

# Task-Adaptive Analytical Affordance Estimation for Feature-Based Manipulation of Soft Tissues in Robotic Surgery

Sihang Yang<sup>1</sup>, Yuheng Wang<sup>1</sup>, Yiwei Wang<sup>1,2,3\*</sup>, Huan Zhao<sup>1,3</sup>, and Han Ding<sup>1,3</sup>

**Abstract**—Robotic soft tissue manipulation in surgery presents significant challenges due to the tissue’s high deformability and the spatial constraints of the surgical environment. While data-driven methods for affordance estimation are common, they often face challenges with data requirements and generalization in surgical scenarios. To address these challenges, a novel framework is proposed that combines a deformation model-based shape controller with an analytical affordance estimation approach for multi-contact scenarios by employing a manipulability metric. The method leverages the deformation Jacobian matrix derived from a linearized deformation model to evaluate the manipulability of candidate multi-contact points, providing a robust and data-efficient solution. For tissue manipulation, a differentiable deformation model is employed to efficiently compute forward and backward deformation processes in real-time. Point-based visual features are constructed to represent and track the tissue deformation, enabling precise control through visual feedback. The proposed framework has been validated through simulations and physical experiments. These results demonstrate the real-time ( $\sim 60\text{Hz}$ ) ability to achieve targeted configuration with high accuracy ( $< 1\text{mm RMSE}$  for each marker position) and a strong correlation (near-zero p-value) between the predicted affordance and the observed manipulation efficiency, confirming the effectiveness of our approach in soft tissue manipulation and affordance estimation.

**Note to Practitioners**— This paper addresses a critical challenge in automation: how to enable robots to intelligently and efficiently manipulate soft, deformable objects like biological tissue. For engineers and roboticists working on applications in robotic surgery or manufacturing with compliant materials, the conventional approach of using data-driven, machine learning models often proves impractical due to the immense data requirements and poor generalization to new objects or scenarios. This work presents a practical, physics-based alternative that significantly reduces the reliance on large datasets. The core contribution is a framework that allows a robot to analytically determine the optimal multi-contact grasp points on a soft object to achieve a targeted shape. This score directly predicts how efficiently a given grasp point location will deform the object, allowing the robot to make intelligent choices. The key benefits of this approach for practitioners are: 1) It bypasses the need for costly

This work was supported in part by the National Science Foundation of China under Grant 62203180, in part by the Major Program (JD) of Hubei Province (2023BAA016), in part by Hubei Science and Technology Major Program under Grant 2023BCA002 and 2024BCB009, and in part by the Taihu Lake Innovation Fund for Future Technology, HUST: 2023-A-2. (Corresponding author: Yiwei Wang)

<sup>1</sup>The authors are with the State Key Laboratory of Intelligent Manufacturing Equipment and Technology, Huazhong University of Science and Technology, Wuhan 430074, China sihang@hust.edu.cn; wang\_yuheng2000@163.com; wang\_yiwei@hust.edu.cn; huanzhao@hust.edu.cn; dinghan@hust.edu.cn

<sup>2</sup>Yiwei Wang is with the School of Medical Equipment Science and Engineering, Huazhong University of Science and Technology, Luoyu Road 1037, Wuhan 430074, China.

<sup>3</sup>Yiwei Wang, Huan Zhao, and Han Ding are with the School of Future Technology, Huazhong University of Science and Technology, Luoyu Road 1037, Wuhan 430074, China.

and time-consuming data collection and annotation for training machine learning models. 2) The model-based approach provides robust and predictable performance, making it easier to validate and deploy in critical applications like surgery. 3) The system is demonstrated to be fast ( $\sim 60\text{Hz}$ ) and precise (achieving sub-millimeter accuracy), meeting the demands of real-world manipulation tasks. Ultimately, this research provides a blueprint for creating more autonomous and efficient robotic systems for handling deformable objects, moving beyond data-intensive methods toward more robust, physics-informed solutions.

**Index Terms**—Medical Robots and Systems, Surgical Robotics: Laparoscopy, Dexterous Manipulation, Surgical Robotics: Planning

## I. INTRODUCTION

Robotic-assisted surgery has transformed minimally invasive procedures by enhancing dexterity, precision, and visualization [1]. Central to leveraging these advantages is the dexterous manipulation of soft tissues, a task that is ubiquitous in interventions ranging from retracting the liver in laparoscopy to repositioning delicate tissues in neurosurgery [2]. However, achieving human-like dexterity manipulation in these tasks remains a formidable challenge [3]. The irregular geometry, high deformability, and nonlinear properties of soft tissues pose significant challenges for robotic control [4], especially within the constrained surgical workspace.

Achieving task-oriented goals in soft tissue manipulation presents two coupled challenges: selecting an optimal contact point and precisely controlling the tissue’s shape. The choice of contact location is critical; a well-chosen point enables efficient deformation towards the desired shape, whereas a poor choice can lead to inefficient manipulation or even iatrogenic damage [5]. Furthermore, rather than attempting to control the tissue’s full shape, clinical objectives often depend on the precise positioning of specific anatomical features, such as aligning incision edges or exposing a blood vessel [6]. This feature-based approach requires a control algorithm that can manage the tissue’s complex, nonlinear dynamics in real time, a task for which purely data-driven methods are often unsuitable due to the need for extensive patient-specific training data [7].

To address these challenges, a novel two-layer framework that integrates an affordance-based planner for contact selection with a model-based controller for feature-based manipulation is proposed, as illustrated in Fig. 1. For the first layer, the planner builds upon the concept of affordance [8] to formalize contact point selection. Originating from cognitive science, affordance describes the action possibilities an object offers—for example, a doorknob affords turning—and has been widely

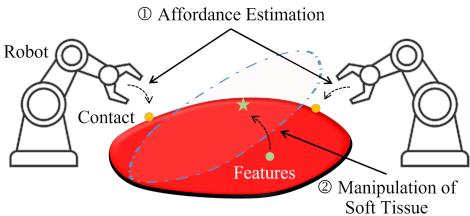


Fig. 1: Schematic illustrating the process of affordance estimation and feature-based manipulation of soft tissue.

applied to various rigid body manipulation tasks [9]. Building on this concept, the planner introduces a task-adaptive affordance metric for soft tissue manipulation inspired by the robotics concept of manipulability. This metric, formulated using the deformation Jacobian, quantifies how effectively an action at a specific point can induce desired shape changes, allowing for the selection of optimal contacts that satisfy both task-specific requirements and physical constraints. Second, this planner guides a real-time shape controller built upon a differentiable Projective Dynamics (PD) model within a Model Predictive Control (MPC) framework. The differentiability of the PD model is key, as it permits the analytical, real-time computation of the deformation Jacobian, enabling the controller to precisely map robot actions to tissue deformation without relying on offline training. The proposed framework is demonstrated on sparse feature point control but is directly extensible to full shape control. The primary contributions of this work are:

- 1) A novel analytical method for robust and interpretable affordance estimation in task-oriented manipulation of soft tissue under multi-contact scenarios, utilizing manipulability metrics from the deformation Jacobian.
- 2) The integration of the differentiable PD model for efficient real-time soft tissue shape control, bypassing the need for extensive pre-training.
- 3) Comprehensive validation of the proposed framework through simulations and experimental studies, demonstrating its efficiency in completing task-specific tissue manipulations and affordance estimation with high accuracy.

The remainder of this paper is structured as follows: Section II summarizes the related work. Section III details the proposed control framework, deformation model, and manipulability-based affordance estimation. Sections IV and V present simulation analyses and experimental results, validating the method's accuracy and efficiency in 2D and 3D scenarios. Finally, Section VI concludes with a summary and directions for future research. The program code is publicly available at: <https://anonymous.4open.science/r/Deformable-Object-Manipulation-Affordance-2025-2422>.

## II. RELATED WORK

### A. Affordance Estimation

Current methods largely focus on rigid objects, using geometric features [9], grasp quality metrics [10], and friction cone analysis [11] to identify optimal grasps. However, these techniques are ill-suited for deformable objects due to complex contact interactions and deformability.

Data-driven approaches have gained traction [12], [13], learning affordance from visual [14] or tactile inputs [15] to predict grasp locations or manipulation strategies. However, applying these methods to real-time surgical settings is particularly challenging. For soft tissues, the infinite degrees of freedom greatly expand the state-action space, making data-driven learning intractable. Furthermore, collecting diverse, high-quality, and annotated surgical data for training is often clinically and ethically infeasible [16].

Recent research has explored the affordance for soft object manipulation, with existing methods analytically defining affordance based on task requirements. For instance, Han et al. [17] investigated contact stability and manipulation efficiency using a simplified deformable object model. Huang et al. [18] focused on the challenges of non-fixed robot-deformable object contacts. Liang et al. [19] explored maximizing independently controllable deformation directions during soft tissue manipulation. Cuiral-Zueco et al. [20] applied multi-scale centroid path error vectors relative to the target shape to define the optimal positions of grippers.

While existing research provides valuable insights, its applicability is often limited by a reliance on predefined target shapes. This approach typically fails to holistically integrate the interdependent factors of contact conditions, tissue deformation, and control objectives. In contrast, the proposed work addresses this limitation by introducing a framework that derives task-adaptive affordances, establishing an analytical link between control requirements, contact interactions, and tissue deformation in task space.

### B. Manipulation of Deformable Object

The manipulation of deformable objects is a significant challenge due to high dimensionality and complex dynamics. The literature can be broadly categorized into two main paradigms: full shape control and task-oriented manipulation.

For full shape control, these methods often leverage dense visual information and sophisticated models to track and manipulate the full object geometry. Recent approaches have used techniques like ARAP servoing [21], lattice-based tracking [22], and surface mapping [23] in controlling the overall shape of elastic objects. While powerful, these methods can be computationally intensive and often rely on having a complete view of the object's surface.

In contrast, task-oriented manipulation focuses on achieving a specific functional goal by controlling a smaller, more targeted set of features. This is often more practical in constrained environments like surgery. Many of these methods are adaptive: Navarro et al. [24] used contour-based visual servoing, while Han et al. [25] adapted model parameters directly from visual feedback. While often computationally efficient, purely adaptive methods can sometimes converge slowly and may lack stability guarantees. Model-based approaches have also been used for both paradigms. Yang et al. [26], [27] applied modal analysis for boundary control, while Afshar et al. [28] combined tissue simulation with a Kalman Filter for precise internal point tracking in brachytherapy. Despite offering precision and stability, these methods face trade-offs between accuracy and computational efficiency.

