2026 IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob 2026)
August 1-4, 2026. Edmonton, Canada

Using a Surgical Physics Simulator as the Predictive Model in Model Predictive Path Integral Control for Soft-Tissue Manipulation

Sajjad Hussain[1,2], Mahdi Tavakoli[2,3], Bruno Siciliano[1], and Fanny Ficuciello[1]

Abstract— **Tissue manipulation is a fundamental and frequently performed sub-task in robotic surgery, but its automation remains challenging because soft tissue exhibits nonlinear deformation and complex contact interactions with rigid instruments. This work presents a simulator-in-the-loop, sampling-based MPC framework that uses *CRESSim*, a GPU-accelerated Unity-PhysX 5 surgical simulation environment, directly as a black-box predictive model inside a Model Predictive Path Integral (MPPI) controller. At each control step, MPPI samples candidate gripper command sequences, evaluates them through forward rollouts in the simulator, and updates the control sequence without requiring analytical tissue dynamics, Jacobians, or local linearizations. We evaluated the approach in planar and volumetric tissue-manipulation tasks, including target-point placement and tissue triangulation. The results show target-point positioning errors within the millimeter range, with representative trials reaching below 2 mm, demonstrating the feasibility of using a surgical physics simulator as the predictive model for closed-loop autonomous soft-tissue manipulation.**

I. INTRODUCTION

Robot-assisted minimally invasive surgery (RAMIS) enables dexterous manipulation through small incisions, reducing tissue trauma relative to open surgery while improving visualization and instrument dexterity compared with conventional laparoscopy [1]. These advantages have driven widespread clinical adoption; however, RAMIS remains largely teleoperated, leaving surgeons responsible for demanding subtasks that contribute to cognitive load [2] and longer operative times [3]. This motivates targeted autonomy for well-defined surgical subtasks [4].

While autonomy has been explored for tasks such as suturing [5], [6], [7] and cutting [8], many procedures also require controlled tissue repositioning to expose anatomy, establish tension, or align internal targets. A representative formulation is Indirect Simultaneous Positioning (ISP) [9], which requires fine control of tissue deformation so that internal points are positioned accurately while minimizing local strain and stress on surrounding regions. ISP is challenging because soft tissues exhibit nonlinear and viscoelastic behavior, patient-specific parameters with frictional contacts, and time-varying boundary conditions, making it substantially more complex than generic soft-object manipulation. Existing approaches for tissue manipulation through ISP falls into two categories: model-based and data-driven (model-free) methods.

Model-based controllers compute control commands using an explicit tissue model, either in feedback or predictive form. Early work often used mass-spring-damper (MSD) models because they are computationally efficient and easy to integrate with classical control. For instance, Hirai et al. [10] used an MSD approximation for multi-finger indirect target positioning, Wada et al. [11] proposed a PID controller robust to modeling uncertainty, and Torabi et al. [12] used a 2D MSD model for tissue manipulation and needle steering. Despite their practicality, MSD models often rely on manually tuned stiffness and damping parameters, which limits their physical fidelity under large deformations, nonlinear stress-strain responses, and contact-rich interactions [13].

Finite element method (FEM)-based approaches improve physical fidelity by modeling tissue as a discretized continuum. Patil et al. [14] used FEM model to compute safe tissue retraction trajectories under strain constraints. Ficuciello et al. [15] proposed an inverse-FEM scheme to compute grasp forces that drive selected points to desired positions. Likewise, Afshar et al. [16] embedded a nonlinear FEM formulation, solved via ADMM, within MPC for multi-point tissue manipulation. However, these approaches remain computationally demanding: real-time nonlinear FEM-based control often requires linearization, model reduction, or other approximations, which can introduce model mismatch.

In contrast to model-based approaches, data-driven methods avoid explicit tissue mechanics by learning local deformation models, surrogate dynamics, or control policies from interaction data. For instance, Navarro et al. [17] estimated a local deformation Jacobian online from visual feedback and used adaptive control for shape regulation, while Hu et al. [18] trained an online deep neural network that maps end-effector motions to shape changes without an analytic dynamics model. Shin et al. [19] combined reinforcement learning (RL) and learning from demonstrations (LfD) to learn image-space tissue dynamics and used MPC to position surgeon-selected tissue points. With the rise of RL and LfD, more recent work has learned direct manipulation policies for ISP and related tasks. Ou and Tavakoli [9] trained a deep RL policy for indirect simultaneous positioning and demonstrated sim-to-real transfer; later, they added human

*This work is supported by ERC grant EndoTheranostics, 101118626. Funded by the European Union. Views and opinions expressed are however, those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

[1]Sajjad Hussain, Fanny Ficuciell and Bruno Siciliano are with the ICAROS Center (PRISMA Lab), University of Naples Federico II, Naples, Italy sajjad.hussain, fanny.ficuciell, bruno.siciliano@unina.it

[2]Sajjad Hussain and Mahdi Tavakoli is with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Canada mahdi.tavakoli@ualberta.ca

[3]Mahdi Tavakoli is also with the Department of Biomedical Engineering, University of Alberta, Edmonton, Canada

supervision and adversarial imitation learning to improve safety and data efficiency [20]. Shahkoo et al. [21] proposed a hybrid DRL-evolutionary method to optimize tensioning in surgical pattern cutting.

