(\mathbf{X}^*)}$ and $p(\mathbf{X}^*) = \hat{p}|_{\Omega(\mathbf{X}^*)}$. We now show that $(\mathbf{u}(\mathbf{X}^*), p(\mathbf{X}^*))$ solves the weak formulation (3.6) over $\Omega(\mathbf{X}^*)$.

To this end, define

$$\begin{aligned}\mathbf{W} = \{&\mathbf{v} \in C^\infty(\Omega_{\text{ALL}}) : \mathbf{v} = \mathbf{v}_{\text{B}} + \boldsymbol{\vartheta}_{\text{B}} \times (\mathbf{x} - \mathbf{x}_c) \text{ on an open subset}\\ &\text{containing } \Omega_{\text{B}}(\mathbf{X}^*), \text{ and } \mathbf{v}_{\text{B}} \in \mathbb{R}^3, \boldsymbol{\vartheta}_{\text{B}} = (0,0,a), a \in \mathbb{R}, \text{ and } \mathbf{v} = \mathbf{0}, \text{ on } \Gamma_{\text{O}}\}\end{aligned} \tag{4.24}$$

and take $\boldsymbol{\varphi}$ in $\mathbf{W}$. Because $\mathbf{X}_n \to \mathbf{X}^*$ uniformly, $\boldsymbol{\varphi}$ is in $H^1(\Omega_m)$ for m sufficiently large. So we can take $\mathbf{v} = \boldsymbol{\varphi}$ in the first equation in (3.6) on Ω_m (for m sufficiently large) to get

$$\frac{1}{2}\int_{\Omega_m} D(\mathbf{u}_m) : D(\boldsymbol{\varphi}) - \int_{\Omega_m} p_m \nabla\cdot\boldsymbol{\varphi} = (\boldsymbol{\vartheta}_{\text{B}}\cdot\mathbf{e}_z)\tau_{\text{B},z}. \tag{4.25}$$

Taking the limit of the first term, we have (by extension)

$$\begin{aligned}\int_{\Omega_m} D(\mathbf{u}_m) : D(\boldsymbol{\varphi}) &= \int_{\Omega_{\text{ALL}}} D(\hat{\mathbf{u}}_m) : D(\boldsymbol{\varphi}) \to \int_{\Omega_{\text{ALL}}} D(\hat{\mathbf{u}}) : D(\boldsymbol{\varphi})\\ &= \int_{\Omega(\mathbf{X}^*)} D(\mathbf{u}(\mathbf{X}^*)) : D(\boldsymbol{\varphi}),\end{aligned}$$

where the last equality follows because $\hat{\mathbf{u}}$ is a rigid motion on $\Omega_{\text{B}}(\mathbf{X}^*)$. Likewise,

$$\int_{\Omega_m} p_m \nabla\cdot\boldsymbol{\varphi} \to \int_{\Omega(\mathbf{X}^*)} p(\mathbf{X}^*)\nabla\cdot\boldsymbol{\varphi},$$

because $\hat{p}$ vanishes on $\Omega_{\text{B}}(\mathbf{X}^*)$. Combining these results reduces (4.25) to

$$\frac{1}{2}\int_{\Omega(\mathbf{X}^*)} D(\mathbf{u}(\mathbf{X}^*)) : D(\boldsymbol{\varphi}) - \int_{\Omega(\mathbf{X}^*)} p(\mathbf{X}^*)\nabla\cdot\boldsymbol{\varphi} = (\boldsymbol{\vartheta}_{\text{B}}\cdot\mathbf{e}_z)\tau_{\text{B,z}} \tag{4.26}$$

for all $\boldsymbol{\varphi}$ in $\mathbf{W}$. By density of $\mathbf{W}$ in $\mathbb{V}_0(\mathbf{X}^*)$, we know that $(\mathbf{u}(\mathbf{X}^*), p(\mathbf{X}^*))$ satisfies the first equation in (3.6) on $\Omega(\mathbf{X}^*)$. Similarly, one can show that $(\mathbf{u}(\mathbf{X}^*), p(\mathbf{X}^*))$ also satisfies the divergence constraint.

Therefore, $(\mathbf{X}^*, \mathbf{u}(\mathbf{X}^*))$ is in $\mathcal{V}$, which implies that $\mathcal{V}$ is weakly closed. Moreover, J_{mag} is lower semicontinuous by Corollary 4.1. Ergo, we apply the following argument [14, 36]: since $\{(\mathbf{X}_n, \mathbf{u}_n)\}$ is a minimizing sequence, we have

$$\begin{aligned}\inf_{(\mathbf{X},\mathbf{u}(\mathbf{X}))\in\mathcal{V}} J_{\text{mag}}(\mathbf{X}, \mathbf{u}(\mathbf{X})) &= \liminf_{n\to\infty} J_{\text{mag}}(\mathbf{X}_n, \mathbf{u}_n) \\ (4.19) \to &\geq J_{\text{mag}}(\mathbf{X}^*, \mathbf{u}(\mathbf{X}^*)) \geq \inf_{(\mathbf{X},\mathbf{u}(\mathbf{X}))\in\mathcal{V}} J_{\text{mag}}(\mathbf{X}, \mathbf{u}(\mathbf{X})),\end{aligned}$$

by the definition of infimum. Thus, we get $J_{\text{mag}}(\mathbf{X}^*, \mathbf{u}^*) = \inf_{(\mathbf{X},\mathbf{u}(\mathbf{X}))\in\mathcal{V}} J_{\text{mag}}(\mathbf{X}, \mathbf{u}(\mathbf{X}))$, so $(\mathbf{X}^*, \mathbf{u}(\mathbf{X}^*))$ is a minimizer for (4.15). ∎

Remark 6. The admissible set $\mathcal{X}$ in (4.11) can easily be modified to include a cargo constraint (see section 5.2.2), but we avoided this in section 4.2.4 in order to have a simpler presentation. The proof goes through, essentially the same, in this case as well.

5. Optimization approach and algorithm. Deriving a gradient-based optimization algorithm requires the sensitivity of the cost functionals with respect to perturbations of the shape [15, 18, 52, 61]. Some applications of shape optimization in fluids can be found in [27, 68, 69]. Some related references on optimization of nonlinear elastic curves are [4, 11, 12, 48].

The following sections state the formulas that give the sensitivity information for our problem. In section 5.3, we pose a variational method for computing descent directions for the centerline curve parametrization $\mathbf{X}$, followed by a description of the complete optimization algorithm. We then prove well-posedness of the descent direction solution (5.16) in section 5.4.

5.1. Existence of shape derivative of cost functionals. The following theorem states that J_{ts} and J_{diss} have a shape derivative in terms of an appropriate adjoint problem under certain smoothness assumptions.

THEOREM 5.1. *Suppose Γ is C^2, and let $\mathbf{V}_\Gamma$ be a C^2 shape perturbation of Γ defined on Ω_{ALL}, such that $\mathbf{V}_\Gamma = \mathbf{0}$ on Γ_{O}. Then $\mathbf{u}$ is $H^2(\Omega)\cap\mathbb{V}_0$ and p is $H^1(\Omega)\cap\mathbb{Q}$ (see* [59, 66]*), and it has a shape derivative $(\mathbf{u}', p')$ in the direction $\mathbf{V}_\Gamma$, which satisfies* (B.3) *(see Appendix* B.1.1*). In addition, the functionals J_{ts} and J_{diss} are differentiable with respect to $\mathbf{V}_\Gamma$ (i.e., $\delta J_{\text{ts}}(\Omega_{\text{B}}; \mathbf{V}_\Gamma)$ and $\delta J_{\text{diss}}(\Omega_{\text{B}}; \mathbf{V}_\Gamma)$ exist) and are given by*

$$\begin{aligned}\delta J_{\text{ts}}(\Omega_{\text{B}}; \mathbf{V}_\Gamma) &= ((\mathbf{M}\boldsymbol{\tau}_{\text{B}}) \times \boldsymbol{\omega}_{\text{B}}) \cdot \dot{\mathbf{x}}_c + \int_\Gamma (\mathbf{V}_\Gamma \cdot \boldsymbol{\nu})(\mathbf{S}\boldsymbol{\nu} - \mathbf{H}\boldsymbol{\nu}) \cdot [\mathbf{I} - \boldsymbol{\nu}\otimes\boldsymbol{\nu}]\boldsymbol{\sigma}\boldsymbol{\nu}, \\ \delta J_{\text{diss}}(\Omega_{\text{B}}; \mathbf{V}_\Gamma) &= \int_\Gamma (\mathbf{V}_\Gamma \cdot \boldsymbol{\nu})(\mathbf{K}\boldsymbol{\nu}) \cdot [\mathbf{I} - \boldsymbol{\nu}\otimes\boldsymbol{\nu}]\boldsymbol{\sigma}\boldsymbol{\nu},\end{aligned} \tag{5.1}$$

where "$\otimes$" denotes the tensor product, $\boldsymbol{\nu}$ points out of Ω (into Ω_{B}), and $\dot{\mathbf{x}}_c$ is given by

$$\dot{\mathbf{x}}_c = -\int_\Gamma (\mathbf{x} - \mathbf{x}_c)(\mathbf{V}_\Gamma \cdot \boldsymbol{\nu}) \,/\, |\Omega_{\text{B}}|. \tag{5.2}$$

The adjoint stress tensor $\mathbf{S}$ *is defined by* $\mathbf{S}(\mathbf{r},\varrho) = -\varrho\mathbf{I} + D(\mathbf{r})$ *with* $(\mathbf{r},\varrho)$ *solving the following adjoint Stokes problem:*

$$
\begin{aligned}
-\nabla\cdot\mathbf{S}(\mathbf{r},\varrho) &= \mathbf{0} \ \textit{in}\ \Omega,\\
\nabla\cdot\mathbf{r} &= 0 \ \textit{in}\ \Omega,\\
\mathbf{r} &= \mathbf{r}_{\mathrm{B}} + \boldsymbol{\eta}_{\mathrm{B}}\times(\mathbf{x}-\mathbf{x}_c) \ \textit{on}\ \Gamma,\\
\mathbf{r} &= \mathbf{0} \ \textit{on}\ \Gamma_{\mathrm{O}},\\
\int_\Gamma \mathbf{S}\boldsymbol{\nu} = \mathbf{M}\boldsymbol{\tau}_{\mathrm{B}} =: \mathbf{g}_{\mathrm{B}}\ (\text{given}),&\quad \int_\Gamma (\mathbf{x}-\mathbf{x}_c)\times(\mathbf{S}\boldsymbol{\nu}) = \boldsymbol{\xi}_{\mathrm{B}},
\end{aligned}
\tag{5.3}
$$

where $\mathbf{r}_{\mathrm{B}}$, $\boldsymbol{\eta}_{\mathrm{B}}$, *and* $\boldsymbol{\xi}_{\mathrm{B}}$ *have the form*

$$
\begin{aligned}
\mathbf{r}_{\mathrm{B}} &= \begin{pmatrix} r_{\mathrm{B,x}}\ (\text{unknown})\\ r_{\mathrm{B,y}}\ (\text{unknown})\\ r_{\mathrm{B,z}}\ (\text{unknown})\end{pmatrix},\ \boldsymbol{\eta}_{\mathrm{B}} = \begin{pmatrix} 0\\ 0\\ \bar{\eta}z\ (\text{unknown})\end{pmatrix},\\
\boldsymbol{\xi}_{\mathrm{B}} &= \begin{pmatrix} \xi_{\mathrm{B,x}}\ (\text{unknown})\\ \xi_{\mathrm{B,y}}\ (\text{unknown})\\ 0\end{pmatrix}.
\end{aligned}
\tag{5.4}
$$

Similarly, $\mathbf{H}$, $\mathbf{K}$ *are the stress tensors for adjoint problems* (B.9), (B.12) *(see Appendix* B.1*).*

Proof. The existence of $(\mathbf{u}', p')$ can be proved by straightforward modification of [27, Theorems 6.15, 6.18], [35, section 11.3.4, Lemmas 11.5, 11.6], or [47]; these references contain results on shape differentiation of the Navier–Stokes equations. The existence of unique solutions to the adjoint equations is clear. A formal derivation of (5.1) is provided in Appendix B.1. ∎

5.2. Sensitivities of constraints. We need sensitivities of the constraints for the optimization algorithm in section 5.3. The inextensibility and cargo constraints are purely geometric, so the sensitivity calculations are straightforward (see subsequent sections). The global radius of curvature constraint is also geometric, but it is nonlocal (see Remark 8).

5.2.1. Local inextensibility constraint. We require that $\|\mathbf{X}'(t)\| = L/2$ for $-1 \le t \le 1$, which is more conveniently written as an integral. To this end, let

$$
L_{\mathrm{loc}}(\mu;\Sigma) = \int_\Sigma \mu(\mathbf{X}_0\circ\mathbf{X}^{-1}) - \int_{\Sigma_0}\mu \quad \text{for all } \mu:\Sigma_0\to\mathbb{R},
\tag{5.5}
$$

where $\mathbf{X}_0$ is a parametrization of a reference curve Σ_0, $\mathbf{X}^{-1}$ is the inverse map of $\mathbf{X}$, and μ is any scalar valued integrable function. One can think of μ as a density distribution. Thus, the inextensibility constraint is now written as $L_{\mathrm{loc}}(\mu;\Sigma) = 0$ for all μ. Using the parametrization, we rewrite this as

$$
\int_{-1}^{1} \mu(\mathbf{X}_0(t))(\|\mathbf{X}'(t)\| - \|\mathbf{X}_0'(t)\|)\,dt = 0 \quad \text{for all } \mu,
$$

where the reference curve satisfies $\|\mathbf{X}_0'(t)\| = L/2$ (local constraint). Since μ is arbitrary, this clearly recovers the differential constraint $\|\mathbf{X}'(t)\| = L/2$. Note that a global constraint is *not* adequate because the surface shape Γ depends on the way Σ is parameterized (see section 2.1).

