Optimal Control of a Soft CyberOctopus Arm

Tixian Wang[1,2], Udit Halder[2], Heng-Sheng Chang[1,2], Mattia Gazzola[1,3,4], Prashant G. Mehta[1,2]

Abstract — In this paper, we use the optimal control methodology to control a flexible, elastic Cosserat rod. An inspiration comes from stereotypical movement patterns in octopus arms, which are observed in a variety of manipulation tasks, such as reaching or fetching. To help uncover the mechanisms underlying these observed morphologies, we outline an optimal control-based framework. A single octopus arm is modeled as a Hamiltonian control system, where the continuum mechanics of the arm is modeled after the Cosserat rod theory, and internal, distributed muscle forces and couples are considered as controls. First order necessary optimality conditions are derived for an optimal control problem formulated for this infinite dimensional system. Solutions to this problem are obtained numerically by an iterative forward-backward algorithm. The state and adjoint equations are solved in a dynamic simulation environment, setting the stage for studying a broader class of optimal control problems. Trajectories that minimize control effort are demonstrated and qualitatively compared with experimentally observed behaviors.

Index Terms — Cosserat rod, optimal control, maximum principle, soft robotics, octopus, Hamiltonian systems

I. INTRODUCTION

A. Background and Objectives

Over the past few decades, the optimal control paradigm has been increasingly used to explain and understand dynamic phenomena in biological systems. Examples range from game theoretic models of population dynamics [1], [2] to testing optimality hypotheses for collective motion in starling murmurations [3]–[5], or the minimum-jerk hypothesis for movement planning [6]–[9]. Through a mixture of experimental data analysis and theoretical modeling, these approaches often reveal deep insights into the underlying mechanisms at play [9], [10]. In this work, we take a similar route to examine the problem of octopus arm movement.

Flexible octopus arms are excellent candidates for studying the intricate interplay between continuum mechanics and sensorimotor control. As opposed to articulated limbs in humans, octopus arms are soft and possess a complex muscular architecture that provides exquisite manipulation

control. The muscles are independently innervated by motor neurons along the arm enabling a rich repertoire of deformations – stretch, shear, bend, and twist. However, despite their virtually infinite degrees of freedom – and thus having many options to carry out a single task – octopuses are observed to engage in certain (task-specific) stereotypical movement strategies. In experimental studies [11]–[13], these strategies are broadly categorized into two groups.

(i) Reaching pattern – bend propagation: For the task of reaching to a fixed target (Fig. 1a), the arm creates a bend at the base of the arm and propagates that bend toward the tip [11]. It was later showed that these waves are not mere whip-like mechanical waves [14], [15] due to the flexible arm structure, rather the bend propagation is achieved by actively creating waves of muscle activation signals [12]. Electromyogram (EMG) recordings of muscle activation reveals association of muscle contraction with the traveling bend. Ex-vivo experiments seem to suggest that these movement patterns may actually be encoded in the neural circuitry of the arm itself [16].

(ii) Fetching pattern – creation of pseudo-joints: The octopus typically employs a different strategy for the scenario of fetching food to its mouth. In this case, the arm behaves like an articulated limb [13], [17] (see Fig. 1a), creating dynamic pseudo-joints at three locations along the arm – proximal, medial, and distal. The medial joint is the most prominent one, and forms at the location where two waves of propagating muscle activation collide.

The objective of the present paper is to introduce an optimal control framework, associated numerical algorithms, and software tools to systematically investigate potential optimality bases of these stereotypical movement strategies. We are particularly interested in understanding the traveling wave phenomena observed in experimental studies. The framework introduced here is seen as a first step towards an inverse optimality analysis of the observed behaviors.

B. Contributions

The dynamics of a soft arm are modeled using the Cosserat rod theory [18]–[20]. Internal muscle forces and couples, when considered as control inputs, give rise to a control system in an infinite-dimensional state space setting. Since the observed stereotypical arm movements occur primarily in-plane [11], we restrict our modeling to planar settings, leading to a control system described by six nonlinear PDEs. We propose an optimal control problem associated with this control system. The Pontryagin's Maximum Principle (PMP) is used to derive the six adjoint PDEs for the costate
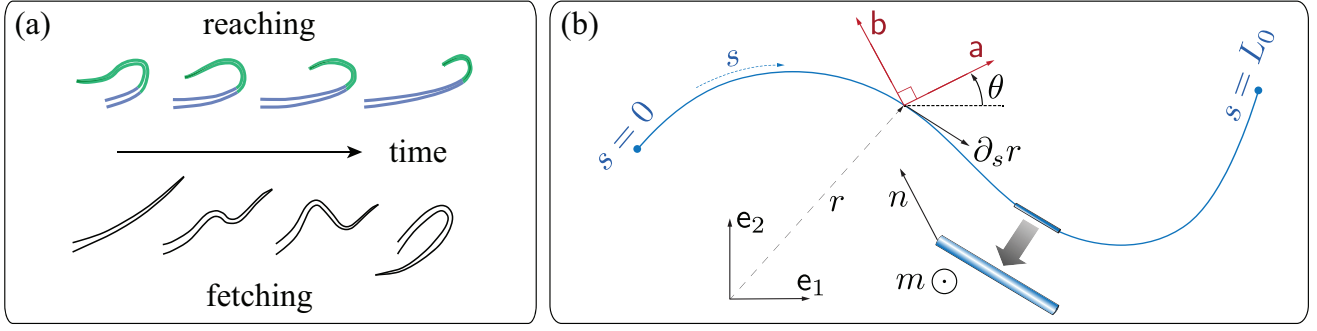
Fig. 1: (a) The octopuses have been observed to exhibit bend propagation (for reaching) and elbow formation (for fetching). The bend propagation is actively achieved by propagating muscle actuation, illustrated by blue color; green represents the unactuated portion of the arm. (b) A schematic of the planar Cosserat rod model.

variables. The PMP is also used to obtain the (open-loop) optimal control input.

The resulting two-point boundary value problem is numerically solved in an iterative manner, referred to here as the forward-backward algorithm. The forward path, or the Cosserat dynamical equations are solved using the existing software tool *Elastica* [19], [21], [22]. A custom solver is implemented to simulate the backward path or the costate equations. The deviation from optimality is utilized to adjust the control in an iterative manner so as to achieve optimality.

The numerical solver is applied to three test cases related to the reaching and the fetching movement patterns. Simulation results are used to qualitatively compare with observed wave propagations or elbow forming.

C. Paper Outline

The remainder of this paper is organized as follows: In Sec. II, the Cosserat rod model and dynamics in the planar case are introduced and an optimal control problem is formulated. The solution to the optimal control problem, including the forward-backward algorithm and the numerical methods are described in Sec. III. Results of numerical experiments appear in Sec. IV. The paper is concluded in Sec. V.