Learning-based tissue manipulation often demands extensive interaction data, yet collecting surgical data is costly and safety-critical. This has motivated the use of physics-based surgical simulators both for data generation and closed-loop evaluation. AMBF [22] and related PBD environments have been used for RL-based surgical manipulation [23], [24] and perception-based deformable control [25]. Although PBD is computationally efficient, its rule-based updates can limit physical fidelity and interpretability, motivating continuum-mechanics-based simulators such as SOFA/LapGym [26], MPM-based extensions to SurRoL [27], and *CRESSim* [28]. *CRESSim* is a GPU-accelerated, visually rich, multi-physics surgical VR environment built in Unity with NVIDIA PhysX 5 that integrates rigid bodies, deformables, collision handling, grasping constraints, and fluid effects, supporting contact-rich interactions such as grasping, cutting, and suction. These capabilities motivate inference-time optimization, where actions are recomputed online from the current state using the simulator as a predictive model.

Building on this idea, we use *CRESSim* not as an offline data generator, but directly as a black box predictive model inside a receding-horizon controller. In this paper, we study autonomous tissue manipulation for internal target-point placement and tissue triangulation, where selected tissue points must be positioned through boundary grasps on volumetric deformable tissue. This setting moves beyond thin-shell or quasi-planar assumptions by considering 3D deformation, grasping constraints, tool-tissue contact, and surrounding soft-tissue attachments within the same simulated surgical scene.

The controller is based on Model Predictive Path Integral (MPPI) control, which is well matched to this black-box setting because it requires only forward rollouts rather than analytical dynamics, Jacobians, or local linearizations [29]. At each control step, MPPI samples candidate gripper command sequences, evaluates them through *CRESSim* rollouts, converts the resulting trajectory costs into importance weights, and updates the nominal sequence in a receding-horizon loop. Thus, the framework requires only forward simulation, task objectives, and constraints, while avoiding manual derivation of contact-rich multi-physics dynamics. Compared with lightweight MSD/PBD controllers or trained policies, this approach is not intended to be the fastest or most computationally lightweight; rather, it trades rollout cost for modeling flexibility in contact-rich volumetric scenes where the relevant dynamics are difficult to derive analytically or learn from limited data.

A. Contributions

The main contributions of this paper are:

1) We demonstrate closed-loop autonomous tissue manipulation using a high-fidelity surgical simulator directly as the predictive model within sampling based MPC (MPPI), establishing a general simulator-in-the-loop control paradigm for contact-rich surgical tasks.
2) We show that MPPI combined with GPU-accelerated physics rollouts achieves millimeter-scale accuracy on volumetric soft tissue without explicit mechanics models.
3) We validate robustness to modeling mismatch and scalability to multi-point, multi-gripper manipulation.

II. Problem Statement

Our objective is to drive a set of target points to desired positions by manipulating boundary control points, subject to kinematic, workspace, and tissue-safety constraints. We assume rigid attachment between each gripper and its grasped boundary nodes, so commanding the gripper pose is equivalent to moving the corresponding control points. The task is therefore naturally posed as a finite-horizon trajectory-optimization (optimal control) problem:

$$\min_{\{u_t\}_{t=0}^{N-1}} \quad \Phi(x_N) \; + \; \sum_{t=0}^{N-1} \Big(\|y_t - y^\star\|_Q^2 \; + \; \|\Delta u_t\|_R^2 \Big) \tag{1a}$$

$$\text{s.t.} \quad x_{t+1} = f(x_t, u_t), \qquad t = 0, \dots, N-1, \tag{1b}$$

$$\tag{1c}$$

$$x_0 \text{ given}, \qquad x_{\mathrm{fp}}(t) = x_{\mathrm{fp}}(0), \tag{1d}$$

$$y_t = h(x_t), \qquad y_t \in \mathbb{R}^{3p}, \tag{1e}$$

$$x_t \in \mathcal{X}, \;\; u_t \in \mathcal{U}, \;\; g(x_t) \le 0, \tag{1f}$$

$$\Delta u_t := u_t - u_{t-1}, \;\; \Delta u_0 := \mathbf{0}.$$

Here, f represents the tissue dynamics, h extracts the positions of the p target points that determine the target deformation, $Q \succeq 0$ weights tracking error, and $R \succeq 0$ penalizes control rate. The terminal term Φ encourages convergence at the horizon. The admissible sets $\mathcal{X}$ and $\mathcal{U}$ encode workspace and rate/velocity limits, while $g(\cdot) \le 0$ aggregates safety constraints such as strain/stress and contact-force bounds, formulated based on the tissue failure data reported in [30], [31]. Fixed anchors (e.g., in triangulation tasks) are enforced by (1c).

III. Simulation environment

Solving (1) with high deformation fidelity requires a predictive model that captures large-strain volumetric mechanics, intermittent frictional contacts, and grasp constraints. While mass-spring-damper (MSD) and position-based dynamics (PBD) models are computationally efficient, they only approximate continuum behavior and often exhibit limited accuracy in force transmission and deformation under contact-rich interaction [9], [32]. Moreover, *in vivo* response depends strongly on surrounding fascia, organ attachments, and tool-tissue interaction, making it difficult to hand-craft a sufficiently accurate analytic model for closed-loop control. To address this, we replace the dynamics map $f(\cdot)$ in (1b) with a GPU-accelerated one-step simulation rollout implemented in *CRESSim* (Unity + NVIDIA PhysX 5).

Specifically, we treat the simulator as a black-box transition operator

$$x_{t+1} = \tilde{f}(x_t, u_t), \tag{2}$$

where $\tilde{f}(\cdot)$ advances the full surgical scene forward by one control step under the commanded gripper action. Unity provides the scene graph and authoring tools, while PhysX 5 provides rigid/deformable solvers and contact handling, enabling interactive-rate simulation of contact-rich surgical scenarios. The organs in simulation environment are discretized as FEM soft bodies with tunable material parameters (e.g., Young's modulus E, Poisson's ratio ν) and are mechanically coupled to adjacent deformable structures via PhysX 5 soft–soft attachment constraints. CRESSim's Unity-native tooling (plugin plus inspector-driven configuration) allows straightforward specification of tissue meshes, material tuning, grasp constraints, and real-time tracking of internal target-point error, yielding high-fidelity deformation and stable contact at interactive rates suitable for simulator-in-the-loop experimentation.

