

# Optimal Design of Continuum Robots With Reachability Constraints

Hyunmin Cheong , Mehran Ebrahimi, and Timothy Duggan

**Abstract**—While multi-joint continuum robots are highly dexterous and flexible, designing an optimal robot can be challenging due to its kinematics involving curvatures. Hence, the current work presents a computational method developed to find optimal designs of continuum robots, given reachability constraints. First, we leverage both forward and inverse kinematic computations to perform reachability analysis in an efficient yet accurate manner. While implementing inverse kinematics, we also integrate torque minimization at joints such that robot configurations with the minimum actuator torque required to reach a given workspace could be found. Lastly, we apply an estimation of distribution algorithm (EDA) to find optimal robot dimensions while considering reachability, where the objective function could be the total length of the robot or the actuator torque required to operate the robot. Through three application problems, we show that the EDA is superior to a genetic algorithm (GA) in finding better solutions within a given number of iterations, as the objective values of the best solutions found by the EDA are 4–15% lower than those found by the GA.

**Index Terms**—Kinematics, methods and tools for robot system design, optimization and optimal control, soft robot applications.

## I. INTRODUCTION

CONTINUUM robots, also called soft robots, are composed of joints that bend continuously along their lengths (Fig. 1). The design of such manipulators are inspired by animal appendages such as elephant trunks or octopus tentacles. They are highly dexterous and flexible, thus ideal for working in cluttered environments. Furthermore, compared to traditional rigid manipulators, continuum robots can better adapt to and interact with their surroundings. However, due to their flexibility, analyzing their reachability requires a more complex kinematic model than traditional manipulators and it is difficult to conceptualize the effect of design changes on the workspace of these robots. Even with 3D computer-aided design tools, designing a continuum robot that can reach the entire target workspace can take weeks of an engineering team's time, and the chance of

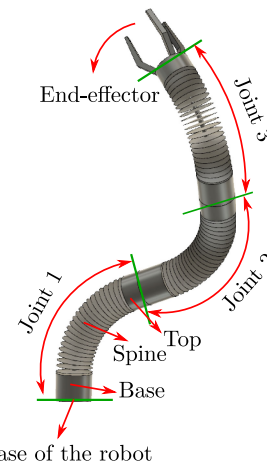


Fig. 1. A three-joint spatial continuum robot. Each joint consists of a base, top and spine. The spine can bend continuously along its length about two orthogonal axes.

finding an optimal design (e.g., with the minimal total length) is low.

To address this challenge, we present a method for optimizing the design of multi-joint continuum robots while satisfying reachability constraints given a desired workspace. The method involves two main parts: 1) kinematic computations to evaluate the reachability of potential designs and 2) an optimization algorithm to find the optimal robot design.

For the first part, we combine both forward and inverse kinematics to perform a reachability analysis in an efficient manner. Since the former can be executed considerably faster than the latter, a large set of randomly sampled robot configurations with forward kinematics is used to quickly estimate the robot's reachability. Then, for the points not deemed to be reachable, inverse kinematics is employed to check whether those points can actually be reached or not, further improving the accuracy of the reachability analysis. In addition, we integrate torque minimization as part of the inverse kinematics computation such that given a target point, the robot configuration with the minimal actuator torque to reach that point can be identified.

The optimization problem considering reachability is challenging because computing the sensitivity of the reachability function with respect to the design variables involved is not possible. Also, the feasible region is likely non-contiguous since different combinations of robot dimensions could satisfy the reachability constraint. To solve such a problem, derivative-free

Manuscript received October 15, 2020; accepted February 17, 2021. Date of publication March 17, 2021; date of current version April 2, 2021. This letter was recommended for publication by Associate Editor L. Wen and Editor K.-J. Cho upon evaluation of the reviewers' comments. This work was supported in part by the Office of Naval Research Contract #N68335-17-C-0045, in part by the ARPA-E Contract #DE-AR0001241, and in part by NASA Contract #80NSSC19C0637. (Corresponding author: Hyunmin Cheong.)

Hyunmin Cheong and Mehran Ebrahimi are with the Autodesk Research, Autodesk Inc., Toronto, ON M5G 1M1, Canada (e-mail: hyunmin.cheong@autodesk.com; mehran.ebrahimi@autodesk.com).

Timothy Duggan is with the Otherlab Inc., San Francisco, CA 94017 USA (e-mail: trd44@cornell.edu).

Digital Object Identifier 10.1109/LRA.2021.3066978

or black-box optimization algorithms can be used, e.g., evolutionary algorithms. For the current work, we use an estimation of distribution algorithm (EDA), which finds optimal solutions by estimating and sampling a probability model of promising designs. EDA is chosen because it has been shown to outperform a genetic algorithm, the most widely used evolutionary algorithm, in a number of prior studies [1]–[5]. Also, the problem knowledge in the form of probabilistic models learned as part of the optimization could be conveyed to the user, which is not an inherent feature of most evolutionary algorithms.

The rest of the paper is organized as follows. First, we present related work on continuum robots and EDAs. Next, a problem formulation and the method developed to solve the problem are presented. We then provide the results of the experiments conducted to validate our optimization method, followed by a summary and conclusions.

## II. RELATED WORK

### A. Kinematics of Continuum Robots

Unlike traditional robot manipulators, continuum robots are made of sequentially actuated deformable links (joints) exhibiting some levels of compliance depending on their design and the applications destined for them. Therefore, the conventional parameters (e.g., joint lengths and joint angles) used for describing the kinematics of rigid-link robots are not applicable to continuum robots. A thorough review of different techniques for modeling the kinematics of such robots can be found in [6]–[8]. A common practice adopted in analyzing continuum robots is to assume that each of their joints deforms as a constant-curvature arc (e.g. [9]–[14]). This turns the robot's configuration space from an infinite to a finite dimensional space and facilitates fast closed-form computations of the robot kinematics. Following this presumption, several kinematic models have been proposed. In [9], the well-known Denavit-Hartenberg procedure is modified to determine the kinematics of planar robots resembling an elephant's trunk. Another approach is to model each continuum joint as a combination of prismatic and revolute joints [10]. In [12], the motion of planar robots is characterized by splitting it into pure bending and extension. The robot's kinematics can also be configured by incorporating quaternions to define the rotation of the robot's joints [13]. Despite the inherent differences among the various kinematic models, it is demonstrated that they often lead to identical results in many different scenarios [8]. In the current work, the robot's three-dimensional kinematics is determined using its joints' length, radius of curvature and angle, thus bearing some similarities to what is used in [8].

