

Integrated planning and control of robotic surgical instruments for task autonomy

The International Journal of
Robotics Research
2023, Vol. 42(7) 504–536
© The Author(s) 2023
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/02783649231179753
journals.sagepub.com/home/ijrr



Fangxun Zhong  and Yun-Hui Liu 

Abstract

Agile maneuvers are essential for robot-enabled complex tasks such as surgical procedures. Prior explorations on surgery autonomy are limited to feasibility study of completing a single task without systematically addressing generic manipulation safety across different tasks. We present an integrated planning and control framework for 6-DoF robotic instruments for pipeline automation of surgical tasks. We leverage the geometry of a robotic instrument and propose the nodal state space to represent the robot state in $SE(3)$ space. Each elementary robot motion could be encoded by regulation of the state parameters via a dynamical system. This theoretically ensures that every in-process trajectory is globally feasible and stably reached to an admissible target, and the controller is of closed-form without computing 6-DoF inverse kinematics. Then, to plan the motion steps reliably, we propose an interactive (instant) goal state of the robot that transforms manipulation planning through desired path constraints into a goal-varying manipulation (GVM) problem. We detail how GVM could adaptively and smoothly plan the procedure (could proceed or rewind the process as needed) based on on-the-fly situations under dynamic or disturbed environment. Finally, we extend the above policy to characterize complete pipelines of various surgical tasks. Simulations show that our framework could smoothly solve twisted maneuvers while avoiding collisions. Physical experiments using the da Vinci Research Kit validates the capability of automating individual tasks including tissue debridement, dissection, and wound suturing. The results confirm good task-level consistency and reliability compared to state-of-the-art automation algorithms.

Keywords

Surgical task autonomy, robot planning and control, goal-varying manipulation, surgical robotics

Received 21 April 2022; Revised 10 January 2023; Accepted 6 March 2023

1. Introduction

Assisting multi-step dexterous manipulation tasks like minimally invasive surgical procedures is one of the primary directions in robotic applications. The well-identified advantages of augmented precision and stabilization, compared to manual handling, inherently offer advancement for positioning of surgical instruments. However, the up-to-date paradigm of robot-assisted minimally invasive surgery (RAMIS) still requires the surgeon to fully teleoperate the instruments throughout the surgery, or could only automate a single positioning step in ex-vivo environment (Kwoh et al., 1988). The surgeon's mental concentration remains greatly demanded during the surgery, where the inconsistency in skill proficiency appears among novices and experts. Once they are automatically planned and executed under surgeon's supervision, the human workload could then be effectively liberated and benefits task effectiveness.

The main goal of surgical task operation is to perform instrument motions and target contact safely according to task guidelines. Many surgical tasks in MIS (e.g., tissue

dissection and suturing) comprise sequences of highly interactive motion sub-steps towards different targets. To cater for the minimally invasive setup, the instruments usually own a long, slender tool shaft with a wrist-like distal structure to enhance motion dexterity. However, task motions commonly encounter highly twisted motions (e.g., consecutive contacts to tissue/needle with inverted orientation). One or more states in $SO(3)$ space might experience large-scale variation(s) ($\sim 180^\circ$) within a small R^3 space. The instruments should also avoid inadvertent intrusion to surrounding tissues and other instruments. In RAMIS, these are done by the surgeon using a control interface that transfers his/her delicate hand motions to robot actions. To

T Stone Robotics Institute and Department of Mechanical and Automation Engineering, The Chinese University of Hong Kong, Hong Kong, China

Corresponding author:

Fangxun Zhong, T Stone Robotics Institute and Department of Mechanical and Automation Engineering, The Chinese University of Hong Kong, 1/f, AB1, CUHK, Shatin, N.T., HKSAR, Hong Kong, China.
Email: fxzhong@cuhk.edu.hk

ensure safe interaction to the target (e.g., a firm grasp of a tissue/needle via a specific contact pose), the surgeon also stays aware of the robot trajectory and will only proceed once the tool-target alignment becomes suitable. Automating a surgical task with a sequence of such motions demands a smooth and reactive planning control policy to pave the way to reliable task execution.

The challenges are then to define and ensure motion safety throughout the pipelines across different tasks. The primary concern is that there must be theoretical proof to support that the robot trajectory is feasible and also reachable to admissible target states. Each motion should also own re-planning strategy upon specific motion constraints to ensure proper contact to a (moving) target during multi-arm coordination. Other features including avoidance of obstacle collision and joint limits, real-time implementation, etc. should be solved as well. More importantly, the above features should be applicable to every elementary task motion throughout the pipeline of surgical tasks to avoid execution-level uncertainties. Planning trajectories for robot manipulators has been widely studied in industrial and/or domestic applications. Prevailing methods include generation of roadmaps (Siméon et al., 2004; Amato and Wu, 1996) and sampling-based methods (Berenson et al., 2009; Berenson et al., 2009) to directly explore the 6-DoF feasible trajectories. Meanwhile, on-the-fly planning/control schemes could be more efficient and reactive, for example, the potential field (Hwang and Ahuja, 1992) to attract/repel workspace constraints, but is subject to local minima that hinders target reachability. Aiming for surgery autonomy, most works adopt the above approaches to execute specified motions (Marinho et al., 2019) and/or a certain task (Sen et al., 2016) in surgery. Learning from demonstration is also utilized to deal with complex maneuvers like knot tying (Osa et al., 2017) and suturing (Schulman et al., 2013) to avoid direct path planning. We are unaware of any existing works that define generic safety motion constraints emerged from surgical tasks and apply them using a unified framework.

In this article, we present an integrated planning and control framework for 6-DoF robotic instruments for surgical task automation. The framework covers a globally stable manipulation controller, a reactive manipulation planning policy, and a generic motion primitive to characterize different surgical tasks. We first classify task-relevant surgical motions into two typical types. The first type is called *tool-centric actions*, where the robot end-effector is manipulated between (feasible) initial and final configurations without tentative contact to the target but might avoid obstacles. The second type is the *target-centric actions*, where the end-effector is delicately guided upon certain motion constraints to zero the remaining distance to the target for final contact. By leveraging the kinematic property of the instrument's wrist geometry, the nodal vectors could directly parametrize the robot's present-to-goal situation by a new set of parameter space, namely, the *nodal state space (NSS)*. This avoids generating the

(coupled) 6-DoF poses as a high-dimensional nonconvex manifold analysis that hinders globally stable manipulation. The robot trajectory is then generated and executed upon regulation of the states in NSS encoded by a dynamical system. The leading robot motion could be rigorously guaranteed stable and smooth under Lyapunov stability theory, and is further globally viable once the state regulation is coordinated by sequential motion allocation (SMA). We verify its capability to smoothly operate twisted maneuvers, which include drastic orientation changes inevitably occurred in surgical motions. Moreover, elementary motions within the task should be coordinated with situation awareness. To this end, we introduce a dynamic attraction state (DAS), whose dynamics is designed by NSS, as a “tunable” instant goal for active guidance of the robot's movements. This transforms motion autonomy into a *goal-varying manipulation (GVM)* problem, which we could further detail the robot's in-process trajectory through extra motion constraints for path versatility. We apply GVM as a reactive planning policy to ensure target-centric actions (e.g., repetitive needle grasping in suturing). By defining a bounded and robust DAS, GVM could adaptively proceed or rewind the contact motion based on on-the-fly situations. This could avoid hazardous movements while yielding a constrained path (e.g., unnecessary space intrusion or premature target collision). The framework provides a close-form motion output for tackling online motion constraints with guaranteed trajectory feasibility and target reachability. Such features are available for each elementary motion without using iterations. Finally, to elevate task-level generality, we extend such scheme to form different modes of behavior (MoBs), which facilitates depicting and automating pipelines of different tasks.

The main contributions of this work are summarized as follows:

- A robot state model with skeleton nodes and the NSS for direct formulation of the present-to-goal situation for a robotic instrument based on its kinematic constraints.
- A DS-based robot controller with synthesized SMA which provides global asymptotic stability proof that guarantees in-process trajectory feasibility and target reachability to an arbitrary (admissible) target.
- The GVM reactive planning policy to facilitate bilateral transition between target contact and path following to smoothly deal with disturbed targets.
- The motion primitive and its five MoBs which could define pipelines of different surgical tasks.
- Simulations and physical experiments on different tasks are conducted for comprehensive validation.

Our previous works have addressed a few individual problems that target surgery autonomy. We have proposed an efficient and autonomous robot-camera calibration approach to compute the robotic instrument's pose from a monocular camera for vision-guided instrument manipulation (Zhong et al., 2020). Aiming for autonomous

suturing, a monocular-based 6-degree-of-freedom (6-DoF) pose estimation algorithm of a surgical needle has been developed (Zhong et al., 2016; Zhong & Liu, 2018), which was later implemented to a dual-arm needle insertion control scheme to increase needle insertion accuracy under deformation (Zhong et al., 2019).

The structure of this article is summarized as follows. **Section 2** reviews the related works from manipulation planning and control approaches to surgery-specific explorations. **Section 3** shows the problem formulation for this work. **Section 4** introduces the new robot kinematics model for the robotic instrument using nodal state space. **Section 5** elaborates the motion-level planning and control strategy and guaranty of safety and in-process constraints. In **Section 6**, we elevate the strategy to task-oriented modelling by characterizing different MoBs to form pipelines of different surgical tasks. **Section 7** demonstrates the simulation results concentrating on motion-level performance study, and the experimental results in **Section 8** show the overall task-level performance. Finally, discussions and conclusions are presented in **Section 9**.