II. PROBLEM FORMULATION

A. Dynamic modeling of an arm as a Cosserat rod

Let $\{e_1, e_2\}$ denote a fixed orthonormal basis for the two-dimensional laboratory frame. Time $t \in \mathbb{R}$ and arc-length $s \in [0, L_0]$, L_0 being the length of the undeformed rod, represent the two independent variables. The partial derivatives with respect to t and s will be denoted by the subscripts $(\cdot)_t$ and $(\cdot)_s$, respectively.

The *state* of the rod is described by the vector-valued function $q(t, s) = (r(t, s), \theta(t, s))$ where $r = (x, y) \in \mathbb{R}^2$ denotes the position vector of the centerline, and the angle $\theta \in \mathbb{R}$ defines the material frame spanned by the orthonormal pairs $\{a, b\}$, where $a = \cos\theta\, e_1 + \sin\theta\, e_2$, $b = -\sin\theta\, e_1 + \cos\theta\, e_2$ (see Fig. 1b). The vector a is normal to the cross section. The deformations $w = (\nu_1, \nu_2, \kappa)$, stretch, shear, and curvature, are related to the local frame $\{a, b\}$ through $r_s = \nu_1 a + \nu_2 b$ and $\theta_s = \kappa$. Finally, $p(t, s) = \mathcal{M} q_t(t, s)$

is used to denote the momentum variable where $\mathcal{M}$ is the mass-inertia density matrix.

The Hamiltonian formulation requires specification of the kinetic energy $\mathcal{T}$ and the potential energy $\mathcal{V}$ of the rod as follows:

$$\mathcal{T}(p) = \frac{1}{2}\int_0^{L_0} p^\mathsf{T} \mathcal{M}^{-1} p \; \mathrm{d}s, \quad \mathcal{V}(q) = \int_0^{L_0} W(w)\, \mathrm{d}s$$

where $W : w \mapsto \mathbb{R}$ is referred to as the stored energy function of the rod. A quadratic stored energy function, which leads to a linear stress-strain relationship, is used in this work. The total energy function or the Hamiltonian $\mathcal{H}(q, p) := \mathcal{T}(p) + \mathcal{V}(q)$ yields the Hamilton's equations of the rod dynamics in the classical Cosserat theory [18], [20].

The *generalized state* of the rod is denoted as

$$z(t) := (q(t, \cdot), p(t, \cdot)) \in \mathrm{Z}, \; t \in [0, T]$$

An appropriate choice of function space is $\mathrm{Z} = \mathrm{H}^1([0, L_0]; \mathbb{R}^3) \times \mathrm{L}^2([0, L_0]; \mathbb{R}^3)$ equipped with the appropriate boundary conditions. The dynamics of the Hamiltonian control system are expressed as follows:

$$\frac{\mathrm{d}z}{\mathrm{d}t}(t) = (J - R)\frac{\delta\mathcal{H}}{\delta z} + \mathsf{G}(z(t))u(t) =: f(z(t), u(t)) \quad (1)$$

where $z(0)$ is the initial condition, J is the skew-symmetric structure matrix $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, and $R = \begin{pmatrix} 0 & 0 \\ 0 & \zeta 1 \end{pmatrix}$ is the dissipation matrix, $\zeta > 0$ is a damping coefficient, modeling viscoelastic effects in the rod [19]. The term $\mathsf{G}(z(t))u(t)$ on the right hand side is used to model the effect of the distributed internal muscle forces and couples. The functions $u(\cdot) \in \mathfrak{U}$ are called control inputs. Here $\mathfrak{U}$ is the set of all measurable functions $u(\cdot) : [0, T] \to \mathrm{U}$, where U is a suitable function space called the control space. We take this as the $\mathrm{L}^2([0, L_0]; \mathbb{R}^3)$ space. The modeling of G is complicated and depends on the muscle type details of the octopus. In this paper, we make the simplifying assumption $\mathsf{G}(z(t)) \equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

The explicit form of the six partial differential equations in the model (1) appears in Appendix I.

B. An optimal control problem

Both sterotypical movement patterns introduced in Sec. I involve reaching a given target point $q^{\text{target}} \in \mathbb{R}^3$. Even if realistic muscle constraints were considered (they are ignored

here), there would exist a large number of potential strategies to achieve the objective. Optimal control appears to be a natural choice to obtain a unique strategy. This is done through formulating the following free endpoint optimal control problem:

$$\underset{u}{\text{minimize}} \ \mathcal{J}(u) = \int_0^T \mathcal{L}(z(t), u(t)) \, \mathrm{d}t + \Phi(z(T)) \tag{2}$$
$$\text{subject to (1) and a given } z(0, s)$$

Here the end point $z(T) = (q(T), p(T))$ is free and penalizes the cost Φ associated with the underlying task, for example the distance from the arm tip to the designated target point. Note that a free endpoint problem is considered as opposed to a fixed endpoint problem due to the ease in algorithmic implementation as described in Sec. III-B.

The choice of the cost function is problem dependent. In this paper, a quadratic model is assumed for the control cost and the elastic potential energy is assumed for the state-dependent cost

$$\mathcal{L}(z, u) = \frac{1}{2}\|u\|_{\mathrm{L}^2}^2 + \chi_1 \mathcal{V}(q) \tag{3}$$

where the weighting parameter $\chi_1 > 0$ is used to penalize the deformation of the arm. The terminal cost is used in place of a fixed endpoint constraint

$$\Phi(z(T)) = \chi_2 \Phi_{\text{tip}}(q(T, L_0), q^{\text{target}}) \tag{4}$$

where the function Φ_{tip} measures the distance between the arm tip and the target point q^{target}, and $\chi_2 > 0$ is a suitably chosen regularization parameter.

Remark 1: Careful analysis is needed regarding the controllability aspect of this infinite dimensional system. The Lie algebra rank condition or otherwise known as the Chow-Rashevsky theorem for finite dimensional systems [23]–[25] typically does not hold for infinite dimensional systems, and one needs additional assumptions, e.g. [26], [27]. Moreover, existence of the first order Pontryagin's Maximum Principle (PMP) type optimality conditions in the infinite dimensional settings is non-trivial. A few attempts have been made to show generalized PMP conditions for infinite dimensional systems with additional assumptions [5], [28], [29]. However, the scope of this paper is not to address these questions, rather to characterize optimal trajectories for a soft arm manipulation task, in a quest to explain experimentally observed behaviors. We will therefore proceed assuming that the controllability and PMP optimality conditions hold.