The sensitivity of (5.5) can be computed by standard methods in the calculus of variations [16, 36], which we write as a bilinear form (useful later)

$$b(\mathbf{V},\lambda) := \delta L_{\text{loc}}(\lambda;\mathbf{V}) = \int_\Sigma \lambda \boldsymbol{\tau}\cdot\partial_s\mathbf{V} \qquad \text{for all } \mathbf{V}\in\mathbb{X}, \quad \text{for all } \lambda\in\mathbb{M}, \tag{5.6}$$

where ∂_s is the derivative with respect to arc-length, $\boldsymbol{\tau}$ is the oriented unit tangent vector of $\mathbf{X}$, $\mathbf{V}$ is a perturbation of $\mathbf{X}$ (i.e., Σ), λ is a Lagrange multiplier (defined on Σ), and $\mathbb{X}\equiv\mathbb{X}(\mathbf{X}) := H^2(\Sigma)$ and $\mathbb{M}\equiv\mathbb{M}(\mathbf{X}) := (H^1(\Sigma))^*$ (see section 5.3.1). Alternatively, we could try to build inextensibility explicitly into the parametrization, but it is more conveniently dealt with by Lagrange multipliers.

5.2.2. Cargo constraint. For studying the practical aspects of optimal locomotor shapes, we impose an additional obstacle or cargo constraint to simulate the case where the locomotor is transporting a fixed payload. Assume the cargo is a rigid body (set denoted Ω_{cargo}) whose shape is described by the zero level-set of ϕ_{cargo}. We want to ensure that Ω_{B} lies strictly outside the cargo, i.e.,

$$\phi_{\text{cargo}}(\mathbf{x}) \le -C_{\text{cargo}}, \quad \text{for all } \mathbf{x}\in\Omega_{\text{B}}, \quad C_{\text{cargo}}>0 \text{ (fixed constant)}, \tag{5.7}$$

where ϕ_{cargo} is the signed distance function for Ω_{cargo} (positive inside Ω_{cargo}). With this, we modify the model (2.3) by replacing Ω_{B} with $\widetilde{\Omega_{\text{B}}} = \Omega_{\text{B}}\cup\Omega_{\text{cargo}}$, i.e., the swimmer consists of the tail and cargo rigidly attached to each other. The tail shape is still parameterized by (2.1).

We include this as a penalty function (added to the cost functional) for the entire curve Σ. Let $g:(0,\infty)\to\mathbb{R}$ be a C^∞ barrier function satisfying the following: $g(t)\ge 0$, $g'(t)\le 0$, $g(t)\to\infty$ as $t\to 0^+$, and $g(t)=0$ for all $t\ge\gamma_0 > 2C_{\text{cargo}}$ where γ_0 is a chosen parameter. Then the extra penalty term is

$$H_{\text{cargo}}(\Sigma) = \int_\Sigma g(-\phi_{\text{cargo}}(\mathbf{x})) = \int_{-1}^{1} g(-\phi_{\text{cargo}}(\mathbf{X}(t)))\|\mathbf{X}'(t)\|dt \tag{5.8}$$

with sensitivity given by standard calculus of variations [16, 36]

$$\begin{aligned}\delta H_{\text{cargo}}(\Sigma;\mathbf{V}) &= \int_{-1}^{1} -g'(-\phi_{\text{cargo}}(\mathbf{X}(t)))(\mathbf{V}(t)\cdot\nabla\phi_{\text{cargo}}(\mathbf{X}(t)))\|\mathbf{X}'(t)\|dt \\ &\quad + \int_{-1}^{1} g(-\phi_{\text{cargo}}(\mathbf{X}(t)))\left(\frac{\mathbf{X}'(t)}{\|\mathbf{X}'(t)\|}\cdot\mathbf{V}'(t)\right)dt \\ &= -\int_\Sigma g'(-\phi_{\text{cargo}}(\mathbf{x}))(\mathbf{V}\cdot\nabla\phi_{\text{cargo}}(\mathbf{x})) + \int_\Sigma g(-\phi_{\text{cargo}}(\mathbf{x}))(\boldsymbol{\tau}\cdot\partial_s\mathbf{V}).\end{aligned} \tag{5.9}$$

In addition, we also have an equality constraint for the positive end-point $\partial\Sigma^+\equiv\mathbf{X}(+1)$, that is, we want the end-point to stay a fixed distance from the surface of the cargo: $-\phi_{\text{cargo}}(\mathbf{X}(+1)) = C_{\text{cargo}}$, where $C_{\text{cargo}}>0$ is a fixed constant. This models the cargo as being rigidly attached to the locomotor. For reasons of scaling, we impose

$$E^+(\Sigma) := g(-\phi_{\text{cargo}}(\mathbf{X}(+1))) - g(C_{\text{cargo}}) \equiv 0 \tag{5.10}$$

with sensitivity given by

$$\delta E^+(\Sigma;\mathbf{V}) = -g'(-\phi_{\text{cargo}}(\mathbf{X}(+1)))\mathbf{V}(+1)\cdot\nabla\phi_{\text{cargo}}(\mathbf{X}(+1)). \tag{5.11}$$

We assume that g is such that $-g'(C_{\text{cargo}}) = \alpha_0 > 0$ for some fixed constant α_0; this is done to ensure the sensitivity does not vanish. Similarly to (5.6), we define the following bilinear form for later use:

$$c(\mathbf{V}, \zeta) = \zeta\, \delta E^+(\Sigma; \mathbf{V}), \qquad \text{for all } \mathbf{V} \in \mathbb{X}, \quad \text{for all } \zeta \in \mathbb{R}. \tag{5.12}$$

Remark 7. For simplicity, we model a passive payload as a rigid sphere with one end-point of Σ (i.e., the "attachment point") constrained to be a fixed distance from the surface of the sphere, i.e., (5.10). Of course, there is a gap between the tail and cargo, but the entire configuration (tail and cargo) is considered as one rigid body $\widetilde{\Omega_{\text{B}}} = \Omega_{\text{B}} \cup \Omega_{\text{cargo}}$. Note that only the tail shape Σ and attachment point are optimized here; cargo shape is not being optimized. Having a small gap between the tail and cargo allows for flexibility in defining the cargo/tail shape parametrization. This changes the problem slightly from what was given in sections 2–4 but does not pose any serious difficulties (recall Remark 6). The set of admissible shapes (4.11) is easily modified to include cargo constraints.

5.3. Variational method.

5.3.1. Lagrangian. For simplicity, we formulate a descent method for the functional $J_{\text{mag}}(\mathbf{X}) = -J_{\text{ts}}(\mathbf{X}, \mathbf{u}(\mathbf{X})) + \epsilon H_{\text{cargo}}(\mathbf{X})$, where $\epsilon > 0$ is a small penalty parameter for the cargo to ensure that the locomotor does not intersect it. The minus sign means that we are actually maximizing J_{ts}. Note that we can reformulate (4.15) over the set (4.11) because $\mathcal{X}(d, L)$ is equivalent to $\mathcal{O}$. (Recall that the surface parametrization is fixed.) Therefore, to determine a solution of (4.15), we first define a Lagrangian functional to handle the constraints:

$$\mathcal{L}(\mathbf{X}, \lambda, \zeta) = J_{\text{mag}}(\mathbf{X}) + L_{\text{loc}}(\lambda; \mathbf{X}) + \zeta E^+(\mathbf{X}), \tag{5.13}$$

where $\lambda \in \mathbb{M}$, $\zeta \in \mathbb{R}$ are Lagrange multipliers. Note that $\Sigma \equiv \mathbf{X}([-1, 1])$, so we replaced Σ by $\mathbf{X}$ in (5.13). The function spaces needed to make sense of the descent method are

$$\begin{aligned} &\mathbb{X} \equiv \mathbb{X}(\mathbf{X}) = H^2(\Sigma) \text{ (space of curve perturbations)},\\ &\mathbb{M} \equiv \mathbb{M}(\mathbf{X}) = (H^1(\Sigma))^* \text{ (local inextensibility Lagrange multiplier space)}. \end{aligned} \tag{5.14}$$

Remark 8. We choose $H^2(\Sigma)$ as the base function space for perturbations of $\mathbf{X}$ in order to maintain control of the curvature $\boldsymbol{\kappa}$. Recall that the admissible set (4.11) had a lower bound on the global radius of curvature which also served to impose the excluded volume constraint (see section 4.2). But including this in our (numerical) method would require an extra penalty term (or inequality constraint) that involves a nonlocal computation. In lieu of this, we take a convenient compromise by "smoothing" the curve perturbation with $H^2(\Sigma)$. Of course, because of the bilinear form (5.6), this induces the space for the multiplier to be a dual space.

The first-order optimality conditions (KKT system [35]) associated with (5.13) are

$$\begin{aligned} \delta_{\mathbf{X}}\mathcal{L}(\mathbf{X}^*, \lambda^*, \zeta^*; \mathbf{V}) &= \delta J_{\text{mag}}(\mathbf{X}^*; \mathbf{V}) + \delta L_{\text{loc}}(\lambda^*; \mathbf{V}) + \zeta^*\, \delta E^+(\mathbf{X}^*; \mathbf{V}) = 0,\\ \delta_{\lambda}\mathcal{L}(\mathbf{X}^*, \lambda^*, \zeta^*; \mu) &= L_{\text{loc}}(\mu; \mathbf{X}^*) = 0,\\ \delta_{\zeta}\mathcal{L}\left(\mathbf{X}^*, \lambda^*, \zeta^*; \tilde{\zeta}\right) &= \tilde{\zeta} E^+(\mathbf{X}^*) = 0 \end{aligned} \tag{5.15}$$

for all $\mathbf{V}$ in $\mathbb{X}(\mathbf{X}^*)$, μ in $\mathbb{M}(\mathbf{X}^*)$, $\tilde{\zeta}$ in $\mathbb{R}$, i.e., we seek a solution $(\mathbf{X}^*, \lambda^*, \zeta^*)$ of (5.15).

5.3.2. Constrained gradient descent. Applying Newton's method to (5.15) yields an iterative method for obtaining an extremal solution. If we replace the Hessian term in Newton's method by a positive definite inner product, we obtain the following variational problem, whose solution gives a descent perturbation $\boldsymbol{\varphi}$ of the centerline Σ at each step of the iteration.

VARIATIONAL FORMULATION 5.1 (constrained gradient flow). *Let $\widehat{\mathbf{X}}$ be a known curve parametrization. Find $\boldsymbol{\varphi}$ in $\mathbb{X}(\widehat{\mathbf{X}})$, λ in $\mathbb{M}(\widehat{\mathbf{X}})$, and ζ in $\mathbb{R}$ such that*

$$\begin{aligned} \langle \boldsymbol{\varphi}, \mathbf{V} \rangle_{H^2(\Sigma)} + b(\mathbf{V}, \lambda) + c(\mathbf{V}, \zeta) &= -\delta J_{\text{mag}}(\widehat{\mathbf{X}}; \mathbf{V}) \text{ for all } \mathbf{V} \text{ in } \mathbb{X}(\widehat{\mathbf{X}}), \\ b(\boldsymbol{\varphi}, \mu) &= -L_{\text{loc}}(\mu; \widehat{\mathbf{X}}) \text{ for all } \mu \text{ in } \mathbb{M}(\widehat{\mathbf{X}}), \\ c(\boldsymbol{\varphi}, \tilde{\zeta}) &= -E^{+}(\widehat{\mathbf{X}})\tilde{\zeta} \text{ for all } \tilde{\zeta} \text{ in } \mathbb{R}. \end{aligned} \tag{5.16}$$

Note that $\boldsymbol{\varphi}$ and $\mathbf{V}$ are defined on $\Sigma \equiv \widehat{\mathbf{X}}([-1,1])$. Assuming that the constraints are satisfied, setting $\mathbf{V} = \boldsymbol{\varphi}$ gives $\delta J_{\text{mag}}(\widehat{\mathbf{X}}; \boldsymbol{\varphi}) = -\langle \boldsymbol{\varphi}, \boldsymbol{\varphi} \rangle_{H^2(\Sigma)} < 0$, provided $\boldsymbol{\varphi} \neq \mathbf{0}$. So $\boldsymbol{\varphi}$ is a descent perturbation of the parametrization $\widehat{\mathbf{X}}$. Evaluating $\delta J_{\text{mag}}(\widehat{\mathbf{X}}; \mathbf{V})$ requires computing $\delta J_{\text{ts}}(\widehat{\mathbf{X}}; \mathbf{V}_\Gamma(\mathbf{V}))$, where $\mathbf{V}_\Gamma$ is defined by extending $\mathbf{V}$ from Σ to Γ by the surface parametrization in (2.1) *(see Appendix* B.2*).*

By iterating the system (5.16), we obtain a *steepest descent* method for optimizing the centerline curve Σ. Intuitively, one can view the optimization process as the deformation of an "elastic beam" driven by a body force given by the shape derivative δJ_{mag}. The algorithm essentially consists of solving a sequence of linear problems and is iterated until the shape converges; see Algorithm 1 for a detailed description. One can find other function space based optimization methods in [10, 32, 33, 34, 35].

ALGORITHM 1. SEMIDISCRETE SHAPE FLOW.