2. Related works

Manipulation planning and control of 6-DoF manipulators subject to in-process constraints has been widely exploited in industrial and domestic applications. Most works target general workspace constraints, mainly for obstacle avoidance (Siméon et al., 2004; Stilman et al., 2007) and/or optimizing characteristics (e.g., avoiding singularities and deadlocks) (Li and Latombe, 1997; Ferbach and Barraquand, 1997), or to decide the sequence of motion profiles for a task (e.g., pushing a box and fetching an object) (Billard and Kragic, 2019; Fang et al., 2020). On the other hand, the motion pipeline required by a surgical task is complicated but clearly given in clinical practice (Cao et al., 1996). However, the wristed structure of the instrument and limited joint ranges render a smaller size of feasible space manifold and makes it challenging to coordinate the end-effector and whole-body motions. Only a few existing works tackle 6-DoF motion control of a robotic instrument, including learning from demonstration to perform a complex but single motion step like looping for knot tying (Osa et al., 2017). However, it requires collection of totally 4 datasets with > 50 individual demonstrations to cover different conditions for a suture looping motion step. In Chiu et al. (2021), grasping a surgical needle is planned by reinforcement learning, which also depends on 1000 pre-recorded simulated grasps to learn this particular motion step. Aiming at task-level autonomy, there are works that introduce convex optimization (Sen et al., 2016) and visual servoing (Pedram et al., 2020) to plan the motions required for needle manipulation in wound suturing. However, how to ensure feasible trajectories for twisted maneuvers, which commonly occur when re-orienting a target (e.g., a tissue or needle), are not rigorously provided. Similar issues also exist in repetitive target contact like debridement (Narazaki

et al., 2006), where poorly planned motions could easily lead to premature collisions. As the majority of previous works only focus on a certain step, we will summarize the state-of-the-art approaches for each subproblem in surgical motion autonomy first, and then the works of task-level automation strategies, and their limitations.

2.1. Automating basic motions

Pioneer works of robot-enabled instrument manipulation in surgical applications have targeted positioning of an instrument to registered poses, for example, the tissue cutter in prostatectomy (Kwoh et al., 1988) and the biopsy probe in neurosurgery (Davies et al., 1991). Combining clinical CT data, planning an optimal port placement and/or admissible path of the instrument becomes available to provide a safe referenced position for localized treatment (Schweikard et al., 1993; Adhami and Coste-Manière, 2003). Among these works, the concept of motion planning is limited to generating a deterministic path using an industrial robot arm for the instrument through admissible space without actual manipulation process or constraints. Funda et al. (1996) computed the optimal robot motion of for adjusting the gaze of the laparoscope subject to the remote center-of-motion constraint, which is widely adopted in MIS procedures. To include online feedback, Wei et al. (1997) applied visual servoing to adjust the robot-mounted laparoscope by tracking the instruments from the images. The method is also adopted in Krupa et al. (2003); Osa et al. (2010) to actuate a surgical instrument to a feature-defined target by minimizing image-based errors. However, the main disadvantage is that it requires continuous monitoring of visual features during manipulation, which could not robustly define a 6-DoF goal pose during the motion or provide re-planning strategy. To deal with multi-instrument environment, Preda et al. (2015) used sampling-based planning scheme (RRT-connected algorithm) for planning collision-free motion step for multiple instruments. Sozzi et al. (2019) proposed a DS-based approach with waypoint selection to actuate a non-wristed instrument with collision-free path to another instrument. The work in Marinho et al. (2019) proposed vector fields to guide the motion of a regular instrument with RCM constraint. Aiming for task-specific motion constraints, a few works adopt the method of learning from demonstration (LfD) to automate complex multi-arm coordinated trajectories like suture looping in knot tying or needle insertion (Osa et al., 2017; Schulman et al., 2013; Schwaner et al., 2021) to avoid modelling complex trajectories. However, it requires pre-recorded motion data for each type of task and is unable to deal with unexpected changes of task process. Notably, most above works used robot-mounted regular surgical instrument with only 4-DoF motions instead of a wristed robotic instrument. More recently, the work in Chiu et al. (2021) further introduced reinforcement learning (RL) policy to generate robotic instrument motions. To feed the online

robot states to the system, continuous visual tracking of the instrument features is required via stereo images.

All previously mentioned works only focused on feasibility study completing a certain motion/task successfully, but failed to define generic safety features to be required by surgery automation to avoid execution-level uncertainties, which is of highest priority in surgeon's perspective (Narazaki et al., 2006). There also lacks a unified solution to deploy those safety features across different tasks as well. In this work, we provide a DS-based framework with rigorous proof of global motion stability and trajectory reachability, such that it could handle stable motions to all task-relevant motions.

2.2. Automating contact motions

Interacting surgical instruments to the environment also constitutes a majority of motion steps in surgical tasks. In MIS, legitimate contact include non-invasive contact (e.g., palpation for tumor localization), grasping (e.g., toward a piece of tissue debris or surgical needle), and invasive contact (e.g., inserting a needle into tissue). To this end, Patil and Alterovitz (2010) addressed manipulation required for tissue retraction by selecting optimal grasping point, but didn't specify the exact motion to safely grasp the tissue in advance. Since then, there are works addressing 6-DoF instrument-based manipulation of soft tissues for tissue dissection (Murali et al., 2015; Nagy et al., 2018) and deformation control (Li et al., 2020). The works in Kehoe et al. (2014); Hwang et al. (2020) automated target grasping in a peg transfer task, which is commonly used in surgical training for novice surgeons. Among them, the instrument targeting and grasping process were executed separately to avoid unsafe tool-target contact, and the grasping orientations highly resembled each other. Meanwhile, efforts have been made to needle grasping in wound suturing. Schulman et al. (2013) introduced LfD to automate manipulation and grasping of a needle with pre-recorded trajectories. The process had also been automated in terms of direct pick-up (D'Ettorre et al., 2018) and needle hand-off (Chiu et al., 2021; Pedram et al., 2020) using VS and LfD, respectively. A static needle was targeted and grasped using visual feedback as the motion input, which could reduce reliability when dealing with awkward grasping poses.

One mutual limitation of these works is that they all regarded grasping as separated steps with a fixed sequence rather than a reactive process, which could be re-planned under sensing/motion disturbances. The motion constraint during target contact should be embedded as well with the aforementioned safety features, but remains an unmet problem in surgery autonomy. We interleave manipulation and contact motion together as a GVM model, such that the robot will adaptively transit between tool-centric or target-centric actions with global stability.

2.3. Automating task-level procedures

Completing the pipeline of a surgical task based on step-level automation is the main goal to achieve task-level autonomy. Researchers have approached various individual surgical tasks with different sub-steps and different controllable DoFs required for the tasks. Works include autonomous field of view control of endoscope in Agustinos et al. (2014); Nageotte et al. (2006); Voros et al. (2006); Yang et al. (2019) that used VS technique to track a pre-defined visual feature to allow hands-free adjustment intraoperatively. In Kehoe et al. (2014) and Hwang et al. (2020), the authors targeted simulated surgical debridement with the instrument trajectories being performed using model predictive control and pose-to-pose interpolations, respectively. Given the targets' positions, the instrument autonomously manipulated, grasped, and cleared the debris sequentially. Meanwhile, there are efforts targeting autonomy of tissue-based procedures, including non-invasive tasks like tissue palpation (McKinley et al., 2015; Nichols and Okamura, 2015) for tumor localization, and invasive tasks like tumor ablation (Hu et al., 2015), tissue dissection (Murali et al., 2015; Nagy et al., 2018), blood area detection, and suctioning from vessel rupture (Richter et al., 2021). Targeting more complex tasks like knot tying, Osa and Van Den Berg (Osa et al., 2010; Van Den Berg et al., 2010) automated single knot tying in suturing based on LfD.

Suturing is one of the most common but complicated tasks to be performed in MIS. It contains a long sequence of motion steps, involving dexterous manipulation of a needle and dual-arm coordination. Automating such tedious task could significantly reduce human workload, and thus has gained attention from researchers. For example, the works in Kapoor and Taylor (2008); Nageotte et al. (2005) have planned an optimal trajectory to automate the insertion process of a half-circle surgical needle into the tissue. A complete stitching automation process was given by Nageotte et al. (2009) including entry point planning, pose sensing, manipulation, and stitching of the suturing needle. However, these only constitute incomplete steps required by the suturing guideline. There are also studies addressing hand-off of the needle between two instruments using human–robot collaborative approach (Mikada et al., 2020; Watanabe et al., 2016, 2017), and also autonomous approach (Varier et al., 2020; Chiu et al., 2021) using RL-generated motions. Zhong et al. (Zhong et al., 2019) further addressed active tissue deformation during needle insertion to improve insertion accuracy. Lenard and Shademan (Leonard et al., 2014; Shademan et al., 2016) proposed a suturing device to inherently simplify the suturing task while ensuring a firm stitch. However, the needle size and stitch width could not be customized restricted by the tool. Meanwhile, only a few works have completed a throw of wound suturing. Sen et al. (Sen et al., 2016) provided an optimization-based approach to automate a four-throw suturing task, but does not support online trajectory

re-planning due to tool-target motion disturbances with theoretical proof for motion stability. Pedram et al. (Pedram et al., 2020) completed an automatic single-throw suturing process using dual-arm coordination using a VS-based control framework. The instruments are visually tracked during the manipulation, which was reported to own accurate targeting accuracy. However, continuous visual sensing might hinder its applicability to deal with large orientation change. We are unaware of any existing works that provide systematic safety features for complex instrument manipulation in one or more surgical tasks.

3. Preliminaries

3.1. Problem formulation

Performing a surgical task is essentially to manipulate a robotic instrument from pose to pose in $SE(3)$ space across all motion steps. Under any (instant) robot configurations, a 6-DoF goal pose of the robot's end-effector (i.e., the instrument's distal tool) is assumed to be available. It could be either computed from online sensoring feedback or from user input. One common method is to add image-based markers on the display (e.g., labeling the tissue's dissecting trajectory or the wound position for suturing), as a surgeon's decision and supervision remains critical to reduce uncertainties of target recognition in supervised robotic surgery (Haidegger, 2019). Although computerized surgical image analysis has gained considerable attention to assist target localization (Loukas, 2018), we emphasize that our aim in this work is an autonomous framework to perform robot motions safely and reliably instead of comprehensive detection algorithms.