### B. Optimal Design of Continuum Robots

Optimal design of continuum robots has been actively researched in the past few years. Early on, much of the work was on optimizing concentric tube robots for biomedical applications. These robots are similar to the robots considered in the current work as they also require constant-curvature kinematics and can involve multiple joints. For example, generalized pattern search is used to optimize robots with the minimal curvature and length

while being able to reach the workspace [15], [16]. In another work, the Nelder-Mead algorithm is implemented to find optimal designs while considering volume-based workspaces [17]. Also, a particle swarm algorithm is employed to perform design optimization for dual-arm concentric tube robots [18].

Recently, evolutionary algorithms have been widely applied for optimal design of continuum robots. For instance, Hiller *et al.* uses a genetic algorithm to optimally distribute soft materials in a compliant structure [19]. Runge *et al.* [20] incorporates a genetic algorithm to optimize a single-joint soft robot while considering its mechanics. Cheney *et al.* [21] uses a variant of the NEAT algorithm (NeuroEvolution of Augmenting Topologies) to find the optimal volumetric structure and material choice for a contiguous soft robot. Finally, most similar to our work, Bodily *et al.* [22] applies a genetic algorithm to optimize a multi-joint continuum robot considering reachability, dexterity, and manipulability.

### C. Estimation of Distribution Algorithms

EDAs are population-based, derivative-free optimization methods that use probability distributions estimated from a population of candidate solutions to sample new solutions at each iteration of optimization [23], [24]. They have been shown to be more effective at finding optimal solutions with a fewer number of function evaluations than genetic algorithms for several benchmark problems [1]–[4], hence motivating the application for the current work.

EDAs have been applied to solve various engineering design problems [25], including vehicle suspensions [5], [26] and multi-speed gearboxes [27], [28]. In robotics, EDAs have been used to solve problems such as inverse displacement [29], gait generation [30], [31], and cable-driven parallel robot design [32]. However, no prior work has explored applying an EDA for optimal design of multi-joint continuum robots.

For the current work, we have considered various continuous EDAs [23] because the design variables involved in optimization, e.g., the dimensions of joints in a continuum robot, are continuous. Specifically, we have applied a univariate normal distribution algorithm inspired by [33].

## III. PROBLEM FORMULATION

The design of a multi-joint continuum robot is defined as the following optimization problem.

$$\begin{aligned} \min_{\mathcal{D}} \quad & f(\mathcal{D}) \\ \text{s.t.} \quad & d_{i,lb} \leq x_i \leq d_{i,ub} \quad \forall x_i \in \mathcal{D} \\ & \Theta(\mathcal{D}, \mathcal{S}) \geq \alpha \end{aligned} \quad (1)$$

Here,  $\mathcal{D}$  represents a set of design variables  $x_i$ , such as the dimensions of a robot, and  $\mathcal{S}$  denotes a set of state variables that define the configuration of a robot, such as each joint's radius of curvature and angle. For the objective function  $f(\mathcal{D})$ , the current work considers two different types – the total length of the robot and the minimum actuator torque required to reach a given workspace. Each design variable  $x_i$  has lower and upper

bounds,  $d_{i,lb}$  and  $d_{i,ub}$ , respectively, indicating the limits on robot dimensions. Finally,  $\Theta(\mathcal{D}, \mathcal{S})$  computes the percentage of the workspace that can be reached by a robot, and this value must be greater than the threshold  $\alpha$ .

In order to compute  $\Theta(\mathcal{D}, \mathcal{S})$ , a given workspace is discretized into a set of target points in the 3D space. Then, its value can be calculated by

$$\Theta(\mathcal{D}, \mathcal{S}) = \frac{N_{\text{reached}}}{N_{\text{target}}} \quad (2)$$

where  $N_{\text{target}}$  represents the number of target points in the workspace and  $N_{\text{reached}}$  is the number of such points that can be reached by a robot.  $N_{\text{reached}}$  is determined based on

$$\|p_j - \vec{r}_e(\mathcal{D}, \mathcal{S}_j)\| \leq \epsilon \quad \text{for } j = 1, \dots, m \quad (3)$$

in which  $p_j$  is each of the  $m$  target points considered and  $\vec{r}_e(\mathcal{D}, \mathcal{S}_j)$  computes the corresponding position of the robot's end-effector as shown later.  $\epsilon$  is the tolerance allowed.

#### IV. METHODS

First presented is the formulation of kinematics used to perform a reachability analysis, followed by the optimization method using an estimation of distribution algorithm.

##### A. Reachability Analysis

As shown in [22], to estimate the reachability of a robot, one could randomly sample a large number of robot configurations and use forward kinematics to compute the end-effector position for each configuration. Then, the set of end-effector positions is compared against the set of target points in the workspace to estimate the reachability. Note that computing forward kinematics is quite fast and therefore one could sample a large number of configurations within a given time budget. With a large enough sample size, a reasonable estimate of the reachability could be obtained.

On the other hand, one could employ inverse kinematics to assess the reachability of a robot. For instance, for each target point in the workspace, one could solve for the robot configuration that would result in the end-effector position coinciding with the target point. The percentage of the target points for which solutions are found would then indicate the reachability. While this approach can provide a more accurate assessment of the reachability, computing inverse kinematics for all target points can take a significant amount of time.

The current work therefore uses a hybrid approach. First, random sampling of robot configurations with forward kinematics is performed. Then, for target points that have not been reached during this first step, inverse kinematics is used to check whether those points can actually be reached or not. In this section, we present the forward and inverse kinematic models employed, while the exact implementation of the hybrid approach is described in the Experiments section.

We assume that a robot consists of several non-extensible mutually tangent constant-curvature arcs that can spatially deform. Each joint is composed of a base, top and spine as depicted in Fig. 1. The base and top sections are straight and rigid links,

whereas the spine can bend about two orthogonal axes, thus giving each joint two degrees of freedom (DOF). Depending on the joint design, one may need to model the spine as a spatial beam including its longitudinal, bending and torsional deformations altogether [34].