Let α be the step size for updating the domain shape.
Let $\mathbf{X}^0$ be an initial parametrization of the centerline such that $\Sigma^0 := \mathbf{X}^0([-1,1])$ which has a given length, i.e., $|\Sigma^0| = L$. Note that this induces a shape Ω_{B}^0, $\Gamma^0 := \partial\Omega_{\text{B}}^0$.
Define $\Omega^0 = \Omega_{\text{ALL}} \setminus \Omega_{\text{B}}^0$.
for $k = 0, 1, 2, \ldots$ **do**
  Solve Stokes: Let $(\mathbf{u}^k, p^k, \boldsymbol{\tau}_{\text{B}}^k)$ solve (2.3) on Ω^k. Let $\boldsymbol{\sigma}^k$ be the associated stress tensor.
  Solve Adjoint Stokes: Let $(\mathbf{r}^k, \varrho^k, \boldsymbol{\xi}_{\text{B}}^k)$ solve (5.3) on Ω^k. Let $\mathbf{S}^k$ be the associated stress tensor.
  Evaluate Sensitivities: Compute $\delta J_{\text{mag}}(\mathbf{X}^k; \mathbf{V})$ for all $\mathbf{V}$.
  Solve for Descent Direction: Let $\boldsymbol{\varphi}^{k+1}$ solve (5.16) on $\mathbf{X}^k$.
  Update Shape: Let $\mathbf{X}^{k+1}(t) := \mathbf{X}^k(t) + \alpha\boldsymbol{\varphi}^{k+1}(\mathbf{X}^k(t))$ for all t in $[-1,1]$; α is obtained via a backtracking line-search. This yields Σ^{k+1}, which induces a shape Ω_{B}^{k+1}, $\Gamma^{k+1} := \partial\Omega_{\text{B}}^{k+1}$.
  Define $\Omega^{k+1} = \Omega_{\text{ALL}} \setminus \Omega_{\text{B}}^{k+1}$.
end for

The fully discrete algorithm follows directly from Algorithm 1 by applying a spatial discretization (see section 6.1). A convergence criteria can be based on $\|\mathbf{X}^j - \mathbf{X}^{j+1}\|_{L^2([-1,1])}$ and whether the constraints are suitably satisfied.

Remark 9. nonlocal computation. One could include the global radius of curvature constraint in the Lagrangian (5.13), but it is difficult to compute with because it is nonlocal. Alternatively, one can perform the global radius of curvature calculation (or an approximation of it) within the line-search process. In other words, if

$\widetilde{\mathbf{X}}$ is a candidate curve, and if $\mathcal{R}_{\text{global}}(\widetilde{\mathbf{X}}) < d$ (where d is some thickness threshold), then the step size should be reduced. Of course this could lead to small steps in the optimization if the shape really wants to have self-contact. In all the numerical cases we explored for the microswimmer problem, we never witnessed any tendency for self contact. However, it is not clear how to avoid this constraint when proving the existence of a minimizer (section 4).

5.4. Well-posed descent method. Because the inner product $\langle\cdot,\cdot\rangle_{H^2(\Sigma)}$ is trivially coercive over $H^2(\Sigma)$, the well-posedness of computing a descent direction with (5.16) is guaranteed if the so-called inf-sup condition is satisfied [6, 8]. This is proved in the following lemma.

LEMMA 5.1 (continuous inf-sup). *Let Σ be a parameterized 3-D curve and assume Σ has bounded curvature (see Remark 8). Define $\mathbb{X}^- = \{\mathbf{v} \in \mathbb{X} : \mathbf{v} = 0$ at $\partial\Sigma^- \equiv \mathbf{X}(-1)\}$; $\mathbb{X}^+$ is defined similarly with $\partial\Sigma^+ \equiv \mathbf{X}(+1)$. Then there exists a constant $\beta > 0$ that* only depends *on $L = |\Sigma|$ and $\|\kappa\|_{L^2(\Sigma)}$, where κ is the scalar curvature of $\mathbf{X}$, such that*

$$\sup_{\mathbf{v}\in\mathbb{X}^-} \frac{\int_\Sigma \mu[\boldsymbol{\tau}\cdot(\partial_s\mathbf{V})]}{\|\mathbf{V}\|_{H^2(\Sigma)}} \geq \beta\|\mu\|_{\mathbb{M}} \text{ for all } \mu \in \mathbb{M}. \tag{5.17}$$

Note that $\mathbb{X}^-$ can be replaced with $\mathbb{X}^+$.

Proof. Without loss of generality, we will consider the $\mathbb{X}^-$ case only. Let μ be an arbitrary function in $\mathbb{M}$. Let $s : [-1, 1] \to [0, L]$ be the arc-length function associated with $\mathbf{X}$, i.e., $s(t) = \int_{-1}^{t} \|\mathbf{X}'(r)\|\, dr$, which implies that $t = -1 \Leftrightarrow s = 0$ and $t = +1 \Leftrightarrow s = L$.

First, note the definition of the dual norm:

$$\|\mu\|_{\mathbb{M}} = \sup_{\omega\in H^1(\Sigma)} \frac{\int_\Sigma \mu(s)\omega(s)}{\|\omega\|_{H^1(\Sigma)}}, \tag{5.18}$$

where the integral is understood in the sense of duality pairing [1, 21]. Ergo, there exists a ϱ in $H^1(\Sigma)$ such that

$$\int_\Sigma \mu\varrho = \|\mu\|^2_{\mathbb{M}}, \quad \|\varrho\|_{H^1(\Sigma)} = \|\mu\|_{\mathbb{M}}. \tag{5.19}$$

Next, let $\mathbf{W}$ be defined on Σ in terms of $s(\cdot)$ by $\mathbf{W}(\hat{s}) := \int_0^{\hat{s}} \varrho(s)\boldsymbol{\tau}(s)ds$. This gives

$$\int_\Sigma \mu[\boldsymbol{\tau}\cdot(\partial_s\mathbf{W})] = \int_0^L \mu(s)\boldsymbol{\tau}(s)\cdot(\varrho(s)\boldsymbol{\tau}(s))ds = \int_0^L \mu(s)\varrho(s)ds = \|\mu\|^2_{\mathbb{M}}. \tag{5.20}$$

Next, by standard inequalities and noting $|\boldsymbol{\tau}| = 1$, we have

$$\begin{aligned}
\|\mathbf{W}\|^2_{L^2(\Sigma)} &= \int_0^L |\mathbf{W}(s)|^2 ds \leq \int_0^L \left(\int_0^s |\varrho(r)|dr\right)^2 ds \\
&\leq L^2 \int_0^L \varrho^2(s)ds = L^2\,\|\varrho\|^2_{L^2(\Sigma)}, \\
\|\partial_s\mathbf{W}\|^2_{L^2(\Sigma)} &= \|\varrho\|^2_{L^2(\Sigma)}, \\
\|\partial_s^2\mathbf{W}\|^2_{L^2(\Sigma)} &\leq 2(\|\partial_s\varrho\|^2_{L^2(\Sigma)} + \|\kappa\|^2_{L^2(\Sigma)}\|\varrho\|^2_{L^2(\Sigma)}).
\end{aligned} \tag{5.21}$$

Thus, $\mathbf{W}$ is in $\mathbb{X}^-$ and we obtain

(5.22)
$$\|\mathbf{W}\|^2_{H^2(\Sigma)} \leq (L^2 + 1 + 2\|\kappa\|^2_{L^2(\Sigma)})\|\varrho\|^2_{L^2(\Sigma)} + 2\|\partial_s\varrho\|^2_{L^2(\Sigma)} \leq C_0^2\|\varrho\|^2_{H^1(\Sigma)} = C_0^2\|\mu\|^2_{\mathbb{M}},$$

where $C_0 = \max((L + 1 + \sqrt{2}\|\kappa\|_{L^2(\Sigma)}), \sqrt{2})$. Combining (5.22) with (5.20) gives

$$\frac{\int_\Sigma \mu[\boldsymbol{\tau}\cdot(\partial_s\mathbf{W})]}{\|\mathbf{W}\|_{H^2(\Sigma)}} \geq \beta\|\mu\|_{\mathbb{M}}, \tag{5.23}$$

where $\beta = 1/C_0$. Replacing $\mathbf{W}$ by the supremum over all $\mathbb{X}^-$ gives the assertion. ∎

Remark 10. In Lemma 5.1, we allowed one end-point to be constrained. However, it is not possible to satisfy (5.17) in general if both are constrained. Consider the case where Σ is straight ($\boldsymbol{\tau}$ is constant) and $\mathbf{V}$ is required to vanish at $\partial\Sigma$. Then for $\mu = 1$, we have

$$\int_\Sigma \mu[\boldsymbol{\tau}\cdot(\partial_s\mathbf{V})] = \mu\boldsymbol{\tau}\cdot\int_\Sigma \partial_s\mathbf{V} = \mu\boldsymbol{\tau}\cdot\mathbf{V}\Big|_{\partial\Sigma} = 0,$$

which implies that (5.17) cannot be true.

To conclude, we state the full well-posedness result.

THEOREM 5.2 (well-posedness). *Assume that $\partial\Sigma^+ \equiv \mathbf{X}(+1)$ satisfies the end-point constraint* (5.10)*, and the barrier function g satisfies $-g'(C_{\text{cargo}}) = \alpha_0 > 0$ for some fixed constant α_0 (see section* 5.2.2*). Under the hypothesis of Lemma* 5.1*, there exists a constant $\beta > 0$ that* only depends *on $L = |\Sigma|$, $\|\kappa\|_{L^2(\Sigma)}$, α_0, such that*

$$\sup_{\mathbf{V}\in\mathbb{X}} \frac{b(\mathbf{V},\mu) + c(\mathbf{V},\tilde{\zeta})}{\|\mathbf{V}\|_{H^2(\Sigma)}} \geq \beta(\|\mu\|_{\mathbb{M}} + |\tilde{\zeta}|), \text{ for all } \mu\in\mathbb{M}, \quad \tilde{\zeta}\in\mathbb{R}. \tag{5.24}$$

Moreover, there is a unique solution $(\boldsymbol{\varphi},\lambda,\zeta)$ of (5.16) *that depends continuously on the data.*

Proof. Let $\mu\in\mathbb{M}$ and $\tilde{\zeta}\in\mathbb{R}$ be arbitrary but fixed. By Lemma 5.1, there exists $\mathbf{W}\in\mathbb{X}^+$ satisfying

$$b(\mathbf{W},\mu) = \int_\Sigma \mu[\boldsymbol{\tau}\cdot\partial_s\mathbf{W}] = \beta_0\|\mu\|_{\mathbb{M}}, \quad \|\mathbf{W}\|_{H^2(\Sigma)} = 1, \quad \mathbf{W}|_{\partial\Sigma^+} = \mathbf{0}.$$

Let $\mathbf{Z}\in\mathbb{X}$ be the constant vector given by $\mathbf{Z} = \nabla\phi_{\text{cargo}}(\mathbf{X}(+1))\,\text{sgn}(\tilde{\zeta})$, which satisfies

$$\|\mathbf{Z}\|_{H^2(\Sigma)} = |\nabla\phi_{\text{cargo}}(\mathbf{X}(+1))|\;\|1\|_{L^2(\Sigma)} = |\Sigma|^{1/2},$$

because ϕ_{cargo} is a distance function. By hypothesis, $\mathbf{X}(+1)$ satisfies $-\phi_{\text{cargo}}(\mathbf{X}(+1)) = C_{\text{cargo}}$. Thus, by (5.10), (5.11), and (5.12), we have

$$c(\mathbf{Z},\tilde{\zeta}) = \tilde{\zeta}\,(-g'(C_{\text{cargo}}))\mathbf{Z}(+1)\cdot\nabla\phi_{\text{cargo}}(\mathbf{X}(+1)) = \tilde{\zeta}\text{sgn}(\tilde{\zeta})\alpha_0 = \alpha_0|\tilde{\zeta}|,$$

where $\alpha_0 > 0$ is a fixed constant.

Now define $\mathbf{V} = \mathbf{W} + \mathbf{Z}$. Then $\|\mathbf{V}\|_{H^2(\Sigma)} \leq 1 + |\Sigma|^{1/2}$ and

$$b(\mathbf{V},\mu) + c(\mathbf{V},\tilde{\zeta}) = b(\mathbf{W},\mu) + \underbrace{b(\mathbf{Z},\mu)}_{=0} + \underbrace{c(\mathbf{W},\tilde{\zeta})}_{=0} + c(\mathbf{Z},\tilde{\zeta}) = \beta_0\|\mu\|_{\mathbb{M}} + \alpha_0|\tilde{\zeta}|.$$

The statement (5.24) follows by forming the quotient and taking the supremum with $\beta = \min(\beta_0,\alpha_0)$. Well-posedness follows from [6], [8, Theorem 1.1]. ∎

6. Numerical discretization. This section describes the finite element/spline method we use to approximate the system (5.16). We then prove well-posedness of a discrete version of (5.16). As for the approximation of the 3-D Stokes equations, many numerical schemes already exist. The method we use in our numerical demonstrations is described in section 7.

6.1. Spatial discretization. The discretization of the infinite dimensional optimization problem requires a set of discrete spaces to replace the function spaces given earlier. We begin by partitioning the interval $[-1,1]$ into a set of subintervals: $\mathcal{I} := \{I_k\}_{k=1}^{N-1}$, i.e., the mesh. Moreover, we use the symbol "h" to denote discretization of the domain. If $\mathbf{X} : [-1,1] \to \mathbb{R}^3$ is a parametrization of Σ, then $\mathbf{X}$ induces a partition of edge segments $\{E_k\}$ on Σ, i.e., $E_k = \mathbf{X}(I_k)$.

Next, we need (at least) C^1 type basis functions to have a conforming approximation of the space $H^2(\Sigma)$. Since $\mathbf{X}$ is a 1-D curve, it is advantageous to use a cubic spline basis which has better continuity: C^2 [55, 63]. Let $\{t_k\}_{k=1}^N$ be the nodal points in the interval $[-1,1]$, i.e.,

$$t_1 = -1, \qquad t_{k+1} = I_k \cap I_{k+1}, \quad k = 1, 2, \ldots, N-2, \qquad t_N = 1.$$

Let $\{\eta_k\}_{k=0}^{N+1}$ be the cardinal cubic spline basis functions [55, 63] such that

$$\eta_i(t_j) = \delta_{ij} \quad \text{for all } 1 \le i, j \le N, \tag{6.1}$$

and η_0 and η_{N+1} are the end-slope basis functions

$$\begin{aligned} &\eta_0(t_j) = \eta_{N+1}(t_j) = 0, \quad 1 \le j \le N, \quad \eta_0'(t_1) = 1, \eta_0'(t_N) = 0, \\ &\eta_{N+1}'(t_1) = 0, \eta_{N+1}'(t_N) = 1. \end{aligned} \tag{6.2}$$

Then the cubic spline space is defined by

$$\mathbb{S}_h := \left\{ v \in C^2([-1,1]) : v(t) = \sum_{k=0}^{N+1} \alpha_k \eta_k(t) \quad \text{for all } \alpha_k \in \mathbb{R} \right\} \subset H^2([-1,1]), \tag{6.3}$$

where h is the mesh size.