III. OPTIMAL CONTROL SOLUTION

A. The maximum principle

The *costate* is denoted as $\xi(t) := (\mu(t), \gamma(t)) \in \mathrm{Z}^*, \ t \in [0, T]$. The control Hamiltonian function[1] $H : \mathrm{Z} \times \mathrm{U} \times \mathbb{R} \times \mathrm{Z}^* \to \mathbb{R}$ is defined as

$$H(z(t), u(t), \xi_0, \xi(t)) := \xi_0 \mathcal{L}(z(t), u(t))$$
$$+ \langle \xi(t), f(z(t), u(t)) \rangle \tag{5}$$

[1]Notice the difference between the Hamiltonian function H in the optimal control theory and the Hamiltonian $\mathcal{H}$ in the elastic rod theory.

The Hamilton's equations in the infinite-dimensional settings are as follows:

Proposition 3.1 (Maximum Principle [5], [28]): Let $\bar{u} \in \mathfrak{U}$ be an optimal control for problem (2) and $\bar{z}(t)$ be the corresponding optimal trajectory. Then, there exists a pair $(\bar{\xi}_0, \bar{\xi}(t)) \in \mathbb{R} \times \mathrm{Z}^*, \ t \in [0, T]$, such that $(\bar{\xi}_0, \bar{\xi}) \not\equiv 0, \bar{\xi}_0 \le 0$, $\bar{\xi}$ satisfies the differential equation

$$\frac{\mathrm{d}\bar{\xi}}{\mathrm{d}t}(t) = -\left(\frac{\delta f}{\delta z}\right)^\dagger (\bar{z}(t), \bar{u}(t)) \, \bar{\xi}(t) - \bar{\xi}_0 \frac{\delta \mathcal{L}}{\delta z}(\bar{z}(t), \bar{u}(t)) \tag{6}$$

where $(\cdot)^\dagger$ denotes the adjoint operator. The pointwise maximization of the pre-Hamiltonian holds, i.e.

$$H(\bar{z}(t), \bar{u}(t), \bar{\xi}_0, \bar{\xi}(t)) \ge H(\bar{z}(t), v, \bar{\xi}_0, \bar{\xi}(t)) \tag{7}$$

for all $v \in \mathrm{U}$ and for all $t \in [0, T]$. Moreover, $\bar{z}$ and $\bar{\xi}$ satisfy Hamilton's canonical equations

$$\frac{\mathrm{d}\bar{z}}{\mathrm{d}t}(t) = \frac{\delta H}{\delta \xi}(\bar{z}(t), \bar{u}(t), \bar{\xi}_0, \bar{\xi}(t))$$
$$\frac{\mathrm{d}\bar{\xi}}{\mathrm{d}t}(t) = -\frac{\delta H}{\delta z}(\bar{z}(t), \bar{u}(t), \bar{\xi}_0, \bar{\xi}(t)) \tag{8}$$

Furthermore, the vector $\bar{\xi}(T)$ satisfies the transversality condition

$$\bar{\xi}(T) = -\frac{\delta \Phi}{\delta z}(\bar{z}(T)) \tag{9}$$

In the remainder of this paper, we will restrict ourselves in studying only the normal extremals, i.e. where $\bar{\xi}_0 \ne 0$ and can be normalized to -1. The explicit form of the Hamilton's equations as a set of six (forward) PDEs and six (adjoint) PDEs appears in Appendix I.

B. Computing optimal control – the forward-backward algorithm

A solution to the optimal control problem (2) necessarily has to satisfy the PMP conditions (7), (8), and (9). This calls for solving the resulting two point boundary value problem in a function space. This is a challenging task even for a finite-dimensional nonlinear problem, for which various numerical techniques have been proposed [30]–[32].

An alternate approach is to employ an iterative algorithm (here referred to as forward-backward algorithm) to compute the optimal control. The idea is to start with an initial guess of the control $u^{(1)}$ in the first iteration. (This guess may be zero.) In each subsequent iteration, the control is modified so as to achieve the maximization of the control Hamiltonian H [33], [34].

Suppose the state, costate and control at iteration k is denoted as $z^{(k)}, \xi^{(k)}$, and $u^{(k)}$, respectively. At k-th iteration the steps of this algorithm are as follows:

1) *Run forward path:* The state equation (1) is integrated forward in time from $t = 0$ to T, to obtain the state $z^{(k)}$.
2) *Calculate terminal condition of the costate* from the transversality condition (9).
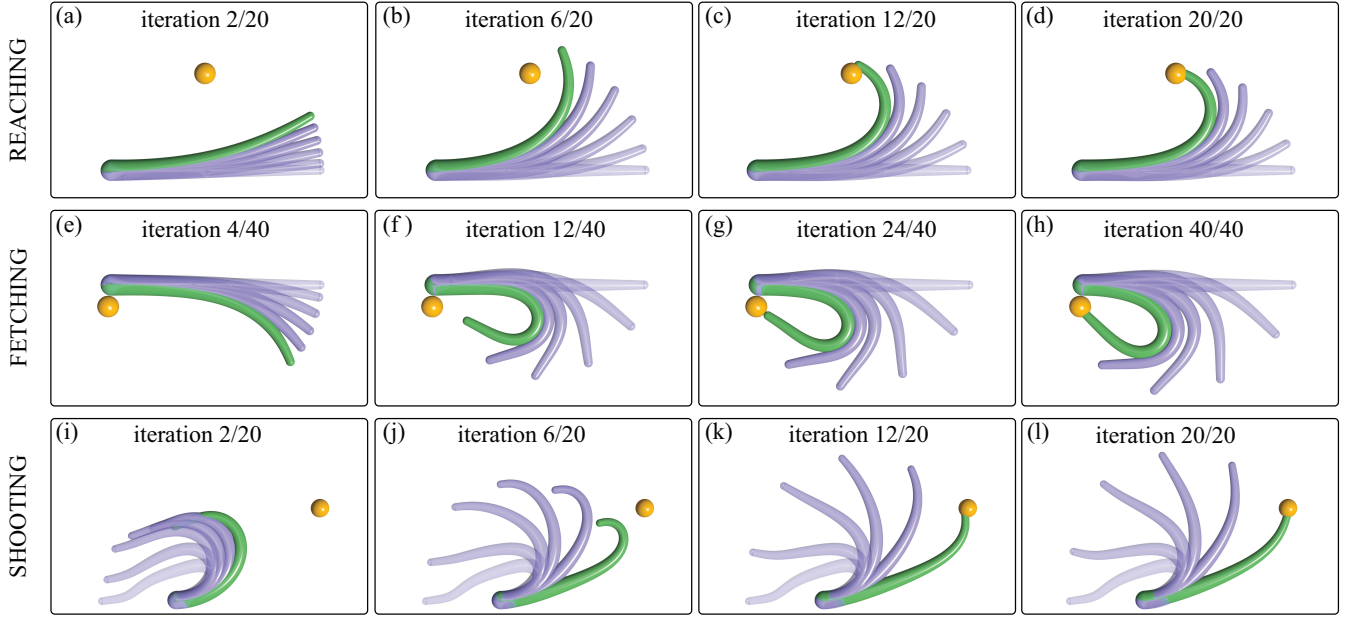
Fig. 2: Summary of the numerical experiments: We select four iterations for each experiment. Six time instances, including the initial time $t = 0$ and the terminal time $t = T$, are illustrated for each iteration. The rod at the terminal time is depicted in green while other time instances are depicted in fade-in purple. The target is represented by an orange ball. (a)-(d) The arm is initialized with straight, undeformed configuration and is tasked to reach the target located in the first quadrant at $r^{\text{target}} = (9, 9)$ [cm] with the tip. Simulation time is $T = 0.5$ s for all 20 iterations. (e)-(h) The arm is initialized with straight, undeformed configuration and is tasked to reach the target located in the first quadrant at $r^{\text{target}} = (0, -2)$ [cm] with the tip. Simulation time is $T = 0.6$ s for all 40 iterations. (i)-(l) The arm is initialized with bent, deformed configuration and is tasked to reach the target located at $r^{\text{target}} = (16, 10)$ [cm] with the tip. Simulation time is $T = 0.8$ s for all 20 iterations.