To overcome these limitations, a model-based framework is presented that is specifically designed for task-oriented surgical applications, which is characterized by a differentiable model that offers both computational efficiency and accurate deformation prediction, enabling real-time control.

### C. Deformation Model

Accurate and efficient modeling of soft tissue deformation is crucial for both realistic surgical simulation and effective robotic manipulation. The literature presents a well-established trade-off in this domain. On one hand, the Finite Element Method (FEM) offers high physical fidelity, but its computational cost is prohibitive for real-time applications [29]. On the other hand, methods like the Mass-Spring Model are computationally efficient but lack physical realism and correlate poorly with biomechanical material parameters [30]. Alternatively, data-driven approaches, such as Interaction Networks, can learn complex deformation behaviors from data but struggle to generalize to novel surgical scenes [7], [31].

Despite these valuable advancements, a persistent challenge remains: the formulation of a deformation model that is simultaneously physically-plausible, fully differentiable, and computationally tractable for real-time control applications. Our method achieves this by uniquely leveraging a differentiable, physics-based deformation model that is fast enough for real-time control, with its inherent inaccuracies compensated for by a closed-loop visual feedback system.

## III. METHOD

### A. Overview

*1) Problem Statement:* This article addresses the problem of task-oriented tissue manipulations under unknown material parameters and unspecified contact. The soft tissue is discretized into a mesh consisting of small elements. The model update step time is denoted by  $\Delta t$ . The vertices of these elements form a node system, representing the shape and structure of the soft tissue. These nodes' positions are denoted as  $\mathbf{q} \in \mathbb{R}^{3n}$ , where  $n$  represents the total number of nodes. Meanwhile, this study assumes that the tissue is manipulated using a position-controlled robot tool. The contact points are fixed once selected, and the applied action is defined as the change in the position of the contact points, denoted as  $\mathbf{a} = \Delta \mathbf{q}^c \in \mathbb{R}^{3l}$ , where  $l$  is the number of contact points.

Building upon the approach [32], the position of an arbitrary node  $i$  on the soft tissue at time step  $k_0 + k_1$ , denoted as  $q_i(k_0 + k_1)$ , results from actions applied at the contact points at time step  $k_0$ . These actions are assumed to remain constant between time steps  $k_0$  and  $k_0 + 1$ , enabling the application of impulse dynamics to capture the resulting motion:

$$q_i(k_0 + k_1) = q_i(k_0) + \mathbf{J}_i(k_0) \Delta \mathbf{q}^c(k_0) + \chi_i(k_0 + k_1), \quad (1)$$

where  $\Delta \mathbf{q}^c(k_0)$  represents the applied actions. The term  $\mathbf{J}_i(k_0) \Delta \mathbf{q}^c(k_0)$  corresponds to the stable deformation of the soft tissue in response to the applied actions.  $\chi_i(k_0 + k_1)$  captures the transient deformation component, which diminishes over time, as  $\lim_{k_1 \rightarrow \infty} \chi_i(k_0 + k_1) = 0$ . Under the assumption of quasi-static deformation, the transition part  $\chi_i(k_0 + k_1)$  can

be neglected. The deformation Jacobian matrix  $\mathcal{J}(k_0)$  at time step  $k_0$  is defined as:

$$\mathcal{J}(k_0) = [\mathbf{J}_1(k_0), \dots, \mathbf{J}_n(k_0)]^T, \quad (2)$$

where  $\mathbf{J}_i(k_0) \in \mathbb{R}^{3 \times 3l}$  is the  $i$ -th row block of  $\mathcal{J}(k_0) \in \mathbb{R}^{3n \times 3l}$ , capturing the relationship between the position change of all nodes and input actions.

Unlike conventional deformation control tasks that rely on predetermined contact points, our approach tackles the challenge of identifying suitable contact points by explicitly evaluating the affordance of candidate contact points. The proposed manipulability-based affordance estimation framework is tailored for multi-contact scenarios, enabling the selection of optimal contact points for control requirements.

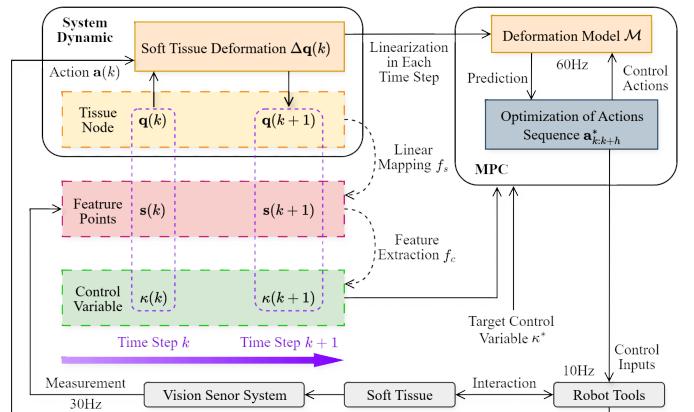


Fig. 2: The control framework and system dynamics.

*2) The Control Framework & System Dynamics:* The system dynamics governing the soft tissue deformation are depicted in Fig. 2.  $\mathbf{q}(k)$  denote the nodal positions representing the soft tissue shape at time step  $k$ , while  $\mathbf{s}(k) \in \mathbb{R}^{3e}$  denote the spatial coordinates of  $e$  designated feature points on the soft tissue surface. These feature points are tracked in real-time via a vision system using reflective markers affixed to the tissue surface. To capture the deformation characteristics relevant to the task, a control variable  $\kappa(k) \in \mathbb{R}^o$  is defined as a low-dimensional feature representation constructed from  $\mathbf{s}(k)$ . This compact representation enables efficient and effective feedback control. The overall system evolution can be expressed in state-space form as:

$$\begin{aligned} \mathbf{q}(k+1) &= \mathcal{M}(\mathbf{q}(k), \mathbf{a}(k)), \quad \mathbf{q}(0) = \mathbf{q}_{init} \\ \kappa(k) &= \Phi(\mathbf{q}(k)), \end{aligned} \quad (3)$$

where  $\mathcal{M}(\cdot)$  denotes the deformation model used to update tissue configuration under action  $\mathbf{a}(k)$ , and  $\Phi(\cdot)$  provides a reduced-order representation of the full nodal state  $\mathbf{q}(k)$ , which serves as the control variable  $\kappa(k)$ .

In practice,  $\kappa(k)$  is obtained from vision-tracked feature points via a two-stage mapping process. First, the feature positions  $\mathbf{s}(k)$  are interpolated from the mesh nodes  $\mathbf{q}(k)$  using barycentric weights:

$$s_j(k) = f_s(\mathbf{q}(k)) = \sum_i \varphi_i q_i, \quad j = 1, \dots, e, \quad (4)$$

where  $\sum_i \varphi_i = 1$  and  $\varphi_i \geq 0$ .  $q_i \in \mathbb{R}^3$  are the positions of the neighboring nodes influencing point  $s_j$ . This mapping ensures geometric consistency and differentiability across the mesh.

Second, the control variable is abstracted from the feature points using a geometric descriptor function  $f_c(\cdot)$ :

$$\kappa(k) = f_c(\mathbf{s}(k)). \quad (5)$$

During control execution, the measured feature positions  $\hat{\mathbf{s}}(k)$  yield an observed control variable  $\hat{\kappa}(k) = f_c(\hat{\mathbf{s}}(k))$ , which is used as real-time feedback for MPC. By formulating the system in this hierarchical structure—from mesh nodes to feature positions to control variables—it efficiently integrates visual feedback into a model-based control loop.

To achieve the targeted shape, we formulate a predictive control framework that optimizes the input sequence over a finite horizon  $h$ . The dynamics of the soft tissue are governed by a deformation model  $\mathcal{M}(\cdot)$ , and the deformation objective is expressed via  $\kappa(k) = \Phi(\mathbf{q}(k))$ , as introduced previously. The optimal action sequence  $\mathbf{a}_{k:k+h}^* = [\mathbf{a}^*(k), \dots, \mathbf{a}^*(k+h)] \in \mathbb{R}^{3l(h+1)}$  is computed by solving the following constrained optimization problem:

$$\begin{aligned} \mathbf{a}_{k:k+h}^* &= \arg \min_{\mathbf{a}_{k:k+h}} \sum_{\tilde{k}=k}^{k+h} \alpha^{\tilde{k}-k} \mathcal{L}\left(\Phi\left(\mathcal{M}(\mathbf{q}(\tilde{k}), \mathbf{a}(\tilde{k}))\right), \kappa^*\right) \\ \text{s.t. } \mathbf{q}(0) &= \mathbf{q}_{init} \\ \|\mathbf{a}(\tilde{k})\|_2 &\leq a_{max} \\ \varepsilon_{min} &\leq \varepsilon_i(k) \leq \varepsilon_{max}, \end{aligned} \quad (6)$$

where only the first action  $\mathbf{a}^*(k)$  from the solution sequence is applied to the controller.  $\kappa^*$  denotes the target deformation control variable. The discount factor  $\alpha < 1$  ensures that future errors are weighted less heavily. The constraint  $\|\mathbf{a}(\tilde{k})\|_2 \leq a_{max}$  limits the actuation magnitude for safety. In addition, the element-wise strain  $\varepsilon_i(k)$ , derived from the deformation model, is constrained within physiologically safe bounds  $[\varepsilon_{min}, \varepsilon_{max}]$  to prevent over-compression or over-stretching of the tissue.