Due to the delicate nature of surgical tasks, we define safety for task autonomy to be decomposed into the following two parts: 1) There must be theoretical proof to globally guarantee that the robot motion is admissible in every (future) time instant, and the goal state can be smoothly and stably arrived. The term "global" indicates it should apply for any admissible initial/final robot states, and 2) the robot motions must be able to deal with obstacle avoidance (both tool-tool and tool-tissue), joint limit avoidance, etc. The first part is of primary concern which has not been systematically addressed. The second part consists of basic safety properties. We especially target "twisted" motions where the distal articulated joints experience large-range adjustment ($>90^\circ$), and the robot state especially of $SO(3)$ is significantly tuned ($\sim 180^\circ$), which is normally used in surgery for flipping the grasping orientation of a suturing needle and/or tissue. Such motion might approach the edge of admissible robot workspace manifold, which is nonconvex due to small joint motion range of the two distal joints ($\pm 90^\circ$). To inherently maximize the performance of system response, the workspace constraints should be embedded into the control model rather than iterative approaches. The subproblems of instrument non-contact manipulation and contact motion control are

processed under the same core in an integrated DS, and must be applicable to all motions appeared in one or more tasks.

3.2. Nomenclature

We regard the entire structure of the detachable tube-like instrument and its proximal motorized joints (totally 7 DoFs) as a surgical robot (or just "robot" if not further clarified). We call the distal structure of the instrument to achieve clamping, grasping, etc. as the "tool," which is also the end-effector in robotics. To detail the workspace control problem, we use the robot's 6-DoF end-effector pose (denoted by a 3D Cartesian coordinate frame \mathcal{F}) or its corresponding state \mathcal{X} (computed from a newly proposed set of space parameters) to represent the robot's present or goal configuration. The acronyms appeared in this article have been listed out in Table 1.

4. Robotic instrument model

We begin by derive the kinematics of a robotic surgical instrument based on its mechanical properties. A typical serial-link manipulator generally owns six joints to fully determine the 6-DoF pose of the end-effector ($SE(3)$ space) in 3D Cartesian space modelled by a coordinate frame $\mathcal{F}_e(\mathcal{O}_e, x_e, y_e, z_e)$, which could also describe the robot's state space $\mathbf{x} \in \mathbb{R}^6$ in robot kinematics with the following well known form:

$$\mathbf{x} = \mathbf{f}(\mathbf{q}) \quad (1)$$

with $\mathbf{q} = [q_1 \dots q_6]^\top$ denotes the generalized coordinate vector in joint space that describes the robot joint angles. As robot joints are mostly revolute, the mapping $\mathbf{f}(\cdot)$ is usually nonlinear and indicates a coupled relationship between the end-effector's position and orientation while being regulated. Such property remains in velocity mapping using the Jacobian matrix $\mathbf{J}(\mathbf{q}) \in \mathbb{R}^{6 \times 6}$:

$$\dot{\mathbf{x}} = \mathbf{J}(\mathbf{q})\dot{\mathbf{q}} \quad (2)$$

Then, adjustment of the end-effector's pose will inevitably lead to movements for all six joints. Considering a complex motion step, the more "twisted" trajectory the end-effector undergoes, the more details are to be planned for describing its profile. Moreover, the limited joint motion ranges ($\pm 90^\circ$ for the last two joints) results in a nonconvex reachable workspace manifold, where the path should be further refined to avoid all infeasible regions. Existing approaches deal with such trajectory complexity using a priori workspace analysis (Mirrazavi Salehian et al., 2018) or iterative optimization (Zucker et al., 2013), which could not guarantee real-time and/or theoretically feasible output in such high-dimensional space, and are unsuitable for surgical applications where the environment is usually temporarily set up and must not sacrifice efficiency. Thus, a new parametrization method is to be developed to reformulate the kinematic relationship between the robot's

Table 1. List of acronyms used in this article (in alphabetical order).

Acronym	Definition
DAS	Dynamic attraction state
DHc	Denavit–Hartenberg convention
DoF	Degree-of-freedom
DS	Dynamical system
GVM	Goal-varying manipulation
LfD	Learning from demonstration
MIS	Minimally invasive surgery
MoB	Mode of behavior
NSS	Nodal state space
RAMIS	Robot-assisted minimally invasive surgery
RCM	Remote center-of-motion
RL	Reinforcement learning
SMA	Sequential motion allocation
SMP	Surgical motion primitive
VS	Visual servoing

proximal parts and its end-effector. Here, we introduce the concept of *skeleton node analysis* that regards each junction (or the “elbow”) between two consecutive links (or the “limbs”) along the robot skeleton as a point of interest. Specifically, we define a 3D point $p \in \mathbb{R}^3$ in Cartesian space as a skeleton node of the robot if it satisfies either of the following criteria:

- It is the intersection point of the rotary axes of two consecutive robot joints.
- It is the intersection point of the rotary axis of a robot joint and the geometric centerline of a link (regarding straight rigid links).
- It is the origin of the end-effector’s 3D Cartesian coordinate frame.

Meanwhile, we disqualify a candidate if it overlaps an existing node which has already been defined from proximal robot parts. The general form of a node set considering a 6-DoF serial robot manipulator could be given as follows:

$$\mathcal{N} = \{n_1 \ n_2 \ \dots \ n_k\} \quad (3)$$

where the dimension of the node set \mathcal{N} being $k \in [2, 12] : k \in \mathbb{Z}^+$, which is associated with the number of non-zero link length/offsets due to different kinematic design and does not necessarily correlates to the number of joints (see Figure 1 for the examples of \mathcal{N} on different types of robots). This implies the skeleton node analysis emphasizes an explicit 3D geometry of the robot’s kinematic chain relative to the end-effector, which is not achievable by using the Denavit–Hartenberg convention (DHc). Meanwhile, one can denote n 3D position unit vectors as a set $\mathcal{V} = \{v_1 \ v_2 \ \dots \ v_k\}$ referenced from the corresponding n skeleton nodes and head to arbitrary directions. Each vector, named as the *nodal vector*, is described in robot base frame and is virtually fixed to the skeleton and thus is adjustable by

all joints proximal to the attached node. Such group of nodal vectors are then denoted by $\Xi \in \mathbb{R}^{k \times 3}$, with

$$\Xi = [\xi_1 \ \xi_2 \ \dots \ \xi_k] \quad (4)$$

Normally, for an arbitrary node $n_i, i \in [1, k]$, there always exists a nonlinear and variable transformation between it and the end-effector (except v_k) regarding its distal joints. No consistent geometric relationship between the nodal vectors and the end-effector could be derived.

We then use the skeleton nodes to analyze a robotic surgical instrument implemented to MIS. Without loss of generality, we consider the structure of the EndoWrist (by Intuitive Surgical Inc) which has been widely adopted for designing robotic surgical systems, and regard it as a typical example to study surgery autonomy.¹ The instrument owns a slender cylindrical shaft that provides 4-DoF RCM-constrained motions to yield the minimally invasive setup, a 2-DoF wristed joint set to provide wristed motions, and one extra DoF for tool actuation. Such kinematics leads its skeleton node set to be degenerated into the following form sufficed by $k = 3$:

$$\mathcal{N} = \{n_1 \ n_2 \ n_3\} \quad (5a)$$

$$\Xi = \{\xi_1 \ \xi_2 \ \xi_3\} \quad (5b)$$

where $\Xi \in \mathbb{R}^{3 \times 3}$, and the nodes n_1, n_2 , and n_3 geometrically represent the shaft-to-wrist junction $\mathbf{x}_s \in \mathbb{R}^3$, the wrist-to-tool junction $\mathbf{x}_w \in \mathbb{R}^3$, and the tool tip $\mathbf{x}_t \in \mathbb{R}^3$, respectively, as shown in Figure 2. Then, due to the 3D geometric constraint of a wristed robotic instrument, one can define the following orthogonality that unconditionally holds:

$$\begin{bmatrix} \xi_1^\top & \dots & \mathbf{0}_{1 \times 3} \\ \vdots & \xi_2^\top & \vdots \\ \mathbf{0}_{1 \times 3} & \dots & \xi_3^\top \end{bmatrix} \begin{bmatrix} \mathbf{l}_h \\ \mathbf{l}_v \\ \mathbf{m} \end{bmatrix} = \mathbf{0}_{9 \times 1} \quad (6)$$

where \mathbf{m} indicates the longitudinal unit vector (or heading) and \mathbf{l}_h and \mathbf{l}_v being two lateral unit vectors with respect to the tool’s heading orientation (refer to Figure 2 for illustration). If one considers a typical end-effector frame $\mathcal{F}_e(\mathcal{O}_e, x_e, y_e, z_e)$ using DHc whose origin position overlaps \mathbf{x}_t , the following factorization could also be derived:

$$[\mathbf{l}_h \ \mathbf{l}_v \ \mathbf{m}] = \mathbf{H}[\vec{x}_e \ \vec{y}_e \ \vec{z}_e] \quad (7)$$

where $\mathbf{H} \in \mathbb{R}^{3 \times 3}$ is the constant orthogonal matrix computed by a series of fundamental transformations. Notably, this reveals for any wristed robot manipulators, its end-effector orientation could be interpreted by intermediate body parts (characterized by nodal vectors) via a constant geometric relationship.² The nodal vectors could also be depicted as $\xi_1 = k_1 \vec{x}_e, \xi_2 = k_2 \vec{z}_e, \xi_3 = k_3 \vec{y}_e$. This is an important kinematic property for us to study the robot’s whole-body motion coordination subject to task-relevant constraints. Here, to yield (6), we could select one group of nodal vectors for the instrument:

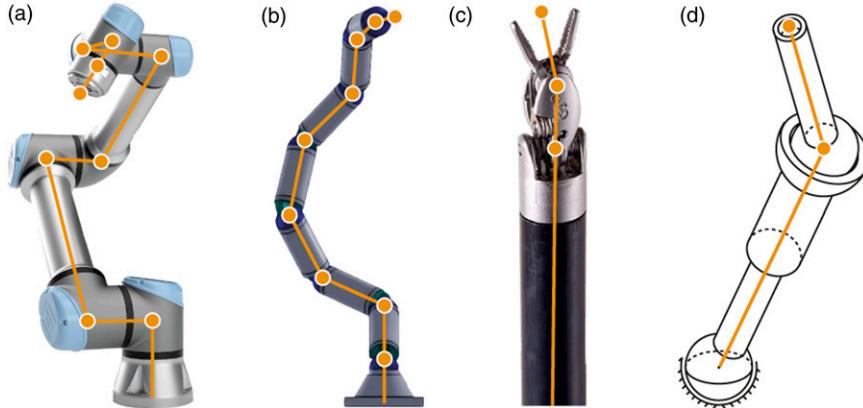


Figure 1. The definition of “nodes” on different types of serial robot manipulators: a) the UR5 as a typical robot arm; b) a hyper-redundant robot arm (Li and Chen, 2015); c) the EndoWrist as a representative of robotic surgical instrument; and d) a highly wristed non-redundant robot with only three links (with DoFs being 3R-P-2R).