Given the robot's base position  $\vec{r}_b$ , tangent  $\vec{t}_b$  and normal  $\vec{n}_b$  (Fig. 2 a and Fig. 2 b), the robot kinematics can be fully configured using its joints' radius of curvature  $R$  ( $R > 0$ ) and rotation  $\theta$  ( $0 \leq \theta \leq 2\pi$ ). The latter specifies the rotation of a joint with respect to its respective base normal, so it determines the bending plane of each joint. These parameters, also called state variables, are illustrated in Fig. 2. In order to assess the reachability of a robot, one needs to solve the forward and inverse kinematic equations laid out as follows.

1) *Forward Kinematics*: In a forward kinematic (FK) problem, the goal is to find the spatial position and orientation at different points on the robot, particularly the end-effector, given a set of state variables (i.e. joint radii and rotations). Employing the parametrization introduced earlier and utilizing the Rodrigues' rotation formula [35], the vectors  $\vec{t}_i$  and  $\vec{n}_i$  depicted in Fig. 2(c) for Joint  $i$  ( $i \geq 2$ ) read

$$\begin{aligned} \vec{t}_i &= \vec{t}_{i-1} \cos(\phi_{i-1}) + \vec{n}_{i-1} \sin(\phi_{i-1}) \\ \vec{n}_i &= \vec{n}_{i-1} \cos(\phi_{i-1}) - \vec{t}_{i-1} \sin(\phi_{i-1}) \end{aligned} \quad (4)$$

in which  $\phi_{i-1} = l_{s,i-1}/R_{i-1}$ ; and  $l_{b,i}$ ,  $l_{s,i}$  and  $l_{t,i}$  represent base, spine and top lengths of Joint  $i$ , respectively. Therefore, referring to Fig. 2(c) and using (4) one can express  $\vec{n}_i$  by

$$\begin{aligned} \vec{n}_i &= \vec{n}_i \cos(\theta_i) + (\vec{t}_i \times \vec{n}_i) \sin(\theta_i) \\ &= (\vec{n}_{i-1} \cos(\phi_{i-1}) - \vec{t}_{i-1} \sin(\phi_{i-1})) \cos(\theta_i) \\ &\quad + (\vec{t}_{i-1} \times \vec{n}_{i-1}) \sin(\theta_i) \end{aligned} \quad (5)$$

Defining  $\vec{b}_i := \vec{t}_i \times \vec{n}_i$ , the three vectors  $\vec{t}_i$ ,  $\vec{n}_i$  and  $\vec{b}_i$  form a Cartesian coordinate frame at the base of each Joint  $i$ . Through employing (4) and (5), vector  $\vec{b}_i$  is stated as

$$\begin{aligned} \vec{b}_i &= \vec{t}_i \times \vec{n}_i = (\vec{t}_{i-1} \sin(\phi_{i-1}) - \vec{n}_{i-1} \cos(\phi_{i-1})) \sin(\theta_i) \\ &\quad + \vec{b}_{i-1} \cos(\theta_i) \end{aligned} \quad (6)$$

Also, for Joint 1 connected to the ground we have

$$\vec{t}_1 = \vec{t}_b, \quad \vec{n}_1 = \vec{n}_b, \quad \vec{b}_1 = \vec{n}_1 \cos(\theta_1) + (\vec{t}_1 \times \vec{n}_1) \sin(\theta_1) \quad (7)$$

The position of the bottom node of each joint  $\vec{r}_{b,i}$  ( $i \geq 2$ ), considering Fig. 2, (4) and (5) can be formulated as

$$\begin{aligned} \vec{r}_{b,i} &= \vec{r}_{b,i-1} + l_{b,i-1} \vec{t}_{i-1} + R_{i-1} (\vec{n}_{i-1} - \vec{n}_i) + l_{t,i-1} \vec{t}_i \\ &= \vec{r}_{b,i-1} \\ &\quad + (l_{b,i-1} + R_{i-1} \sin(\phi_{i-1}) + l_{t,i-1} \cos(\phi_{i-1})) \vec{t}_{i-1} \\ &\quad + (R_{i-1} - R_{i-1} \cos(\phi_{i-1}) + l_{t,i-1} \sin(\phi_{i-1})) \vec{n}_{i-1} \end{aligned} \quad (8)$$

with  $\vec{r}_{b,1} = \vec{r}_b$ . Given the joint radii and rotations, (4)–(8) provide the FK model of a multi-joint continuum robot that upon solving them recursively, vectors  $\vec{t}_i$ ,  $\vec{n}_i$ ,  $\vec{b}_i$  and  $\vec{r}_{b,i}$  for all the points on the robot including the end-effector are found. For a three-joint robot, as the one in Fig. 2, vector  $\vec{r}_{b,4}$  determines

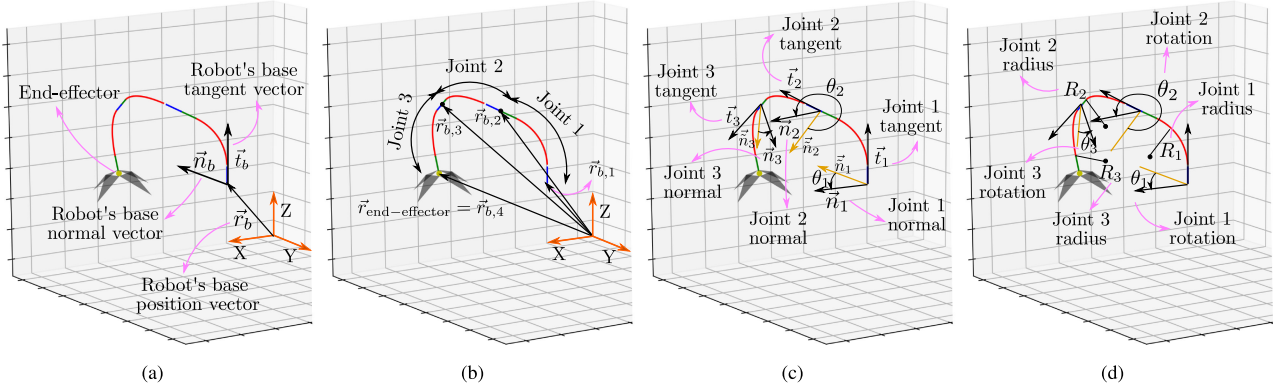


Fig. 2. Simplistic representation of a three-joint continuum robot with the proposed parametrization. (a) The tangent vector at the base defines the direction towards which the first joint is pointing at, and the normal vector determines the bending plane of that joint. (b) The tip of each joint coincides with the bottom of the subsequent joint. (c)  $\theta$  of each joint is defined as the rotation of  $\vec{n}$  about that joint's tangent  $\vec{t}$  leading to  $\vec{n}$  determining the bending plane of that joint. (d) Given each joint's radius  $R$  and  $\theta$ , the robot can be fully configured in space.