For the remainder, let $\mathbf{X}_h : [-1,1] \to \mathbb{R}^3$ such that $\mathbf{X}_h \in \mathbb{S}_h$. In other words, $\Sigma_h = \mathbf{X}_h([-1,1])$ is a (parametric) cubic spline approximation of Σ. Note that $\mathbf{v}$ in L^2, for example, means each component of $\mathbf{v}$ is in L^2. The same notation holds for all the other spaces (including discrete).

Remark 11. Note that the statements of Lemma 5.1 and Theorem 5.2 are true if $\mathbf{X}$ is replaced with $\mathbf{X}_h$. Of course, having a curve with kinks in it would cause problems with the surface parametrization; recall (2.1). Thus, the use of a spline space is advantageous because of the extra level of differentiability.

We state the polynomial spaces [6, 7] needed in approximating $H^2(\Sigma_h)$ and $\mathbb{M}(\Sigma_h)$: the mapped spline space

$$\mathbb{X}_h := \{ v \in C^2(\Sigma_h) : (v \circ \mathbf{X}_h) \in \mathbb{S}_h \} \subset H^2(\Sigma_h) \tag{6.4}$$

and the piecewise constant space

$$\mathbb{M}_h := \{ v : v \circ \mathbf{X}_h|_I \in \mathcal{P}_0(I) \text{ for } I \in \mathcal{I} \} \subset L^2(\Sigma_h) \subset \mathbb{M}(\Sigma_h), \tag{6.5}$$

where $\mathcal{P}_k(I)$ is the space of polynomials of degree $\le k$ on the domain I. We now state the discrete version of (5.16).

VARIATIONAL FORMULATION 6.1 (discrete gradient flow). *Let $\widehat{\mathbf{X}}_h$ in $\mathbb{S}_h$ be a known curve parametrization. Find $\boldsymbol{\varphi}$ in $\mathbb{X}_h(\widehat{\mathbf{X}}_h)$, λ in $\mathbb{M}_h(\widehat{\mathbf{X}}_h)$, and ζ in $\mathbb{R}$ such that*

$$\begin{aligned} \langle \boldsymbol{\varphi}, \mathbf{V} \rangle_{H^2(\Sigma_h)} + b(\mathbf{V}, \lambda) + c(\mathbf{V}, \zeta) &= -\delta J_{\mathrm{mag}}(\widehat{\mathbf{X}}_h; \mathbf{V}) \text{ for all } \mathbf{V} \text{ in } \mathbb{X}_h(\widehat{\mathbf{X}}_h), \\ b(\boldsymbol{\varphi}, \mu) &= -L_{\mathrm{loc}}(\mu; \widehat{\mathbf{X}}_h) \text{ for all } \mu \text{ in } \mathbb{M}_h(\widehat{\mathbf{X}}_h), \\ c(\boldsymbol{\varphi}, \tilde{\zeta}) &= -E^+(\widehat{\mathbf{X}}_h)\tilde{\zeta} \text{ for all } \tilde{\zeta} \text{ in } \mathbb{R}. \end{aligned} \tag{6.6}$$

6.2. Stable descent scheme. The well-posedness of (6.6) follows by the same criteria as in section 5.4, the main result being the following.

LEMMA 6.1 (discrete inf-sup, piecewise constant). *Let Σ_h be parameterized by $\mathbf{X}_h \in \mathbb{S}_h$ and assume Σ_h has bounded curvature. Assume the hypothesis of Lemma 5.1. Let $\mathbb{X}_h^- = \mathbb{X}_h \cap \mathbb{X}^-$. Then there exists a constant $\beta > 0$ independent of $h_0 = \max_{E \subset \Sigma_h} |E|$ such that*

$$\sup_{\mathbf{V} \in \mathbb{X}_h^-} \frac{\int_{\Sigma_h} \mu[\boldsymbol{\tau} \cdot (\partial_s \mathbf{V})]}{\|\mathbf{V}\|_{H^2(\Sigma_h)}} \geq \beta \|\mu\|_{\mathbb{M}(\Sigma_h)} \text{ for all } \mu \in \mathbb{M}_h, \tag{6.7}$$

provided the mesh size h_0 is sufficiently small. Note that $\mathbb{X}_h^-$ can be replaced by $\mathbb{X}_h^+$.

Proof. Let us consider the $\mathbb{X}_h^-$ case only and let μ be an arbitrary function in $\mathbb{M}_h$. By Lemma 5.1 (and Remark 11), there exists an $\mathbf{F}$ in $\mathbb{X}^-$ such that

$$\|\mathbf{F}\|_{H^2(\Sigma_h)} = \|\mu\|_{\mathbb{M}}, \quad \int_{\Sigma_h} \mu[\boldsymbol{\tau} \cdot \partial_s \mathbf{F}] = \beta \|\mu\|_{\mathbb{M}}^2. \tag{6.8}$$

Next, let $\mathbf{W}$ be in $\mathbb{X}_h^-$ with nodal values given by interpolating $\mathbf{F}$ at the nodes of the mesh $\mathcal{I}$ and setting the end-slopes to zero; this is possible because of the Sobolev embedding: $H^2(\Sigma_h) \subset C^0(\Sigma_h)$. By basic approximation theory [58, Theorem 6.25, p. 230]

$$\|\mathbf{W}\|_{H^2(\Sigma_h)} \leq c_0 \|\mathbf{F}\|_{H^2(\Sigma_h)} \tag{6.9}$$

for some independent constant c_0. Also, let $\hat{\boldsymbol{\tau}}$ in $\mathbb{M}_h$ be a piecewise constant approximation of the unit tangent vector $\boldsymbol{\tau}$ of Σ_h defined by

$$\hat{\boldsymbol{\tau}}|_E = \frac{1}{|E|} \int_E \boldsymbol{\tau}, \quad \Rightarrow \|\hat{\boldsymbol{\tau}} - \boldsymbol{\tau}\|_{L^\infty(E)} \leq K_0 |E|, \tag{6.10}$$

for each edge segment E contained in Σ_h. Note that K_0 is a fixed constant that depends on the curvature of Σ_h (or its Lipschitz constant).

We must show that $\mathbf{W}$ satisfies a similar relation as (5.23). We start by considering an edge segment $E \subset \Sigma_h$. Note that $\int_E \partial_s \mathbf{W} = \int_E \partial_s \mathbf{F}$ because $\mathbf{W}$ is the nodal interpolant of $\mathbf{F}$. Then,

$$\begin{aligned} \int_E \mu[\boldsymbol{\tau} \cdot (\partial_s \mathbf{W})] &= \mu \int_E (\boldsymbol{\tau} - \hat{\boldsymbol{\tau}}) \cdot \partial_s \mathbf{W} + \mu \hat{\boldsymbol{\tau}} \cdot \int_E \partial_s \mathbf{W} \\ &= \mu \int_E (\boldsymbol{\tau} - \hat{\boldsymbol{\tau}}) \cdot \partial_s \mathbf{W} + \mu \hat{\boldsymbol{\tau}} \cdot \int_E \partial_s \mathbf{F}. \end{aligned} \tag{6.11}$$

Continuing, we get

$$\int_{\Sigma_h} \mu[\boldsymbol{\tau} \cdot (\partial_s \mathbf{W})] = \int_{\Sigma_h} \mu(\boldsymbol{\tau} - \hat{\boldsymbol{\tau}}) \cdot \partial_s (\mathbf{W} - \mathbf{F}) + \int_{\Sigma_h} \mu \boldsymbol{\tau} \cdot \partial_s \mathbf{F}, \tag{6.12}$$

where we summed over all edges. Therefore,

(6.13)
$$\begin{aligned}\int_{\Sigma_h} \mu[\boldsymbol{\tau}\cdot(\partial_s\mathbf{W})] &\geq -\|\boldsymbol{\tau}-\hat{\boldsymbol{\tau}}\|_{L^\infty(\Sigma_h)}\|\mu\|_{\mathbb{M}}(\|\partial_s\mathbf{W}\|_{H^1(\Sigma_h)}+\|\partial_s\mathbf{F}\|_{H^1(\Sigma_h)})\\ &\quad+\int_{\Sigma_h}\mu\boldsymbol{\tau}\cdot\partial_s\mathbf{F}\\ \text{by (6.10)} &\geq -c_1h_0\left[\|\mu\|_{\mathbb{M}}^2+\|\partial_s\mathbf{W}\|_{H^1(\Sigma_h)}^2+\|\partial_s\mathbf{F}\|_{H^1(\Sigma_h)}^2\right]+\int_{\Sigma_h}\mu[\boldsymbol{\tau}\cdot\partial_s\mathbf{F}],\end{aligned}$$

where $h_0=\max_{E\subset\Sigma_h}|E|$ and c_1 is an independent constant. Using (6.8), (6.9) gives

(6.14)
$$\int_{\Sigma_h}\mu[\boldsymbol{\tau}\cdot(\partial_s\mathbf{W})]\geq -c_2h_0\|\mu\|_{\mathbb{M}}^2+\int_{\Sigma_h}\mu[\boldsymbol{\tau}\cdot\partial_s\mathbf{F}]=(\beta-c_2h_0)\|\mu\|_{\mathbb{M}}^2\geq\frac{\beta}{2}\|\mu\|_{\mathbb{M}}^2$$

for some constant $c_2>0$ and h_0 sufficiently small. Thus, we obtain the quotient

$$\frac{\int_{\Sigma_h}\mu[\boldsymbol{\tau}\cdot(\partial_s\mathbf{W})]}{\|\mathbf{W}\|_{H^2(\Sigma_h)}}\geq\frac{\beta}{2c_0}\|\mu\|_{\mathbb{M}}. \tag{6.15}$$

The assertion follows by taking the supremum. ∎

Using splines gives an effective way to ensure smooth tangent and normal fields (when using piecewise approximations), so as to have a well-defined surface parametrization for Γ (i.e., no artificial self-intersections). Note that updating the centerline curve $\mathbf{X}_h$ throughout the optimization process is completely consistent because $\mathbf{X}_h$ and $\boldsymbol{\varphi}\circ\mathbf{X}_h$ are both in $\mathbb{S}_h$ (recall Algorithm 1). We close with the full stability result for computing discrete descent directions.

THEOREM 6.1 (well-posedness). *Assume that $\partial\Sigma^+\equiv\mathbf{X}(+1)$ satisfies the endpoint constraint, i.e., the last line of* (5.15)*. Under the hypothesis of Lemma* 6.1*, there exists a constant $\beta>0$ dependent on $L=|\Sigma_h|$, $\|\kappa\|_{L^2(\Sigma_h)}$, and α_0 (see section* 5.2.2*) but independent of $h_0=\max_{E\subset\Sigma_h}|E|$, such that*

$$\sup_{\mathbf{V}\in\mathbb{X}_h}\frac{b(\mathbf{V},\mu)+c(\mathbf{V},\tilde{\zeta})}{\|\mathbf{V}\|_{H^2(\Sigma_h)}}\geq\beta(\|\mu\|_{\mathbb{M}(\Sigma_h)}+|\tilde{\zeta}|),\ \textit{for all}\ \mu\in\mathbb{M}_h,\quad\tilde{\zeta}\in\mathbb{R}, \tag{6.16}$$

provided the mesh size h_0 is sufficiently small. Moreover, there is a unique solution $(\boldsymbol{\varphi}_h,\lambda_h,\zeta_h)$ of (6.6) *that depends continuously on the data.*

Proof. The proof is similar to the proof of Theorem 5.2. ∎

7. Computational results. We present numerical results of our discrete optimization algorithm, which solves (6.6) to obtain a descent direction at each iteration. Another important part of the algorithm is the solution of the 3-D Stokes equations (2.3) (and adjoint equations) in order to evaluate the cost sensitivities. We use a boundary integral approach, which easily accommodates the swimmer shapes described by (2.1) when $\{\mathbf{N}_1,\mathbf{N}_2\}$ are determined using parallel transport (see section 2.1). Details of the boundary integral method are in [38, 40]. Subsequent sections describe the numerical optimization results.

Remark 12. For each numerical example, we *choose* the desired length L and a constant C such that $L/(2A)=C$, where C is the aspect ratio. This determines A, which appears in (2.2) and fixes $\mathcal{O}$ in (4.12).