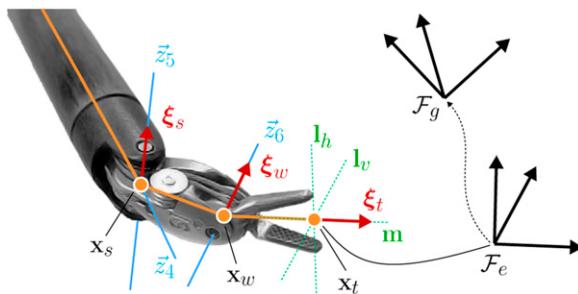


Figure 2. Kinematics of the wristed instrument and the illustration of the nodes and nodal vectors according to our definition.

$$\xi_1 = \vec{z}_6, \xi_2 = \vec{z}_5, \xi_3 = \vec{x}_6 \quad (8)$$

where \vec{z}_5, \vec{z}_6 denote the axial direction of the respective joint q_5 and q_6 from robot base, and \vec{x}_6 being along the tool’s heading direction.

We decompose the robot joints correspondingly into the following:

$$\mathbf{q} = [\mathbf{q}_s \quad q_4 \quad q_5 \quad q_6]^\top \quad (9)$$

where $\mathbf{q}_s = [q_1, q_2, q_3] \in \mathbb{R}^3$ denotes the joints that control the shaft pose, and the rest regulate the wristed part of the instrument. Then, combining DHc to the skeleton nodes, we further give the derivation of the nodal vectors computed with respect to the robot base frame:

$$\xi_i = \mathbf{R}_s(\mathbf{q}_s) \prod_{i=1}^{3} R_{i+2}(\mathbf{q}_{i+3}) \bar{\xi}_i, i = 1, 2, 3 \quad (10)$$

where $\mathbf{R}(\cdot) \in \mathbb{R}^{3 \times 3}$ denotes the rotation matrices and $\bar{\xi}_{1,2,3}$ are the nodal vectors described in local frames. In the next section, we will implement ξ_i as a new media for the robot’s present-to-goal model and for versatile motion control of the robot’s end-effector as well as the body skeleton.

5. Motion-level autonomy architecture

Based on the above modelling, we first aim to regulate the robot’s end-effector to reach an arbitrary (but feasible) goal pose by proposing a basic control architecture to provide safety features to the robot’s motions. We then extend it to yield more task-relevant motion constraints and utilize the derived properties to integrate a series of dexterous maneuvers into a single motion step.

5.1. DS-guided instrument manipulation

5.1.1. End-effector control system. The DS is an effective mathematical approach to generate stable, convergent robot motions which are efficiently reactive to instant state variations and/or disturbances. This endows DS-guided manipulation with powerful adaptability in motion re-planning without the need to analyze the entire trajectory (compared to optimization-based approaches), which is superior to tackle robot manipulation once the motion constraints are embedded into the system model.

We aim to incorporate the instrument’s kinematic property using NSS analysis to the DS to enable globally guaranteed safe instrument manipulation planning. We start by giving the general form of a time-invariant nonlinear dynamical system as follows:

$$\dot{\zeta}(t) = \mathbf{g}(\zeta(t), \mathbf{u}(t)) \quad (11)$$

which indicates the evolution of the system’s states $\zeta(t)$ subject to the control input \mathbf{u} , and $\mathbf{g}(\cdot)$ denotes the differentiable vector-valued function. To apply such technique to robot’s motion generation, the robot states $\zeta(t)$ must be appropriately defined to enable its dynamics (via $\mathbf{g}(\cdot)$) to guide the robot with demanded performances as the system evolves over time. The states are usually defined as the 6-DoF end-effector pose or configuration space (Mirrazavi

Salehian et al., 2016) and thus, $\mathbf{g}(\cdot) : \mathbb{R}^{6l} \mapsto \mathbb{R}^6$. Here, unlike the above methods, we use the nodal vectors to encode the relationship between the robot's present and goal configurations based on its body parts. Consider an arbitrarily defined $\mathcal{F}_g(\mathcal{O}_g, x_g, y_g, z_g)$ as the target pose to reach, we particularly introduce a new parametrization space \mathcal{H} which is defined as follows:

$$\boldsymbol{\eta}(\mathbf{q}) = \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{bmatrix} := \begin{bmatrix} d^*(\mathbf{x}_s, \mathbf{x}_{s_g}) \\ \langle \xi_1, \vec{x}_g \rangle \\ \langle \xi_2, \vec{z}_g \rangle \\ \langle \xi_3, \vec{y}_g \rangle \end{bmatrix} \subseteq \mathcal{H} \subseteq \mathbb{R}^4 \quad (12)$$

where the parameters η_2, η_3, η_4 depict via inner product the orientation alignment of the end-effector from its target, \mathbf{x}_{s_g} is the goal position of \mathbf{x}_s computed from \mathcal{F}_g . The scalar $d^*(\cdot) = \mathbf{x}_s^\top (\mathbf{x}_s - \mathbf{x}_{s_g}) / \|\mathbf{x}_s - \mathbf{x}_{s_g}\|^2$ defines the signed distance of \mathbf{x}_s from \mathbf{x}_{s_g} via vector projection. Note that getting \mathbf{x}_{s_g} requires computation of inverse kinematics which could not be analytically solved due to the coupled position and orientation. Here, we compute a “naive” but also closed-form $\tilde{\mathbf{x}}_{s_g}$

$$\tilde{\mathbf{x}}_{s_g} = \mathbf{x}_g - \mathbf{x}_t + \mathbf{x}_s \quad (13)$$

from the knowledge of distal joints (encoded by $\mathbf{x}_t(\mathbf{q})$ which are uncontrollable by $\mathbf{x}_s(\mathbf{q}_s)$). Note that in (13), $\tilde{\mathbf{x}}_{s_g}$ is derived based on the instant forward kinematics, which provides “coarse” guidance towards the end-effector's goal position using only \mathbf{q}_s . However, we will later show that such $\tilde{\mathbf{x}}_{s_g}$ does not affect the performance of the target DS and could converge to its genuine value subject to robot control. We name \mathcal{H} as the *nodal state space (NSS)* which characterizes the robot's goal-oriented spatial relationship using the skeleton features of the nodes. The robot's current state \mathcal{X} and goal state \mathcal{X}_g could now be reformulated by the components in (12) as follows

$$\begin{aligned} \mathcal{X} &:= [\mathbf{x}_s \quad \xi_1 \quad \xi_2 \quad \xi_3]^\top, \\ \mathcal{X}_g &:= [\mathbf{x}_{s_g} \quad \vec{x}_g \quad \vec{y}_g \quad \vec{z}_g]^\top. \end{aligned} \quad (14)$$

Now, we will use “state” in (14) to describe the robot's status which is equivalent to a 3D Cartesian space, as we aim to solve automation via Cartesian workspace analysis. Several properties that NSS could contribute to robot motion guidance are highlighted.

Proposition 1. $\boldsymbol{\eta}(\mathbf{q})$ has a globally unique equilibrium $\boldsymbol{\eta}^*$ with $\dot{\boldsymbol{\eta}}^* = \mathbf{0}$.

Proof. Regarding η_1 , any given (feasible) \mathbf{x}_{s_g} will correspond to a globally unique solution of \mathbf{q}_{s_g} , as the local kinematics $\mathbf{x}_s = \mathbf{f}_s(\mathbf{q}_s)$ Cartesian workspace manifold being a solid sphere (From, 2013). This shows that $\mathbf{f}_s(\cdot)$ is convex and non-singular for regulating η_1 . Meanwhile, according to (12), there exists multiple solutions of ξ_{i+1} to satisfy $\eta_i = 0$ ($i = 2, 3, 4$). However, we will prove in Proposition 2 that due to the joint limits of the

instrument, only one ξ_{i+1} is physically reachable subject to $\mathbf{x}_s \rightarrow \mathbf{x}_{s_d}$, which further corresponds to a globally unique q_{i+3} . Thus, there exists a unique q to meet $\eta_i = 0$ for $i = 1, 2, 3$ for any feasible \mathcal{X}_g interpreted by \mathcal{F}_g .

Proposition 2. Given a 3D unit vector \mathbf{v}_r based on \mathcal{F}_g and a nodal vector $\xi_i(\mathbf{q}_s, q_{i+3}^*)$ of node n_i ($i = 1, 2, 3$), there exists a unique solution q_{i+3} for ξ_i such that $\mathbf{v}_r^\top \xi_i(\mathbf{q}_{i+3}) = 0$ subject to $\mathbf{q}_s \rightarrow \mathbf{0}$.