the end-effector position. Also, vectors  $\vec{t}_4$ ,  $\vec{n}_4$  and  $\vec{b}_4$  define the orientation at that location. Note that when  $R_i$  tends to infinity,  $\phi_i = l_{s,i}/R_i$  moves towards zero meaning that Joint  $i$  becomes a straight line. In this case, utilizing the aforementioned equations, we conclude

$$\begin{aligned} \vec{t}_i &= \vec{t}_{i-1}, \\ \vec{n}_i &= \vec{n}_{i-1} \cos(\theta_i) + \vec{b}_{i-1} \sin(\theta_i) \\ \vec{b}_i &= \vec{b}_{i-1} \cos(\theta_i) - \vec{n}_{i-1} \sin(\theta_i) \\ \vec{r}_{b,i} &= \vec{r}_{b,i-1} + \vec{t}_{i-1} (l_{b,i-1} + l_{s,i-1} + l_{t,i-1}) \end{aligned} \quad (9)$$

2) *Inverse Kinematics*: Inverse kinematics (IK) is the inverse of FK where the goal is to find a set of state variables such that the desired end-effector's position and orientation are achieved. In other words

$$\vec{q} = \vec{f}_{FK}^{-1}(\vec{r}_e) \quad (10)$$

where  $\vec{f}_{FK}$  is the forward kinematic equations presented earlier and  $\vec{q}$  denotes the vector of state variables (joints' radii and rotations). Since in this work only the reachability of the robot is considered, vector  $\vec{r}_e$  is a three-dimensional vector determining the end-effector's position. Referring to the previous section, for a robot containing  $m$  joints, since each joint has two DOF, the robot has  $2 \sim m$  DOF in total. This means that a robot with only two joints is a redundant manipulator, so (10) may have an infinite number of solutions. Many numerical techniques can be applied to solve (10) such as the pseudo-inverse, Jacobian transpose or damped least-square method [36]. In this work, the last approach is adopted.

A significant advantage of having redundancy in a robot is that other types of performance criteria (e.g., collision avoidance, actuator torque or energy minimization) can be incorporated in the IK routine. In this paper, minimizing the total actuator torque applied at the base of each continuum joint is selected

as the secondary performance measure. Therefore, (10) is reformulated as

$$\begin{aligned} \min_{\vec{q}} \quad & \left( \frac{1}{2} \sum_i \|\vec{\tau}_i(\vec{q})\|^2 \right) \\ \text{s.t.} \quad & \vec{f}_{FK}(\vec{q}) = \vec{r}_e \\ & \vec{l}b \leq \vec{q} \leq \vec{u}b \end{aligned} \quad (11)$$

where  $\vec{\tau}_i$  is the static actuator torque at the base of Joint  $i$ . Also,  $\vec{l}b$  and  $\vec{u}b$  specify respectively the desired lower and upper bounds for the joints' radius and rotation. The problem in (11) can be solved using a gradient-based constrained optimization method [37]. In this work, the constrained trust region algorithm implemented in Python's Scipy package is used. The objective function in (11) is the total torque required to hold a given payload at the end-effector location.

### B. Optimization Using an EDA

As stated earlier, an EDA uses a probability model to estimate and sample promising candidate solutions at each iteration. The detailed procedure of the algorithm is as follows.

1) *Generation of Initial Population*: An initial population of solutions  $\mathbf{P}$  is randomly generated at the start. A solution is a particular instantiation of the design variables  $x_i \in \mathcal{D}$  while respecting the dimensional limits per  $d_{i,lb} \leq x_i \leq d_{i,ub}$ .

2) *Evaluation and Selection*: Solutions in the population are evaluated by computing the objective and constraint functions. Here, we use a penalty method [38] to quantify the constraint violation and combine it with the objective function to formulate a single fitness function as follows

$$\text{fitness} := f(\mathcal{D}) + \sigma \max(0, \alpha - \Theta(\mathcal{D}, \mathcal{S})) \quad (12)$$

where  $\sigma$  is a penalty coefficient. Evaluated solutions are ranked based on their fitness values. Then, a subset of the population,  $\mathbf{S}$ , representing the top  $N$  solutions is selected from  $\mathbf{P}$ . The truncation rate is defined as  $|\mathbf{S}|/|\mathbf{P}|$ .

3) *Estimation of Probability Distribution*: From  $\mathbf{S}$ , the probability distribution of promising solutions are estimated. For the



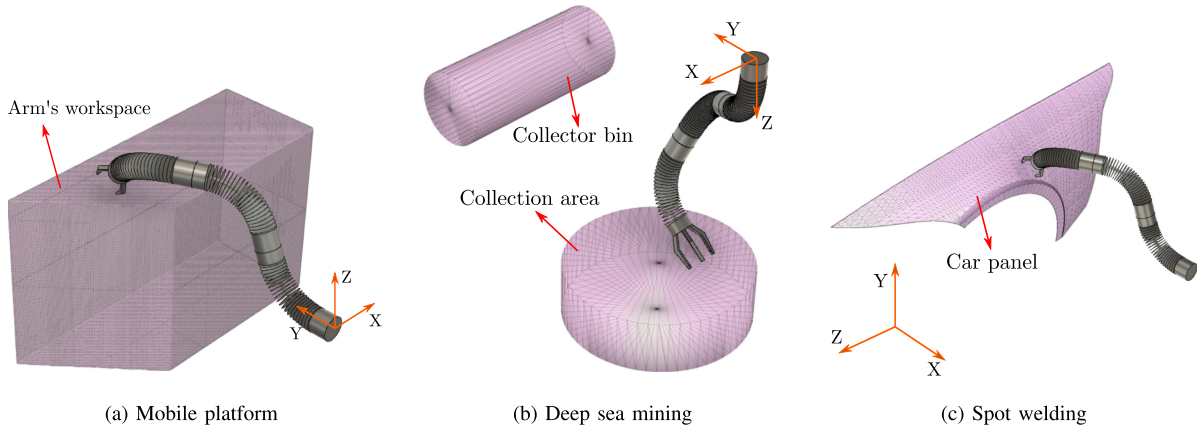


Fig. 3. The robot's location/orientation and the workspaces that must be reached for each application problem. For (a) and (b), the end-effector must reach the entire workspace volumes while for (c), it must reach a series of points along the periphery of the car panel.