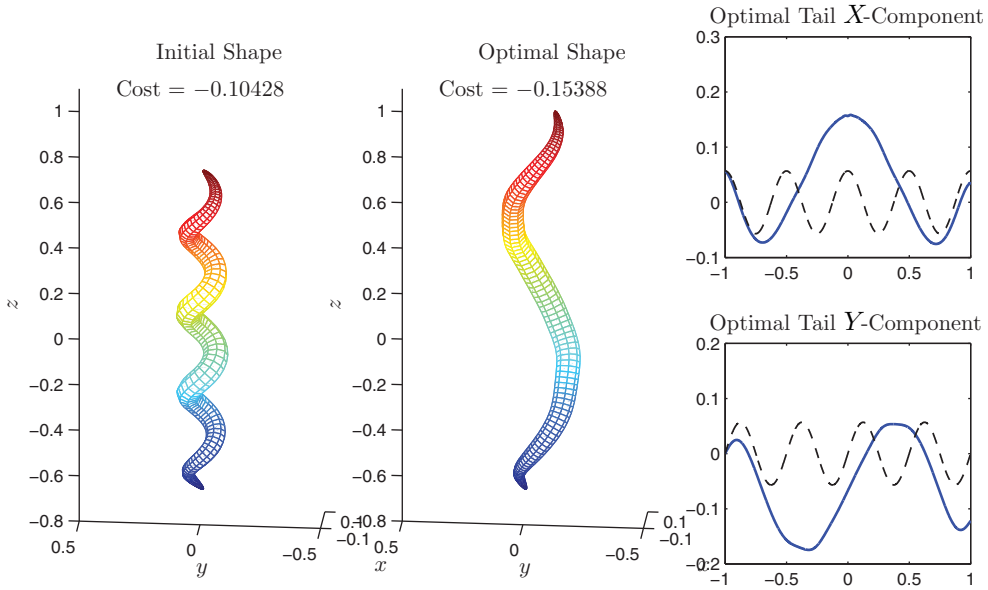


FIG. 7.1. *Optimal locomotor shape. The far left plot shows the initial shape (Σ is four full turns of a helix) with an aspect ratio of* 20; *the centerline length is constrained to be* $L = 2.0$. *The next plot shows the optimized result from our algorithm. On the right, we see the* X, Y *components of the parametrization of* Σ *versus the parameter variable:* $-1 \leq t \leq +1$; Z *component is just linear increasing. The dashed curve is the initial shape and the solid curve is the optimal shape. The surface grid is only for illustration; the computational grid for solving Stokes is of much higher resolution.*

7.1. Examples without cargo. We first present optimizations of swimmers that do not possess an attached cargo. Figure 7.1 shows the initial and final configurations for such a swimmer that has aspect ratio $L/(2A) = 20$ with $L = 2$. The centerline of the initial shape is given by $\mathbf{X}_0(t) = b\cos(kt)\mathbf{e}_x + b\sin(kt)\mathbf{e}_y + \alpha t\mathbf{e}_z$, where $\alpha = 0.7$, $k = 4\pi$, $b = k^{-1}\sqrt{1-\alpha^2}$, and $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$ are the canonical basis vectors of $\mathbb{R}^3$. This corresponds to a simple helix with four turns and pitch $\lambda = 2\pi\alpha/k = 0.35$. During the optimization, the body evolves into a shape that closely resembles a helix with 3/2 turns. Figure 7.2 shows the cost, J_{ts}, and sensitivity, δJ_{ts}, over the course of the optimization. We see the cost improve by a factor of 1.475, and the sensitivity approach zero, indicating that the shape predicted by our algorithm is a (local) minimizer.

Due to the scale invariance of the Stokes equations, the optimal shape will not change if L and A are scaled by the same factor. This is not the case, however, if one varies the aspect ratio. We demonstrate the effect of aspect ratio on the optimal shape by performing an optimization of a swimmer with $L/(2A) = 50$. The results from this optimization are provided in Figure 7.3. To compare with the $L/(2A) = 20$ case, we set the initial centerline shape to be given by $\mathbf{X}_0(t)$. We immediately see that the thinner body leads to higher speeds (lower cost values). We also see that the resulting optimal shape is quite different than that of the $L/(2A) = 20$ case. Here, the optimal shape retains the four turns but is straighter than the initial condition, namely, the projection of the centerline tangent onto the swimming direction is greater for the optimal shape. We find that the aspect ratio clearly has an impact on the optimal shape. It is important to note that these effects are not captured by drag-

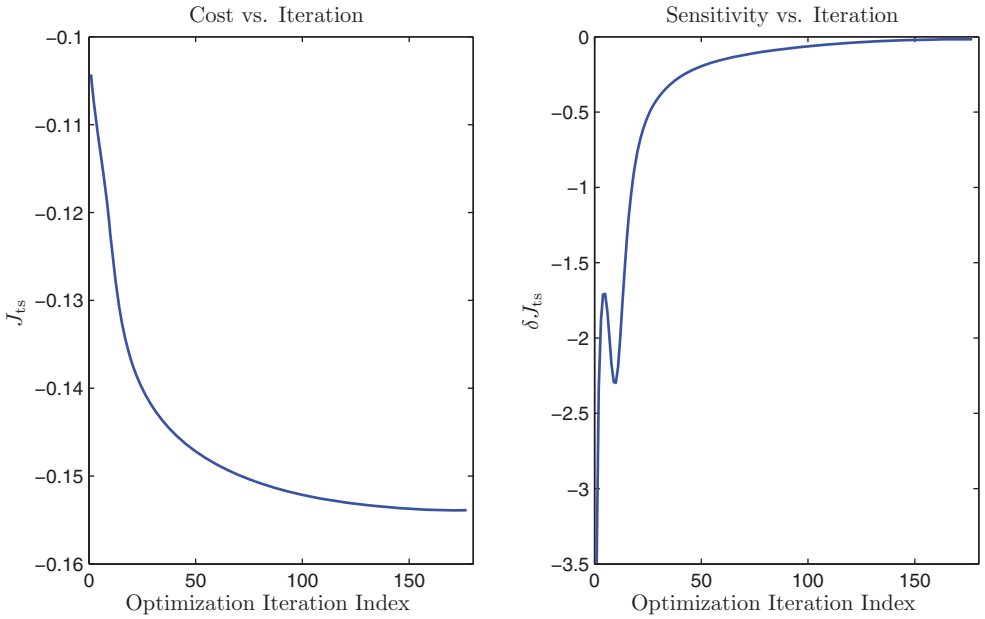


FIG. 7.2. *Cost functional and sensitivity versus optimization iteration index; data corresponds to that shown in Figure* 7.1. *In the sensitivity plot,* δJ_{ts} *is evaluated along the descent direction and is normalized such that the optimization step size is* 1.

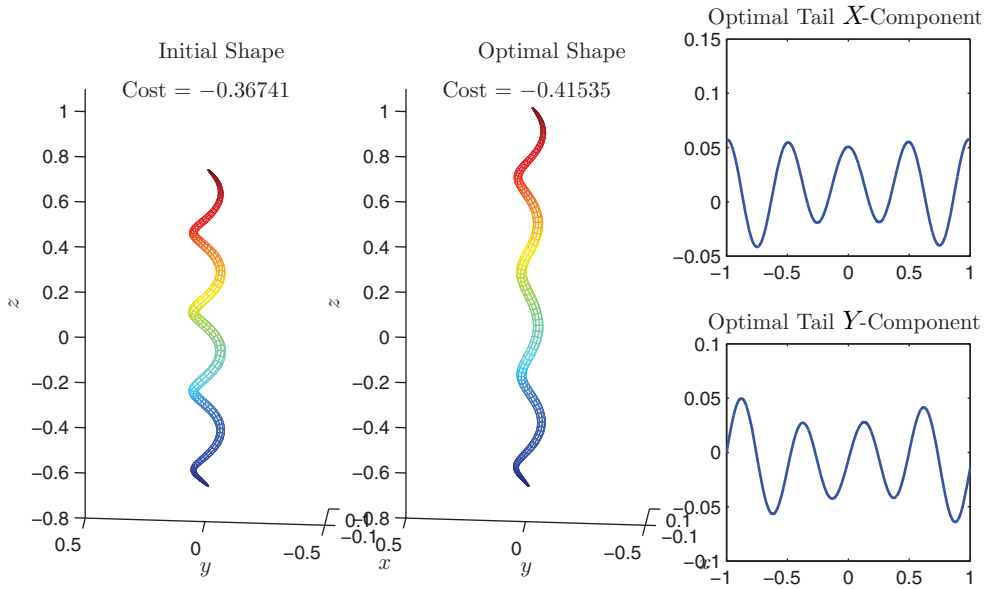


FIG. 7.3. *Optimal locomotor shape. The far left plot shows the initial locomotor shape (initial* Σ *is four full turns of a helix) with an aspect ratio of* 50; *the centerline length is constrained to be* $L = 2.0$. *The format is similar to Figure* 7.1.

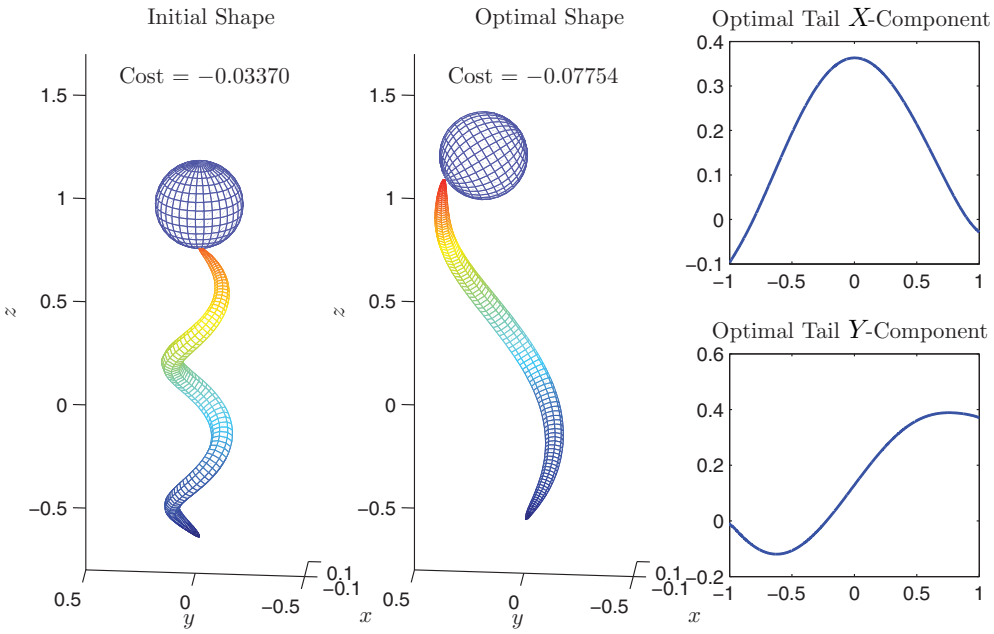


FIG. 7.4. *Optimal locomotor shape with spherical cargo. The far left plot shows the initial locomotor shape and cargo placement (initial* Σ *is two full turns of a helix); the centerline length is constrained to be* $L = 2.0$. *The format is similar to Figure* 7.1.

based models such as resistive force theory [44] that treat the hydrodynamic problem in the limit of $L/(2A) \to \infty$.

7.2. Example with cargo. The presence of a passive payload attached to one end of the swimmer affects the optimal shape. Figure 7.4 shows the result of an optimization of a swimmer with $L = 2$ and aspect ratio $L/(2A) = 20$ that is "attached" to a sphere with radius $R = 0.1L$ (recall section 5.2.2). This particular choice of aspect ratio, cargo size, and shape corresponds directly to the microswimmer geometries realized experimentally in [23]. With the payload, the optimal tail shape has a much greater radius (relative to the main swimming axis) and only about 3/4 of a turn, i.e., the tail must "fan out" from the payload in order to maximize its swimming speed. The additional drag caused by having the cargo greatly slows down the swimmer by a factor of 2. Moreover, the attachment point moves from being directly underneath the spherical cargo to along its side, i.e., the attachment point is affected by the optimization. Compared to the payload-free case, the optimization yields a greater (relative) decrease in the total cost, lowering it by a factor of 2.3. Therefore, introduction of the cargo significantly affects the optimal tail shape.

8. Conclusion. We have shown the existence of a minimizer to an infinite dimensional shape optimization problem related to microlocomotors in 3-D Stokes flow, which used a calculus of variations type framework while taking advantage of theoretical tools in self-contact of curves. We then described a variational gradient based optimization algorithm that uses shape differential calculus to obtain the functional sensitivities. Moreover, we proved well-posedness of a variational formulation used to compute descent directions for the locomotor centerline curve $\mathbf{X}$ (both for the continuous and discrete formulations). Last, we presented numerical results illustrating the effectiveness of our method.

We emphasize the importance of the generality of our method. It allows us to capture the complete optimal tail configuration and accommodate cargo with various shapes. In addition, our method readily incorporates more complex cargo geometries (such as a payload at both ends) which can dramatically alter the optimal Σ. Our approach also utilizes the complete Stokes flow problem to obtain the fluid forces experienced by the swimmers. Relying on approximations such as resistive-force theory [44] can lead to degenerate optimizing shapes [62] (i.e., helices with decreasing amplitude as the number of turns per unit length increases). Accordingly, we are now employing our method to further the results presented in the previous section and to determine experimentally realizable optimal swimmer shapes [39]. We are examining the dependence of the optimal shape and associated cost on the aspect ratio of the tail, thereby extending Lighthills results [44] to finite-sized swimmers. We are also applying our approach to assess the effects of cargo size on the results of the optimization.

Appendix A. Basic estimates.