Proof. According to (10), the derivation of the nodal vector ξ_i is simplified to the following:

$$\xi_i = \mathbf{C}^{i+2} \mathbf{R}_{i+3}(q_{i+3}) \bar{\xi}_i \text{ s.t. } \mathbf{q}_s \rightarrow \mathbf{0} \quad (15)$$

where the variation of the rotation matrix $\mathbf{C} \in \mathbb{R}^{3 \times 3}$ is negligible due to the settled proximal robot parts. \mathbf{R}_{i+3} is the fundamental z-axis rotation matrix which acts as a peridical factor applied to $\bar{\xi}_i$ and is solely regulated by q_{i+3} . Thus, ξ_i has a period of 2π subject to change of $[q_{i+3}, q_{i+3} + 2\pi]$, as its components own the following form:

$$\xi_{i,j} = a_1 + a_2 \cos(q_{i+3}) + a_3 \sin(q_{i+3}), \quad j = 1, 2, 3 \quad (16)$$

where a_1, a_2 , and a_3 are constant scalars. Therefore, there are two possible solutions of q_{i+3} for $\xi_i = \mathbf{v}_r^\top \xi_i(\mathbf{q}_{i+3}) = 0$ within 2π . For surgical robots, the motion ranges of the two distal joints \mathbf{q}_w and q_t are limited to $[-\pi/2, \pi/2]$. Thus, it is impossible to generate two toggled directions of ξ_i . There exists only one q_{i+3} that satisfies $\mathbf{v}_r^\top \xi_i = 0$ within the robot's feasible workspace (i.e., $[q_{i+3}^-, q_{i+3}^+]$) to yield $\eta_i = 0$ (shown in Figure 3(b)). That corresponds to the unique solution of ξ_i , denoted by ξ_i^* . Meanwhile, the direction of vector ξ_i periodically changes through $[q_{i+3}, q_{i+3} + 2\pi]$, which sweeps ξ_i through a circle whose center being exactly the position of n_i within a plane π_{ξ_i} (see Figure 3 for geometrical interpretation).³

Now, we establish a dynamical system to guide the instrument end-effector to its target pose. The NSS $\mathcal{H} \subseteq \mathbb{R}^4$ are defined as the state variables, and then due to Proposition 1, there exists a globally unique equilibrium at origin, that is, $\boldsymbol{\eta}^* = \mathbf{0}_{4 \times 1}$. \mathbf{u} is of $\dot{\mathbf{q}}$ or \mathbf{q} only. Thus, we propose the following controller:

$$\mathbf{u} = - \left(\frac{\partial \boldsymbol{\eta}(\mathbf{q})}{\partial \mathbf{q}} \right)^{-1} \boldsymbol{\Gamma} \boldsymbol{\eta}(\mathbf{q}) \quad (17)$$

where \mathbf{M} is the $\mathbf{M} \in \mathbb{R}^{6 \times 6} > 0$ is the diagonal gain matrix applied to joint velocities. Then, by applying the control input $\mathbf{u} = \dot{\mathbf{q}}$ and substituting (17) to $\dot{\boldsymbol{\eta}}(\mathbf{q})$ leads to the following:

$$\dot{\boldsymbol{\eta}} = -\boldsymbol{\Gamma} \boldsymbol{\eta} \quad (18)$$

which clearly indicates an asymptotically stable DS (with sole negative real parts in phase plane analysis) whose the attractor is exactly $\boldsymbol{\eta}^* = \mathbf{0}_{4 \times 1}$. Its evolution guides the instrument motions via $\dot{\mathbf{q}}(t)$ to reach the end-effector \mathcal{F}_e to any feasible \mathcal{F}_g . As $\boldsymbol{\eta}$ is also bounded due to (12), from (18), the DS that regulates $\boldsymbol{\eta}$ is globally asymptotically stable

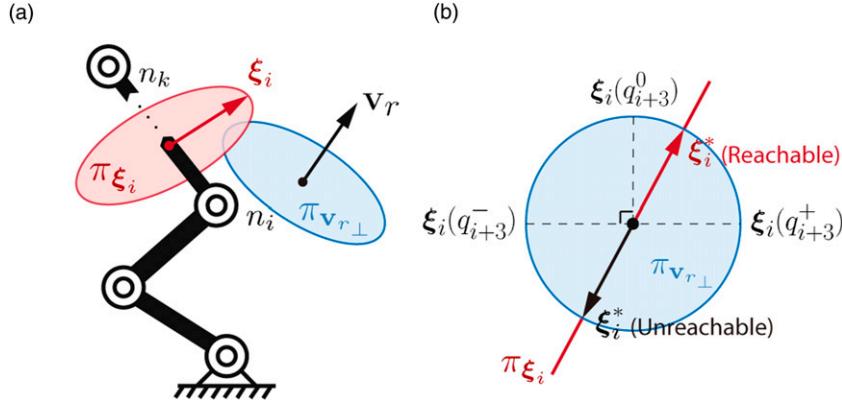


Figure 3. Geometric representation of the variability of Ξ subject to q_{i+3} in a), and the solution uniqueness to achieve $v_r^\top \xi_i(q_{i+3}) = 0$ when considering the physical joint motion of q_{i+3} in 3D workspace in b). Respectively, q_{i+3}^- , q_{i+3}^0 , and q_{i+3}^+ denotes the negative limit, zero position, and positive limit.

from Lyapunov stability theory as each equilibrium in η^* is globally unique (refer to Proposition 2) (Slotine and Li, 1991). The system is also of C^1 continuity as (1) and (10) are all differentiable regarding both $\dot{\eta}(t)$ and $\eta(t)$.

Remark 1. Deploying the controller (17) to robot's target-reaching control requires interpretation of a \mathcal{F}_g to η . \mathcal{F}_g is normally provided as a 3D Cartesian coordinate frame for task-oriented robot manipulation. Then, according to (6) and (7), $\eta_{2,3,4}$ are all computable. Meanwhile, we use \tilde{x}_{sg} as in (13) as approximation of x_{sg} to avoid computing inverse kinematics, which is not differentiable for on-the-fly manipulation guidance.

5.1.2. Sequential motion allocation (SMA). The current form of DS only addresses the convergent motion of the end-effector to a goal pose and does not necessarily guarantee the feasibility of the in-process trajectory. Surgical tasks commonly require grasping the needle and tissue from specific poses to minimize potential trauma. When transitioning between awkward poses, the robot might encounter joint limits and/or near-singular configurations. In terms of a serial robot manipulator, one robot joint could parallelly tune different η . Therefore, regulating a specific η_i might cause disturbances to other (distal) $\eta_j, j > i, j \in \mathbb{N}^+$, which could decrease the overall control effectiveness and target reachability.

To solve the above issues, new constraints must be added to the system to improve the instrument's manipulation delicacy. Here, we introduce the *sequential motion allocation (SMA)* that progressively allocates the regulation output from proximal to distal robot joints. This will facilitate "rhythmic" movements of the robot's body parts and avoid "clumsy" joint motions, which could be harmful to both the robot and the environment. We achieve SMA by proposing a new set of state-based parameters $\lambda \in \mathbb{R}$ as the nodal allocation coefficients to dynamically weight the control output deployed to η for each node based on (12) and (17). The definition of λ is chosen by considering the

following properties to meet the task-relevant motion characteristics:

1. We set all coefficients to $\lambda \in [0, 1] \subseteq \mathbb{R}$, $\lambda \in \mathcal{C}^0$ to normalize the scales of all η among different types of robot joints and/or their different ranges of motion during regulation. $\lambda = 1$ and $\lambda = 0$ indicates the motion output that regulates the corresponding η has been fully deployed and suspended, respectively.
2. To achieve sequential actuation of the skeleton nodes, the λ_i that allocates the regulating motions to the current η_i should initiate posterior to the elevation of the previous (proximal) λ_{i-1} . This is interpreted as $\lambda_{i-1} \geq \lambda_i \forall \lambda_{i-1}, \lambda_i \in [0, 1]$.
3. To incorporate SMA to whole-body coordination, λ_i needs to open a gap to the rising λ_{i-1} ahead, but also needs to catch up with λ_{i-1} as $\lambda_{i-1} \rightarrow 1$ to eventually enable full regulation. This indicates the boundary conditions of $\lambda_i(x)$ to yield $\dot{\lambda}_i(x)|_{x=0} = 0 \forall i$ and $\dot{\lambda}_i(x)|_{x=1} = 1, \forall i > 1$ with $x = \lambda_{i-1}$.

All λ will constitute a space \mathcal{L} whose dimension corresponds to that of \mathcal{H} , with $\lambda = [\lambda_1, \lambda_2, \lambda_3, \lambda_4]^\top \in \mathcal{L} \subseteq \mathbb{R}^4$.

A feasible mathematical form of λ to achieve SMA could be given as follows:

$$\lambda_i = \begin{cases} \exp(-k_i \|\eta_1\|), & i = 2 \\ \exp(-k_i(1/\lambda_{i-1} - 1)), & i > 2 \end{cases} \quad (19)$$

where k_i is the preset gain for each λ_i . We set $k_i = 1 \forall i > 1$ to meet $\lim_{\lambda_{i-1} \rightarrow 0} \dot{\lambda}_i(\lambda_{i-1}) = 0 \forall i$. The definitions in (28) shows that to initiate SMA, λ_1 needs to be exclusively modelled apart from the others. Once $\eta_1 \rightarrow 0$, the control output will be smoothly allocated to the next node with η_2 by elevating λ_2 , and vice versa. The subsequent λ_i are tuned directly based on λ_{i-1} to ensure a robust and flexible SMA process regardless of the robot's initial/final state or its control input.⁴

$$\bar{\mathbf{u}} = \mathbf{L}\mathbf{u} \quad (20)$$

with

$$\mathbf{L} = \begin{bmatrix} 1 & \mathbf{0}_{1 \times 3} \\ \mathbf{0}_{3 \times 1} & \text{diag}(\lambda)_{3 \times 3} \end{bmatrix} \in \mathbb{R}^{4 \times 4} \quad (21)$$

Remark 2. In (28), λ_1 is only used as a reference to proceed the SMA under $\|\eta_1\| \rightarrow 0$, and at any time instant, deviation of η_1 off 0 will suspend and restart the whole SMA process until η_1 converges again. This implies the instrument's position (via η_1) is always regulated prior to the orientation (via $\eta_{2,3,4}$) due to the “weak coupling” effect of the instrument's end-effector pose. This is a demanding motion constraint in instrument manipulation, as the position misalignment of the end-effector caused by η_1 becomes significantly larger than that by $\eta_{2,3,4}$, also thanks to the “weak coupling” effect (see [Appendix A](#) for mathematical explanations how “weak coupling” appears on a wristed robotic instrument).