The objective function  $\mathcal{L}(\cdot)$  quantifies the deviation between the current and target control variable as:

$$\mathcal{L}(k) = \frac{1}{2} (\kappa(k) - \kappa^*)^T \mathbf{Q} (\kappa(k) - \kappa^*), \quad (7)$$

where  $\mathbf{Q} \succeq 0$  is a symmetric positive semi-definite weighting matrix that encodes task-specific priorities. The action penalty term is discarded to ensure faster convergence to the targeted shape outcomes. Potential excessive actions are managed by the explicit constraints on actuation magnitude in the MPC formulation. To address the MPC problem denoted by (6), a linearized deformation model is utilized. The specifics of this model are detailed in the subsequent section, and its discrete-time state evolution is governed by:

$$\mathbf{q}(k+1) = \mathcal{J}(k)\mathbf{q}(k). \quad (8)$$

### B. The Deformation Model of Soft Tissue

1) *Dynamics of Soft Tissue Deformation:* Soft tissue is represented by a discrete element mesh. The state of the system at time step  $k$  is defined by the nodal positions  $\mathbf{q}(k) \in \mathbb{R}^{3n}$  and

velocities  $\mathbf{v}(k) \in \mathbb{R}^{3n}$ . The temporal evolution of the states is governed by Newtonian mechanics, with numerical integration performed using the implicit Euler method:

$$\begin{aligned} \mathbf{q}(k+1) &= \mathbf{q}(k) + \Delta t \cdot \mathbf{v}(k+1) \\ \mathbf{v}(k+1) &= \mathbf{v}(k) + \Delta t \cdot \mathbf{M}^{-1}(\mathbf{f}_{int}(\mathbf{q}(k+1)) + \mathbf{f}_{ext}(k+1)), \end{aligned} \quad (9)$$

where  $\mathbf{M}$  denotes the mass matrix.  $\mathbf{f}_{int}(\cdot) \in \mathbb{R}^{3n}$  represents the internal forces arising from tissue deformation, and  $\mathbf{f}_{ext} \in \mathbb{R}^{3n}$  represents the external forces applied to the tissue.

Based on the minimal deformation energy principle, the update equations in (9) can be reformulated as the optimization problem:

$$\begin{aligned} \mathbf{q}(k+1) &= \min_{\mathbf{q}} g(\mathbf{q}) \\ g(\mathbf{q}) &= \frac{1}{2\Delta t^2} \|\mathbf{M}^{\frac{1}{2}}(\mathbf{q} - \mathbf{y}(k))\|_F^2 + \sum_i E_i(\mathbf{q}) \\ \mathbf{f}_{int}(\mathbf{q}) &= -\sum_i \nabla E_i(\mathbf{q}), \end{aligned} \quad (10)$$

where  $\mathbf{y}(k) = \mathbf{q}(k) + \Delta t \mathbf{v}(k)$  represents the predicted unconstrained position.  $\|\cdot\|_F$  is the Frobenius norm. Each  $E_i(\mathbf{q})$  corresponds to a local deformation energy function, which accounts for internal forces arising from tissue deformation. The matrix  $\mathbf{M}^{1/2}$  is defined such that  $(\mathbf{M}^{1/2})^2 = \mathbf{M}$ , with element-wise square roots.

For contact nodes subjected to external forces, the unconstrained position update is given by  $\mathbf{y}_c(k) = \mathbf{q}^c(k) + \Delta t^2 \mathbf{M}_c^{-1} \mathbf{f}_{ext}(k+1)$ , where  $\mathbf{M}_c \in \mathbb{R}^{3l \times 3l}$  selects the contact-point-relevant components from the mass matrix.

To solve (10) efficiently, the PD method [33] is adopted, which is a fast simulation technique well-suited for real-time applications since it produces physically plausible and stable deformations without the high cost of traditional FEM. A key advantage of PD is that it is derived from continuous energy functions rather than heuristic constraints, as in mass-spring or position-based dynamics models. This allows for a physically consistent modeling of soft tissue governed by standard material parameters (e.g., Young's Modulus and Poisson's ratio), avoiding the non-intuitive, object-specific stiffness tuning required by methods with a less direct physical basis.

Each energy term  $E_i(\cdot)$  in PD is reformulated as a projection-based distance minimization:

$$\begin{aligned} E_i(\mathbf{q}) &= W_i(\mathbf{q}, \mathbf{p}^*) \\ \mathbf{p}^* &= \arg \min_{\mathbf{p}} W_i(\mathbf{q}, \mathbf{p}) = d_i(\mathbf{q}, \mathbf{p}) + \delta_i(\mathbf{p}), \end{aligned} \quad (11)$$

where  $d(\mathbf{q}, \mathbf{p})$  is a quadratic penalty function that measures deviation between  $\mathbf{q}$  and the auxiliary variable  $\mathbf{p}$ .  $\delta(\cdot)$  is an indicator function that enforces the projection of  $\mathbf{p}$  onto the constraint manifold, which is undeformed configurations. The solution proceeds in two alternating steps. *Local step:* Each auxiliary variable  $\mathbf{p}_i$  is updated independently and in parallel by minimizing (11). *Global step:* A convex quadratic problem, formulated using the updated  $\mathbf{p}$ , is solved to find  $\mathbf{q}$  that minimizes the total energy  $g(\mathbf{q})$ , with the condition  $\nabla g(\mathbf{q}) = 0$  in (10).

The internal forces  $\mathbf{f}_{int}$  in the system come from the deformation of the soft tissue, which is modeled using both shear

strain and volume-preservation constraints to maintain realistic behavior. Boundary conditions are incorporated using positional constraints at fixed nodes. For controlled nodes, velocity constraints ensure that robotic manipulations are accurately translated into tissue deformation. A complete description of constraint formulations is provided in the **supplementary**.

*2) Differentiable Deformation Model:* To derive the Jacobian deformation matrix  $\mathcal{J}$  in MPC, it is essential to compute accurate derivatives of the system's deformation response with respect to external actions. This section presents a differentiable formulation of the soft tissue deformation model based on the PD, following [34].

By implicitly differentiating the first-order optimality condition  $\nabla g(\mathbf{q}) = \mathbf{0}$  in (10) with respect to  $\mathbf{y}_c$ , which encodes the robot-induced displacement at contact nodes, yields (with time indices omitted for clarity):

$$\frac{\partial \mathbf{q}}{\partial \mathbf{y}_c} = \frac{1}{\Delta t^2} \underbrace{\left[ \frac{1}{\Delta t^2} \mathbf{M} + \nabla^2 E(\mathbf{q}) \right]^{-1}}_{\nabla^2 g(\mathbf{q})} \mathbf{M}_c, \quad (12)$$

where  $E(\mathbf{q}) \in \mathbb{R}$  denotes the total internal deformation energy.  $\partial \mathbf{q} / \partial \mathbf{y}_c$  is the Hessian of the system energy. The matrix  $\mathbf{M}_c$  selects the sub-block of the mass matrix corresponding to the contact points. Equation (12) provides the sensitivity of the full nodal positions  $\mathbf{q}$  with respect to the external input  $\mathbf{y}_c$ , can be expressed as:

$$\frac{\partial \mathbf{q}}{\partial \mathbf{y}_c} = \Delta t^2 \frac{\partial \mathbf{q}}{\partial \mathbf{f}_{ext}} \mathbf{M}_c = \Delta t^2 \mathbf{K} \mathbf{M}_c, \quad (13)$$

where  $\mathbf{K} \in \mathbb{R}^{3n \times 3l}$  denotes the **tangent stiffness matrix**, which characterizes the linearized response of the soft tissue to infinitesimal external forces, which is crucial for predicting the tissue's deformation. Under position control of the robot tool, the displacement  $\mathbf{y}_c(k)$  can be written as:  $\mathbf{y}_c(k) = \mathbf{q}^c(k) + \mathbf{a}(k+1)$ , enabling the definition of the Jacobian of the nodal configuration with respect to the action input.

$$\frac{\partial \mathbf{q}}{\partial \mathbf{a}} = \frac{\partial \mathbf{q}}{\partial \mathbf{y}_c} = \mathcal{J}, \quad (14)$$

where  $\mathcal{J} \in \mathbb{R}^{3n \times 3l}$  is referred as the **deformation Jacobian matrix**.

Direct inversion of the large matrix  $\nabla^2 g(\mathbf{q})$  in (12) is computationally prohibitive in real time. To address this, an efficient iterative method based on conjugate updates [34] is adopted. This method enables efficient evaluation of the gradient of the loss function  $\mathcal{L}$  with respect to the displacement input:

$$\frac{\partial \mathcal{L}}{\partial \mathbf{y}_c} = \frac{\partial \mathcal{L}}{\partial \mathbf{q}} \frac{\partial \mathbf{q}}{\partial \mathbf{y}_c} = \frac{1}{\Delta t^2} \underbrace{\frac{\partial \mathcal{L}}{\partial \mathbf{q}} [\nabla^2 g(\mathbf{q})]^{-1}}_{\beta^T} \mathbf{M}_c, \quad (15)$$

where  $\beta \in \mathbb{R}^{3n}$  is an intermediate variable. Building on the results from Sec. III-B1 and [35], the expression of Hessian matrix  $\nabla^2 g(\mathbf{q})$  can be decomposed into a fixed and variable component:

$$\nabla^2 g(\mathbf{q}) = \mathbf{D} - \Delta \mathbf{D}(\mathbf{q}), \quad (16)$$

where  $\mathbf{D}, \Delta \mathbf{D} \in \mathbb{R}^{3n \times 3n}$  are symmetrical matrix. The constant matrix  $\mathbf{D}$  only depends on the undeformed configuration, while  $\Delta \mathbf{D}(\mathbf{q})$  varies with the current state. The variable  $\beta$  is computed iteratively by combining (15) and (16) as:

$$\mathbf{D} \cdot \beta^{j+1} = \Delta \mathbf{D}(\mathbf{q}) \cdot \beta^j + \left( \frac{\partial \mathcal{L}}{\partial \mathbf{q}} \right)^T, \quad (17)$$

where  $j$  is the iteration step. This formulation allows efficient updates by solving a linear system with a constant matrix  $\mathbf{D}$ , which can be pre-factorized offline. Similar to the PD, a local solver computes the value of  $\Delta \mathbf{D}(\mathbf{q})$  based on the current position of the nodes, while the global solver updates  $\beta$ .

To compute individual elements of the deformation Jacobian  $\mathcal{J}$ , the loss gradient  $\partial \mathcal{L} / \partial \mathbf{q}$  is constructed as a one-hot encoding vector  $[\dots, 0, 1, 0, \dots]^T \in \mathbb{R}^{3n}$  targeting a specific nodal dimension. In this manner,  $\mathcal{J}$  can be built incrementally, row by row, enabling parallel computation. This differentiable formulation of the deformation model is critical to enabling real-time, model-based control and affordance estimation under complex tissue dynamics.

### C. Analytical Framework for Affordance Estimation

Effective soft tissue manipulation requires selecting contact configurations that maximize the ability to induce target deformations. We present an analytical **manipulability-based affordance estimation** framework, evaluating how effectively a given multi-contact point enables control over feature-based deformation objectives.

Conventional approaches for estimating affordance, such as demonstration learning and reinforcement learning [14], heavily rely on extensive training data and empirical assumptions, limiting their applicability in dynamic surgical scenarios. In contrast, the proposed analytical framework eliminates the need for data-driven training while achieving computational efficiency, which leverages two core innovations.