3) *Run backward path:* The costate, or the adjoint equation (6) is integrated backward in time from $t = T$ to 0 to obtain the costate $\xi^{(k)}$.

4) *Update control:* The triad $(z^{(k)}, \xi^{(k)}, u^{(k)})$ will typically not satisfy the Hamiltonian maximization criterion (7). Therefore, the control is updated in the direction of steepest ascent of the control Hamiltonian. Denoting the gradient of H with respect to the control u as $\frac{\delta H}{\delta u}$, the control update law is expressed as

$$u^{(k+1)} = u^{(k)} + \eta_k \frac{\delta H}{\delta u^{(k)}} \qquad (10)$$

where $\eta_k > 0$ is the learning rate at iteration k.

Then we repeat steps $1) - 4)$ until either of the two convergence criteria is met: i) the absolute change in control update becomes lower than a threshold ϵ; ii) the number of iterations exceeds a predefined value.

C. Numerical solver

Both the forward and backward path equations (1), (6) are systems of nonlinear PDEs that need to be propagated forward (or backward) in time given initial data. For the forward path, the specialized software *Elastica* [19] is used. The software is designed for high-fidelity simulations of three dimensional Cosserat rods. A custom numerical solver is implemented for the backward adjoint equation.

Both forward and backward dynamics solvers use finite difference techniques to discretize the spatial dimension.

For the backward dynamics, certain spatial discretization operators are employed [35], [36], the details of which appear in the Appendix II. As for the time discretization, the forward dynamics are evolved via a position Verlet scheme. Such a scheme is commonly used to simulate a mechanical system where the state is decomposed into (q, p) pair [37]. As explicit calculations show in Appendix I, the costate ξ is decomposed into a (μ, γ) pair which can be interpreted as velocity-position variables. Hence, the position Verlet scheme is also used for costate dynamics to integrate backward in time.

IV. SIMULATION RESULTS

In this section, we demonstrate the numerical results of the optimal control on a single CyberOctopus arm of rest length L_0. In all our experiments, the intrinsic strains are chosen so that the arm is intrinsically straight, i.e. $\nu^\circ = (1, 0)$ and $\kappa^\circ = 0$. The variable diameter $\phi(s) = \phi_{\text{base}}(L_0 - s) + \phi_{\text{tip}} s$ models the tapering of the arm. The cross sectional area and the second moment of area are given by $A = \frac{\pi \phi^2}{4}$ and $I = A^2/4\pi$. The effective shear modulus is given by $G = \frac{4}{3} \cdot \frac{E}{2(1+\text{Poisson's ratio})}$ [19], where we take the Poisson's ratio to be 0.5 by assuming a perfectly incompressible isotropic material. Parameters like density, modulus of elasticity, and physical dimensions are taken from [20], [38]. Simulation parameters are tabulated in Table I.

TABLE I: Parameters for Numerical Simulation

Parameter	Description	Numerical value
Rod model		
L_0	length of the undeformed rod [cm]	20
ϕ_{base}	rod base diameter [cm]	2
ϕ_{tip}	rod tip diameter [cm]	0.8
ρ	density [kg/m^3]	1042
ζ	damping coefficient [kg/s]	0.01
E	Young's modulus [kPa]	10
Numerics		
Δt	Discrete time step-size [s]	10^{-5}
N	number of discrete segments	100
ϵ	threshold for control convergence	10^{-8}

Fig. 3: Learned optimal controls for the reaching task: Control inputs $u^F = (u^{F_1}, u^{F_2})$ and u^C along the arm are illustrated for the last iteration. Nine time snapshots are shown in orange from $t = 0$ s to $t = 0.4$ s, and sixteen time snapshots are shown in blue from $t = 0.4$ s to $t = 0.5$ s (the most transparent lines correspond to the beginning of the time interval). The orange lines indicate the global profile of the optimal controls. The blue lines indicate the distinguishable traveling waves in optimal controls.

A. Numerical experiments

We test our solver to find the optimal trajectories for three different test cases. We set the terminal tip orientation free and only penalize the distance between the terminal tip position and the target position $r^{\text{target}} \in \mathbb{R}^2$, i.e. for $q^{\text{target}} = (r^{\text{target}}, \theta^{\text{target}})$, we use the following formula for $\Phi_{\text{tip}}(\cdot, \cdot)$

$$\Phi_{\text{tip}}\left(q(T, L_0), q^{\text{target}}\right) = \frac{1}{2}\left\| r(T, L_0) - r^{\text{target}} \right\|^2 \quad (11)$$

where the norm is the usual Euclidean distance in $\mathbb{R}^2$.

1) Reaching task: Our first experiment is a simple reaching problem. The arm is initialized to be straight and undeformed. Our goal is to control the arm to reach the target with the tip at time $T = 0.5$ s. We consider the optimal control problem (2)-(4), (11) with weight parameter $\chi_1 = 10$ and regularization parameter $\chi_2 = 2 \times 10^4$. We ran the forward-backward algorithm for 20 iterations with fixed learning rate $\eta_k = 3 \times 10^{-5}$.

We select four different iterations to demonstrate the control results. As we see in Fig 2a-d, the reaching capability of the arm improves over iterations due to control updates. In the 2nd iteration, the arm does not bend much yet but shows the trend of moving towards the target. In the 6th iteration, the arm tip already gets close to the target. The controls converge quickly, and in the last iteration, the time snapshots show that the learned optimal control drives the arm to smoothly bend towards the target and the tip reaches the target at the terminal time. Fig. 3 depicts the control inputs in the last iteration. We can see the emergence of a wave propagation in control inputs.