One main advantage of introducing SMA to the DS-guided instrument manipulation is that the whole-body robot trajectory subject to (18) becomes always feasible. This is led by (28) which enforces the regulation of a

specific $\eta_i, i > 1$ to be always accompanied by a settling η_{i-1} . This further transforms the state space representation (12) into:

$$\eta_i = (\mathbf{C}_{i+2}^{i+2} \mathbf{R}_{i+3}(q_{i+3}) \bar{\xi}_{i-1})^\top \mathbf{e}_i \leftarrow \lambda_{i-1} \rightarrow 1 \quad (22)$$

with $i = 1, 2, 3$, where $\mathbf{C}_{3,4,5}$ are matrices whose variations are negligible due to $\lambda_{1,2,3} \rightarrow 1$, $\mathbf{e}_1 = \vec{x}_g$, $\mathbf{e}_2 = \vec{y}_g$, $\mathbf{e}_3 = \vec{z}_g$. Note that the regulation of each η is now dominated by only one robot joint. Then, based on the property revealed in [Proposition 2](#) and the controller (17), the process of SMA is transformed to a sequence of globally stable control subsystems coordinated by \mathbf{L} . As long as \mathcal{F}_g is reachable, \mathcal{H} must be reachable and thus $\boldsymbol{\eta} = \mathbf{0}$ could be reached by using (17). Examples of state evolution upon SMA and its performance when dealing with disturbances from $\boldsymbol{\eta}$ are demonstrated in [Figure 4](#).

Another explicit advantage of SMA is that the motion control of the robot's end-effector and body parts does not rely on computing inverse kinematics and thus allows a differentiable system, which could guarantee efficiency and smoothness when dealing with complicated motions. Recall the goal position of \mathbf{x}_s is coarsely computed as $\tilde{\mathbf{x}}_{s_g}$ in (13). During the manipulation, movements of the distal joints (e.g., $q_{5,6}$) might change \mathbf{x}_e and also $\tilde{\mathbf{x}}_{s_g}$, which generates disturbance to η_1 and also the whole SMA process. However, this facilitates $\tilde{\mathbf{x}}_{s_g} \rightarrow \mathbf{x}_{s_g}$ as “self-propagating” goal

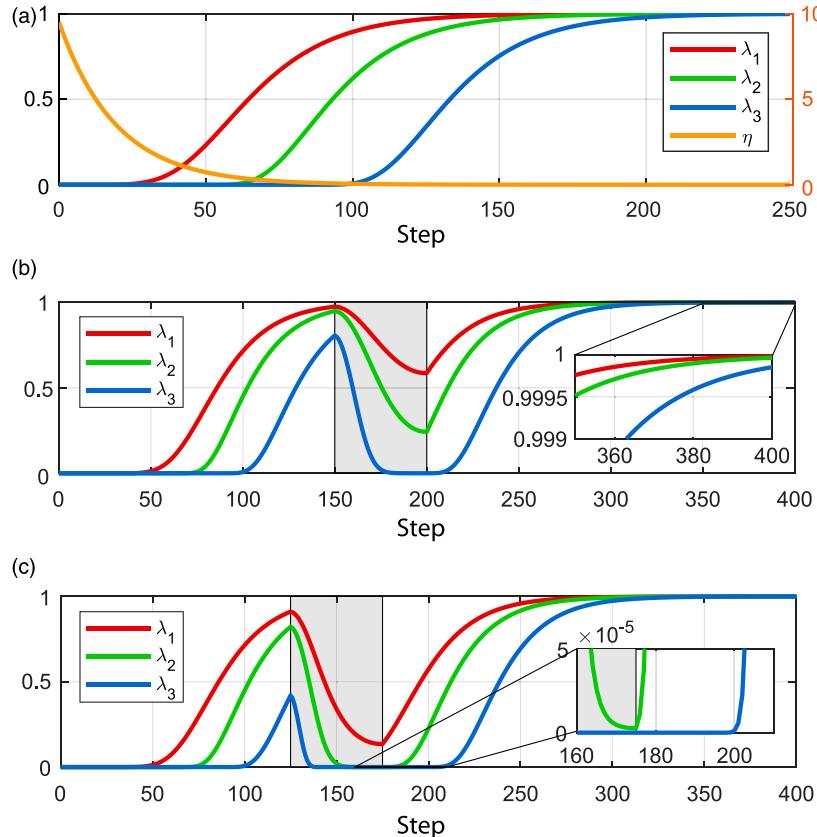


Figure 4. Evolution of λ subject to $\|\eta\|$, where a) denotes an undisturbed evolution with exponentially descending $\|\eta\|$; b) evolution upon slight increase of $\|\eta\|$ during $t \in [150, 200]$; and c) evolution upon large increase of $\|\eta\|$ during $t \in [125, 175]$. Note that the proximal λ stays greater than the distal one to enforce sequential coordination.

pose reaching process among \mathcal{H} , as $\eta_{2,3,4}$ are strictly bounded and globally converges to 0. This together with the “weak coupling” effect of the wrist kinematics leads \mathbf{x}_t to a bounded disturbance which does not affect our control system.

5.1.3. Collision avoidance. In this part, we briefly introduce our framework to deal with obstacle avoidance. In robotic surgery, obstacle avoidance mainly solves two scenarios: tool-tool collision when performing dual-arm surgical tasks within a shared workspace, and tool-tissue collision to prevent trauma to surrounding tissues. As the obstacles might be described as a single obstacle \mathcal{O} or a set of obstacles $\mathcal{O} = \{\mathcal{O}_1, \mathcal{O}_2, \dots, \mathcal{O}_n\}$ from image-based measurements, without loss of generality, we start by considering a single obstacle identity \mathcal{O} as a sphere (which could be easily extended to lines or surfaces) with 3D position \mathbf{x}_o . From robot kinematics and eye-to-hand data, the closest point \mathbf{x}_{clo} on the robot skeleton relative the obstacle, and its local Jacobian matrix $\mathbf{J}_{clo} \in \mathbb{R}^{3 \times 3}$ can be computed. Note that the position of \mathbf{x}_{clo} on the robot might be changing upon robot motions. Then, by giving a minimum tolerant distance r_o when avoiding collision (i.e., the radius of the sphere), the avoidance could be done by finding the tangent direction for the sphere \mathcal{O} from \mathbf{x}_{clo} . There are infinite directions to satisfy such, but one could choose the closest one to $\mathbf{x}_{sg} - \mathbf{x}_{clo}$ to reduce trajectory length. Thus, an instant target position $\mathbf{x}_{g,o}$ for \mathbf{x}_{clo} could be derived by letting

$$\mathbf{x}_{g,clo} = \mathbf{x}_{clo} + a_e \mathbf{e}_g \quad (23)$$

where $\mathbf{e}_g \in \mathbb{R}^3$ is a unit vector that denotes the selected avoiding direction and a_e is a scalar denoting the target motion speed. This indicates that we directly seek for obstacle avoidance by deriving the escaping velocity to be followed by the \mathbf{x}_{clo} . This allows that once \mathbf{x}_{clo} goes through $\mathbf{x}_{g,clo}$, the tangent direction changes and then varies $\mathbf{x}_{g,clo}$ upon (23) such that the collision avoidance motion continues. Meanwhile, we monitor whether the obstacle still obstructs the robot to reach the target by computing the distance between \mathbf{x}_{obs} and the line formed by $\mathbf{x}_{sg} - \mathbf{x}_{g,clo}$. Finally, we decide whether the robot completes obstacle avoidance by the following

$$\mathcal{X}_g = \begin{cases} \mathcal{X}_o, & f_o(\mathcal{X}, \mathcal{X}_p, \mathbf{p}_o, r_o) \leq 0 \\ \mathcal{X}_p, & f_o(\mathcal{X}, \mathcal{X}_p, \mathbf{p}_o, r_o) > 0 \end{cases} \quad (24)$$

where $f_o(\cdot)$ denotes the scalar function that compute the aforementioned point-to-line distance subtracted by r_o . \mathcal{X}_o is the end-effector state for collision avoidance constructed by \mathbf{x}_g with unchanged orientation, as the local Jacobian only regulates the 3D position of \mathbf{x}_{clo} . For practicality, to avoid abrupt change of the instant target state led by (24), one could apply linear fusion of two states to achieve smooth state transition.

5.2. Integrated planning and control

Apart from instrument manipulation, a majority of surgical instrument tasks (e.g., dissection and suturing) involve reaching and contact a specific target. The end-effector is normally manipulated to a proper candidate pose first (i.e., the tool-centric action) and then proceed to the final contact phase (i.e., the target-centric action) to could avoid unnecessary collisions or hazardous movements (Cao et al., 1996; Jun et al., 2012). Coordination between these two actions should be based on online awareness of the tool-target situation such that it could adaptively decide whether the robot is ready for contact or requires further adjustment. To this end, instead of regarding them as separated steps, we propose a unified strategy that integrates the tool-centric and target-centric actions to a single motion step to facilitate safer, smoother, and more efficient motion performance.