A. Tissue Model

PhysX 5 employs a fixed-corotated FEM formulation with a robust constraint-based contact/attachment solver, yielding realistic deformation under manipulation and stable contact dynamics. Let $\mathbf{x} \in \mathbb{R}^{3n}$ denote the stacked nodal positions of an n-node tetrahedral mesh with lumped mass matrix $M \in \mathbb{R}^{3n\times 3n}$, then system dynamics are modeled using Newton's law:

$$M\,\ddot{\mathbf{x}} \;=\; \mathbf{F}_{\text{ext}} \;-\; \mathbf{F}_{\text{int}} \;-\; \mathbf{D}\,\dot{\mathbf{x}}, \tag{3}$$

where $\mathbf{F}_{\text{ext}}$ collects external loads (e.g., gravity, contact/attachment reactions), $\mathbf{F}_{\text{int}} = -\partial\Psi(\mathbf{x})/\partial\mathbf{x}$ is the internal force derived from the strain-energy potential. In PhysX 5, this elastic response is modeled using a fixed-corotated FEM energy density, which accounts for large rotations while retaining the efficiency of a linear elastic model in the corotated frame and is mathematically given as [33]:

$$\psi(\mathbf{F}) \;=\; \mu\,\|\mathbf{F}-\mathbf{R}\|_F^2 \;+\; \tfrac{\lambda}{2}\left(\det(\mathbf{F})-1\right)^2, \tag{4}$$

where $\mathbf{F}$ is the deformation gradient, $\mathbf{R}$ is the rotational component from the polar decomposition $\mathbf{F} = \mathbf{RS}$, and μ, λ are the Lamé parameters:

$$\mu = \tfrac{E}{2(1+\nu)}, \qquad \lambda = \tfrac{E\,\nu}{(1+\nu)(1-2\nu)}, \tag{5}$$

In implementation, PhysX 5 advances the deformable-body constraints using an iterative XPBD solver with Temporal Gauss-Seidel substepping; for details, see [33], [34].

a) Soft–soft attachment (replicating fascia): To attach two deformable bodies A and B, PhysX 5 anchors an embedded material point in A to an embedded material point in B inside their local tetrahedron collision mesh. The points are connected through the following hard constraint:

$$\mathbf{C}(\mathbf{v},\mathbf{w}) = \mathbf{x}_A - \mathbf{x}_B = 0, \tag{6}$$

with $\mathbf{x}_A = \sum_i \alpha_i \mathbf{v}_i$, and $\mathbf{x}_B = \sum_i \beta_i \mathbf{w}_i$, where $\mathbf{v}_i$ and $\mathbf{w}_i$ are the vertices of the collision tetrahedral meshes of the two bodies, and α and β are the corresponding barycentric weights computed once (at the time of creating the attachment) by locating the containing tetrahedra and evaluating barycentric coordinates of the selected surface/contact points. This constraint ensures that the two bodies can move and deform under forces while still connected.

b) Soft–rigid (grasper) attachment: To grasp and move a specific tissue point, we have used explicit attachment constraints, which anchor a vertex $\mathbf{x}_s$ (the points that are considered to be grasped) to a point on the rigid tool and is given as:

$$\mathbf{C}(\mathbf{x}_s, \mathbf{p}_R, \mathbf{R}_R) \;=\; \mathbf{x}_s \;-\; \left(\mathbf{p}_R + \mathbf{R}_R\,\mathbf{a}\right) \;=\; \mathbf{0}, \tag{7}$$

where $\mathbf{p}_R, \mathbf{R}_R$ represents the tool's position and orientation, and $\mathbf{a}$ is the anchor expressed in the tool's actor space. In our implementation, $\mathbf{x}_s$ is a collision-mesh vertex.

IV. Model Predictive Path Integral Control

Unlike gradient-based MPC, which requires a differentiable model and struggles with nonconvexities or contact discontinuities, MPPI is a derivative-free stochastic optimal control method [35], [29]. It evaluates noisy open-loop control sequences through forward simulation and steers a nominal sequence toward low-cost regions by importance reweighting of the sampled perturbations.

Consider a discrete-time system

$$\mathbf{x}_{t+1} \;=\; F(\mathbf{x}_t, \mathbf{v}_t), \tag{8}$$

where $\mathbf{x}_t \in \mathcal{X} \subset \mathbb{R}^{n_x}$ is the state and $\mathbf{v}_t = \mathbf{u}_t + \boldsymbol{\epsilon}_t \in \mathcal{U} \subset \mathbb{R}^{n_u}$ is the applied control, with $\mathbf{u}_t$ the nominal input and $\boldsymbol{\epsilon}_t \sim \mathcal{N}(\mathbf{0}, \Sigma)$ the sampled perturbation.