2) Fetching task: During a fetching motion, the arm is observed to form several pseudo-joints [17]. To investigate this behavior, optimal trajectories are computed where the static target r^{target} is close to the base of the arm and is thought of as the mouth of the octopus. The arm is initialized to be straight and undeformed. The forward-backward algorithm is run for 40 iterations with parameters $\chi_1 = 10$, $\chi_2 = 2 \times 10^4$ and $\eta_k = 4 \times 10^{-5}$. The terminal time $T = 0.6$ s is fixed for all iterations. Fig. 2e-h depicts the fetching movement where the arm forms a bend as it tries to get close to the target point.
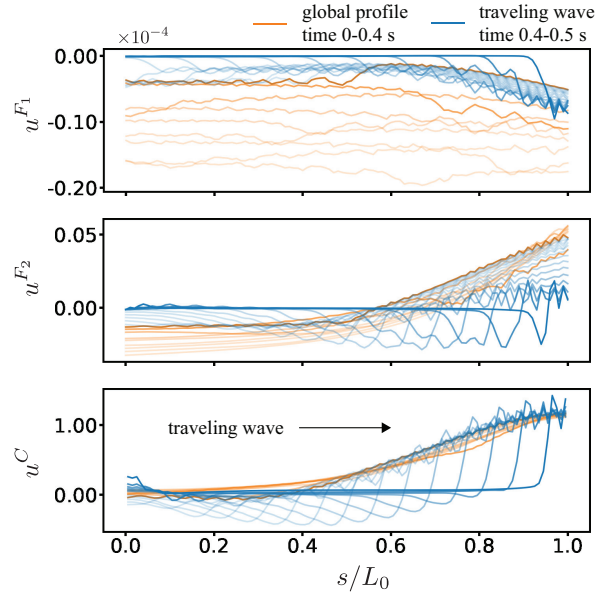
3) Shooting task (reaching from bent position): Octopuses are known to curl up their arms while at rest, and when they try to catch food from a distance, they 'shoot' one of the arms towards the target [16]. During this, the bend propagation is most prominently observed. Inspired by these observations, in our last experiment the arm is initialized at a bent position according to the initial curvature

$$\kappa(0, s) = \sum_{i=1}^{4} M_i \exp\left(-\frac{(s - s_i)^2}{2 \times \sigma_i^2}\right)$$

where M_i's are [20, 78, 10, -30], s_i's are [0, $0.3L_0$, $0.7L_0$, $0.85L_0$], and σ_i's are [0.015, 0.015, 0.012, 0.008]. Our goal is to reach the target at time $T = 0.8$ s. We ran the forward-backward algorithm for 20 iterations with parameters $\chi_1 = 100$, $\chi_2 = 2 \times 10^4$ and $\eta_k = 3 \times 10^{-5}$.

The control results of four forward-backward iterations are demonstrated in Fig. 2i-l. Even though the arm reaches the target at the final iteration, the stereotypical bend propagation [16] is absent. This is suggestive of the potential importance of environmental effects such as drag forces.

B. Characteristics of optimal control

In our simulations, the optimal control solutions exhibit the following patterns. There is an initial global profile for the control. Starting from $t = 0$ s, a localized wave travels back and forth along the global profile and the magnitude of this wave increases as t increases. At first, the wave is not

discernible due to its small magnitude and thus, the global profile is dominant as indicated by the orange lines in Fig. 3. As time t nears the final time T, the wave traveling from the base to the tip of the arm becomes more visible and it dominates the control as shown by the blue lines in Fig. 3.

We vary different parameters to investigate how they affect the optimal control solution, especially the wave propagation.

1) Wave speed: We observed that the parameters of the optimal control problem, e.g. $T, \chi_1, \chi_2, r^{\text{target}}$ (see discussion in Sec. IV-B.2 about the parameter χ_1), geometry of the arm (the length and tapered diameter profile), dissipation constant ζ, and numerical integration constants (e.g. $N, \Delta t, \eta$) do not affect the speed c of the wave. However, Young's modulus (E) and density (ρ) of the arm do affect the wave speed. We calculate the speeds for different sets of E and ρ values, which shows the linear relationship (the graphic is omitted due to lack of space)

$$c = 0.653\sqrt{\frac{E}{\rho}}$$

This experiment indicates that the wave in the optimal control solution is actually a fundamental property of the elastic arm. Further study is required to draw connections to the stereotypical bend propagation waves observed in octopuses [16].

2) Parameter χ_1: In the cost function (3), we penalize the deformation of the arm with the parameter χ_1. Even though the parameter χ_1 does not affect the wave speed, increasing χ_1 leads to an interesting observation. When χ_1 is high enough, a visible second wave appears in the control solution which propagates in the opposite direction of the original wave. Moreover, these two waves meet exactly at the middle point of the arm (Fig. 4). The resemblance of this behavior with the observation of [17] demands further analysis.

V. CONCLUSION AND FUTURE WORK

In this paper, we investigate an optimal control problem for a single CyberOctopus arm modeled as a planar Cosserat rod. A free endpoint optimal control problem is formulated to minimize the control energy and a weighted potential energy of the rod. To reach a target point, the proximity of the arm's tip to the target point is penalized at the terminal time. The necessary first order optimality conditions yield two systems, the Cosserat rod dynamics (forward) and the adjoint dynamics (backward), both described by nonlinear PDEs. To numerically solve these PDEs, specific spatial and temporal discretization techniques are used. The optimal controls are found by updating the controls in an iterative manner called the forward-backward algorithm. This framework is used to solve several biologically motivated control tasks. These numerical experiments reveal emergence of propagating waves in the optimal controls. However, the stereotypical bend propagation along the arm is not discovered under our current problem formulation. This motivates us to consider environmental effects like drag, and constraints of muscle actuation into our optimal control framework.
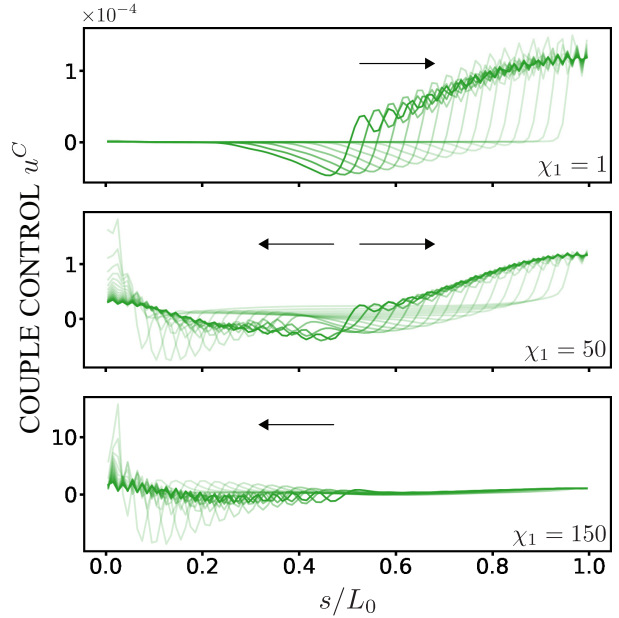