Recall the instrument’s goal configuration described by frame \mathcal{F}_g . It now also needs to be used to guide the end-effector to direct contact a target besides manipulation (mostly grasping an object). To this end, we extend \mathcal{F}_g to two poses, the pre-contact pose \mathcal{F}_p and a contact pose \mathcal{F}_c to be reached by the instrument’s end-effector. It should be noted that \mathcal{F}_c is usually computed via online sensing feedback or assigned by a user, and \mathcal{F}_p is computed based on \mathcal{F}_c which should be reached in advance in order to proceed to the contact safely. Then, we define a known end-effector trajectory $\rho(\tau(t))$ from \mathcal{F}_p to \mathcal{F}_c to be tracked by the end-effector to clear the final tool-target distance for finishing the contact, which also exactly defines the target-centric action. The parameter $\tau(t) = [0, 1] \in \mathbb{R}$ denotes the progress of the trajectory and is of class C^1 . Thus, the appearance of τ enables a dynamic \mathcal{F}_g and leads the differentiation of $\boldsymbol{\eta}$, currently of $\mathcal{X}(\mathbf{q})$ and $\mathcal{X}_g(\tau)$, to the following extended form:

$$\dot{\boldsymbol{\eta}}(\mathcal{X}(\mathbf{q}), \mathcal{X}_g(\tau)) = \frac{\partial \boldsymbol{\eta}(\cdot)}{\partial \mathcal{X}(\mathbf{q})} \frac{\partial \mathcal{X}(\mathbf{q})}{\partial \mathbf{q}} \dot{\mathbf{q}} + \frac{\partial \boldsymbol{\eta}(\cdot)}{\partial \mathcal{X}_g(\tau)} (\mathcal{X}_c - \mathcal{X}_p) \dot{\tau} \quad (25)$$

whose evolution performance could be proved globally stable as of (18) by deploying the following controller

$$\mathbf{u} = \left(\frac{\partial \mathcal{X}(\mathbf{q})}{\partial \mathbf{q}} \right)^{-1} \left(\frac{\partial \boldsymbol{\eta}(\cdot)}{\partial \mathcal{X}(\mathbf{q})} \right)^{-1} \left(\frac{\partial \boldsymbol{\eta}(\cdot)}{\partial \mathcal{X}_g(\tau)} (\mathcal{X}_p - \mathcal{X}_c) \dot{\tau} - \boldsymbol{\Gamma} \boldsymbol{\eta}(\cdot) \right) \quad (26)$$

which not only stabilizes $\boldsymbol{\eta}$ to $\mathbf{0}$ as in (18) but further guarantees that the robot state \mathcal{X} will converge to the following

$$\lim_{t \rightarrow \infty} \mathcal{X} = \mathcal{X}_g = \tau \mathcal{X}_c + (1 - \tau) \mathcal{X}_p \quad (27)$$

which indicates that the robot state will now arrive exactly at the prescribed trajectory $\rho(\tau)$ that connects \mathcal{X}_p and \mathcal{X}_c . The instant \mathcal{X}_g to be reached by the robot is determined by the value of τ whose online adjustment will specifically enable

$$\lim_{t \rightarrow \infty} \|\boldsymbol{\eta}(\cdot)\| = \mathbf{0} \Rightarrow \mathcal{X} = \begin{cases} \mathcal{X}_c, & \lim_{t \rightarrow \infty} \tau = \infty \\ \mathcal{X}_p, & \lim_{t \rightarrow \infty} \tau = 0 \end{cases} \quad (28)$$

where the performance of regulating $\boldsymbol{\eta}$ will not be affected by the change of τ , but meanwhile facilitates a self-tunable end-effector goal pose \mathcal{X}_g . We define this as a goal-varying manipulation (GVM) problem that facilitates the versatile robot motions via prescribed trajectories and/or states, by setting \mathcal{X}_g to a set of target states $\{\mathcal{X}_{g_1}, \mathcal{X}_{g_2}, \dots, \mathcal{X}_{g_n}\}$ that characterize the goal-relevant manipulation process. Here, τ is also regarded as the goal varying parameter (GVP) and will be used to tune the robot's instant goal state adaptively based on its on-the-fly configuration $\boldsymbol{\eta}$, in order to guide the robot through specific motion patterns via a controllable process upon the task's needs. To make GVM contribute to our case, we propose the following dynamics for τ :

$$\dot{\tau}(\boldsymbol{\eta}) = \gamma(2 \exp(-\kappa \|\boldsymbol{\eta}\|) - 1) \quad (29)$$

where γ and κ are the tuning parameters of the dynamical performance. Particularly, the dynamics (30) owns the following properties that contribute to our application in surgical task autonomy.

- **Property 1.** The alignment of \mathcal{X} to \mathcal{X}_g , characterized by $\|\boldsymbol{\eta}\|$, tunes both the magnitude and the direction of $\dot{\tau}$:

$$\dot{\tau} = \begin{cases} -\gamma, & \lim_{t \rightarrow \infty} \|\boldsymbol{\eta}\| = \infty \\ \gamma, & \lim_{t \rightarrow \infty} \|\boldsymbol{\eta}\| = 0 \end{cases} \quad (30)$$

where alignment of the robot to its instant goal state will elevate τ upon the tuning step up to γ , which carries forward the prescribed GVM process, while misalignment (with $\eta \rightarrow 0$ renders $\dot{\tau} < 0$ which rewinds the GVM process back as $\dot{\tau} < 0$ drags τ back to zero). This enables the manipulation process to be regulated in an adaptive and bidirectional manner guided by τ .

- **Property 2.** Based on Property 1, τ will reach 1 within finite time. This can be proved that for any instant $\boldsymbol{\eta}$, $\dot{\tau} > 0$ holds once

$$\boldsymbol{\eta}(T_1) < -\frac{1}{\kappa} \ln \frac{1}{2} \quad (31)$$

where T is the time instant when τ starts to elevate, and obviously $T < \infty$ as $\lim_{t \rightarrow \infty} \boldsymbol{\eta} = \mathbf{0}$. Then, since $\dot{\tau}$ is monotonically increasing, $\dot{\tau} > 0$ will hold and further lead to

$$\int_T^{T+T'} \dot{\tau} dt > 1 \quad (32)$$

where $T' < \infty$ as well if we further saturate τ in discrete time as

$$\tau(t_{k+1}) = \begin{cases} 1, & \tau(t_k) + \dot{\tau}(t_k) \Delta t > 1 \\ 0, & \tau(t_k) + \dot{\tau}(t_k) \Delta t < 0 \end{cases} \quad (33)$$

to enforce $\tau \in [0, 1]$. This property facilitates \mathcal{X}_g to land on the final goal state (thanks to τ (30)) prior to $\mathcal{X} \rightarrow \mathcal{X}_g$ to ensure efficient and accurate guidance towards the target.

- **Property 3.** The boundary conditions of $\dot{\tau}$ subject to (30) evaluated with respect to $\boldsymbol{\eta}$ are

$$\left. \frac{\partial \dot{\tau}}{\partial \boldsymbol{\eta}} \right|_{\boldsymbol{\eta} \rightarrow \infty} \rightarrow 0, \quad \left. \frac{\partial \dot{\tau}}{\partial \boldsymbol{\eta}} \right|_{\boldsymbol{\eta}=0} = -2\kappa\gamma \quad (34)$$

which indicates that the tuning of τ is sensitive to the change of $\boldsymbol{\eta}$ around the alignment situation (i.e., $\boldsymbol{\eta} \rightarrow 0$), and becomes relaxed under misalignment with $\boldsymbol{\eta} \rightarrow \infty$. This will make τ rapidly descend away from 1 once encountering misalignment between \mathcal{X} and \mathcal{X}_g .

During the adjustment of τ , we regard $\boldsymbol{\eta}$ as an external input whose regulation is solely subject to (26). Note that the introduction of GVM using τ on the basis of DS-guided manipulation in (25) further integrates the trajectory planning and the motion control of versatile robot manipulation steps to a unified solver. Selection of individual \mathcal{X}_g and their connected trajectory ρ will enable versatile motion characteristics, which could be either predefined or solved on-the-fly.

The above GVM framework could be used to automate typical instrument motions in robotic surgery. First, it could solve delicate motions required for target contact. We can simply define \mathcal{X}_p whose orientation is identical to that of \mathcal{F}_c but with a constant position difference d along the tool's pointing direction, with

$$d > l_w + l_t \quad (35)$$

to ensure the tool-centric action keeps the manipulation clear from the target before it reaches \mathcal{X}_p . It degenerates ρ to a line segment $\overrightarrow{\mathcal{O}_p \mathcal{O}_c}$, such that the robot will first reach \mathcal{X}_p with the orientation already suitable for contact, and then clears the remaining distance d upon $\tau \rightarrow 1$ to finally reach \mathcal{X}_c . Here, the tool-centric action (reaching \mathcal{X}_p) and target-centric action (reaching \mathcal{X}_c) are independent but smoothly coordinated. Second, GVM can be applied for tracking trajectories required by specific task-relevant motions (e.g., needle insertion path must follow its own circular curvature to minimize tissue trauma). Here, $\rho(\tau)$ denotes a general path which starts from \mathcal{X}_{g_1} and ends at \mathcal{X}_{g_2} . The GVP τ will attract the robot to enter \mathcal{X}_{g_1} and then guide \mathcal{X} to follow $\rho(\tau)$ until \mathcal{X}_{g_2} is reached. Due to $\dot{\tau}$ in (25) and Property 3, τ will smoothly elevates to 1 while ensuring that \mathcal{X} could catch up the varying \mathcal{X}_g . The overall model of the integrated planning and control framework is shown in Figure 5.