*Task-adaptive Affordance:* An analytical framework is presented for modeling the relationship between feature-based control variables and actions, enabling the real-time evaluation of manipulability-based affordance.

*Efficient Manipulability Aggregation via Polygon Approximations:* An approach that combines multi-contact effects through geometric set operations and convex polygon approximations is proposed, yielding  $O(n)$  complexity.

The proposed affordance has a direct physical meaning: it quantifies the efficiency with which a targeted shape can be reached for a given pair of contact points. A higher score indicates that a desired deformation can be achieved with less control effort, thus providing an intuitive and actionable insight for both surgeons and autonomous planners. In the framework, this entire process, from candidate sampling to final selection, is completed offline prior to task execution. The affordance metric is derived from the linearized deformation Jacobian matrix, which captures the sensitivity of feature

points' movement to applied actions at contacts. In multi-contact scenarios, the cumulative effect of actions is modeled through the Minkowski addition of manipulability sets. This approach aggregates the influence of individual contact points into a composite manipulability shape, representing the feasible deformation space for target features. The affordance is estimated by maximizing the projected manipulability set along the targeted deformation axis in feature space (26). This direction is derived from the gradient vector of the loss function with respect to features, which is adaptive in various tasks. The comprehensive implementation pipeline of the method is illustrated in Fig. 3 and detailed computational steps are provided in **supplementary**.

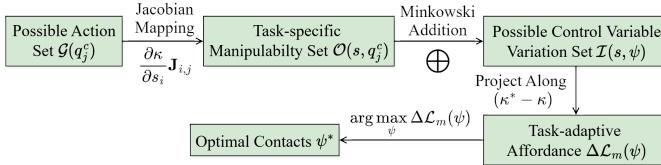


Fig. 3: Overview of the contact selection optimization process.

1) *Overview of the Analytical Framework:* Given a set of uniformly sampled candidate contact points along the tissue boundary:

$$\mathcal{P} = \{q_j^c, j = 1, \dots, e\}, \quad (18)$$

where  $q_j^c \in \mathbb{R}^3$  represents the position of contact point.  $e$  is the number of all potential contact points. A contact point selection  $\psi$  is defined as a subset of  $\mathcal{P}$  of size  $l$ , corresponding to the number of robot tools:

$$\psi = \{q_1^c, \dots, q_l^c | q_m^c \in \mathcal{P}, \forall m, q_m^c \neq q_n^c \text{ for } m \neq n\}. \quad (19)$$

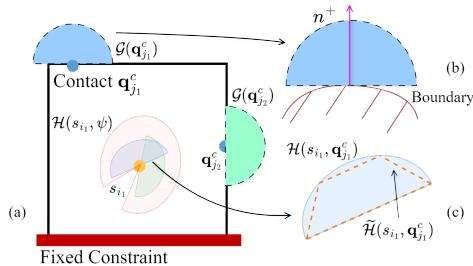


Fig. 4: (a). Illustrating the manipulability set of the feature point  $s_{i1}$  when all possible actions are applied at contact point  $q_{j1}^c$  and  $q_{j2}^c$ .  $\mathcal{H}(s_i, \psi)$  is the aggregated manipulability set under contact selection  $\psi = \{q_{j1}^c, q_{j2}^c\}$ , represented by red shape. (b). The shape with dashed borders represents the set of possible actions. The red curve represents the object boundary. (c). The convex polygon approximation  $\tilde{\mathcal{H}}(s_{i1}, q_{j1}^c)$  of a convex set  $\mathcal{H}(s_{i1}, q_{j1}^c)$ .

For each contact point  $q_j^c$ , an action set  $\mathcal{G}(q_j^c)$  is defined to describe feasible displacement vectors, as illustrated in Fig. 4-(a) and (b). The deformation response at a feature point  $s_i$  is obtained through the Jacobian  $\mathbf{J}_{i,j} \in \mathbb{R}^{3 \times 3}$ , extracted from the global deformation Jacobian  $\mathcal{J}$ , such that:

$$\mathcal{H}(s_i, q_j^c) = \{v | v = \mathbf{J}_{i,j} u, u \in \mathcal{G}(q_j^c)\}, \quad (20)$$

where  $\mathcal{H}(s_i, q_j^c)$  is the manipulability set, as illustrated in Fig. 4-(c), representing the range of achievable displacements of

feature point  $s_i$  resulting from actions at  $q_j^c$ . Representing  $\mathcal{H}(s_i, q_j^c)$  in polar coordinates enables a clear analysis of direction-dependent deformation capabilities:

$$\mathcal{H}(s_i, q_j^c) = \{(r, \theta) | \theta_1 \leq \theta \leq \theta_2 \wedge 0 \leq r \leq r(\theta)\}, \quad (21)$$

where  $\theta_1, \theta_2$  define the angular range of deformation directions.  $r(\theta)$  describes the magnitude of the deformation in every direction.

2) *Manipulability-based Affordance Metrics:* Through singular value decomposition, the local deformability is characterized by the singular values  $\lambda_m$  of  $\mathbf{J}_{i,j}$ , which represent the relative stiffness along the corresponding deformation axis.  $\lambda_m$  less than 1 indicates reduced stiffness in the corresponding directions.

Inspired by the concept of robotic manipulability measure, the manipulability  $Y_{i,j}$  can quantify the deformability of the region between the contact point  $q_j^c$  and the feature point  $s_i$ . It provides a scalar representation of how effectively the system can produce deformation on the feature point  $s_i$  under the action applied at the contact point  $q_j^c$ .

$$Y_{i,j} = \sqrt{\det(\mathbf{J}_{i,j} \cdot \mathbf{J}_{i,j}^T)} = \sqrt{\prod \lambda_m}, \quad (22)$$

where  $\det(\cdot)$  represent the determinant of a matrix.  $\prod$  is the product operator. The **manipulability-based affordance** metric for a feature-contact pair  $(s_i, q_j^c)$  is defined as the area enclosed by the manipulability set:

$$\mathcal{Y}_{pt}(s_i, q_j^c) = \mathcal{A}(\mathcal{H}(s_i, q_j^c)) = Y_{i,j} \cdot \mathcal{A}(\mathcal{G}(P_j^c)), \quad (23)$$

where the function  $\mathcal{A}(\cdot) \in \mathbb{R}$  denotes the size of the region enclosed by the set boundary, which must form a continuous region. The proposed manipulability-based affordance metric uniquely accounts for directional constraints on applied action, which is a common consideration in surgical settings, extending the classic notion of manipulability to deformation control.

3) *Task-adaptive Affordance:* Given the contact selection  $\psi$ , the variation of feature-based control variable  $\kappa$ , due to applied actions, is derived from (5) by applying total derivative (with time indices omitted):

$$\begin{aligned} \Delta\kappa(\mathbf{u}, \psi) &= \sum_i \frac{\partial\kappa}{\partial s_i} \Delta s_i = \sum_i \frac{\partial\kappa}{\partial s_i} \left( \sum_j \mathbf{J}_{i,j} u_j \right) \\ &= \sum_{q_j^c \in \psi} \left( \sum_i \frac{\partial\kappa}{\partial s_i} \mathbf{J}_{i,j} \right) u_j, \quad u_j \in \mathcal{G}(q_j^c), \end{aligned} \quad (24)$$

where  $\mathbf{u} = [u_1, \dots, u_l]^T \in \mathbb{R}^{3l}$  is the applied actions.  $\partial\kappa/\partial s_i \in \mathbb{R}^{o \times 3}$ .

The set of all possible  $\Delta\kappa(\mathbf{u}, \psi)$  is then represented by collecting each contact influence:

$$\begin{aligned} \mathcal{I}(\mathbf{s}, \psi) &= \bigoplus_{q_j^c \in \psi} \mathcal{O}(\mathbf{s}, q_j^c) \\ \mathcal{O}(\mathbf{s}, q_j^c) &= \left\{ v | v = \left( \sum_i \frac{\partial f_c}{\partial s_i} \mathbf{J}_{i,j} \right) \cdot u_j, \quad u_j \in \mathcal{G}(q_j^c) \right\}, \end{aligned} \quad (25)$$

where  $\bigoplus$  denotes the addition of elements across sets, which is accomplished by the Minkowski addition method.

The **task-adaptive affordance** for a given contact combination  $\psi$  is defined as the maximal directional progress toward the target deformation, denoted as  $\Delta\mathcal{L}_m(\psi)$ :

$$\Delta\mathcal{L}_m(\psi) = \max_{u_j \in \mathcal{G}(q_j^c), q_j^c \in \psi} -2(\kappa - \kappa^*)^\top \cdot \Delta\kappa(\mathbf{u}, \psi), \quad (26)$$

where  $\kappa$  is the control variable in the current step.  $\Delta\mathcal{L}_m(\psi)$  corresponds to the maximal projection of  $\mathcal{I}(\mathbf{s}, \psi)$  onto the vector  $(\kappa^* - \kappa)$ , indicating the steepest descent in the loss.

The optimal contact combination  $\psi^*$  is then selected as the one that yields the highest task-adaptive affordance among all candidates' multi-contacts:

$$\psi^* = \arg \max_{\psi} \Delta\mathcal{L}_m(\psi). \quad (27)$$

*4) Efficient Affordance Aggregation via Polygon Approximation:* To ensure computational tractability, each manipulability set  $\mathcal{H}(s_i, q_j^c)$  is approximated by a convex polygon  $\tilde{\mathcal{H}}(s_i, q_j^c)$ , shown in Fig. 4-(c).

*Assumption 1:* All manipulability sets are convex sets.

*Theorem 1:* The Minkowski sum of convex polygons  $\mathcal{Q}_1$  and  $\mathcal{Q}_2$  is a convex polygon with at most  $|\mathcal{Q}_1| + |\mathcal{Q}_2|$  vertices, where  $|\cdot|$  denotes the number of vertices of the respective polygon.

We formulate *Assumption 1*, which assumes that all manipulability sets are convex, a condition that holds in most practical cases. Under the *Assumption 1* and applying *Theorem 1*, the approximate sum is iteratively computed:

$$\tilde{\mathcal{H}}(s_i, \psi) = \bigoplus_{q_j^c \in \psi} \tilde{\mathcal{H}}(s_i, q_j^c). \quad (28)$$

The multi-contact affordance metric can also be defined as:

$$\mathcal{Y}_{pt}(s_i, q_j^c) = \mathcal{A}\left(\tilde{\mathcal{H}}(s_i, q_j^c)\right). \quad (29)$$

Samely, the task-adaptive affordance  $\Delta\mathcal{L}_m(\psi)$  is then obtained by projecting  $\tilde{\mathcal{I}}(s_i, \psi)$  onto the task required direction.