Let $l(\mathbf{x}_t, \mathbf{u}_t)$ denote the running cost and $\phi(\mathbf{x}_T)$ the terminal cost. MPPI frames the optimal control problem as the minimization of the expected cost under the sampling distribution:

$$\mathbf{U}^* = \arg\min_{\mathbf{U}\in\mathcal{U}^T} \; \mathbb{E}_{\boldsymbol{\epsilon}_{0:T-1}\sim\mathcal{Q}}\left[\phi(\mathbf{x}_T) + \sum_{t=0}^{T-1} l(\mathbf{x}_t, \mathbf{u}_t)\right], \tag{9}$$

where $\mathbf{U} := [\mathbf{u}_0^\top, \ldots, \mathbf{u}_{T-1}^\top]^\top$ and T is the prediction horizon. As discussed in [29], solving (9) is equivalent to minimizing the Kullback–Leibler divergence between the optimal distribution $\mathcal{Q}^*$ and the controlled distribution $\mathcal{Q}_{\mathbf{U},\Sigma}$:

$$\mathbf{U}^* \;=\; \arg\min_{\mathbf{U}\in\mathcal{U}^T} D_{\mathrm{KL}}(\mathcal{Q}^* \,\|\, \mathcal{Q}_{\mathbf{U},\Sigma})\,. \tag{10}$$

Direct minimization of (10) is intractable, as the optimal distribution $\mathcal{Q}^*$ is defined only up to a normalizing constant, and its density depends on all possible trajectory rollouts through the dynamics. MPPI therefore employs importance sampling from a proposal $\mathcal{Q}(\mathbf{U}, \Sigma)$ centered at the current nominal control sequence $\mathbf{U}$. Concretely, At each control instant, draw M i.i.d. perturbation sequences

$$\{\boldsymbol{\epsilon}_t^{(m)}\}_{t=0}^{T-1} \sim \mathcal{Q}(0, \Sigma), \qquad m = 1, \ldots, M,$$

and form M perturbed control sequences

$$\mathbf{U}^{(m)} = [\,\mathbf{u}_0 + \boldsymbol{\epsilon}_0^{(m)},\; \mathbf{u}_1 + \boldsymbol{\epsilon}_1^{(m)},\; \ldots,\; \mathbf{u}_{T-1} + \boldsymbol{\epsilon}_{T-1}^{(m)}\,]. \tag{11}$$

Rolling each $\mathbf{U}^{(m)}$ through (8) (in our case the CREESim scene) yields trajectories $\mathbf{X}^{(m)} = (\mathbf{x}_0^{(m)}, \ldots, \mathbf{x}_T^{(m)})$ with cost

$$J^{(m)} = \phi\big(\mathbf{x}_T^{(m)}\big) \; + \; \sum_{t=0}^{T-1} l\big(\mathbf{x}_t^{(m)}, \mathbf{u}_t^{(m)}\big). \tag{12}$$

Given these accumulated costs, MPPI computes importance weights

$$w_m \;=\; \frac{\exp\big(-\,(J^{(m)} - J_{\min})/\lambda\big)}{\sum_{j=1}^{M} \exp\big(-\,(J^{(j)} - J_{\min})/\lambda\big)}, \tag{13}$$

where $\lambda > 0$ is the inverse temperature and $J_{\min} = \min_j J^{(j)}$ stabilizes the exponent. The control is then updated by a weighted average of perturbations:

$$\mathbf{u}_t^{\star} \;=\; \mathbf{u}_t \;+\; \sum_{m=1}^{M} w_m \, \boldsymbol{\epsilon}_t^{(m)}, \qquad t = 0, 1, \ldots, T-1, \tag{14}$$

which steers the nominal sequence $\mathbf{U}$ toward lower-cost sampled trajectories. After updating $\mathbf{U}$, apply only the first control $\mathbf{u}_0^{\star}$, shift the horizon forward and repeat with the new state measurement.

V. Results

To evaluate the proposed simulator in the loop sampling-based control algorithm, we conducted two sets of experiments: **(i) planar patch manipulation** and **(ii) 3D volumetric tissue manipulation**, each instantiating a representative instance of the Indirect Simultaneous Positioning (ISP) problem. In both scenes, the objective is to drive a set of internal target points to the prescribed goal position by manipulating a set of boundary control points (through gripper-attached nodes). The planar patch scene represents surface-like deformation where targets lie on a surface sheet/patch, while the volumetric scene represents bulk tissue deformation where internal point motion depends strongly on 3D material response and tool-tissue interaction. These two scenarios provide complementary testbeds for assessing controller performance across deformation regimes.

In both tasks, Unity-CRESSim is used as the predictive dynamics model inside MPPI: each sampled control sequence is forward simulated to generate a trajectory, from which we evaluate target-point tracking costs and safety-related penalties/constraints. Rollouts are evaluated in parallel, leveraging MPPI sampling and GPU-accelerated simulation to enable efficient trajectory assessment.

Because this study is simulation-only, we emulate a sim-to-real gap via an explicit model-plant mismatch. The simulator instance used as the predictor inside MPPI is assigned mechanical parameters that differ from those of the "plant" scene being controlled (e.g., stiffness-related settings), yielding systematic dynamics uncertainty while preserving identical sensing and actuation interfaces. This setup allows us to evaluate the robustness of simulator-in-the-loop MPPI when the predictive model is imperfect.