current work, a univariate normal distribution is used.

$$p(\mathcal{D}) = \prod_{i=1}^{|\mathcal{D}|} p(x_i), \quad x_i \in \mathcal{D} \quad (13)$$

where  $p(x_i) \sim \mathcal{N}(\mu_i, \sigma_i^2)$ , which assumes that the probability of an optimal design can be independently computed with normally distributed  $p(x_i)$ . Hence, the estimation of the probability distribution involves computing the mean  $\mu_i$  and the variance  $\sigma_i^2$  for each variable  $x_i$ . While more complex probability distributions such as multivariate normal or Gaussian mixture models could be used [33], preliminary experiments showed that a simple model worked better for our case.

4) *New Population Generation Via Sampling*: A new population of the size  $|\mathbf{P}|$  is generated by sampling from the probability distribution estimated from the previous generation of population. This can be performed in a straightforward manner using the Box-Muller transform [39] and the values of  $\mu_i$  and  $\sigma_i^2$ . Note that if necessary, we clip the values of sampled solutions such that the dimensional limits per  $d_{i,lb} \leq x_i \leq d_{i,ub}$  in (1) are respected.

If the objective function can be computed quickly, such as the total length of the robot, we could repeat sampling until the objective value of each solution in the new population is less than the best objective value found so far. In other words, we only accept a sampled solution for the next generation if

$$f(\mathcal{D}_{\text{sampled}}) < f_{\text{best}} \quad (14)$$

We call this strategy as *select generation* and evaluate its effect on the algorithm's convergence rate in the Experiments section.

5) *Iterate Steps 2-4*: Steps 2 to 4 are repeated with the newly generated population until meeting a convergence criterion or a maximum number of iterations allocated.

## V. EXPERIMENTS

The optimization method is evaluated using three application problems. The first two are problems given to Otherlab Inc. (the last author's affiliate) by its clients while the last problem is created to test the torque minimization capability of the method.

The workspaces of each problem can be seen in Fig. 3. All dimensional units in this section are in *cm*.

### A. Application Problems

The first application is a two-arm dexterous manipulation on a mobile platform. Two arms are symmetrically mounted on the sides of a vehicle with a camera equidistant away from them along the line of symmetry. The workspace is in the camera's field of view. Fig. 3(a) represents a portion of the workspace that must be reached by one arm, while the workspace for the second arm (not shown) would be the mirror image of the workspace shown. The envelop dimensions of the workspace including the robot arm's origin are  $80 \times 32 \times 51$ . The arm's base is positioned at  $(0, 0, 0)$  with the tangent direction of the base pointing at  $(0, 1, 0)$ .

The second application is deep sea mining. The workspace can be seen in Fig. 3(b). The robot arm is on a vehicle orientated such that it is pointing towards the ground. The vehicle traverses the ocean floor and the arm picks up nodules and deposits them in a collector. For this application, we represent the workspace as two disjointed volumes; the wide short cylinder being a space above the ocean floor and the thin long cylinder being the collector. The robot only needs to be able to reach each of the volumes but not necessarily the space in between them. The envelop dimensions of the workspace including the both workspace volumes and the arm's base are  $51 \times 86 \times 61$ . The arm's origin is positioned at  $(0, 0, 0)$  with the tangent direction of the base pointing along  $(0, 0, 1)$ .

The third application is an industrial automation application in which there is a series of points that need to be welded on a car panel. The 3D rendering of the car panel can be seen in Fig. 3(c). This application differs from the previous applications because the target workspace is a series of points located on the periphery of the car panel boundary. Thirty-two evenly spaced points are identified along the boundary, with their positions in the ranges of  $([13.7, 37.7], [10.9, 70.3], [0.0, -120.8])$ . The robot's base is positioned at  $(75.0, 45.0, -70.0)$  with its tangent direction at  $(-1, 0, 0)$ .

For all three application problems, we consider a three-joint continuum robot. The following set of design variables are considered – the lengths of the first bottom base  $l_{b,1}$ , first spine  $l_{s,1}$ , second bottom base  $l_{b,2}$ , second spine  $l_{s,2}$ , third bottom base  $l_{b,3}$ , third spine  $l_{s,3}$ , and third top base  $l_{t,3}$  representing the end-effector. The imposed dimensional constraints on the design variables are

$$\begin{aligned} 4 \leq l_{b,1} \leq 30; \quad 2.675 \leq l_{s,1} \leq 32.1 \\ 4 \leq l_{b,2} \leq 30; \quad 2.173 \leq l_{s,2} \leq 26.076 \\ 4 \leq l_{b,3} \leq 30; \quad 2.173 \leq l_{s,3} \leq 26.076; \quad 36 \leq l_{t,3} \leq 60 \end{aligned} \quad (15)$$

Note that the top bases for the first and second joints are ignored and assigned zero lengths for optimization, as they are redundant design variables with respect to the bottom bases of the second and third joints.

In addition, minimum bend radius constraints must be imposed on each joint's radius of curvature ( $R_1$ ,  $R_2$  and  $R_3$ ).

$$10.22 \leq R_1; \quad 8.3 \leq R_2; \quad 8.3 \leq R_3 \quad (16)$$

A minimum bend radius is the smallest allowed radius that a joint can take while bending, due to its physical limits. These constraints are to be respected during the reachability analysis. For instance, when sampling random configurations with forward kinematics, the values of  $R_i$  should be greater than those imposed above. Also, when finding the robot configuration that can reach a particular point using inverse kinematics, the above constraints restrict the possible values of  $R_i$  considered.

For the first two problems, the objective function is the total length of the arm, i.e., the sum of all design variables, which is equivalent to having a lightweight robot. The threshold for the reachability constraint is set to 0.95, i.e., the robot must be able to reach 95% of the target points representing the workspace. For the last problem, the objective function is the total minimum actuator torque required to reach all welding points. The torque is for resisting the gravity load due to the welder device's mass (1 kg) at the end-effector and joints' mass (1 kg each) at their centroids. This torque minimization problem is solved through (11) and summing up the minimum actuator torque values  $\sum_i \|\vec{\tau}_i(\vec{q})\|$  for all the welding points. We require that all 32 welding points must be reachable by the robot, i.e., 100% reachability.