#### D. Controller Design

*1) Deformation Limits of Soft Tissue:* The deformation strain of each element in soft tissue is constrained within safety thresholds, established based on clinical expertise, to ensure the soft tissue remains within its physiological tolerance, thus forming the constraint component of the MPC framework.

*2) Model-based Controller:* The MPC framework applies the optimal action  $\mathbf{a}^*(k)$  as the controller input. This action is determined by solving the constrained optimization problem (6), which ensures both the convergence of the control objectives and the satisfaction of deformation constraints.

The deformation Jacobian matrix in (14) provides the linearized system dynamics required by MPC formulation at each control step to predict the effect of robot actions on the tissue shape.

*3) Visual Feedback:* At each time step, the vision system measures the current 3D positions of the feature points on the tissue  $\hat{\mathbf{s}}(k)$ . This measurement is then compared against the goal configuration  $\mathbf{s}^*$ , to compute the task loss  $\mathcal{L}(k)$  in (7). This loss value  $\mathcal{L}(k)$  serves as the optimization objective in the MPC framework, which then calculates the optimal robot actions to minimize this loss over the prediction horizon.

#### E. Overview of Each Module

To provide a cohesive overview of the proposed architecture, the key computational modules, their respective inputs and outputs, and their approximate runtimes are summarized in Tab. I. In summary, the presented framework combines a differentiable deformation model with an analytical affordance metric. A series of simulations and physical experiments designed to rigorously validate the performance and effectiveness of this approach is presented in the next section.

TABLE I: Overview of the computational modules in the framework.

Module	Inputs	Outputs	Runtime (ms)
Visual Perception	Infrared Images	3D Positions of Markers $s(k)$	~25ms
Differentiable Deformation Model	Robot Actions $\mathbf{a}(k)$	Deformation Jacobian Matrix $\mathcal{J}(k)$	~16ms
Affordance Estimation	Jacobian $\mathcal{J}(k)$ , Target $\kappa^*$ , Current States $\mathbf{q}(k)$ ,	Affordance score for candidate points	/
MPC Controller	Target $\kappa^*$ , Jacobian $\mathcal{J}(k)$	Optimal Robot Actions $\mathbf{a}^*(k+1)$	~75ms

## IV. SIMULATION ANALYSIS

#### A. Simulation Setup

Having detailed the framework in the previous section, we now evaluate the effectiveness and accuracy of the proposed method. To this end, two simulation environments were developed. These include a 3D anatomical liver model and a simplified 2D soft tissue mesh, as shown in Fig. 5.

The first simulation environment is constructed using a 3D liver-shaped object, which is defined by a tetrahedral mesh with a minimum bounding box of [0.376, 0.245, 0.176]m. Anatomical constraints are simulated by rigidly fixing the base of the liver, highlighted in pink. A displacement vector is applied to a designated surface node (blue sphere), emulating the effect of a surgical tool. The tissue deformation mechanics are governed by Simulation Open Framework Architecture (SOFA) [36].

The second environment consists of a planar soft tissue represented as a triangular mesh. The bottom boundary of the tissue was fixed, and the remaining boundaries were discretized into 27 potential contact points. The material properties of the objects in two environments are set to be:  $E = 500$ ,  $v = 0.4$ .

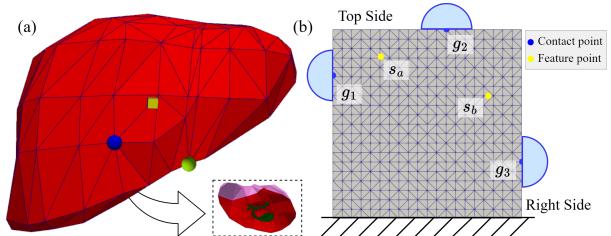


Fig. 5: (a). 3D liver-shaped object simulation environment, where the image in the lower right represents a rear view. (b). 2D soft tissue setup for manipulability-based affordance evaluation, with predefined contact points (dark blue dots).

## B. Method Validation

1) *Deformation Model Validation with SOFA*: To validate the accuracy of our PD-based deformation algorithm, a numerical comparison was conducted with SOFA physics engine. Both systems were configured with the same physical parameters, including material properties (e.g., elasticity, density) and mesh structure of the liver. A controlled displacement of  $[0.025, 0, 0]$ m was applied to node {85} on the liver surface shown in Fig. 5-(a).

Figure 6-(a) presents the normalized mean squared error (NMSE) between the PD model and SOFA across all nodes. The majority of nodes exhibit low NMSE values, indicating high consistency in predicted deformation. Fixed nodes (orange circles) serve as references and remain invariant in both simulations.

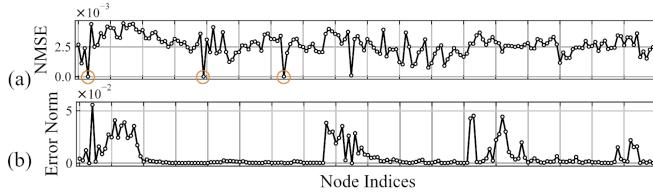


Fig. 6: (a). The simulation error between the SOFA and the PD algorithm. The orange circles indicate the fixed nodes. (b). The model gradient error between numerical differentiation and the analytic Jacobian method.

2) *Differentiable Model Validation with Finite Difference Method*: The Finite Difference method is employed to compute the model gradient  $\mathcal{E}$ , serving as a benchmark to validate the accuracy of the analytic gradient calculation from the differentiable model. The gradient of the model  $\mathcal{E}$  with respect to the action at the node {85} was calculated by applying a small action and subsequently computing the resulting displacement of all nodes.

$$\mathcal{E} = \frac{\mathbf{q}_f - \mathbf{q}_0}{\Delta q_{cnt}} \quad (30)$$

Figure 6-(b) shows the gradient error between the analytical Jacobian (used in the MPC framework) and numerical differentiation. The small discrepancies validate the correctness of the closed-form Jacobian derivation for control purposes.

3) *Manipulability-based Affordance Computation*: In the 2D square soft tissue simulation environment shown in Fig. 5-(b), two feature points,  $s_a$  and  $s_b$ , and three candidate contact points,  $g_1$ ,  $g_2$ , and  $g_3$  are selected. These three points are chosen purely for the illustrative process of affordance calculation and its purpose. A comprehensive evaluation of the proposed affordance metric across all candidate boundary points is presented in simulation cases (see Fig. 12). These contact points were constrained to move outward from the tissue surface, reflecting typical surgical manipulation constraints. This setup was used to evaluate the influence of different contact locations on the deformation behavior of feature points. Figure 7-(a) presents several manipulability sets corresponding to different contact-feature point pairings. These sets characterize the achievable deformation directions and magnitudes at a given feature point under a specific contact location. Higher affordance values indicate a broader and more

effective range of deformation control at the feature point. Figure 7-(b) and (c) illustrate the computational process for

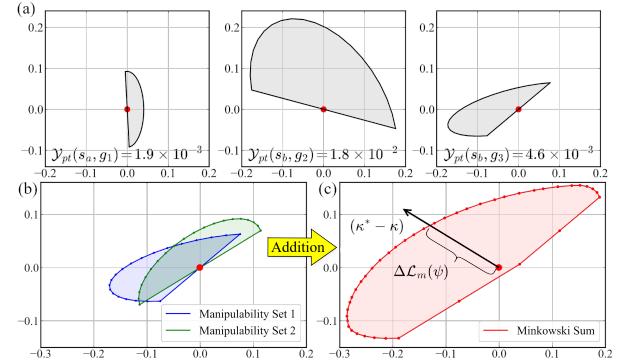


Fig. 7: (a). Manipulability sets obtained from different single-contact points with respect to feature points. (b) Polygonal discretizations of two distinct manipulability sets (shown in blue and green). (c) The Minkowski sum of the discretized manipulability sets from (b). The red region represents the aggregated manipulability set.

aggregating the manipulability sets. The manipulability set for an individual contact point is first discretized into a convex polygon, followed by the application of the Minkowski sum to produce the aggregated set. The task-adaptive affordance  $\Delta \mathcal{L}_m(\psi)$  is the maximum distance from the contour of the manipulability set in the direction  $(\kappa^* - \kappa)$ . As demonstrated in Fig. 7-(c), some directions exhibit significantly reduced manipulability, limiting deformation control over these regions.

4) *Deformation Limits*: To validate the effectiveness of the proposed strain-based safety constraints, a simulation of a pre-tensioning task is conducted. The MPC controller was compared against a standard adaptive controller across two multi-contact configurations, denoted as {55, 98} and {77, 118}. For all tests, the maximum allowable deformation strain was set to  $\epsilon_{max} = 1.4$  to prevent simulated tissue damage.

The results, plotted in Fig. 8, demonstrate a critical difference in the {55, 98} contact scenario. In this case, fully achieving the task's position target was predicted to violate the strain limit. Prioritizing safety, the MPC controller autonomously regulated its control actions, deliberately halting the tissue's movement before the strain constraint was breached. As shown in Fig. 8-(b), the maximum strain under MPC control consistently remained below the 1.4 threshold, even though this meant the final position error (loss) did not converge to zero.

By contrast, the model-agnostic adaptive controller, lacking predictive capability, failed to respect the safety constraint. It continued to drive the tissue toward the target, causing the maximum strain to significantly exceed the prescribed limit—a behavior that could lead to tissue damage in a clinical setting. This comparison clearly highlights the advantage of the presented model-predictive approach in ensuring safe manipulation when task goals conflict with physical constraints.