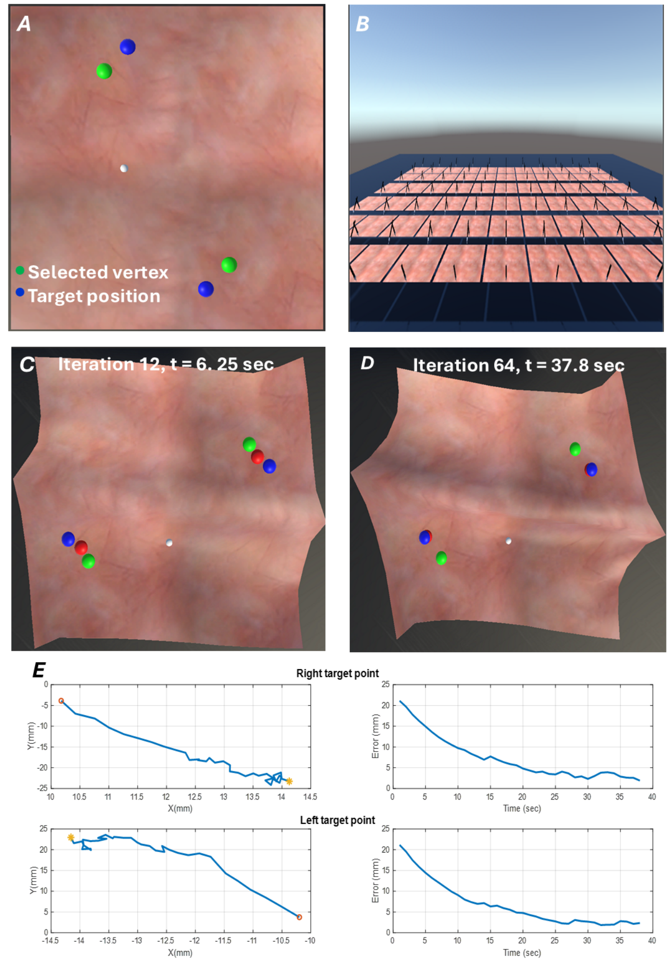

Fig. 1. A sequence of snapshots during tissue manipulation. The target point (green) is manipulated through the control point (blue) to move it to the desired position (red).

A. 2-D planner tissue

As our first subcase, we consider a 2-D *planar* tissue patch (Fig. 1) stabilized by clamping its four corners. The robot end-effectors (PSM grippers), rigidly grasping the boundary points, are constrained to in-plane translations with fixed orientation. The controller then commands planar gripper motions to drive the target points toward the desired position.

For MPPI, the prediction and control horizons are fixed to $N_p = 10$ and $N_c = 1$, respectively. At each control cycle, we draw $M = 300$ zero-mean Gaussian perturbation sequences (with dimension matching the in-plane control DoFs) and add them to the nominal control sequence to form the candidate sequences. These candidates are evaluated in the Unity–CRESSim prediction instance, which executes multiple rollouts in parallel (see Fig. 1B). Before each rollout, the predictor is reinitialized to the current plant ("real-world") state; trajectories are then propagated, and their accumulated costs are used to compute the MPPI update. The controller terminates when either a maximum iteration count is reached or the target-point Euclidean error falls below 2.5 mm. This threshold was chosen as a practical engineering tolerance for millimeter-scale positioning in simulation: it is small relative to the commanded tissue displacements, while avoiding unnecessary additional MPPI iterations after the target has already reached the desired region.

Fig. 1 A–B shows the plant (real) scene and the parallel prediction instance for a thin-shell (cloth-like) tissue model. Figure 1C–D provides representative snapshots as the PSM gripper manipulates the grasped boundary point to drive the current target (green) to its desired position (blue). In this run, MPPI placed the target at the desired location in $\approx 37.8\,\mathrm{s}$ (64 iterations), with a terminal error of $\approx 2.15\,\mathrm{mm}$ (Fig. 1F). To assess robustness, we selected four distinct target points with corresponding desired poses and repeated each configuration 50 times; across all cases, the terminal error remained below $2.5\,\mathrm{mm}$ (see Fig. 2).

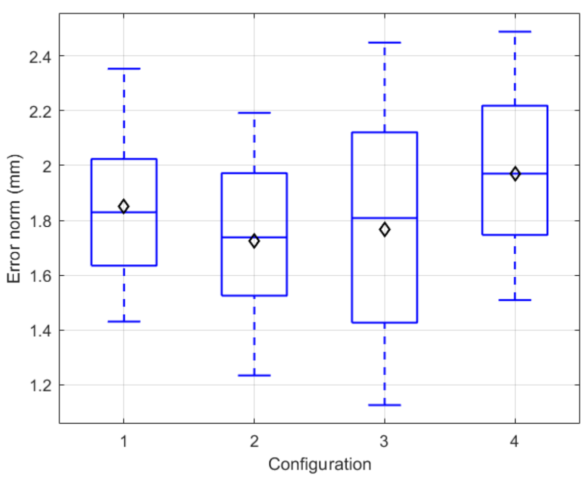

Fig. 2. Positioning error for 4 different target points

We also investigated the effect of the rollout budget M. As expected, increasing M improved success rates and reduced the number of iterations required to complete the task, at the expense of higher per-iteration compute (Fig. 3). Concretely, for $M = 300$ we observed an average iteration count $\mathrm{IT} = 72$, a success rate of 93%, and a per-iteration wall time $t_{\mathrm{iter}} = 0.59\,\mathrm{s}$; for $M = 500$, $\mathrm{IT} = 65$ with $t_{\mathrm{iter}} = 1.29\,\mathrm{s}$; and for $M = 700$, $\mathrm{IT} = 56$ with $t_{\mathrm{iter}} = 2.25\,\mathrm{s}$. Given this trade-off, we used $M = 300$ for the remaining experiments.

B. Volumetric tissue case.

To evaluate the proposed framework on volumetric tissue, we consider a clinically motivated scenario in which the surgeon selects a surface point on the liver and specifies a desired 3D location for that point (e.g., to improve exposure or align an internal target). Fig. 4 illustrates a representative trial for this case. For stability we clamp the liver at one corner to hold the structure in place. Target points were selected on the surface, and the gripper was attached at a preselected grasp point. Using the MPPI and Unity-CRESSim configuration described above, we tested four three-dimensional configurations of target points and desired positions. Fig. 4A–D illustrates a representative volumetric