A.1. Rigid motions. In the proof of Lemma 3.2, we will consider the decomposition of the rigid motion for the body Ω_{B}. Let $\{\eta_1, \eta_2, \eta_3, \eta_4\}$ be defined by

$$\eta_1 = \mathbf{e}_x, \quad \eta_2 = \mathbf{e}_y, \quad \eta_3 = \mathbf{e}_z, \quad \eta_4 = \mathbf{e}_z \times (\mathbf{x} - \mathbf{x}_g) \quad \text{for } \mathbf{x} \in \Gamma, \tag{A.1}$$

where $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$ are the canonical basis vectors of $\mathbb{R}^3$. Note that $\{\eta_i\}_{i=1}^4$ are orthogonal with respect to the $L^2(\Gamma)$ inner product. Moreover, note

$$\|\eta_4\|^2_{L^2(\Gamma)} = \int_\Gamma |\mathbf{e}_z \times (\mathbf{x} - \mathbf{x}_g)|^2 = \int_\Gamma (|\mathbf{x} - \mathbf{x}_g|^2 - |\mathbf{e}_z \cdot (\mathbf{x} - \mathbf{x}_g)|^2) = \mathrm{I}_z,$$

where I_z is the moment of inertia of Γ (about the z-axis) with respect to the geometric center $\mathbf{x}_g$ assuming the "shell mass density" is unity [49, 67]. The smallest that I_z can be is when Σ is a line segment. Thus, there is a constant $C_{\mathrm{I}_z} > 0$ depending only on a_{c} such that

$$\mathrm{I}_z \geq C_{\mathrm{I}_z} |\Sigma|,$$

provided Γ does not intersect itself. Therefore, we have

$$\begin{aligned} &\|\eta_i\|_{L^2(\Gamma)} = |\Gamma|^{1/2}, \ i = 1, 2, 3, \\ &\|\eta_4\|_{L^2(\Gamma)} = \left(\int_\Gamma |\mathbf{e}_z \times (\mathbf{x} - \mathbf{x}_g)|^2 \right)^{1/2} = \sqrt{\mathrm{I}_z} \geq \sqrt{C_{\mathrm{I}_z} |\Sigma|}. \end{aligned} \tag{A.2}$$

A.2. Proof of Lemma 3.2. *Proof.* Setting $\mathbf{v} = \mathbf{u}$ in (3.6) gives

$$\frac{1}{2} \int_\Omega D(\mathbf{u}) : D(\mathbf{u}) = \omega_{\mathrm{B,z}} \tau_{\mathrm{B,z}} = |\omega_{\mathrm{B,z}}| \|\eta_4\|_{L^2(\Gamma)} \frac{|\tau_{\mathrm{B,z}}|}{\|\eta_4\|_{L^2(\Gamma)}} \leq c_1 \|\omega_{\mathrm{B,z}} \eta_4\|_{L^2(\Gamma)} \frac{|\tau_{\mathrm{B,z}}|}{|\Sigma|^{1/2}}, \tag{A.3}$$

where we used (A.2), and c_1 is a constant dependent (at most) on a_{c}. Let $\widetilde{\mathbf{u}_{\mathrm{B}}} = \mathbf{u}_{\mathrm{B}} + \boldsymbol{\omega}_{\mathrm{B}} \times (\mathbf{x}_g - \mathbf{x}_c)$ (a spatially constant vector). Continuing, we have

$$\begin{aligned} \frac{1}{2} \int_\Omega D(\mathbf{u}) : D(\mathbf{u}) &\leq c_1 \left(\|\omega_{\mathrm{B,z}} \eta_4\|^2_{L^2(\Gamma)} + \|\widetilde{\mathbf{u}_{\mathrm{B}}}\|^2_{L^2(\Gamma)} \right)^{1/2} \frac{|\tau_{\mathrm{B,z}}|}{|\Sigma|^{1/2}} \\ &= c_1 \|\widetilde{\mathbf{u}_{\mathrm{B}}} + \omega_{\mathrm{B,z}} \eta_4\|_{L^2(\Gamma)} \frac{|\tau_{\mathrm{B,z}}|}{|\Sigma|^{1/2}}, \end{aligned} \tag{A.4}$$

where the equality follows by the orthogonality of $\{\eta_i\}_{i=1}^4$. Now note

$$\begin{aligned}\widetilde{\mathbf{u}_{\mathrm{B}}} + \omega_{\mathrm{B,z}}\eta_4 &= \mathbf{u}_{\mathrm{B}} + \boldsymbol{\omega}_{\mathrm{B}} \times (\mathbf{x}_g - \mathbf{x}_c) + \boldsymbol{\omega}_{\mathrm{B}} \times (\mathbf{x} - \mathbf{x}_g) \\ &= \mathbf{u}_{\mathrm{B}} + \boldsymbol{\omega}_{\mathrm{B}} \times (\mathbf{x} - \mathbf{x}_c) = \mathbf{u}\Big|_{\Gamma}.\end{aligned} \tag{A.5}$$

Thus, we obtain by (3.10) and a trace theorem

$$\begin{aligned}C\|\mathbf{u}\|_{H^1(\Omega)}^2 \le \frac{1}{2}\int_\Omega D(\mathbf{u}) : D(\mathbf{u}) \le c_1 \frac{|\tau_{\mathrm{B,z}}|}{|\Sigma|^{1/2}}\|\mathbf{u}\|_{L^2(\Gamma)} \le c_2 \frac{|\tau_{\mathrm{B,z}}|}{|\Sigma|^{1/2}}\|\mathbf{u}\|_{H^1(\Omega)} \\ \Rightarrow \|\mathbf{u}\|_{H^1(\Omega)} \le \frac{c_3}{|\Sigma|^{1/2}}|\tau_{\mathrm{B,z}}|,\end{aligned} \tag{A.6}$$

where c_3 only depends on a_{c} and Ω_{ALL}. The bound for the pressure follows similarly. The other inequality for the velocity follows by first solving an auxiliary problem: let $\mathbf{w}$ be the unique velocity solution of

$$\begin{aligned}-\nabla \cdot \boldsymbol{\sigma}(\mathbf{w}) = 0, \quad \nabla \cdot \mathbf{w} = 0 \text{ in } \Omega, \\ \mathbf{w} = \mathrm{sgn}(\omega_{\mathrm{B,z}} \cdot \tau_{\mathrm{B,z}})\boldsymbol{\omega}_{\mathrm{B}} \times (\mathbf{x} - \mathbf{x}_c) \text{ on } \Gamma, \quad \mathbf{w} = \mathbf{0} \text{ on } \Gamma_{\mathrm{O}},\end{aligned} \tag{A.7}$$

where $\boldsymbol{\omega}_{\mathrm{B}}$ comes from the solution $\mathbf{u}$ of (3.6). The PDE in (A.7) is a standard Stokes problem, so we have the following estimate [66, 59, 21]:

$$\begin{aligned}\|\mathbf{w}\|_{H^1(\Omega)} &\le c_4\|\boldsymbol{\omega}_{\mathrm{B}} \times (\mathbf{x} - \mathbf{x}_c)\|_{H^{1/2}(\Gamma)} = c_4|\omega_{\mathrm{B,z}}|\|\mathbf{e}_z \times (\mathbf{x} - \mathbf{x}_c)\|_{H^{1/2}(\Gamma)} \\ &\le c_4|\omega_{\mathrm{B,z}}|\,\mathrm{diam}(\Gamma)\|1\|_{H^{1/2}(\Gamma)} \le c_4|\omega_{\mathrm{B,z}}|\,|\Sigma|\,\|1\|_{H^{1/2}(\Gamma)} \le c_5|\omega_{\mathrm{B,z}}|,\end{aligned} \tag{A.8}$$

where c_5 only depends on Σ, Ω_{ALL}, and a_{c} (note that $|\Gamma| \approx |\Sigma|$ by (3.9)). Because $\mathbf{w}$ is in $\mathbb{V}_0$, we can set $\mathbf{v} = \mathbf{w}$ in (3.6) to get

$$\begin{aligned}|\omega_{\mathrm{B,z}}||\tau_{\mathrm{B,z}}| = \frac{1}{2}\int_\Omega D(\mathbf{u}) : D(\mathbf{w}) \le 2\|\mathbf{u}\|_{H^1(\Omega)}\|\mathbf{w}\|_{H^1(\Omega)} \le 2c_5\|\mathbf{u}\|_{H^1(\Omega)}|\omega_{\mathrm{B,z}}| \\ \Rightarrow \quad |\tau_{\mathrm{B,z}}| \le 2c_5\|\mathbf{u}\|_{H^1(\Omega)}.\end{aligned} \tag{A.9}$$

We have proved (3.11). A similar argument gives that $|\mathbf{u}_{\mathrm{B}}| + |\boldsymbol{\omega}_{\mathrm{B}}| \le c_6|\tau_{\mathrm{B,z}}|$. And the inequality $|\tau_{\mathrm{B,x}}| + |\tau_{\mathrm{B,y}}| \le c_7|\tau_{\mathrm{B,z}}|$ comes from the previous results and the definition $\boldsymbol{\tau}_{\mathrm{B}} := \int_\Gamma \mathbf{x} \times (\boldsymbol{\sigma}\boldsymbol{\nu})$. Ergo, we obtain (3.12). ∎

Appendix B. Shape sensitivity analysis. In deriving a gradient-based method for computing optimal solutions of (4.13), we use shape sensitivity calculations. We recall some basic concepts; details can be found in [15, 31, 46, 47, 52]. Let $\dot{f}(\Omega; \mathbf{V}_\Gamma)(\cdot)$ denote the Lagrangian material derivative of $f : \Omega \to \mathbb{R}$, where $\mathbf{V}_\Gamma$ denotes the velocity field perturbation in a neighborhood of Γ. The *shape* derivative of f (Eulerian partial derivative) is defined by [15, 31]

$$f'(\Omega; \mathbf{V}_\Gamma) := \dot{f}(\Omega; \mathbf{V}_\Gamma) - \nabla_{\mathbf{x}} f(\Omega) \cdot \mathbf{V}_\Gamma. \tag{B.1}$$

Note that the regularity of solution to the state equation (see Theorem 5.1) is sufficient to apply (B.1).

B.1. Shape derivative PDE and adjoints. Recall the cost functionals J_{ts}, J_{diss} given in (2.6), (2.7). Because of the simple form of J_{ts}, J_{diss}, we can compute the variational derivatives by the product rule:

$$\delta J_{\mathrm{ts}}(\Omega_{\mathrm{B}}; \mathbf{V}_\Gamma) = \mathbf{u}_{\mathrm{B}}' \cdot \mathbf{M}\boldsymbol{\tau}_{\mathrm{B}} + \boldsymbol{\tau}_{\mathrm{B}}' \cdot \mathbf{M}^T\mathbf{u}_{\mathrm{B}}, \qquad \delta J_{\mathrm{diss}}(\Omega_{\mathrm{B}}; \mathbf{V}_\Gamma) = \omega_{\mathrm{B,z}}'\tau_{\mathrm{B,z}}. \tag{B.2}$$

However, these formulas are not efficient to evaluate because they depend on the particular perturbation $\mathbf{V}_\Gamma$ used. Thus, we rewrite them using adjoints (recall Theorem 5.1).

B.1.1. Shape derivative PDE. The shape derivative PDE that $(\mathbf{u}', p')$ satisfies is given as follows:

(B.3)
$$\begin{aligned}
-\nabla\cdot\boldsymbol{\sigma}'(\mathbf{u},p) &= \mathbf{0} \text{ in } \Omega,\\
\nabla\cdot\mathbf{u}' &= 0 \text{ in } \Omega,\\
\mathbf{u}' &= \mathbf{u}'_{\mathrm{B}} + \boldsymbol{\omega}'_{\mathrm{B}}\times(\mathbf{x}-\mathbf{x}_c) - \boldsymbol{\omega}_{\mathrm{B}}\times\dot{\mathbf{x}}_c - (\mathbf{V}_\Gamma\cdot\nabla)\mathbf{u} + \boldsymbol{\omega}_{\mathrm{B}}\times\mathbf{V}_\Gamma \text{ on } \Gamma,\\
\mathbf{u}' &= \mathbf{0} \text{ on } \Gamma_{\mathrm{O}},\\
\int_\Gamma \boldsymbol{\sigma}'\boldsymbol{\nu} = \mathbf{0} =: \mathbf{f}'_{\mathrm{B}}, &\quad \int_\Gamma \mathbf{x}\times(\boldsymbol{\sigma}'\boldsymbol{\nu}) = \boldsymbol{\tau}'_{\mathrm{B}},
\end{aligned}$$

where the vectors $\mathbf{u}'_{\mathrm{B}}$, $\boldsymbol{\omega}'_{\mathrm{B}}$, and $\boldsymbol{\tau}'_{\mathrm{B}}$ have the form

(B.4)
$$\mathbf{u}'_{\mathrm{B}} = \begin{pmatrix} u'_{\mathrm{B,x}}\ (\text{unknown})\\ u'_{\mathrm{B,y}}\ (\text{unknown})\\ u'_{\mathrm{B,z}}\ (\text{unknown})\end{pmatrix},\ \boldsymbol{\omega}'_{\mathrm{B}} = \begin{pmatrix} 0\\ 0\\ \omega'_{\mathrm{B,z}}\ (\text{unknown})\end{pmatrix},$$
$$\boldsymbol{\tau}'_{\mathrm{B}} = \begin{pmatrix} \tau'_{\mathrm{B,x}}\ (\text{unknown})\\ \tau'_{\mathrm{B,y}}\ (\text{unknown})\\ 0\end{pmatrix},$$

and $\dot{\mathbf{x}}_c$ is given by (5.2). The derivation of (B.3) is as follows. One can transform (2.3) (by a rigid motion) to an equivalent Stokes problem where all of the nonzero boundary conditions are on the outer boundary. Similarly, one can take the net force and torque conditions to be on the outer boundary (by Gauss' divergence theorem). The advantage here is that the shape perturbation calculation (for perturbing $\Gamma \equiv \partial\Omega_{\mathrm{B}}$) is easier for the translated problem because the outer boundary is fixed. In particular, the perturbation of the normal vector in the net force and torque conditions does not appear. Finally, one maps the perturbed PDE system back using the inverse of the rigid motion transformation. This yields (B.3).