Fig. 4: Comparison of wave behaviors in couple control for different χ_1 parameters in the reaching task: Sixteen time snapshots are shown in green from $t = 0.45$ s (most solid) to $t = 0.5$ s. (most transparent) The black arrows indicate the direction of the dominant wave propagation. For $\chi_1 = 1$, the usual dominating wave travels from base to the tip. For $\chi_1 = 50$, both the original wave and a second wave are visible and they meet at the middle point $0.5L_0$. For $\chi_1 = 150$, the second wave is dominant which travels from tip to the base.

REFERENCES

[1] J. Maynard Smith and G. R. Price, "The logic of animal conflict," *Nature*, vol. 246, no. 5427, pp. 15–18, 1973.

[2] J. Maynard Smith, *Evolution and the Theory of Games*. Cambridge university press, 1982.

[3] A. Attanasi, A. Cavagna, *et al.*, "Information transfer and behavioural inertia in starling flocks," *Nature physics*, vol. 10, no. 9, pp. 691–696, 2014.

[4] E. W. Justh and P. Krishnaprasad, "Optimality, reduction and collective motion," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 471, no. 2177, p. 20140606, 2015.

[5] U. Halder, "Optimality, synthesis and a continuum model for collective motion," Ph.D. dissertation, University of Maryland, College Park, 2019.

[6] T. Flash and N. Hogan, "The coordination of arm movements: an experimentally confirmed mathematical model," *Journal of neuroscience*, vol. 5, no. 7, pp. 1688–1703, 1985.

[7] P. Viviani and T. Flash, "Minimum-jerk, two-thirds power law, and isochrony: converging approaches to movement planning." *Journal of Experimental Psychology: Human Perception and Performance*, vol. 21, no. 1, p. 32, 1995.

[8] E. Todorov and M. I. Jordan, "Smoothness maximization along a predefined path accurately predicts the speed profiles of complex arm movements," *Journal of Neurophysiology*, vol. 80, no. 2, pp. 696–714, 1998.

[9] E. Todorov, "Optimality principles in sensorimotor control," *Nature neuroscience*, vol. 7, no. 9, pp. 907–915, 2004.

[10] E. Todorov and M. I. Jordan, "Optimal feedback control as a theory of motor coordination," *Nature neuroscience*, vol. 5, no. 11, pp. 1226–1235, 2002.

[11] Y. Gutfreund, T. Flash, *et al.*, "Organization of octopus arm movements: a model system for studying the control of flexible arms," *Journal of Neuroscience*, vol. 16, no. 22, pp. 7297–7307, 1996.

[12] ——, "Patterns of arm muscle activation involved in octopus reaching movements," *Journal of Neuroscience*, vol. 18, no. 15, pp. 5976–5987, 1998.

[13] G. Sumbre, G. Fiorito, *et al.*, "Octopuses use a human-like strategy to control precise point-to-point arm movements," *Current Biology*, vol. 16, no. 8, pp. 767–772, 2006.

[14] M. Hines and J. Blum, "Bend propagation in flagella. i. derivation of equations of motion and their simulation," *Biophysical Journal*, vol. 23, no. 1, pp. 41–57, 1978.

[15] B. D. Coleman and E. H. Dill, "Flexure waves in elastic rods," *The Journal of the Acoustical Society of America*, vol. 91, no. 5, pp. 2663–2673, 1992.

[16] G. Sumbre, Y. Gutfreund, *et al.*, "Control of octopus arm extension by a peripheral motor program," *Science*, vol. 293, no. 5536, pp. 1845–1848, 2001.

[17] G. Sumbre, G. Fiorito, *et al.*, "Motor control of flexible octopus arms," *Nature*, vol. 433, no. 7026, p. 595, 2005.

[18] S. S. Antman, *Nonlinear Problems of Elasticity*. Springer, 1995.

[19] M. Gazzola, L. Dudte, *et al.*, "Forward and inverse problems in the mechanics of soft filaments," *Royal Society Open Science*, vol. 5, no. 6, p. 171628, 2018.

[20] H.-S. Chang, U. Halder, *et al.*, "Energy shaping control of a cyberoctopus soft arm," in *2020 59th IEEE Conference on Decision and Control (CDC)*. IEEE, 2020, pp. 3913–3920.

[21] X. Zhang, F. K. Chan, *et al.*, "Modeling and simulation of complex dynamic musculoskeletal architectures," *Nature Communications*, vol. 10, no. 1, pp. 1–12, 2019.

[22] N. Naughton, J. Sun, *et al.*, "Elastica: A compliant mechanics environment for soft robotic control," *IEEE Robotics and Automation Letters*, 2021.

[23] C. Wei-Liang, "Uber systeme von linearen partiellen differentialgleichungen erster ordnung," *Math. Ann*, vol. 117, pp. 98–105, 1939.

[24] P. Rashevsky, "About connecting two points of a completely nonholonomic space by admissible curve," *Uch. Zapiski Ped. Inst. Libknechta*, vol. 2, pp. 83–94, 1938.

[25] H. J. Sussmann and V. Jurdjevic, "Controllability of nonlinear systems." *Journal of Differential Equations*, vol. 12, pp. 95–116, 1972.

[26] E. Heintze and X. Liu, "Homogeneity of infinite dimensional isoparametric submanifolds," *Annals of mathematics*, vol. 149, pp. 149–181, 1999.

[27] M. K. Salehani and I. Markina, "Controllability on infinite-dimensional manifolds: a chow–rashevsky theorem," *Acta Applicandae Mathematicae*, vol. 134, no. 1, pp. 229–246, 2014.

[28] M. I. Krastanov, N. Ribarska, and T. Y. Tsachev, "A pontryagin maximum principle for infinite-dimensional problems," *SIAM Journal on Control and Optimization*, vol. 49, no. 5, pp. 2155–2182, 2011.

[29] X. Li and J. Yong, *Optimal Control Theory for Infinite Dimensional Systems*. Springer Science & Business Media, 2012.

[30] D. D. Morrison, J. D. Riley, and J. F. Zancanaro, "Multiple shooting method for two-point boundary value problems," *Communications of the ACM*, vol. 5, no. 12, pp. 613–614, 1962.