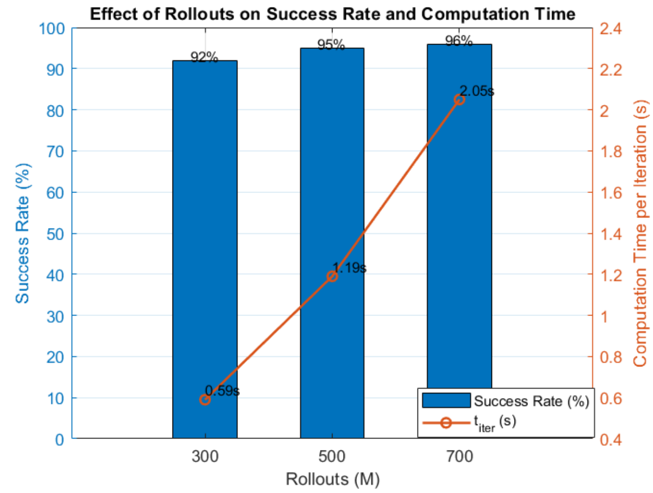

Fig. 3. Effect of the number of rollouts M on success rate and computation time.

manipulation trial. Starting from the initial configuration (A), the controller progressively deforms the liver through the grasp point (white) so that the selected surface vertex (green) approaches the prescribed target (blue), reaching close alignment by $t \approx 130s(D)$. The corresponding control inputs applied at the gripper are shown in Fig. 4E. The commanded signals remain bounded and exhibit mild high-frequency fluctuations ("chatter"), which is expected in sampling-based MPC when operating on contact-rich deformable dynamics and finite-horizon rollouts; importantly, these fluctuations do not destabilize the interaction and the overall control action remains coherent, driving steady error reduction. Fig. 4F shows the 3D trajectory of the selected vertex, which evolves from the initial position (open circle) toward the goal (star). Over the course of the trial, the positioning error decreases from approximately $35mm$ to $<2mm$ by $t \approx 130s$ with a final offset of $\approx 1.7mm$.

To assess efficacy on a multi-point task, we used the proposed framework to automate tissue triangulation: a common subtask in which the surgeon expands a tissue patch to expose and tension the operative field. This task typically forces frequent handovers between main and assistant forceps; here, both grasp points are driven simultaneously by MPPI. For this, we built a scene with a liver model (see Fig. 5) and instrumented it with one fixed boundary point (black) and two movable surface vertices that must be driven to their target locations (blue markers) by two grasp points (white spheres). The current triangle connecting the anchor to the instantaneous vertex positions is drawn with black dashed edges, while the desired triangle formed by the anchor and the targets is shown with solid green edges.

Fig. 5 panels A–C visualize the evolution of the triangle formed by the fixed point and the two grasped vertices: the initial configuration (A) exhibits a clear mismatch between the current triangle (dashed) and the desired triangle (solid). As MPPI updates the two gripper motions, the tissue is progressively retracted and tensioned, reducing the geometric

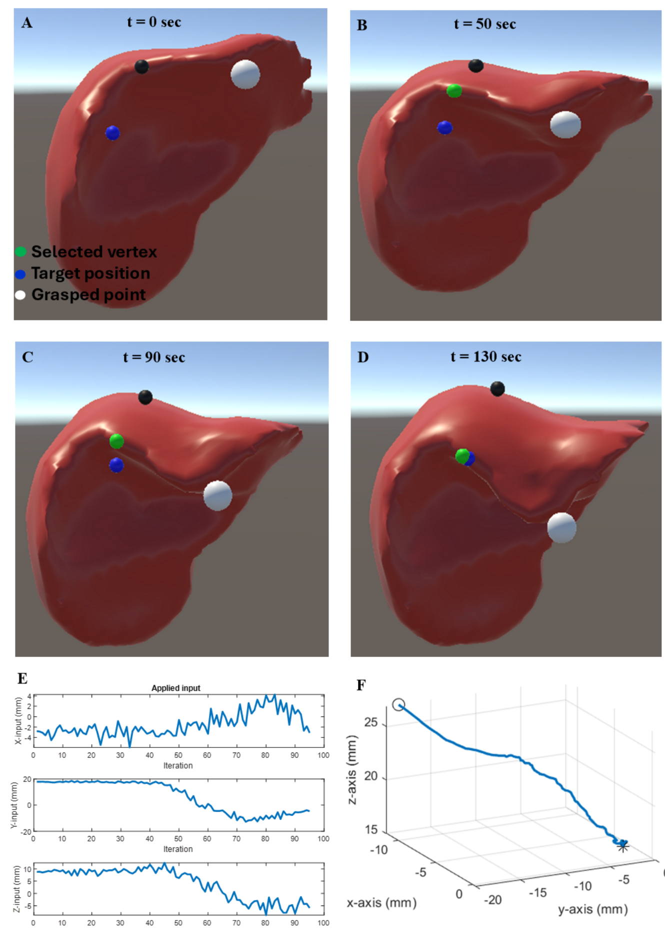

Fig. 4. Manipulation of a point on the liver surface using MPPI with the simulator in the loop. **A–D**: Snapshots as the gripper (white) drives the selected tissue vertex (green) to the target position (blue). **E**: Positioning error of the target point over time **H**: movement of the target point starting at ⋆ ending at ∘.