### B. Reachability Analysis

The 3D workspace for each application problem is created using CAD software and exported as an STL model. For the first two application problems, we convert their STL models into 3D voxels (capturing the interior volumes) with each voxel sized at  $3\text{cm} \times 3\text{cm} \times 3\text{cm}$ . Using forward kinematics, if the end-effector's position is within a particular voxel, we conclude that the target point represented by that voxel can be reached. As for inverse kinematics, we solve for possible robot configurations that would result in the end-effector's position coinciding with the center of each target voxel with the tolerance  $\epsilon$  of  $1\text{cm}$ . The lower tolerance is used for the inverse kinematics because more

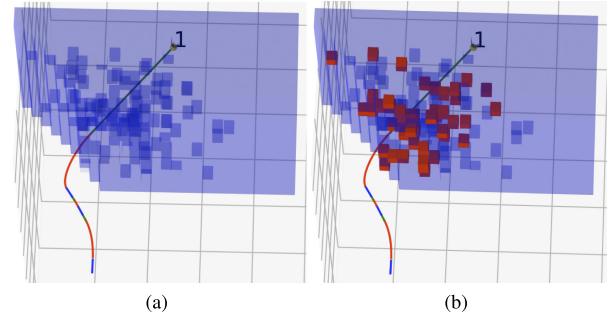


Fig. 4. Reachability analysis performed using the hybrid approach. (a) First, the reachability is estimated using forward kinematics with randomly generated configurations. (b) Then, inverse kinematics is used to check the reachability of those points not reached during the forward kinematics simulation. Dark blue boxes indicate the points deemed unreachable within the workspace by the forward kinematics simulation. Red boxes indicate the missed points confirmed as reachable by the inverse kinematics simulation.

accurate evaluation of reachability is desired for solutions that are close to the feasibility threshold.

For the first two application problems, we use the hybrid approach of incorporating both forward kinematics and inverse kinematics to estimate the reachability (illustrated in Fig. 4). First, three million robot configurations (as recommend in [22]) are generated by randomly sampling the values of  $R_i$  and  $\theta_i$ , and forward kinematics for each configuration is performed to estimate the reachability of a given candidate solution. Then, inverse kinematics is used if the estimated reachability by the forward kinematics approach does not meet but is close to the required threshold. Here, we utilize inverse kinematics to check on the points that are not reached during the forward kinematics simulation if the estimated reachability is between 0.9 and 0.95 (the required threshold). This hybrid approach prevents the algorithm from unnecessarily penalizing a feasible solution, the reachability of which could be underestimated if only the forward kinematics approach is used.

For the last application problem, the 32 welding points are identified along the boundary of the 3D car panel model. We use inverse kinematics to check whether all welding points can be reached or not, while simultaneously solving for the minimum torque required to reach all those points.

### C. Algorithms Compared

We compare the EDA based on a univariate normal distribution against a genetic algorithm (GA) for all three application problems. The GA is considered as the benchmark method because it was used in prior work to optimize a multi-joint continuum robot considering reachability [22]. The cross-over and mutation operators used for the GA are the same as those in [22], with the cross-over and mutation rates of 0.9 and 0.1, respectively. For selection, both algorithms incorporate the same truncation method with the truncation rate of 0.5.

For the first two application problems, we test two versions of each EDA/GA. The first version of each algorithm uses the standard generation procedure at each iteration. On the other hand, the second version of each algorithm incorporates the

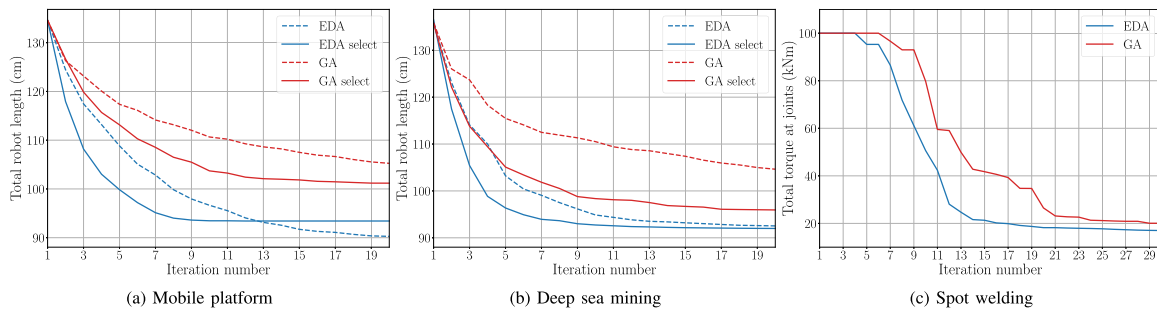


Fig. 5. Summary of optimization results for the application problems. Reported in the y-axis are the objective values of the best feasible solution at each iteration, averaged from 20 runs for each algorithm. The algorithms with “select” notation indicate the usage of select generation. In all cases, the final objective values are significantly lower for the EDA compared to the GA.

TABLE I  
FINAL OBJECTIVE VALUES FOUND FOR EACH SETTING

	Mobile platform		Deep sea mining		Spot welding	
	mean (cm)	SD	mean (cm)	SD	mean (kNm)	SD
EDA	90.3	0.64	92.5	0.69	17.0	1.8
GA	105.3	3.6	104.6	3.0	20.1	2.5
EDA-select	93.4	3.0	92	0.20		
GA-select	101.2	5.2	95.9	2.6		

select generation strategy described in the Methods section. The strategy involves accepting new solutions – sampled from probability distributions for the EDA vs. created using genetic operators for the GA – only if their objective value is smaller than the best objective value found hitherto. Select generation is repeated until the required population size is met or the maximum number of trials (=10 000) is reached, at which point the standard generation procedure is used to fill the remaining population.

For the last application problem, the evaluation of the objective function requires computing inverse kinematics for all target points, which is much more expensive than computing the total length as in the first two problems. Hence, we do not examine the select generation strategy and compare the EDA against the GA only with the standard generation procedure.