[31] H. G. Bock and K.-J. Plitt, "A multiple shooting algorithm for direct solution of optimal control problems," *IFAC Proceedings Volumes*, vol. 17, no. 2, pp. 1603–1608, 1984.

[32] O. Von Stryk, "Numerical solution of optimal control problems by direct collocation," in *Optimal control*. Springer, 1993, pp. 129–143.

[33] A. E. Bryson and W. F. Denham, "A Steepest-Ascent Method for Solving Optimum Programming Problems," *Journal of Applied Mechanics*, vol. 29, no. 2, pp. 247–257, 1962.

[34] K. Fujimoto and T. Sugie, "Iterative learning control of hamiltonian systems: I/o based optimal control approach," *IEEE Transactions on Automatic Control*, vol. 48, no. 10, pp. 1756–1761, 2003.

[35] M. Bergou, M. Wardetzky, *et al.*, "Discrete elastic rods," in *ACM SIGGRAPH 2008 papers*, 2008, pp. 1–12.

[36] H. Lang, J. Linn, and M. Arnold, "Multi-body dynamics simulation of geometrically exact cosserat rods," *Multibody System Dynamics*, vol. 25, no. 3, pp. 285–312, 2011.

[37] L. Verlet, "Computer "experiments" on classical fluids. i. thermodynamical properties of lennard-jones molecules," *Physical review*, vol. 159, no. 1, p. 98, 1967.

[38] Y. Yekutieli, R. Sagiv-Zohar, *et al.*, "Dynamic model of the octopus arm. i. biomechanics of the octopus reaching movement," *Journal of neurophysiology*, vol. 94, no. 2, pp. 1443–1458, 2005.

APPENDIX I
EXPLICIT CALCULATIONS

A. Details of a planar Cosserat rod dynamics

For the planar case of the Cosserat rod, we denote $q = (r, \theta)$ as the state where the position vector along the rod $r(t, s) \in \mathbb{R}^2$ and the angle $\theta(t, s) \in \mathbb{R}$ can be used to measure local strains – stretch (ν_1), shear (ν_2), and curvature (κ). These are defined as follows:

$$r_s = Q\nu, \quad \theta_s = \kappa$$

where $Q = \left(\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix} \right)$ is the planar rotation matrix, and $\nu = \left(\begin{smallmatrix} \nu_1 \\ \nu_2 \end{smallmatrix} \right)$. The internal stresses, i.e. the forces n (represented in the material frame) and couple m are related to the stored energy function W by

$$n = \frac{\partial W}{\partial \nu}, \quad m = \frac{\partial W}{\partial \kappa}$$

We take the following quadratic form of W so that the stress-strain relationship becomes linear

$$W = \frac{1}{2}\left((\nu - \nu^\circ)^\mathsf{T} S(\nu - \nu^\circ) + B(\kappa - \kappa^\circ)^2\right)$$

where the intrinsic strains of the rod are denoted by $(\nu^\circ, \kappa^\circ)$. Here, $S = \mathrm{diag}(EA, GA)$ is the stretch-shear rigidity matrix and $B = EI$ is the bending rigidity. E, G are the Young's modulus and shear modulus, respectively.

Let us denote $p^r = \rho A r_t$ and $p^\theta = \rho I \theta_t$ as the momentum variables $p = (p^r, p^\theta)$, where ρ is the density, A is the cross sectional area and I is the second moment of area. Let '$\cdot$' denote the dot product of two planar vectors, and '$\times$' represent the component of the cross product of two planar vectors along the normal vector that is coming out of the plane, i.e. $\left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix} \right) \cdot \left(\begin{smallmatrix} y_1 \\ y_2 \end{smallmatrix} \right) = x_1 y_1 + x_2 y_2$, and $\left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix} \right) \times \left(\begin{smallmatrix} y_1 \\ y_2 \end{smallmatrix} \right) = x_1 y_2 - x_2 y_1$.

The Cosserat dynamics (1) are written as

$$
\begin{aligned}
r_t &= \frac{1}{\rho A} p^r \\
\theta_t &= \frac{1}{\rho I} p^\theta \\
p^r_t &= (Qn)_s - \frac{1}{\rho A}\zeta p^r + u^F \\
p^\theta_t &= (m)_s + \nu \times n - \frac{1}{\rho I}\zeta p^\theta + u^C
\end{aligned}
\tag{A-1}
$$

where $u = (u^F, u^C)$ denote the force and couple control inputs.

B. Details of the adjoint equations

Denote the costate to $(q, p) = ((r, \theta), (p^r, p^\theta))$ as $(\mu, \gamma) = ((\mu^r, \mu^\theta), (\gamma^r, \gamma^\theta))$. Then, the pre-Hamiltonian (5) is explicitly written as

$$
\begin{aligned}
H = \int_0^{L_0} &\left[\frac{1}{\rho A}\mu^r \cdot p^r + \frac{1}{\rho I}\mu^\theta p^\theta + \gamma^r \cdot \left((Qn)_s - \frac{1}{\rho A}\zeta p^r\right) \right. \\
&+ \gamma^\theta \left((m)_s + \nu \times n - \frac{1}{\rho I}\zeta p^\theta\right) + \gamma^r \cdot u^F + \gamma^\theta u^C \\
&\left. - \frac{1}{2}\left(u^F \cdot u^F + \left(u^C\right)^2\right) - \chi_1 \mathcal{V}(q) \right] \mathrm{d}s
\end{aligned}
\tag{A-2}
$$

Maximizing H with respect to u gives the first order necessary condition for optimal control

$$u^F = \gamma^r, \quad u^C = \gamma^\theta \tag{A-3}$$

Furthermore, the costate evolution equations (6) take the explicit form

$$
\begin{aligned}
\mu_t^r &= -\frac{\delta H}{\delta r} \\
&= -\left(QSQ^\mathsf{T}\gamma_s^r\right)_s - \left[QM_1(GA\nu - EA\sigma)\gamma^\theta\right]_s - \chi_1(Qn)_s \\
\mu_t^\theta &= -\frac{\delta H}{\delta \theta} \\
&= -\left(B\gamma_s^\theta\right)_s + \left[Q\left(M_2 n - SM_2\nu\right)\right] \cdot \gamma_s^r \\
&\quad + \left[(M_2\nu) \times n + \nu \times (SM_2\nu)\right]\gamma^\theta - \chi_1\left((m)_s + \nu \times n\right) \\
\gamma_t^r &= -\frac{\delta H}{\delta p^r} = -\frac{1}{\rho A}\left(\mu^r - \zeta\gamma^r\right) \\
\gamma_t^\theta &= -\frac{\delta H}{\delta p^\theta} = -\frac{1}{\rho I}\left(\mu^\theta - \zeta\gamma^\theta\right)
\end{aligned}
\tag{A-4}
$$

where $\sigma = \nu - \nu^\circ$, $M_1 = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ and $M_2 = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$.