B.1.2. Adjoint problem for $\mathbf{u}'_{\mathrm{B}}\cdot\mathbf{M}\boldsymbol{\tau}_{\mathrm{B}}$. Using (5.3), integration by parts, and (B.3), we have

(B.5)
$$\begin{aligned}
0 &= -\int_\Omega (\nabla\cdot\mathbf{S})\cdot\mathbf{u}' = \int_\Omega \mathbf{S}:\nabla\mathbf{u}' - \int_\Gamma (\mathbf{S}\boldsymbol{\nu})\cdot\mathbf{u}' - \int_{\Gamma_{\mathrm{O}}} (\mathbf{S}\boldsymbol{\nu})\cdot\mathbf{u}'\\
&= \int_\Omega \nabla\mathbf{r}:\boldsymbol{\sigma}'(\mathbf{u}) - \int_\Gamma (\mathbf{S}\boldsymbol{\nu})\cdot\mathbf{u}'\\
&= -\int_\Omega \mathbf{r}\cdot(\nabla\cdot\boldsymbol{\sigma}'(\mathbf{u})) + \int_\Gamma \mathbf{r}\cdot(\boldsymbol{\sigma}'(\mathbf{u})\boldsymbol{\nu}) - \int_\Gamma (\mathbf{S}\boldsymbol{\nu})\cdot\mathbf{u}',
\end{aligned}$$

which simplifies further by (B.3) and (B.4) to give

$$
\begin{aligned}
0 &= \mathbf{r}_{\mathrm{B}} \cdot \int_{\Gamma} \boldsymbol{\sigma}'(\mathbf{u})\boldsymbol{\nu} + \int_{\Gamma} (\boldsymbol{\eta}_{\mathrm{B}} \times (\mathbf{x} - \mathbf{x}_c)) \cdot (\boldsymbol{\sigma}'(\mathbf{u})\boldsymbol{\nu}) - \int_{\Gamma} (\mathbf{S}\boldsymbol{\nu}) \cdot \mathbf{u}' \\
&= \boldsymbol{\eta}_{\mathrm{B}} \cdot \int_{\Gamma} (\mathbf{x} - \mathbf{x}_c) \times (\boldsymbol{\sigma}'(\mathbf{u})\boldsymbol{\nu}) - \int_{\Gamma} (\mathbf{S}\boldsymbol{\nu}) \cdot \mathbf{u}' = \underbrace{\boldsymbol{\eta}_{\mathrm{B}} \cdot \boldsymbol{\tau}_{\mathrm{B}}'}_{=0} - \int_{\Gamma} (\mathbf{S}\boldsymbol{\nu}) \cdot \mathbf{u}' \\
&= -\int_{\Gamma} (\mathbf{S}\boldsymbol{\nu}) \cdot (\mathbf{u}_{\mathrm{B}}' + \boldsymbol{\omega}_{\mathrm{B}}' \times (\mathbf{x} - \mathbf{x}_c) - \boldsymbol{\omega}_{\mathrm{B}} \times \dot{\mathbf{x}}_c - (\mathbf{V}_{\Gamma} \cdot \nabla)\mathbf{u} + \boldsymbol{\omega}_{\mathrm{B}} \times \mathbf{V}_{\Gamma}) \qquad \text{(B.6)} \\
&= -\mathbf{u}_{\mathrm{B}}' \cdot \mathbf{M}\boldsymbol{\tau}_{\mathrm{B}} - \underbrace{\boldsymbol{\omega}_{\mathrm{B}}' \cdot \boldsymbol{\xi}_{\mathrm{B}}}_{=0} + (\boldsymbol{\omega}_{\mathrm{B}} \times \dot{\mathbf{x}}_c) \cdot \mathbf{M}\boldsymbol{\tau}_{\mathrm{B}} + \int_{\Gamma} (\mathbf{S}\boldsymbol{\nu}) \\
&\quad \cdot [(\mathbf{V}_{\Gamma} \cdot \nabla)\mathbf{u}] - \int_{\Gamma} (\mathbf{S}\boldsymbol{\nu}) \cdot (\boldsymbol{\omega}_{\mathrm{B}} \times \mathbf{V}_{\Gamma}).
\end{aligned}
$$

Therefore, we have

$$
\mathbf{u}_{\mathrm{B}}' \cdot \mathbf{M}\boldsymbol{\tau}_{\mathrm{B}} = (\boldsymbol{\omega}_{\mathrm{B}} \times \dot{\mathbf{x}}_c) \cdot \mathbf{M}\boldsymbol{\tau}_{\mathrm{B}} + \int_{\Gamma} (\mathbf{S}\boldsymbol{\nu}) \cdot \{(\mathbf{V}_{\Gamma} \cdot \nabla)\mathbf{u} - (\boldsymbol{\omega}_{\mathrm{B}} \times \mathbf{V}_{\Gamma})\}. \tag{B.7}
$$

With further manipulation, we obtain

$$
(\mathbf{V}_{\Gamma} \cdot \nabla)\mathbf{u} - (\boldsymbol{\omega}_{\mathrm{B}} \times \mathbf{V}_{\Gamma}) = (\mathbf{V}_{\Gamma} \cdot \boldsymbol{\nu})[\mathbf{I} - \boldsymbol{\nu} \otimes \boldsymbol{\nu}]\boldsymbol{\sigma}\boldsymbol{\nu}. \tag{B.8}
$$

B.1.3. Adjoint problem for $\boldsymbol{\tau}_{\mathbf{B}}' \cdot \mathbf{M}^T\mathbf{u}_{\mathbf{B}}$. Let $(\mathbf{r}, \varrho)$ solve

$$
\begin{aligned}
-\nabla \cdot \mathbf{H}(\mathbf{r}, \varrho) &= \mathbf{0} \text{ in } \Omega, \\
\nabla \cdot \mathbf{r} &= 0 \text{ in } \Omega, \\
\mathbf{r} &= \mathbf{r}_{\mathrm{B}} + \boldsymbol{\eta}_{\mathrm{B}} \times (\mathbf{x} - \mathbf{x}_c) \text{ on } \Gamma, \\
\mathbf{r} &= \mathbf{0}, \text{ on } \Gamma_{\mathrm{O}}, \\
\int_{\Gamma} \mathbf{H}\boldsymbol{\nu} = \mathbf{0} =: \mathbf{g}_{\mathrm{B}} \text{ (given)}, &\quad \int_{\Gamma} (\mathbf{x} - \mathbf{x}_c) \times (\mathbf{H}\boldsymbol{\nu}) = \boldsymbol{\xi}_{\mathrm{B}},
\end{aligned} \tag{B.9}
$$

where $\mathbf{r}_{\mathrm{B}}$, $\boldsymbol{\eta}_{\mathrm{B}}$, and $\boldsymbol{\xi}_{\mathrm{B}}$ have the form

$$
\begin{aligned}
\mathbf{r}_{\mathrm{B}} &= \begin{pmatrix} r_{\mathrm{B,x}} \text{ (unknown)} \\ r_{\mathrm{B,y}} \text{ (unknown)} \\ r_{\mathrm{B,z}} \text{ (unknown)} \end{pmatrix}, \quad \boldsymbol{\eta}_{\mathrm{B}} = \begin{pmatrix} (\mathbf{M}^T\mathbf{u}_{\mathrm{B}}) \cdot \mathbf{e}_x \text{ (given)} \\ (\mathbf{M}^T\mathbf{u}_{\mathrm{B}}) \cdot \mathbf{e}_y \text{ (given)} \\ \bar{\eta} z \text{ (unknown)} \end{pmatrix}, \\
\boldsymbol{\xi}_{\mathrm{B}} &= \begin{pmatrix} \xi_{\mathrm{B,x}} \text{ (unknown)} \\ \xi_{\mathrm{B,y}} \text{ (unknown)} \\ 0 \end{pmatrix}.
\end{aligned} \tag{B.10}
$$

Note that for the case where $\mathbf{M}$ is such that $\mathbf{M}_{ij} = 0$, except $\mathbf{M}_{33} = 1$ (recall Definition 2.1), the solution of (B.9) vanishes, i.e., $(\mathbf{r}, \varrho) = (\mathbf{0}, 0)$ so that $\mathbf{H} = \mathbf{0}$. A similar derivation as in section B.1.2 gives

$$
\boldsymbol{\tau}_{\mathrm{B}}' \cdot \mathbf{M}^T\mathbf{u}_{\mathrm{B}} = -\int_{\Gamma} (\mathbf{H}\boldsymbol{\nu}) \cdot \{(\mathbf{V}_{\Gamma} \cdot \nabla)\mathbf{u} - (\boldsymbol{\omega}_{\mathrm{B}} \times \mathbf{V}_{\Gamma})\}. \tag{B.11}
$$

Combining (B.2), (B.7), (B.8), and (B.11) gives the first line of (5.1).

B.1.4. Adjoint problem for $\omega'_{\mathrm{B,z}}\tau_{\mathrm{B,z}}$. Let $(\mathbf{r}, \varrho)$ solve

$$\begin{aligned} -\nabla\cdot\mathbf{K}(\mathbf{r},\varrho) &= \mathbf{0} \text{ in } \Omega,\\ \nabla\cdot\mathbf{r} &= 0 \text{ in } \Omega,\\ \mathbf{r} &= \mathbf{r}_{\mathrm{B}} + \boldsymbol{\eta}_{\mathrm{B}}\times(\mathbf{x}-\mathbf{x}_c) \text{ on } \Gamma,\\ \mathbf{r} &= \mathbf{0} \text{ on } \Gamma_{\mathrm{O}},\\ \int_\Gamma \mathbf{K}\boldsymbol{\nu} = \mathbf{0} =: \mathbf{g}_{\mathrm{B}} \text{ (given)}, &\quad \int_\Gamma (\mathbf{x}-\mathbf{x}_c)\times(\mathbf{K}\boldsymbol{\nu}) = \tau_{\mathrm{B,z}}\mathbf{e}_z, \end{aligned} \tag{B.12}$$

where $\mathbf{r}_{\mathrm{B}}$, $\boldsymbol{\eta}_{\mathrm{B}}$, and $\boldsymbol{\xi}_{\mathrm{B}}$ have the form

$$\begin{aligned} \mathbf{r}_{\mathrm{B}} &= \begin{pmatrix} r_{\mathrm{B,x}}\ (\text{unknown})\\ r_{\mathrm{B,y}}\ (\text{unknown})\\ r_{\mathrm{B,z}}\ (\text{unknown}) \end{pmatrix},\ \boldsymbol{\eta}_{\mathrm{B}} = \begin{pmatrix} 0\\ 0\\ \bar{\eta}z\ (\text{unknown}) \end{pmatrix},\\ \boldsymbol{\xi}_{\mathrm{B}} &= \begin{pmatrix} \xi_{\mathrm{B,x}}\ (\text{unknown})\\ \xi_{\mathrm{B,y}}\ (\text{unknown})\\ 0 \end{pmatrix}. \end{aligned} \tag{B.13}$$

Just as before, we obtain

$$\omega'_{\mathrm{B,z}}\tau_{\mathrm{B,z}} = \int_\Gamma (\mathbf{K}\boldsymbol{\nu})\cdot\{(\mathbf{V}_\Gamma\cdot\nabla)\mathbf{u} - (\boldsymbol{\omega}_{\mathrm{B}}\times\mathbf{V}_\Gamma)\}. \tag{B.14}$$

Combining (B.2), (B.14) with (B.8) gives the second line of (5.1).

B.2. Mapping perturbations of $\boldsymbol{\Sigma}$ to $\boldsymbol{\Gamma}$. The previous sections (of this appendix) presented the sensitivity analysis for the case where $\mathbf{V}_\Gamma$ is a perturbation defined in the neighborhood of the surface Γ. Since the control variable (or optimization parameter) for our problem is the codimension 2 set Σ, we need to map perturbations of Σ to Γ to compute the sensitivities using (5.1). This is needed to compute the right-hand side of (5.16).

Let $\mathbf{V}$ be a parametric perturbation of $\mathbf{X}$ and define $\mathbf{X}_\epsilon = \mathbf{X} + \epsilon\mathbf{V}$. The perturbation of the tangent vector $\boldsymbol{\tau}$ is given by

$$\delta\boldsymbol{\tau} = \frac{1}{\|\mathbf{X}'(t)\|}(\mathbf{I}-\boldsymbol{\tau}\otimes\boldsymbol{\tau})\frac{d\mathbf{V}}{dt} = (\mathbf{I}-\boldsymbol{\tau}\otimes\boldsymbol{\tau})\,\partial_s\mathbf{V}. \tag{B.15}$$

Next, let $\delta\mathbf{N}_i$ be the corresponding perturbation of $\mathbf{N}_i$ (for $i=1,2$) and note that

$$\boldsymbol{\tau} = \mathbf{N}_1\times\mathbf{N}_2, \quad \mathbf{N}_1 = \mathbf{N}_2\times\boldsymbol{\tau}, \quad \mathbf{N}_2 = \boldsymbol{\tau}\times\mathbf{N}_1, \tag{B.16}$$

because $\{\boldsymbol{\tau}, \mathbf{N}_1, \mathbf{N}_2\}$ is an orthogonal frame. By the product rule and (B.16), we have

$$\delta\mathbf{N}_1 = \delta\mathbf{N}_2\times\boldsymbol{\tau} + \mathbf{N}_2\times\delta\boldsymbol{\tau}, \qquad \delta\mathbf{N}_2 = \delta\boldsymbol{\tau}\times\mathbf{N}_1 + \boldsymbol{\tau}\times\delta\mathbf{N}_1. \tag{B.17}$$

By the definition of the surface parametrization, the perturbation of Γ is

$$\mathbf{V}_\Gamma(t,\theta) = \mathbf{V}(t) + a_{\mathrm{c}}(t)\left(\cos\theta\delta\mathbf{N}_1(t) + \sin\theta\delta\mathbf{N}_2(t)\right). \tag{B.18}$$

But we can simplify this further because we assume a circular cross-section (recall (2.2)). In this case, the normal vector (on Γ) is given by

$$\boldsymbol{\nu}(t,\theta) = \alpha(t)(\cos\theta\mathbf{N}_1(t) + \sin\theta\mathbf{N}_2(t)) + \beta(t)\boldsymbol{\tau}(t), \tag{B.19}$$

where $\alpha^2 + \beta^2 = 1$. We now calculate $\mathbf{V}_\Gamma \cdot \boldsymbol{\nu}$ since this appears in (5.1). Combining (B.18) and (B.19), we obtain

$$\begin{aligned}\mathbf{V}_\Gamma \cdot \boldsymbol{\nu} &= \mathbf{V} \cdot \boldsymbol{\nu} + a_c(t) \left(\cos\theta \delta\mathbf{N}_1 + \sin\theta \delta\mathbf{N}_2\right) \cdot \boldsymbol{\nu} \\ &= \mathbf{V} \cdot \boldsymbol{\nu} + a_c(t)\alpha(t) \left(\cos\theta \sin\theta \delta\mathbf{N}_1 \cdot \mathbf{N}_2 + \cos\theta \sin\theta \delta\mathbf{N}_2 \cdot \mathbf{N}_1\right) \\ &\quad + a_c(t)\beta(t) \left(\cos\theta \delta\mathbf{N}_1 \cdot \boldsymbol{\tau} + \sin\theta \delta\mathbf{N}_2 \cdot \boldsymbol{\tau}\right),\end{aligned} \tag{B.20}$$

where we used the fact that $\delta\mathbf{N}_1 \cdot \mathbf{N}_1 = \delta\mathbf{N}_2 \cdot \mathbf{N}_2 = 0$. Now note the following identity:

$$\mathbf{N}_1 \cdot \mathbf{N}_2 = 0 \quad \Rightarrow \quad \delta\mathbf{N}_1 \cdot \mathbf{N}_2 + \mathbf{N}_1 \cdot \delta\mathbf{N}_2 = 0.$$

Hence, (B.20) reduces to

$$\begin{aligned}\mathbf{V}_\Gamma \cdot \boldsymbol{\nu} &= \mathbf{V} \cdot \boldsymbol{\nu} + a_c(t)\beta(t) \left(\cos\theta (\mathbf{N}_2 \times \delta\boldsymbol{\tau}) \cdot \boldsymbol{\tau} + \sin\theta (\delta\boldsymbol{\tau} \times \mathbf{N}_1) \cdot \boldsymbol{\tau}\right) \\ &= \mathbf{V} \cdot \boldsymbol{\nu} + a_c(t) \left(\cos\theta (\mathbf{N}_2 \times \delta\boldsymbol{\tau}) + \sin\theta (\delta\boldsymbol{\tau} \times \mathbf{N}_1)\right) \cdot \boldsymbol{\nu}.\end{aligned} \tag{B.21}$$

Therefore, the perturbation of Γ can be written as (B.18) with $\delta\mathbf{N}_1$, $\delta\mathbf{N}_2$ replaced by

$$\delta\mathbf{N}_1 = \mathbf{N}_2 \times \delta\boldsymbol{\tau}, \qquad \delta\mathbf{N}_2 = \delta\boldsymbol{\tau} \times \mathbf{N}_1, \tag{B.22}$$

i.e., $\delta\mathbf{N}_1$, $\delta\mathbf{N}_2$ are *completely* determined by $\delta\boldsymbol{\tau}$. So $\mathbf{V}_\Gamma$ is fully determined from $\mathbf{V}$.