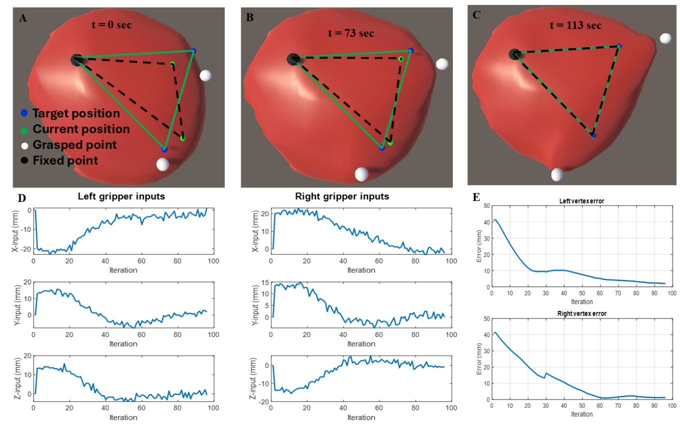

Fig. 5. Tissue triangulation with two grippers and a fixed anchor in Unity-CRESSim. **A-C:** Task progression: Target vertices are shown in blue and current vertices in green; grasp points are indicated by white spheres. **D:** Left and right gripper control inputs over MPPI iterations. **E:** Tracking error of the left and right target vertices, converging below the $2\,\mathrm{mm}$ tolerance.

mismatch (B) and ultimately bringing both grasped vertices into alignment with their targets by $t = 113 sec$). The corresponding commanded inputs for the left and right grippers are shown in the panel D; similar to the previous, the input signal remains bounded with high-frequency chatter. Panel E reports the per-vertex tracking errors, which decrease steadily and fall below the $2.5\,\mathrm{mm}$ tolerance for both vertices, confirming successful triangulation. Overall, the results demonstrate that simulator-in-the-loop MPPI can coordinate two boundary grasps to satisfy a multi-point geometric constraint on deformable tissue.

VI. Conclusion and Future Work

This paper presented a simulator-in-the-loop, sampling-based model predictive control framework for autonomous soft-tissue manipulation. By embedding the Unity-PhysX 5/*CRESSim* simulator directly inside the model predictive path integral (MPPI) rollout loop, the controller computes actions through forward simulation rather than hand-crafted analytical dynamics, Jacobians, or local linearizations, while retaining explicit user-defined objectives and constraints. In simulation tasks involving target-point placement and tissue triangulation, the proposed framework achieved millimeter-scale target-point positioning, demonstrating the feasibility of using a surgical physics simulator as a black-box predictive model for contact-rich deformable-tissue manipulation. More broadly, the approach provides a control paradigm for surgical scenes in which volumetric deformation, grasping constraints, tool-tissue contact, and surrounding attachments are represented in the simulator. Provided that the simulator captures the task-relevant interactions and exposes forward rollouts under candidate control inputs, the same framework can be applied to other contact-rich surgical manipulation tasks supported within the same physics environment.

The study has several limitations. The speed and accuracy required for clinical use are procedure-dependent: slow tissue-positioning or retraction tasks may tolerate lower update rates if the target remains stable, whereas manipulation near critical anatomy or rapidly changing scenes would require higher control bandwidth and stronger safety validation. In the present implementation, the controller update rate is bounded by mesh resolution, contact/attachment complexity, and MPPI rollout budget, yielding updates on the order of a few hertz. Thus, the results should be interpreted as a simulation feasibility study rather than evidence of clinical readiness. We also assumed rigid, slip-free grasps and preselected grasp sites. Because grasp choice strongly affects tissue deformation and convergence, future work will integrate grasp-point selection into the optimization loop to favor configurations that achieve the task with lower overall deformation. Finally, although *CRESSim* captures geometry and constraints, tissue parameters and attachment mechanics remain uncertain and patient-specific; online parameter adaptation will therefore be needed to reduce model-plant mismatch over time.

ACKNOWLEDGMENT

REFERENCES

[1] P. Biswas, S. Sikander, and P. Kulkarni, "Recent advances in robot-assisted surgical systems," *Biomedical Engineering Advances*, vol. 6, p. 100109, 2023.

[2] C. Lim, M. Obuseh, J. Cha, J. Steward, C. Sundaram, and D. Yu, "Neural insights on expert surgeons' mental workload during live robotic surgeries," *Scientific Reports*, vol. 15, no. 1, p. 12073, 2025.

[3] T.-J. Lai, C. Roxburgh, K. A. Boyd, and J. Bouttell, "Clinical effectiveness of robotic versus laparoscopic and open surgery: an overview of systematic reviews," *BMJ open*, vol. 14, no. 9, p. e076750, 2024.

[4] A. Singh, W. Shi, and M. D. Wang, "Autonomous soft tissue retraction using demonstration-guided reinforcement learning," *arXiv preprint arXiv:2309.00837*, 2023.

[5] A. Shademan, R. S. Decker, J. D. Opfermann, S. Leonard, A. Krieger, and P. C. Kim, "Supervised autonomous robotic soft tissue surgery," *Science translational medicine*, vol. 8, no. 337, pp. 337ra64–337ra64, 2016.

[6] P. Marra, S. Hussain, M. Caianiello, and F. Ficuciello, "Mpc for suturing stitch automation," *IEEE Transactions on Medical Robotics and Bionics*, 2024.

[7] H. Saeidi, J. Opfermann, M. Kam, S. Wei, S. Leonard, M. Hsieh, J. Kang, and A. Krieger, "Autonomous robotic laparoscopic surgery for intestinal anastomosis. sci robot. 2022; 7 (62): eabj2908."

[8] T. Nguyen, N. D. Nguyen, F. Bello, and S. Nahavandi, "A new tensioning method using deep reinforcement learning for surgical pattern cutting," in *2019 IEEE international conference on industrial technology (ICIT)*. IEEE, 2019, pp. 1339–1344.

[9] 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.

[10] S. Hirai and T. Wada, "Indirect simultaneous positioning of deformable objects with multi-pinching fingers based on an uncertain model," *Robotica*, vol. 18, no. 1, pp. 3–11, 2000.

[11] T. Wada, S. Hirai, S. Kawamura, and N. Kamiji, "Robust manipulation of deformable objects by a simple pid feedback," in *Proceedings 2001 ICRA. IEEE International Conference on Robotics and Automation (Cat. No. 01CH37164)*, vol. 1. IEEE, 2001, pp. 85–90.