For all algorithms, the population size of 100 is used. The maximum iteration number allocated is 20 for the first two problems and 30 for the last one. The penalty coefficient used for the reachability constraint is 0.33 for all three problems. Each algorithm is repeated 20 times, starting with the same randomly generated initial population for each run. For the results, the average objective values of the best feasible solution (i.e., the solution that meets the reachability constraint with the lowest objective value) at each iteration are reported.

#### D. Results

Fig. 5 summarizes the optimization results and Table I reports the mean objective values and their standard deviations found at the end of optimization runs. For all three application problems, the EDA outperforms the GA in terms of the quality of the best solutions found at the end. The final objective values found are significantly lower with the EDA than the GA ( $p < 0.0001$  for all cases, including with or without select generation). Note that

the standard deviations of the final objective values found by the EDAs are very low, indicating a convergence to optimal solutions at the end.

In general, the select generation strategy seems to improve the convergence rate for both the EDA and the GA, as shown in Fig. 5. However, for the first application problem, it causes the EDA to get stuck in a local minimum and the version with standard generation finds better solutions in the end ( $p < 0.0001$ ) than the version with select generation. For the second application problem, the select generation strategy resulted in significantly lower final objective values for the GA compared to the standard generation strategy ( $p < 0.0001$ ).

In terms of computation time, each optimization run takes about 3.5 hours for the first two problems and 1.7 hours for the last problem, on a single desktop computer with two Intel Xeon CPUs (E5-2650 v2 2.60 GHz) and 32 GB of RAM. Note that because the EDA/GA is a population-based algorithm, we could run 16 evaluations in parallel during optimization runs. The computation time for both the EDA and the GA are roughly the same because the majority of the time is spent on evaluating solutions vs. other procedures in the algorithms.

In summary, the overall results of our experiments demonstrate the effectiveness of the EDA compared to the GA in finding optimal designs of continuum robots considering reachability constraints. In addition, the select sampling strategy has shown the potential in improving the convergence rate, although it could lead to premature convergence in some cases.

## VI. SUMMARY AND CONCLUSIONS

The current work has presented a computational method to find optimal designs of continuum robots while considering reachability constraints. To assess the reachability of a given robot design, the method takes advantage of both forward kinematic and inverse kinematic approaches. The former with randomly sampled robot configurations is used to quickly estimate the reachability, while the latter is used to further assess the reachability so that a feasible solution could be more accurately identified during optimization.

In addition, our implementation of inverse kinematics can incorporate the minimization of secondary performance criteria, specifically in the current work the minimum torque required to reach a target point. This capability allows us to find an optimal



design with the total minimum actuator torque required to reach the workspace.

Lastly, the optimization method is based on the estimation of distribution algorithm (EDA), a population-based, derivative-free optimization method that uses a univariate marginal distribution to estimate and sample promising candidate solutions. It also features a penalty method to handle reachability constraints and a select generation strategy to increase the convergence rate. Through three application problems, the EDA is shown to be superior to the genetic algorithm implemented in finding better solutions within a given number of iterations. In practice, the method could find optimal solutions in 2–4 hours rather than the typical 1–2 weeks taken for the manual work performed by an engineering team. This would drastically decrease the overall design time and also allow the team to consider a greater number of alternatives for the final design.

Future work includes considering other performance criteria of continuum robots, such as their dexterity and manipulability. The scalability of the proposed method should be investigated as well, e.g., optimizing robots with more joints. Also, other derivative-free optimization algorithms or reinforcement learning techniques can be explored for solving the problem. Lastly, the current method using the EDA could be extended to solve other optimal design problems found in robotics.