Acknowledgment. The authors thank Professor Michael Shelley for valuable and insightful discussions during the course of this work.

REFERENCES

[1] R. A. Adams and J. J. F. Fournier, *Sobolev Spaces*, Pure Appl. Math. (N.Y.) 140, 2nd ed., Elsevier, New York, 2003.
[2] F. Alouges, A. DeSimone, and A. Lefebvre, *Optimal strokes for low Reynolds number swimmers: An example*, J. Nonlinear Sci., 18 (2008), pp. 277–302.
[3] F. Alouges, A. Desimone, and L. Heltai, *Numerical strategies for stroke optimization of axisymmetric microswimmers*, Math. Models Methods Appl. Sci., 21 (2011), pp. 361–387.
[4] S. S. Antman, *Nonlinear Problems of Elasticity*, 2nd ed., Springer, New York, 2005.
[5] R. L. Bishop, *There is more than one way to frame a curve*, Amer. Math. Monthly, 82 (1975), pp. 246–251.
[6] D. Braess, *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics*, 2nd ed., Cambridge University Press, Cambridge, 2001.
[7] S. C. Brenner and L. Ridgway Scott, *The Mathematical Theory of Finite Element Methods*, 2nd ed., Springer, New York, 2002.
[8] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer-Verlag, New York, 1991.
[9] J. Burdick, R. Laocharoensuk, P. M. Wheat, J. D. Posner, and J. Wang, *Synthetic nanomotors in microchannel networks: Directional microchip motion and controlled manipulation of cargo*, J. Amer. Chem. Soc., 130 (2008), pp. 8164–8165.
[10] M. Burger, *A framework for the construction of level set methods for shape optimization and reconstruction*, Interfaces Free Bound., 5 (2002), pp. 301–329.
[11] D. Chenais and B. Rousselet, *Dependence of the buckling load of a nonshallow arch with respect to the shape of its midcurve*, RAIRO Model. Math. Anal. Numer., 24 (1990), pp. 307–341.
[12] D. Chenais, B. Rousselet, and R. Benedict, *Design sensitivity for arch structures with respect to midsurface shape under static loading*, J. Optim. Theory Appl., 58 (1988), pp. 225–239.
[13] S. Childress, *Mechanics of Swimming and Flying*, Cambridge University Press, Cambridge, 1981.
[14] M. Chipot, *Elements of Nonlinear Analysis*, Birkhäuser, Basel, Switzerland, 2000.

[15] M. C. Delfour and J.-P. Zolésio, *Shapes and Geometries: Analysis, Differential Calculus, and Optimization*, Adv. Des. Control 4, SIAM, Philadelphia, 2001.
[16] M. P. do Carmo, *Differential Geometry of Curves and Surfaces*, Prentice-Hall, Upper Saddle River, NJ, 1976.
[17] C. Dombrowski, L. Cisneros, S. Chatkaew, R. E. Goldstein, and J. O. Kessler, *Self-concentration and large-scale coherence in bacterial dynamics*, Phys. Rev. Lett., 93 (2004), 098103.
[18] G. Doğan, P. Morin, and R. H. Nochetto, *A variational shape optimization approach for image segmentation with a Mumford-Shah functional*, SIAM J. Sci. Comput., 30 (2008), pp. 3028–3049.
[19] R. Dreyfus, J. Baudry, M. L. Roper, M. Fermigier, H. A. Stone, and J. Bibette, *Microscopic artificial swimmers*, Nature, 437 (2005), pp. 862–865.
[20] G. Duvaut and J. L. Lions, *Inequalities in Mechanics and Physics*, Springer, New York, 1976.
[21] L. C. Evans, *Partial Differential Equations*, AMS, Providence, RI, 1998.
[22] N. Fujii, *Lower semicontinuity in domain optimization problems*, J. Optim. Theory Appl., 59 (1988), pp. 407–422.
[23] A. Ghosh and P. Fischer, *Controlled propulsion of artificial magnetic nanostructured propellers*, Nano Lett., 9 (2009), pp. 2243–2245.
[24] M. Giaquinta and S. Hildebrandt, *Calculus of Variations* II: *The Hamiltonian Formalism*, Grundlehren Math. Wiss. 311, Springer-Verlag, New York, 1996.
[25] V. Girault and P. A. Raviart, *Finite Element Methods for Navier-Stokes Equations: Theory and Algorithms*, Springer-Verlag, Berlin, 1986.
[26] O. Gonzalez, J. H. Maddocks, F. Schuricht, and H. von der Mosel, *Global curvature and self-contact of nonlinearly elastic curves and rods*, Calc. Var., 14 (2002), pp. 29–68.
[27] M. D. Gunzburger, *Perspectives in Flow Control and Optimization*, SIAM, Philadelphia, 2003.
[28] M. D. Gunzburger and H. Kim, *Existence of an optimal solution of a shape control problem for the stationary Navier-Stokes equations*, SIAM J. Control Optim., 36 (1998), pp. 895–909.
[29] A. J. Hanson and H. Ma, *Parallel Transport Approach to Curve Framing*, Technical report, Indiana University, Bloomington, IN, 1995.
[30] J. Happel and H. Brenner, *Low Reynolds Number Hydrodynamics: With Special Applications to Particulate Media*, Prentice-Hall, Upper Saddle River, NJ, 1965.
[31] J. Haslinger and R. A. E. Mäkinen, *Introduction to Shape Optimization: Theory, Approximation, and Computation*, Adv. Des. Control 7, SIAM, Philadelphia, 2003.
[32] M. Hintermüller, K. Ito, and K. Kunisch, *The primal-dual active set strategy as a semismooth newton method*, SIAM J. Optim., 13 (2003), pp. 865–888.
[33] M. Hintermüller and K. Kunisch, *Path-following methods for a class of constrained minimization problems in function space*, SIAM J. Optim., 17 (2006), pp. 159–187.
[34] M. Hintermüller and W. Ring, *A second order shape optimization approach for image segmentation*, SIAM J. Appl. Math., 64 (2004), pp. 442–467.
[35] K. Ito and K. Kunisch, *Lagrange Multiplier Approach to Variational Problems and Applications*, Adv. Des. Control, SIAM, Philadelphia, 2008.
[36] J. Jost and X. Li-Jost, *Calculus of Variations*, Cambridge University Press, Cambridge, 1998.
[37] E. E. Keaveny and M. R. Maxey, *Spiral swimming of an artificial micro-swimmer*, J. Fluid Mech., 598 (2008), pp. 293–319.
[38] E. E. Keaveny and M. J. Shelley, *Applying a second-kind boundary integral equation for surface tractions in Stokes flow*, J. Comput. Phys., 230 (2011), pp. 2141–2159.
[39] E. E. Keaveny, S. W. Walker, and M. J. Shelley, *Optimization of chiral structures for microscale propulsion*, Nano Lett., 13 (2013), pp. 531–537.
[40] S. Kim and S. J. Karrila, *Microhydrodynamics: Principles and Selected Applications*, Dover Publications, New York, 2005.
[41] S. Lang, *Real and Functional Analysis*, Grad. Texts Math., 3rd ed., Springer, New York, 1993.
[42] E. Lauga and T. R. Powers, *The hydrodynamics of swimming microorganisms*, Rep. Progr. Phys., 72 (2009), 096601.
[43] P. D. Lax, *Functional Analysis*, Wiley-Interscience, New York, 2002.
[44] J. Lighthill, *Mathematical Biofluiddynamics*, SIAM, Philadelphia, 1975.
[45] E. J. Lobaton and A. M. Bayen, *Modeling and optimization analysis of a single-flagellum micro-structure through the method of regularized stokeslets*, IEEE Trans. Control Systems Technology, 10 (2008), pp. 1–8.
[46] B. Mohammadi and O. Pironneau, *Applied Shape Optimization for Fluids*, Numer. Math. Sci. Comput., Oxford University Press, New York, 2001.

[47] M. Moubachir and J.-P. Zolésio, *Moving shape analysis and control: Applications to fluid structure interactions*, Pure Appl. Math. 277, Chapman and Hall/CRC, Boca Raton, FL, 2006.
[48] P. Neittaanmäki, J. Sprekels, and D. Tiba, *Optimization of Elliptic Systems: Theory and Applications*, Springer, New York, 2006.
[49] O. M. O'Reilly, *Engineering Dynamics: A Primer*, Springer, New York, 2001.
[50] G. A. Ozin, I. Manners, S. Fournier-Bidoz, and A. Arsenault, *Dream nanomachines*, Advanced Materials, 17 (2005), pp. 3011–3018.
[51] W. F. Paxton, K. C. Kistler, C. C. Olmeda, A. Sen, S. K. St. Angelo, Y. Cao, T. E. Mallouk, P. E. Lammert, and V. H. Crespi, *Catalytic nanomotors: Autonomous movement of striped nanorods*, J. Amer. Chem. Soc., 126 (2004), pp. 13424–13431.
[52] O. Pironneau, *Optimal Shape Design for Elliptic Systems*, Springer Ser. Comput. Phys., Springer-Verlag, New York, 1984.
[53] O. Pironneau and D. F. Katz, *Optimal swimming of flagellated microorganisms*, J. Fluid Mech., 66 (1974), pp. 391–415.
[54] E. M. Purcell, *Life at low Reynolds number*, Amer. J. Phys., 45 (1977), pp. 3–11.
[55] A. Quarteroni, R. Sacco, and F. Saleri, *Numerical Mathematics*, Texts Appl. Math. 37, Springer, New York, 2000.
[56] D. Saintillan and M. J. Shelley, *Instabilities and pattern formation in active particle suspensions: Kinetic theory and continuum simulations*, Phys. Rev. Lett., 100 (2008), 178103.
[57] D. Saintillan and M. J. Shelley, *Instabilities, pattern formation, and mixing in active suspensions*, Phys. Fluids, 20 (2008), 123304.
[58] L. L. Schumaker, *Spline Functions: Basic Theory*, 3rd ed., Cambridge Math. Lib., Cambridge University Press, Cambridge, 2007.
[59] H. Sohr, *The Navier-Stokes Equations: An Elementary Functional Analytic Approach*, Birkhäuser Advanced Texts, Birkhäuser-Verlag, Basel, Switzerland, 2001.
[60] A. Sokolov, R. E. Goldstein, F. I. Feldchtein, and I. S. Aranson, *Enhanced mixing and spatial instability in concentrated bacterial suspensions*, Phys. Rev. E, 80 (2009), 031903.
[61] J. Sokolowski and J.-P. Zolésio, *Introduction to Shape Optimization*, Springer Ser. Comput. Math., Springer-Verlag, New York, 1992.
[62] S. E. Spagnolie and E. Lauga, *The optimal elastic flagellum*, Phys. Fluids, 22 (2010), 031901.
[63] J. Stoer and R. Bulirsch, *Introduction to Numerical Analysis*, Texts Appl. Math. 12, 3rd ed., Springer, New York, 2002.
[64] S. Sundararajan, P. E. Lammert, A. W. Zudans, V. H. Crespi, and A. Sen, *Catalytic motors for transport of colloidal cargo*, Nano Lett., 8 (2008), pp. 1271–1276.
[65] D. Tam and A. Hosoi, *Optimal stroke patterns for Purcell's three-link swimmer*, Phys. Rev. Lett., 98 (2007), 068105.
[66] R. Temam, *Navier-Stokes Equations: Theory and Numerical Analysis*, AMS Chelsea, Providence, RI, 2001.
[67] C. A. Truesdell, *A First Course in Rational Continuum Mechanics*, Pure Appl. Math., a Series of Monographs and Textbooks, Academic Press, New York, 1976.
[68] S. W. Walker and M. J. Shelley, *Shape optimization of peristaltic pumping*, J. Comput. Phys., 229 (2010), pp. 1260–1291.
[69] J. Wilkening and A. Hosoi, *Shape optimization of a sheet swimming over a thin liquid layer*, J. Fluid Mech., 601 (2008), pp. 25–61.
[70] L. Zhang, J. J. Abbott, L. Dong, B. E. Kratochvil, D. Bell, and B. J. Nelson, *Artificial bacterial flagella: Fabrication and magnetic control*, Appl. Phys. Lett., 94 (2009), 064107.