## C. Simulation Cases

1) *Liver Retraction for Gallbladder Exposure*: To validate the capability of the proposed deformation control strategy in 3D soft tissue, the control algorithm was applied to perform

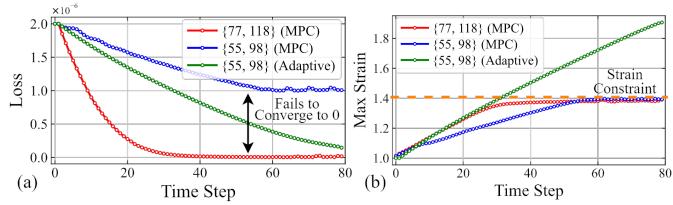


Fig. 8: (a). The loss curves of three cases. (b). Maximum strain experienced by the tissue elements. The adaptive controller ignored constraints and completed the task. Meanwhile, the MPC controller adhered to constraints even when the task failed.

the liver retraction and gallbladder exposure task. The simulation setup is shown in Fig. 5-(a), where a control point (green sphere) was specified with a target displacement of  $[-0.04, 0.06, 0.01]$ m, illustrated as a yellow cube. The control variable  $\kappa$  corresponds to the spatial coordinates of this point.

As shown in Fig. 9-(a), our method successfully achieves gallbladder exposure through controlled liver deformation. Quantitative results presented in Fig. 9-(b) demonstrate effective reduction of the loss function, with the control point converging to its target position. These experimental outcomes confirm the method's capacity to execute precise point deformations in 3D structures.

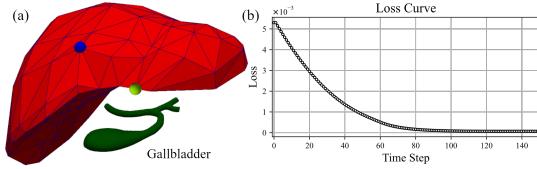


Fig. 9: (a). The final shape of the liver model clearly shows the successful exposure of the gallbladder when compared to the initial shape (Fig. 5). (b). The loss curve demonstrates the reduction in error over time.

**2) Control of a Specified Angle:** The objective of the task is to control the angle formed by three marked points on the soft tissue surface to a targeted value. The angle serves as a compact shape descriptor, enabling low-dimensional but semantically meaningful control. The simulation setup is illustrated in Fig. 5-(b), and the task loss function is defined as follows:

$$\begin{aligned} \mathcal{L}(k) &= (\kappa(k) - \kappa^*)^2 + \sum_{j=1}^2 \omega_j (d_j(k) - d(0))^2 \\ \kappa(k) &= (s_2(k) - s_1(k)) \cdot (s_3(k) - s_1(k)), \end{aligned} \quad (31)$$

where  $s(k) = [s_1(k), s_2(k), s_3(k)] \in \mathbf{R}^{3 \times 3}$  denotes the positions of the three marked points at time step  $k$ .  $\cdot$  represents the dot product of vectors. The angle is implicitly represented by  $\kappa(k)$ , through the dot product of two vectors originating from  $s(k)$ . The second summation term imposes length constraints, where  $d_j(k) = \|q_{j+1}(k) - q_1(k)\|$  represents the Euclidean distance between  $q_{j+1}(k)$  and  $q_1(k)$ , and  $\omega_j$  is the corresponding weighting factors.

Figure 10 shows that the method accurately adjusts the angle while preserving geometric constraints. These results illustrate the controller's capability to control low-dimensional shape

features beyond simple point-wise displacements, which is difficult for other methods.

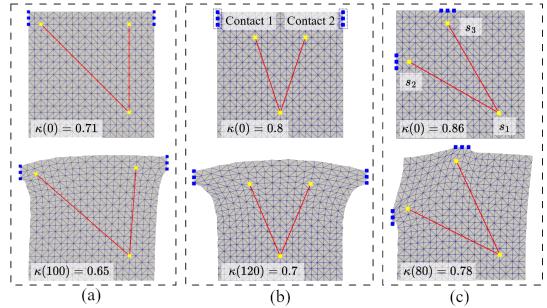


Fig. 10: The initial shape (first row) and final shape (last row) of the soft tissue under control. Blue points represent the robot tool contact, while yellow points are the marked points. Sub-figures are the different cases.

### 3) Control of Edge Points on a 2D Deformable Surface:

The proposed control framework is tested on a challenging 2D surface manipulation task. This surface is characterized by its high degree of deformability, making it difficult to model and control accurately. A physically-based bending energy term is incorporated into the system's total energy (see Appendix for formulation). As shown in the initial state in Fig. 11-(a), the goal is to manipulate the surface by controlling a sparse set of feature points on its boundary. The control error is quantified by a loss function, defined as the squared Euclidean distance between the current and target positions of the edge points:

$$\mathcal{L}(k) = \sum_j^e (s_j(k) - s_j^*)^2, \quad (32)$$

where  $e$  is the number of edge feature points.  $s_j(k)$  is the position of the  $j$ -th feature point at time step  $k$ , and  $s_j^*$  is its corresponding target position.

A qualitative result of this simulation is presented in Fig. 11, which provides snapshots of the surface converging from its initial configuration to the desired target shape. The residual error, visible as a misalignment at contact point 2, stems from a "too small gradient" issue. In this region of the state space, the loss function becomes insensitive to the position of contact point 2, yielding a gradient too small for the optimizer to produce a meaningful control action.

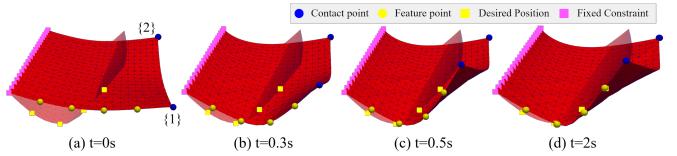


Fig. 11: Snapshots from the 2D surface control simulation: the initial configuration, two intermediate states during the control process, and the final configuration converged to the target. The semi-transparent overlay shows the target configuration.

**4) Task-adaptive Affordance:** To validate the agreement between the estimated and actual affordance, a task was performed in which the distance between two markers on the soft tissue surface was stretched by a factor of 1.05. The actual affordance was quantified based on the total strain experienced

by the soft tissue after task completion, with lower strain values indicating higher affordance.

The simulation experimental setup is illustrated in Fig. 5-(b). The possible action sets for each contact point are shown in Fig. 12-(a), permitting actions to be applied outward from the boundaries, while inward actions were limited in magnitude to prevent tissue buckling. Additionally, candidate contact pairs were filtered to ensure a minimum inter-contact distance of 0.02 m, mitigating potential tool collisions.

For each contact point combination  $\psi$ , the task-adaptive affordance was computed using (26). Figure 12-(b) illustrates the inverse correlation between estimated affordance and resulting tissue strain for different feature point configurations. Spearman's rank correlation coefficients and root-mean-square error (RMSE) values are also reported. The simulation results indicate that higher affordance correlates with lower deformation strain, and the observed negative correlation validates the efficiency of the proposed analytical affordance metric.

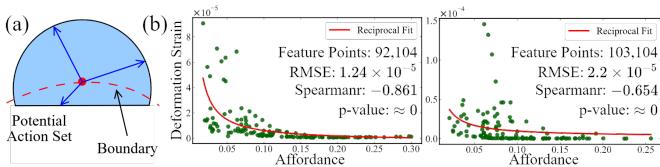


Fig. 12: (a) Schematic representation of a boundary contact point (red dot) and its potential action set, delineated by the black contour. (b) and (c). Scatter plots illustrating the relationship between estimated affordance and deformation strain for selected contact point combinations. The red curves represent reciprocal function fits, demonstrating an inverse correlation between manipulability and deformation strain. Each sub-figure corresponds to a pair of feature points, with associated RMSE and Spearman's rank correlation coefficient values provided.

## V. EXPERIMENTAL VALIDATION & DISCUSSION

### A. Experiment Setup

The experimental platform comprises two UR10e robots with custom-designed, 3D-printed robot tools and a Qualisys Motion Capture System, as illustrated in Fig. 13. The visual measurement system tracks the 3D positions of passive reflective markers attached to the tissue surface with a sub-millimeter precision. One end of the soft tissue prosthesis is securely anchored, while the other is free to move, simulating realistic tissue manipulation conditions. Reflective markers affixed to the surface are tracked in Cartesian space during manipulation.

A PD deformation model was employed to simulate soft tissue behavior with a time step of 0.01s, balancing real-time performance and numerical accuracy. All computations were performed on a desktop computer equipped with an Intel i9-12900 CPU, 32 GB of RAM, and an NVIDIA 3080Ti GPU, using the Taichi library [37] for GPU acceleration. Soft tissue phantoms are made of two materials, silicone and latex, with dimensions of  $0.13 \times 0.13$ m. The representative material parameters were used for a typical silicone phantom ( $E = 5$  KPa and  $\nu = 0.4$ ) without performing a precise calibration for the specific soft tissue. This choice was made to explicitly test

and demonstrate the capability of the proposed framework: that the visual feedback controller is robust enough to compensate for model-plant mismatch and successfully complete the task even with an inaccurate physical model. For the quantitative evaluation of experiments, a trial was formally defined as "successful" if the final loss was below a threshold of  $1 \times 10^{-6}$ .

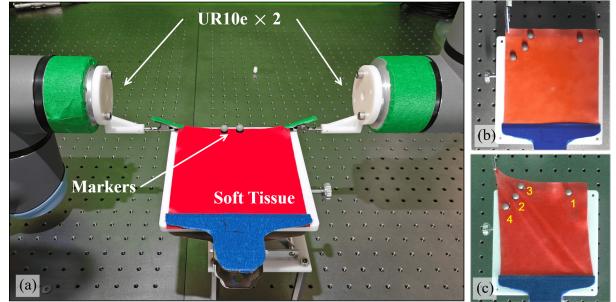


Fig. 13: (a). Experiment platform setup. The initial shape (b) and final shape (c) of the soft tissue upon completion of the multi-point feature deformation task. The feature vector comprises the 3D positions of four optical markers.

### B. Validation Cases

1) *Validation with Multi-point Feature:* The proposed method was validated using a feature vector  $\kappa = [s_1, \dots, s_4]^T \in \mathbb{R}^{12}$ , where each element represents the position of a marked point on the soft tissue surface, as shown in Fig. 13-(b). This feature captures the spatial arrangement of multiple points, enabling global deformation control beyond single marked point manipulation.

The targeted pattern  $\kappa^*$  was generated based on predefined manipulation commands. The loss function error converges to 0, shown in Fig. 14-(a), indicating that the method is capable of handling the global deformation control of soft tissue. The error between the actual feature vector and the predictive feature vector is shown in Fig. 14-(b), demonstrating the ability of the method to minimize the deviation from the target.