## REFERENCES

- [1] M. Pelikan and H. Mühlenbein, “The bivariate marginal distribution algorithm,” in *Proc. Adv. Soft Comput.*, 1999, pp. 521–535.
- [2] S. Shakya and J. McCall, “Optimization by estimation of distribution with deum framework based on markov random fields,” *Int. J. Automat. Comput.*, vol. 4, no. 3, pp. 262–272, 2007.
- [3] S. Shakya, R. Santana, and J. A. Lozano, “A markovianity based optimisation algorithm,” *Genet. Program. Evolvable Machines*, vol. 13, no. 2, pp. 159–195, 2012.
- [4] M. Martins, M. El Yafrani, R. Santana, M. Delgado, R. Lüders, and B. Ahiod, “On the performance of multi-objective estimation of distribution algorithms for combinatorial problems,” in *Proc. IEEE Congr. Evol. Comput.*, 2018, pp. 1–8.
- [5] H. Cheong, M. Ebrahimi, A. Butscher, and F. Iorio, “Configuration design of mechanical assemblies using an estimation of distribution algorithm and constraint programming,” in *Proc. IEEE Congr. Evol. Comput.*, 2019, pp. 2339–2346.
- [6] D. Trivedi, C. D. Rahn, W. M. Kier, and I. D. Walker, “Soft robotics: Biological inspiration, state of the art, and future research,” *Appl. Bionics Biomech.*, vol. 5, no. 3, pp. 99–117, 2008.
- [7] A. A. Transeth, K. Y. Pettersen, and P. Liljeback, “A survey on snake robot modeling and locomotion,” *Robotica*, vol. 27, no. 7, pp. 999–1015, 2009.
- [8] R. J. Webster III and B. A. Jones, “Design and kinematic modeling of constant curvature continuum robots: A review,” *Int. J. Robot. Res.*, vol. 29, no. 13, pp. 1661–1683, 2010.
- [9] M. W. Hannan and I. D. Walker, “Kinematics and the implementation of an elephant’s trunk manipulator and other continuum style robots,” *J. Robot. Syst.*, vol. 20, no. 2, pp. 45–63, 2003.
- [10] B. A. Jones and I. D. Walker, “Kinematics for multisection continuum robots,” *IEEE Trans. Robot.*, vol. 22, no. 1, pp. 43–55, Feb. 2006.
- [11] S. Neppalli and B. A. Jones, “Design, construction, and analysis of a continuum robot,” in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2007, pp. 1503–1507.
- [12] A. D. Kapadia and I. D. Walker, “Self-motion analysis of extensible continuum manipulators,” in *Proc. IEEE Int. Conf. Robot. Automat.*, 2013, pp. 1988–1994.
- [13] A. Garriga-Casanovas and F. Rodriguez y Baena, “Kinematics of continuum robots with constant curvature bending and extension capabilities,” *J. Mechanisms Robot.*, vol. 11, no. 1, 2019, Art. no. 011010.
- [14] T. F. Allen, L. Rupert, T. R. Duggan, G. Hein, and K. Albert, “Closed-form non-singular constant-curvature continuum manipulator kinematics,” in *Proc. 3rd IEEE Int. Conf. Soft Robot.*, 2020, pp. 410–416.
- [15] C. Bedell, J. Lock, A. Gosline, and P. E. Dupont, “Design optimization of concentric tube robots based on task and anatomical constraints,” in *Proc. IEEE Int. Conf. Robot. Automat.*, 2011, pp. 398–403.
- [16] T. Anor, J. R. Madsen, and P. Dupont, “Algorithms for design of continuum robots using the concentric tubes approach: A neurosurgical example,” in *Proc. IEEE Int. Conf. Robot. Automat.*, 2011, pp. 667–673.
- [17] J. Burgner, H. B. Gilbert, and R. J. Webster, “On the computational design of concentric tube robots: Incorporating volume-based objectives,” in *Proc. IEEE Int. Conf. Robot. Automat.*, 2013, pp. 1193–1198.
- [18] M. T. Chikhaoui, J. Granna, J. Starke, and J. Burgner-Kahrs, “Toward motion coordination control and design optimization for dual-arm concentric tube continuum robots,” *IEEE Robot. Automat. Lett.*, vol. 3, no. 3, pp. 1793–1800, Jul. 2018.
- [19] J. Hiller and H. Lipson, “Automatic design and manufacture of soft robots,” *IEEE Trans. Robot.*, vol. 28, no. 2, pp. 457–466, Apr. 2012.
- [20] G. Runge, J. Peters, and A. Raatz, “Design optimization of soft pneumatic actuators using genetic algorithms,” in *Proc. IEEE Int. Conf. Robot. Biomimetics*, 2017, pp. 393–400.
- [21] N. Cheney, J. Bongard, and H. Lipson, “Evolving soft robots in tight spaces,” in *Proc. Genet. Evol. Comput. Conf.*, 2015, pp. 935–942.
- [22] D. M. Bodily, T. F. Allen, and M. D. Killpack, “Multi-objective design optimization of a soft, pneumatic robot,” in *Proc. IEEE Int. Conf. Robot. Automat.*, 2017, pp. 1864–1871.
- [23] P. Larrañaga, “A review on estimation of distribution algorithms,” in *Estimation Distrib. Algorithms*. Berlin, Germany: Springer, 2002, pp. 57–100.
- [24] M. Hauschild and M. Pelikan, “An introduction and survey of estimation of distribution algorithms,” *Swarm Evol. Comput.*, vol. 1, no. 3, pp. 111–128, 2011.
- [25] S. Gao and C. W. de Silva, “Estimation distribution algorithms on constrained optimization problems,” *Appl. Math. Comput.*, vol. 339, pp. 323–345, 2018.
- [26] T. J. Yuen, R. Rahizar, Z. A. M. Azman, A. Anuar, and D. Afandi, “Design optimization of full vehicle suspension based on ride and handling performance,” in *Proc. FISITA World Automot. Congr.*, 2013, pp. 75–86.
- [27] P. Simionescu, D. Beale, and G. V. Dozier, “Teeth-number synthesis of a multispeed planetary transmission using an estimation of distribution algorithm,” *J. Mech. Des.*, vol. 128, no. 1, pp. 108–115, 2006.
- [28] C. Piacentini, H. Cheong, M. Ebrahimi, and A. Butscher, “Multi-speed gearbox synthesis using global search and non-convex optimization,” in *Proc. Int. Conf. Integration Constraint Program., Artif. Intell. Operations Res.*, 2020, pp. 381–398.
- [29] S. Gao and C. W. de Silva, “A univariate marginal distribution algorithm based on extreme elitism and its application to the robotic inverse displacement problem,” *Genet. Program. Evolvable Machines*, vol. 18, no. 3, pp. 283–312, 2017.
- [30] L. Hu and C. Zhou, “Gait generation and optimization using the estimation of distribution algorithm for teensize humanoid soccer robot resr-1,” *Int. J. Humanoid Robot.*, vol. 5, no. 3, pp. 437–456, 2008.
- [31] M. Jiang, Z. Huang, G. Jiang, M. Shi, and X. Zeng, “Motion generation of multi-legged robot in complex terrains by using estimation of distribution algorithm,” in *Proc. IEEE Symp. Ser. Comput. Intell.*, 2017, pp. 1–6.
- [32] E. Hernandez, S. I. Valdez, G. Carbone, and M. Ceccarelli, “Design optimization of a cable-driven parallel robot in upper arm training-rehabilitation processes,” in *Proc. Int. Symp. Multibody Syst. Mechatronics.*, 2017, pp. 413–423.
- [33] P. Larrañaga, R. Etxeberria, J. A. Lozano, and J. M. Pena, “Optimization in continuous domains by learning and simulation of gaussian networks,” in *Proc. Genet. Evol. Comput. Conf.*, 2000, pp. 201–204.
- [34] M. Ebrahimi, A. Butscher, and H. Cheong, “A low order, torsion deformable spatial beam element based on the absolute nodal coordinate formulation and bishop frame,” *Multibody Syst. Dyn.*, vol. 51, pp. 247–278, 2020.
- [35] A. Shabana, *Dynamics of Multibody Systems*. Cambridge, U.K.: Cambridge University Press, 2020.
- [36] S. R. Buss, “Introduction to inverse kinematics with jacobian transpose, pseudoinverse and damped least squares methods,” *IEEE J. Robot. Automat.*, vol. 17, no. 1–19, pp. 681–685, Oct. 2004.
- [37] M. Ebrahimi, A. Butscher, H. Cheong, and F. Iorio, “Design optimization of dynamic flexible multibody systems using the discrete adjoint variable method,” *Comput. Structures*, vol. 213, pp. 82–99, 2019.
- [38] A. E. Smith and D. W. Coit, “Penalty functions,” in *Handbook of Evolutionary Computation*, 1997, vol. 97, no. 1, ch. 5.2, pp. 1–6.
- [39] G. E. Box, “A note on the generation of random normal deviates,” *Ann. Math. Stat.*, vol. 29, pp. 610–611, 1958.