[12] M. Torabi, K. Hauser, R. Alterovitz, V. Duindam, and K. Goldberg, "Guiding medical needles using single-point tissue manipulation," in *2009 IEEE International Conference on Robotics and Automation*. IEEE, 2009, pp. 2705–2710.

[13] H. Yin, A. Varava, and D. Kragic, "Modeling, learning, perception, and control methods for deformable object manipulation," *Science Robotics*, vol. 6, no. 54, p. eabd8803, 2021.

[14] S. Patil and R. Alterovitz, "Automated tissue retraction for robot-assisted surgical procedures," in *IEEE Int. Conf. Rob. and Aut*, 2010.

[15] F. Ficuciello, A. Migliozzi, E. Coevoet, A. Petit, and C. Duriez, "Fem-based deformation control for dexterous manipulation of 3d soft objects," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 4007–4013.

[16] M. Afshar, J. Carriere, T. Meyer, R. S. Sloboda, S. Husain, N. Usmani, and M. Tavakoli, "A model-based multi-point tissue manipulation for enhancing breast brachytherapy," *IEEE Transactions on Medical Robotics and Bionics*, vol. 4, no. 4, pp. 1046–1056, 2022.

[17] 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.

[18] Z. Hu, T. Han, P. Sun, J. Pan, and D. Manocha, "3-d deformable object manipulation using deep neural networks," *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 4255–4261, 2019.

[19] C. Shin, P. W. Ferguson, S. A. Pedram, J. Ma, E. P. Dutson, and J. Rosen, "Autonomous tissue manipulation via surgical robot using learning based model predictive control," in *2019 International conference on robotics and automation (ICRA)*. IEEE, 2019, pp. 3875–3881.

[20] Y. Ou and M. Tavakoli, "Autonomous soft-tissue needle steering using reinforcement learning guided by human input," *Journal of Medical Robotics Research*, vol. 9, no. 03n04, p. 2440010, 2024.

[21] A. A. Shahkoo and A. A. Abin, "Autonomous tissue manipulation via surgical robot using deep reinforcement learning and evolutionary algorithm," *IEEE Transactions on Medical Robotics and Bionics*, vol. 5, no. 1, pp. 30–41, 2023.

[22] A. Munawar and G. S. Fischer, "An asynchronous multi-body simulation framework for real-time dynamics, haptics and learning with application to surgical robots," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2019, pp. 6268–6275.

[23] V. M. Varier, D. K. Rajamani, F. Tavakkolmoghaddam, A. Munawar, and G. S. Fischer, "Ambf-rl: A real-time simulation based reinforcement learning toolkit for medical robotics," in *2022 International Symposium on Medical Robotics (ISMR)*. IEEE, 2022, pp. 1–8.

[24] E. Tagliabue, A. Pore, D. Dall'Alba, E. Magnabosco, M. Piccinelli, and P. Fiorini, "Soft tissue simulation environment to learn manipulation tasks in autonomous robotic surgery," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 3261–3266.

[25] B. Thach, A. Kuntz, and T. Hermans, "Deformernet: A deep learning approach to 3d deformable object manipulation," *arXiv preprint arXiv:2107.08067*, 2021.

[26] P. M. Scheikl, B. Gyenes, R. Younis, C. Haas, G. Neumann, M. Wagner, and F. Mathis-Ullrich, "Lapgym-an open source framework for reinforcement learning in robot-assisted laparoscopic surgery," *Journal of Machine Learning Research*, vol. 24, no. 368, pp. 1–42, 2023.

[27] Z. Yang, Y. Long, K. Chen, W. Wei, and Q. Dou, "Efficient physically-based simulation of soft bodies in embodied environment for surgical robot," *arXiv preprint arXiv:2402.01181*, 2024.

[28] Y. Ou, S. Zargarzadeh, P. Sedighi, and M. Tavakoli, "A realistic surgical simulator for non-rigid and contact-rich manipulation in surgeries with the da vinci research kit," in *2024 21st International Conference on Ubiquitous Robots (UR)*. IEEE, 2024, pp. 64–70.

[29] G. Williams, P. Drews, B. Goldfain, J. M. Rehg, and E. A. Theodorou, "Information-theoretic model predictive control: Theory and applications to autonomous driving," *IEEE Transactions on Robotics*, vol. 34, no. 6, pp. 1603–1622, 2018.

[30] A. R. Kemper, A. C. Santago, J. D. Stitzel, J. L. Sparks, and S. M. Duma, "Biomechanical response of human liver in tensile loading," in *Annals of advances in automotive medicine/annual scientific conference*, vol. 54, 2010, p. 15.

[31] A. F. Khan, M. K. MacDonald, C. Streutker, C. Rowsell, J. Drake, and T. Grantcharov, "Tissue stress from laparoscopic grasper use and bowel injury in humans: establishing intraoperative force boundaries," *BMJ Surgery, Interventions, & Health Technologies*, vol. 3, no. 1, p. e000084, 2021.

[32] S. Patil and R. Alterovitz, "Toward automated tissue retraction in robot-assisted surgery," in *2010 IEEE International Conference on Robotics and Automation*. IEEE, 2010, pp. 2088–2094.

[33] "Soft bodies," https://nvidia-omniverse.github.io/PhysX/physx/5.4.1/docs/SoftBodies.html#softbodies/, 2015, accessed: Sep. 26, 2025.

[34] M. Macklin and M. Muller, "A constraint-based formulation of stable neo-hookean materials," in *Proceedings of the 14th ACM SIGGRAPH conference on motion, interaction and games*, 2021, pp. 1–7.

[35] G. Williams, A. Aldrich, and E. Theodorou, "Model predictive path integral control using covariance variable importance sampling," *arXiv preprint arXiv:1509.01149*, 2015.