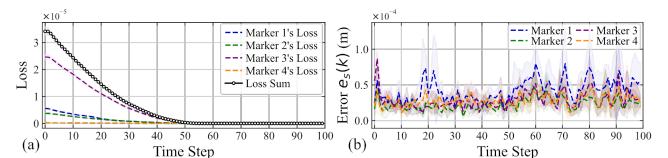


Fig. 14: Performance of the model-based controller, showing the mean (solid/dashed lines) and confidence interval (shaded region,  $\pm$  one standard deviation) over 5 trials. (a). The loss curves for all four markers converge smoothly to values near zero, with the narrow confidence intervals demonstrating high repeatability and successful task completion. (b). The feature tracking error remains low, demonstrating the controller's ability to steadily track feature variations throughout the task.

2) *Pre-Tensioning Prior to Tissue Cutting:* To evaluate the effectiveness of the proposed deformation control strategy in surgical tasks, a pre-tension operation prior to tissue cutting was performed. In this operation, the soft tissue was stretched to increase the distance between two markers positioned on

either side of the intended incision, while preserving the orientation of the line defined by these markers to ensure uniform tension across the cutting region.

The experimental setup is illustrated in Fig. 13-(a). The target distance was specified as 1.15 times the initial distance. Experiments were performed on two materials under varying cutting conditions, with the actual soft tissue topology aligned with the predictive model before each experiment. The task loss function is defined as follows:

$$\mathcal{L}(k) = (\kappa(k) - \kappa^*)^2 + \sum_{j=1}^2 \omega_j d_j^2 \quad (33)$$

$$\kappa(k) = \|\mathbf{s}_2(k) - \mathbf{s}_1(k)\|,$$

where  $\mathbf{s}(k) = [s_1(k), s_2(k)] \in \mathbb{R}^{3 \times 2}$  represents the position of two marked points.  $d_j$  denotes the distance between the marked point  $s_j$  and the line constructed from the initial position of the markers. The experimental results in Fig.

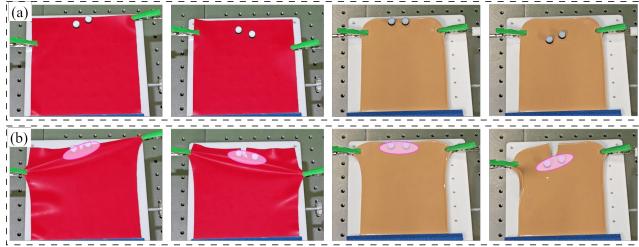


Fig. 15: The initial shape (a) of the soft tissue and its final shape (b) after the control task completion. The final shapes indicate that the presented framework successfully achieves the targeted pretension.

15 demonstrate that the target cutting region (pink area) maintained a tensioned state, with the preceding cutting trajectory visible. This indicates that the proposed control strategy effectively executed the pre-tension operation before soft tissue cutting.

**3) Tissue Retraction:** To evaluate the proposed framework on a complex, clinically-relevant task, two tissue retraction experiments were conducted. The objective was to retract a soft tissue phantom to fully expose a hidden Region of Interest (ROI), as shown in Fig. 16 and Fig. 17. The first experimental setup, depicted in Fig. 16-(a), involved a planar tissue phantom with one edge securely fixed. To provide real-time feedback, four optical markers were affixed to the tissue's free edge, and a robotic tool was positioned to grasp a corner. In the second experiment, a liver-shaped soft tissue phantom was secured to a bottom plate, shown in Fig. 17. In both experiments, the target configuration, representing the desired state where the ROI is completely exposed, was defined by manually performing the retraction and recording the final 3D positions of the optical markers. Figures 16 and 17 present qualitative snapshots of the experiment, showing the tissue successfully retracted from its initial state towards the final target configuration.

### C. Comparative Study

A comparative experiment was conducted to evaluate the performance of our method against the baseline adaptive

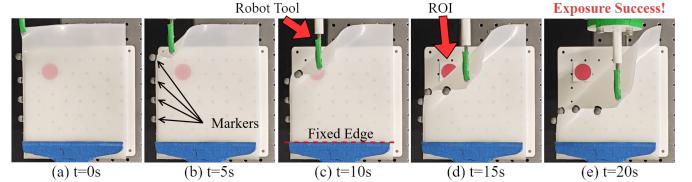


Fig. 16: Snapshots from the tissue retraction control experiment. A robot peeled back a tissue phantom layer to expose an underlying target (red circle). The progression from left to right shows the framework successfully achieving the desired exposure.

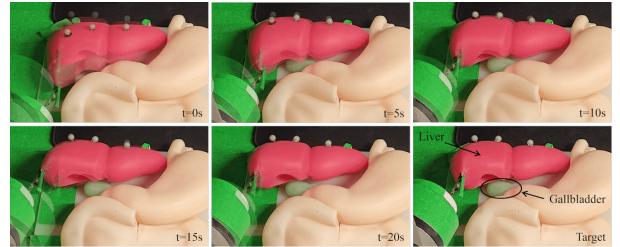


Fig. 17: Snapshots from the liver retraction control experiment. A robot grasps the liver-shaped tissue phantom (red) to expose the gallbladder (green). The semi-transparent layer in the sub-figures represents the target configuration. This process demonstrates that the proposed framework successfully completes the gallbladder exposure task.

control method in [24], which used a 2D deformation adaptive controller based on point features. To compare the two methods in the same setup, the marked points' position was set as the control variable. As shown in Figs. 14 and 18,

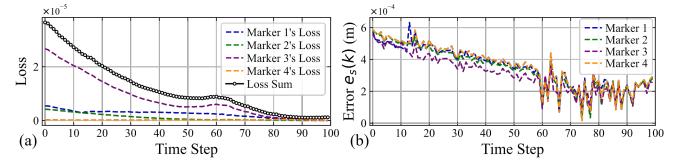


Fig. 18: Performance of the baseline adaptive controller in a multi-point feature deformation task for comparison. (a). The loss curves for all four markers converge smoothly to values near zero. (b). The adaptive controller has non-zero steady-state tracking error  $e_s$ , which is much larger than our method in Fig. 14.

both methods achieve similar final loss values. However, the feature vector tracking error for the proposed model-based controller is consistently lower. This performance improvement is attributed to the precise, analytical formulation of the deformation Jacobian, which enhances prediction accuracy and tracking stability. To validate this result, 5 trials were conducted for both methods across various configurations. A t-test on the mean feature vector tracking error confirmed a statistically significant difference ( $p < 10^{-3}$ ).

### D. Discussion

The collective results from the simulations and physical experiments validate the central hypothesis of this work: that an analytical, manipulability-based affordance metric, when paired with a differentiable, model-based controller, can

achieve precise, safe, and real-time control of soft tissues in complex surgical tasks.

The accuracy of the deformation model was validated through accurate forward and backward simulations of soft tissue deformation, shown in Fig. 6. This bidirectional simulation accuracy enables the prediction of deformation outcomes and the determination of necessary actions for the targeted shape.

The proposed analytical affordance estimation method provides a principled metric for selecting contact points across various soft object manipulation tasks. Spearman correlation analysis revealed a significant correlation between estimated affordance and observed deformation strain, shown in Fig. 12, indicating the metric's effectiveness in guiding contact point selection for higher task efficiency (lower strain). Discrepancies between the predicted affordance and the actual deformation outcomes are attributed to the deformation modeling error and only reliance on the initial tissue configuration.

The proposed framework's shape control capability was validated through a comprehensive set of simulations and physical experiments encompassing both 2D and 3D scenarios. Across all tasks, the framework consistently achieved low final tracking errors, illustrated in Figs. 9, 14 and 16. When compared to a standard adaptive controller, our model-based approach demonstrated significantly lower feature tracking error and highlighting its reliability for affordance estimation, as Figs. 14 and 18 show. Crucially, the proposed framework also excels at ensuring safety. As shown in Fig. 8, the MPC controller successfully adheres to strain constraints during manipulation, whereas the model-free adaptive controller violates these critical safety limits. This demonstrates the proposed framework's dual advantages: high-fidelity shape control and robust enforcement of physical constraints. In particular, the MPC controller can maintain strain constraints during the control process, while the model-free adaptive controller will violate the constraints, as Fig. 8 shows.

Furthermore, the successful control of geometry features, including angle control in simulation and pre-tension in experiments, illustrated in Figs. 10 and 15, demonstrates the framework's adaptability to higher-level task specifications beyond point manipulation. The pre-tension control, simultaneous regulation of distance and orientation, highlights the method's potential for addressing intricate surgical manipulation requirements.

Despite the promising results, the current approach exhibits several limitations. It struggles to handle 1D deformable structures due to minimal internal constraints and pronounced dynamics, introducing large gradient errors in the differentiable model, hindering accurate deformation prediction [38]. The constructed system's reliance on visual feedback is also a potential limitation, as significant occlusion, sensor noise, or marker detachment would compromise control. Enhancing the system's robustness through methods like state estimation remains a critical direction for future research. The system relies on a fixed, linearized deformation model, which is a deliberate trade-off for computational efficiency. However, because this model does not update in real-time to account for the tissue's intrinsic non-linearities, a degree of model-plant mismatch is unavoidable, which may reduce the controller's

predictive accuracy and performance in tasks involving large deformations [39]. The current framework also lacks a predictive mechanism [40] to update the affordance based on evolving tissue states, potentially overlooking critical shifts in manipulability. We anticipate our system's performance would degrade or fail in scenarios involving extremely large deformations that violate the local linearity assumption of the model or highly dynamic manipulation tasks that exceed the system's sensor-to-actuator frequency.

### E. Future Work

Future work will advance the proposed framework along two primary avenues. First, to enhance the model's fidelity by actively closing the gap between the simulation and physical reality, we will integrate richer sensory inputs, including dense visual features from the tissue surface and force feedback from the manipulator. This data fusion is critical for enabling continuous online model refinement, a process where real-time feedback is used to actively update the physical parameters of the model. This online adaptation complements the initial step of creating a patient-specific model from preoperative scans, ensuring a high-fidelity deformation model throughout the procedure. Second, we will focus on expanding the system's operational scope. For its application in robot-assisted surgery, where the framework operates as a semi-autonomous system guided by the surgeon, developing an intuitive human-robot interface becomes crucial to ensure a minimal learning curve. Looking beyond the single-robot setup, a promising extension involves integrating a high-level task planner to enable multi-robot collaboration, leveraging the proposed framework as the robust low-level motion policy for each agent.

## VI. CONCLUSION

This paper proposed a task-adaptive affordance estimation framework for soft tissue manipulation, addressing the limitations of predetermined contact selection. By analytically evaluating manipulability via the deformation Jacobian matrix, the method captures the relationship between contact choice and target deformation, enabling interpretable and efficient shape control beyond the capabilities of data-driven approaches.