These equations are to be accompanied with the transversality condition (9) (with (4), (11))

$$
\begin{aligned}
\mu^r(T, s) &= -\delta(s - L_0)\left[\chi_2(r(t, s) - r^{\text{target}})\right]\Big|_{t=T} \\
\mu^\theta(T, s) &= 0 \\
\gamma^r(T, s) &= 0 \\
\gamma^\theta(T, s) &= 0
\end{aligned}
\tag{A-5}
$$

where $\delta(\cdot)$ denotes the delta function.

C. Control update law

Denoting $u = (u^F, u^C)$ and $\gamma = (\gamma^r, \gamma^\theta)$, we can write the control update law (10) for the forward-backward algorithm at iteration k as

$$u^{(k+1)} = u^{(k)} + \eta_k \frac{\delta H}{\delta u^{(k)}} = u^{(k)} + \eta_k\left(\gamma^{(k)} - u^{(k)}\right) \tag{A-6}$$

APPENDIX II
NUMERICAL METHODS

We use the following spatial and temporal discretization for the backward path that is consistent with the forward path.

A. Spatial discretization

In the software package *Elastica*, the Cosserat rod is decomposed into $N + 1$ nodes for the position r and N segments for the angle θ [19].

We define the following two difference operators for vectors according to finite difference approximation [35], [36]. Let $\{\mathbb{R}^p\}_N$ denote a set of N vectors in $\mathbb{R}^p$. Then, $\tilde{\mathcal{D}}: \{\mathbb{R}^p\}_N \mapsto \{\mathbb{R}^p\}_{N+1}$ and $\bar{\mathcal{D}}: \{\mathbb{R}^p\}_N \mapsto \{\mathbb{R}^p\}_{N-1}$ are defined as follows:

$$
a_{i=1,\ldots,N+1} = \tilde{\mathcal{D}}(b_{j=1,\ldots,N}) = \begin{cases} b_1, & i = 1 \\ b_i - b_{i-1}, & i = 2,\ldots,N \\ -b_N, & i = N+1 \end{cases}
\tag{A-7}
$$

and

$$c_{\ell=1,\ldots,N-1} = \bar{\mathcal{D}}(b_{j=1,\ldots,N}) = b_{\ell+1} - b_\ell, \quad \ell = 1,\ldots,N-1 \tag{A-8}$$

where $a_i \in \mathbb{R}^p$ for $i = 1,\ldots,N+1$, $b_j \in \mathbb{R}^p$ for $j = 1,\ldots,N$ and $c_\ell \in \mathbb{R}^q$ for $\ell = 1,\ldots,N-1$. Note that $\tilde{\mathcal{D}}$ and $\bar{\mathcal{D}}$ operate on a set of N vectors and then return $N+1$ and $N-1$ vectors, respectively.

Now for the rest of this Appendix, we will use specific subscripts $(\cdot)_i$, $(\cdot)_j$ and $(\cdot)_\ell$ to denote the set of discretized variables with the dimension of spatial discretization to be $N+1$, N and $N-1$, respectively.

For the backward path, we discretize the costate into μ_i^r, γ_i^r and μ_j^θ, γ_j^θ. Then the first-order necessary condition for optimal control is

$$
\begin{aligned}
u_i^F &= \gamma_i^r \\
u_j^C &= \gamma_j^\theta
\end{aligned}
\tag{A-9}
$$

where u_i^F and u_j^C are the discretized control inputs to be used in the forward path.

The costate dynamics (A-4) are discretized as follows:

$$
\begin{aligned}
\frac{d\mu_i^r}{dt} &= -\tilde{\mathcal{D}}\left(Q_j S Q_j^\mathsf{T} \bar{\mathcal{D}}(\gamma_i^r)/\Delta s\right) - \tilde{\mathcal{D}}\left(Q_j M_1(GA\nu_j - EA\sigma_j)\gamma^\theta\right) \\
&\quad - \chi_1\tilde{\mathcal{D}}\left(Q_j n_j\right) \\
\frac{d\mu_j^\theta}{dt} &= -\tilde{\mathcal{D}}\left(B\bar{\mathcal{D}}(\gamma_j^\theta)/\Delta s\right) + \left[Q_j\left(M_2 n_j - SM_2\nu_j\right)\right] \cdot \bar{\mathcal{D}}(\gamma_i^r) \\
&\quad + \left[(M_2\nu_j) \times n_j + \nu_j \times (SM_2\nu_j)\right]\gamma_j^\theta \Delta s \\
&\quad - \chi_1\left(\tilde{\mathcal{D}}\left(m_\ell\right) + (\nu_j \times n_j)\Delta s\right) \\
\frac{d\gamma_i^r}{dt} &= -\frac{1}{\rho A}\left(\mu_i^r - \zeta\gamma_i^r\right) \\
\frac{d\gamma_j^\theta}{dt} &= -\frac{1}{\rho I}\left(\mu_j^\theta - \zeta\gamma_j^\theta\right)
\end{aligned}
\tag{A-10}
$$

where $\Delta s = L_0/N$ is the length of each discretized segment of the rod. r_i, Q_j, ν_j, σ_j, n_j and m_ℓ are discretized variables obtained from the forward path. Details of these variables are covered in [19].

The transversality conditions (A-5) are discretized into

$$
\begin{aligned}
\mu_i^r(T) &= -\delta(i - (N+1))\left[\chi_2(r_i - r^{\text{target}})\right]\Big|_{t=T} \\
\mu_j^\theta(T) &= 0 \\
\gamma_i^r(T) &= 0 \\
\gamma_j^\theta(T) &= 0
\end{aligned}
\tag{A-11}
$$

B. Time discretization

We use the second-order position Verlet time integration [19] as follows:

$$
\begin{aligned}
\gamma_i^r\left(t - \frac{\Delta t}{2}\right) &= \gamma_i^r(t) - \frac{\Delta t}{2}\frac{d\gamma_i^r}{dt}(t) \\
\mu_i^r(t - \Delta t) &= \mu_i^r(t) - \Delta t\frac{d\mu_i^r}{dt}\left(t - \frac{\Delta t}{2}\right) \\
\gamma_i^r(t - \Delta t) &= \gamma_i^r\left(t - \frac{\Delta t}{2}\right) - \frac{\Delta t}{2}\frac{d\gamma_i^r}{dt}(t - \Delta t)
\end{aligned}
\tag{A-12}
$$

Similarly for γ_j^θ and μ_j^θ.