Building on this, a model-based controller was developed by combining MPC with a differentiable deformation model. Using task-specific geometric features, the controller achieves accurate manipulation without offline training or explicit parameter identification. Real-time gradient computation supports efficient optimization in complex scenarios. Simulations and physical experiments confirmed improved convergence and tracking performance compared to baseline methods.

By leveraging the affordance metric to select optimal contact points and using closed-loop visual feedback to guide the manipulation, the proposed framework enhances surgical precision while contributing to safety by minimizing required forces and avoiding potentially damaging maneuvers.

Looking toward real-world adoption, the path to clinical use presents substantial challenges, most notably the need for extensive pre-clinical testing and validation to satisfy regulatory and safety requirements.

## REFERENCES

- [1] A. Handa, A. Gaidhane, and S. G. Choudhari, "Role of robotic-assisted surgery in public health: Its advantages and challenges," *Cureus*, vol. 16, no. 6, 2024.
- [2] D. F. Amanatullah, H. N. Shah, A. A. Barrett, and H. W. Storaci, "A small amount of retraction force results in inadvertent piriformis muscle damage during a piriformis-sparing approach to the hip," *JBJS*, vol. 102, no. 19, pp. 1687–1693, 2020.
- [3] J. Zhang, H. Zhao, K. Chen, G. Fei, X. Li, Y. Wang, Z. Yang, S. Zheng, S. Liu, and H. Ding, "Dexterous hand towards intelligent manufacturing: A review of technologies, trends, and potential applications," *Robotics and Computer-Integrated Manufacturing*, vol. 95, p. 103021, 2025.
- [4] M. Kim, Y. Zhang, and S. Jin, "Soft tissue surgical robot for minimally invasive surgery: a review," *Biomedical engineering letters*, vol. 13, no. 4, pp. 561–569, 2023.
- [5] Y. Ou and M. Tavakoli, "Sim-to-real surgical robot learning and autonomous planning for internal tissue points manipulation using reinforcement learning," *IEEE Robotics and Automation Letters*, vol. 8, no. 5, pp. 2502–2509, 2023.
- [6] J. Bodner, F. Augustin, H. Wykypiel, J. Fish, G. Muehlmann, G. Wetscher, and T. Schmid, "The da vinci robotic system for general surgical applications: a critical interim appraisal," *Swiss medical weekly*, vol. 135, no. 4546, pp. 674–674, 2005.
- [7] A. J. Valencia and P. Payeur, "Combining self-organizing and graph neural networks for modeling deformable objects in robotic manipulation," *Frontiers in Robotics and AI*, vol. 7, p. 600584, 2020.
- [8] J. Gibson, *The Ecological Approach to Visual Perception*, ser. Resources for ecological psychology. Lawrence Erlbaum Associates, 1986. [Online]. Available: <https://books.google.com.sg/books?id=DrhCCWmJpWUC>
- [9] Y. Geng, B. An, H. Geng, Y. Chen, Y. Yang, and H. Dong, "Rlafford: End-to-end affordance learning for robotic manipulation," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 5880–5886.
- [10] L. Cavalli, G. Di Pietro, and M. Matteucci, "Towards affordance prediction with vision via task oriented grasp quality metrics," *arXiv preprint arXiv:1907.04761*, 2019.
- [11] F. Veiga, J. Peters, and T. Hermans, "Grip stabilization of novel objects using slip prediction," *IEEE transactions on haptics*, vol. 11, no. 4, pp. 531–542, 2018.
- [12] D. Xu, A. Mandlekar, R. Martín-Martín, Y. Zhu, S. Savarese, and L. Fei-Fei, "Deep affordance foresight: Planning through what can be done in the future," in *2021 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2021, pp. 6206–6213.
- [13] X. Zhu, Y. Li, L. Cui, P. Li, H.-a. Gao, Y. Zhu, and H. Zhao, "Afford-x: Generalizable and slim affordance reasoning for task-oriented manipulation," *arXiv preprint arXiv:2503.03556*, 2025.
- [14] R. Wu, C. Ning, and H. Dong, "Learning foresighted dense visual affordance for deformable object manipulation," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 10947–10956.
- [15] Q. Wu, H. Wang, J. Zhou, X. Xiong, and Y. Lou, "Tars: Tactile affordance in robot synesthesia for dexterous manipulation," *IEEE Robotics and Automation Letters*, 2024.
- [16] R. Xu, F.-J. Chu, C. Tang, W. Liu, and P. A. Vela, "An affordance keypoint detection network for robot manipulation," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 2870–2877, 2021.
- [17] L. Han, Y. Zhang, and H. Wang, "Vision-based contact point selection for the fully non-fixed contact manipulation of deformable objects," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 4368–4375, 2022.
- [18] J. Huang, Y. Cai, X. Chu, R. H. Taylor, and K. S. Au, "Non-fixed contact manipulation control framework for deformable objects with active contact adjustment," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 2878–2885, 2021.
- [19] X. Liang, C.-P. Wang, N. U. Shinde, F. Liu, F. Richter, and M. Yip, "Medic: Autonomous surgical robotic assistance to maximizing exposure for dissection and cauterization," in *2025 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2025, pp. 6889–6895.
- [20] I. Cuiral-Zueco, G. López-Nicolás, and H. Araujo, "Gripper positioning for object deformation tasks," in *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 2022, pp. 963–969.
- [21] M. Shetab-Bushehri, M. Aranda, Y. Mezouar, and E. Özgür, "As-rigid-as-possible shape servoing," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 3898–3905, 2022.
- [22] ———, "Lattice-based shape tracking and servoing of elastic objects," *IEEE Transactions on Robotics*, vol. 40, pp. 364–381, 2023.
- [23] I. Cuiral-Zueco and G. López-Nicolás, "Time consistent surface mapping for deformable object shape control," *IEEE Transactions on Automation Science and Engineering*, 2025.
- [24] D. Navarro-Alarcon, H. M. Yip, Z. Wang, Y.-H. Liu, F. Zhong, T. Zhang, and P. Li, "Automatic 3-d manipulation of soft objects by robotic arms with an adaptive deformation model," *IEEE Transactions on Robotics*, vol. 32, no. 2, pp. 429–441, 2016.
- [25] L. Han, H. Wang, Z. Liu, W. Chen, and X. Zhang, "Visual tracking control of deformable objects with a fat-based controller," *IEEE Transactions on Industrial Electronics*, vol. 69, no. 2, pp. 1673–1681, 2021.
- [26] B. Yang, B. Lu, W. Chen, F. Zhong, and Y.-H. Liu, "Model-free 3-d shape control of deformable objects using novel features based on modal analysis," *IEEE Transactions on Robotics*, 2023.
- [27] B. Yang, C. Sui, F. Zhong, and Y.-H. Liu, "Modal-graph 3d shape servoing of deformable objects with raw point clouds," *The International Journal of Robotics Research*, vol. 42, no. 14, pp. 1213–1244, 2023.
- [28] M. Afshar, J. Carriere, H. Rouhani, T. Meyer, R. S. Sloboda, S. Husain, N. Usmani, and M. Tavakoli, "Accurate tissue deformation modeling using a kalman filter and admmm-based projective dynamics," *IEEE/ASME Transactions on Mechatronics*, vol. 27, no. 4, pp. 2194–2203, 2022.
- [29] J. Zhang, Y. Zhong, and C. Gu, "Deformable models for surgical simulation: a survey," *IEEE reviews in biomedical engineering*, vol. 11, pp. 143–164, 2017.
- [30] L. Xu, Y. Lu, and Q. Liu, "Integrating viscoelastic mass spring dampers into position-based dynamics to simulate soft tissue deformation in real time," *Royal Society open science*, vol. 5, no. 2, p. 171587, 2018.
- [31] P. Battaglia, R. Pascanu, M. Lai, D. Jimenez Rezende *et al.*, "Interaction networks for learning about objects, relations and physics," *Advances in neural information processing systems*, vol. 29, 2016.
- [32] F. Makiyeh, M. Marchal, F. Chaumette, and A. Krupa, "Indirect positioning of a 3d point on a soft object using rgb-d visual servoing and a mass-spring model," in *2022 17th International Conference on Control, Automation, Robotics and Vision (ICARCV)*. IEEE, 2022, pp. 235–242.
- [33] S. Bouaziz, S. Martin, T. Liu, L. Kavan, and M. Pauly, "Projective dynamics: Fusing constraint projections for fast simulation," in *Seminal Graphics Papers: Pushing the Boundaries, Volume 2*, 2023, pp. 787–797.
- [34] T. Du, K. Wu, P. Ma, S. Wah, A. Spielberg, D. Rus, and W. Matusik, "Diffpd: Differentiable projective dynamics," *ACM Transactions on Graphics (TOG)*, vol. 41, no. 2, pp. 1–21, 2021.
- [35] T. Liu, S. Bouaziz, and L. Kavan, "Quasi-newton methods for real-time simulation of hyperelastic materials," *Acm Transactions on Graphics (TOG)*, vol. 36, no. 3, pp. 1–16, 2017.
- [36] F. Faure, C. Duriez, H. Delingette, J. Allard, B. Gilles, S. Marchesseau, H. Talbot, H. Courtecuisse, G. Bousquet, I. Peterlik *et al.*, "Sofa: A multi-model framework for interactive physical simulation," *Soft tissue biomechanical modeling for computer assisted surgery*, pp. 283–321, 2012.
- [37] Y. Hu, T.-M. Li, L. Anderson, J. Ragan-Kelley, and F. Durand, "Taichi: a language for high-performance computation on spatially sparse data structures," *ACM Transactions on Graphics (TOG)*, vol. 38, no. 6, p. 201, 2019.
- [38] F. Xu, H. Wang, W. Chen, and Y. Miao, "Visual servoing of a cable-driven soft robot manipulator with shape feature," *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 4281–4288, 2021.
- [39] X. Liang, F. Liu, Y. Zhang, Y. Li, S. Lin, and M. Yip, "Real-to-sim deformable object manipulation: Optimizing physics models with residual mappings for robotic surgery," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2024, pp. 15 471–15 477.
- [40] A. Dutta, E. Burdet, and M. Kaboli, "Predictive visuo-tactile interactive perception framework for object properties inference," *IEEE Transactions on Robotics*, 2025.