We briefly emphasize the superiority of our architecture from a combined technical and practical perspective that contributes to surgical applications. The introduction of NSS to represent the robot's present-to-goal variation provides distinct modelling advantages. The global asymptotic stability of (18) and (25) guarantees that the robot can be stably maneuvered to any arbitrary feasible state with smooth positions and velocities. Incorporating SMA then

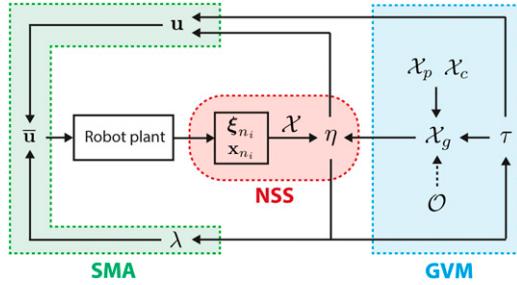


Figure 5. The schematic diagram of our integrated planning and control framework using NSS η to represent and control the robot instead of end-effector pose. GVM serves as an adaptive planning policy that provides active guidance using the goal state X_g via τ , which is reactive to obstacle avoidance.

provides versatile planning of the robot's goal-reaching trajectory as well. They solve the issues of "the robot could reach a goal" and "the robot knows how to reach a goal" simultaneously using our framework via a mathematically determined solver without prior knowledge or online iterations. GVM then plays a role to solve "the robot knows when to act" which picks up the last components required for a motion-level architecture. Its capability to transfer robot motion through different types of motions is illustrated in Figure 6.

For surgical applicability, GVM allows motion constraints applied to the in-process trajectory, which makes it powerful to solve delicate target contact motions, where a constrained path is required to contact tissue via specified poses to avoid unnecessary tissue trauma. Particularly, the tool-centric action guarantees avoidance of premature contact to the target, and the instrument will not approach the target (at X_c) before it reaches X_p (s.t. $\eta \rightarrow 0$). This could avoid entanglement of the tool to the target before the final phase of manipulation (i.e., $X_g \rightarrow X_c$) is settled. In addition, as the tuning of GVP τ is smooth and bounded, the robot actions can be transferred smoothly as well.

6. Task-level autonomy architecture

In this section, we extend our approach to task-level automation such that it could complete the whole pipeline of one or more surgical tasks. The key issue is to further define a powerful motion descriptor that reveals the mutual behavior among individual motion steps while tolerating their underlying differences. GVM has enabled safe target contact by constraining the instrument motion delicately during the final motion phase. However, how to guide the instrument for clearance of the target or interact via specific path are similar but different steps to be considered.

6.1. Surgical motion primitive (SMP)

To systematically entail these motion details, we propose the notion of surgical motion primitive (SMP) based on GVM as a generalized model to define surgery-related

instrument motions. The SMP characterizes a single-instrument manipulation step into three progressive phases:

- **Phase I (Backward guidance):** The robot is guided via evolution of GVM through $\tau : 1 \rightarrow 0$, which is used for target-centric action to escape the instrument escapes from an already contacting object.
- **Phase II (Free control):** The robot is guided via evolution of $\tau = 0$ without GVM, which is used for tool-centric action to reach the instrument to a target state without contacting any target objects.
- **Phase III (Forward guidance):** The robot is guided via evolution of GVM through $\tau : 0 \rightarrow 1$, which is used for target-centric action or trajectory tracking, where the instrument approaches and contacts a new target object or follows specific path.

The current setting of our GVM can only achieve consecutive execution across Phase II and Phase III. Thus, we must modify GVM such that it is applicable to the whole SMP. In this regard, we extend the validity of τ to Phase I such that it must retract the end-effector from the target to a safe position prior to Phase II. We define the initial pose as X_{c_0} where contact still exists, and the retracted state as X_{p_0} , which is also the ending state for Phase I. Without loss of generality, we maintain the relationship between them as that of X_c and X_p , and their in-between trajectory ρ remains a line segment and should be performed rapid by the instrument to avoid unnecessary contact. The GVP τ is also set to 1 at X_{c_0} and 0 at X_{p_0} . Thus, during Phase I, we need to enforce $\dot{\tau} = -c$ (with c being a positive scalar) to lead the instrument to X_{p_0} subject to $\tau : 1 \rightarrow 0$, which exactly reverses the process of Phase III. To smoothly connect Phase I to Phase II, we equip τ with the following properties

$$X_g = \begin{cases} X_{p_0}, & \epsilon < \tau \leq 1 \\ X_p, & 0 \leq \tau < \epsilon \end{cases} \quad (36)$$

where ϵ is the threshold that coordinates Phase I and Phase II. Once X_g switches to X_p , the dynamics of τ is switched by (30) to regulate subsequent SMP phases.

6.2. Modes of behavior (MoBs)

Finally, we utilize SMP to define and automate the entire pipeline of instrument motion sequences for different surgical tasks. Although we define three consecutive phases that form the single SMP, some instrument motion steps might only involve one or two of them. For example, an off-target manipulation can be fully performed by initiating only Phase II. Retracting the instrument from an object to a stand-by contact-free target could be done by using Phase I and Phase II. To make SMP customizable to such motion differences, we combine different motion phases and generate the following five typical modes with their application scenarios:

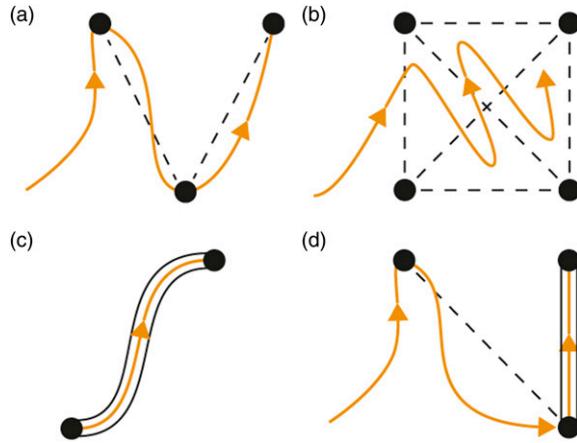


Figure 6. Conceptual illustration of different types of motion behavior achievable by using the GVM structure. The orange curve indicates the guided motion by the dynamic X_g through different types of motion (black dots as the states, dashline as free manipulation motion, and parallel solid lines as motion with determined path).

- **Mode I (Phase II):** The instrument motions is freely adjusted via DS + SMA with no additional workspace constraints and thus no GVM ($\tau = 0$).
- **Mode II (Phase I + II):** The instrument leaves a contacted object under constrained path to an off-contact state guided via GVM with $\tau : 1 \rightarrow 0$.
- **Mode III (Phase II + III):** The instrument reaches and contacts a target from an off-contact state (e.g., grasping a tissue/needle) guided via GVM with $\tau : 0 \rightarrow 1$.
- **Mode IV (Phase I + II + III):** The instrument retracts from the previous object via constrained path ($\tau : 1 \rightarrow 0$), adjusts its pose ($\tau = 0$), and then reaches and contacts a new target object, with $\tau : 0 \rightarrow 1$.
- **Mode V (Phase III):** The instrument is guided through a specific path with determined initial/final state $\tau : 0 \rightarrow 1$.

They are also defined as the MoBs, as they exhibit own motion constraints while sharing the identical control architecture, and are available to each SMP (see Figure 7 for conceptual illustration). To achieve task-level autonomy, we first utilize a chain of SMPs that govern each motion step in the surgical task to construct the whole pipeline. Then, we equip each SMP (or each motion step) with a specific MoB according to its prescribed functionality. As the guidelines for surgical tasks are already mature in clinical practice, the MoBs could be preset for each task such that it could successfully follow the procedure. In Figure 8, we demonstrate the applicability of our framework to characterize many different existing surgical tasks by selecting different SMPs and MoBs, based on clinically adopted surgical guidelines (Cao et al., 1996). The validity covers not only intra-corporeal MIS procedures (e.g., tissue dissection, wound suturing, and knot tying) but also other procedures like robotic palpation and biopsy, which is generalized framework compared to existing task-specific approaches. Although each elementary motion has different constraints,

they could be set by five modes of behavior to form a pipeline. For example, the specification for dual-arm suturing could be easily interpreted as 3-5-0-0-0-4-1 for PSM 1 and 0-0-3-5-2-1-2 for PSM 2 (where 0 means staying idle in this step). This could all be preset in the framework as long as the task type to be automated is determined, which is very easy to specify. We will show later in our experiments that the ability of coordinating motion execution process adapting to the online robot-environment situation is critical to improve the success rate and reliability of task autonomy.

7. Simulations and Results

In this section, we present our simulation setup and results to demonstrate the motion-level performance of robotic instrument manipulation. The simulations will be conducted to reveal the following two main aspects of our approach: How the robot motions are actuated to deal with delicate and drastic manipulation steps, and how the strategy reacts to different tool-target configurations when attempting to perform interaction to a physical target.

7.1. Overview

We implement our algorithm framework on the Virtual Robot Experimentation Platform (V-REP), which is then interfacing to MATLAB R2017a (MathWorks Inc) via remote API on a Core i7 2.8 GHz system without GPU acceleration. In the simulation platform, we use the virtual model of da Vinci Research Kit (dVRK) provided by Fontanelli et al. (2018). The dVRK is a dual-arm surgical robot with two wristed robotic instruments, which are named as the Patient-Side Manipulators (PSMs). Both two attached instruments are Large Needle Drivers (LNDs) which is the typical selection in dVSS-enabled RAMIS to perform delicate intra-corporeal procedures like suturing. Virtual objects, either pegs or a surgical needle, involve pick-and-place operation, which is one common surgical motion type and is frequently used as a proxy for generalized surgical dexterity (Cao et al., 1996). Here, we will use them to define practical task-relevant movement steps for evaluation of instrument's motions. The distance between the RCM of the PSMs to the peg transfer objects are selected to around 150 mm which resembles the genuine robot-target configuration during RAMIS. We do not apply dynamic motion properties and interactions throughout the simulations, as we mainly investigate the motion behavior of the instruments in 3D workspace. The detailed parameter setup of our system is shown in Table 2.

7.2. Scene I: Complex tool manipulation

We first simulate a peg-transfer setup where a plastic ring is to be manipulated by the robot from one peg to another, which is commonly used in surgical skill training for novice

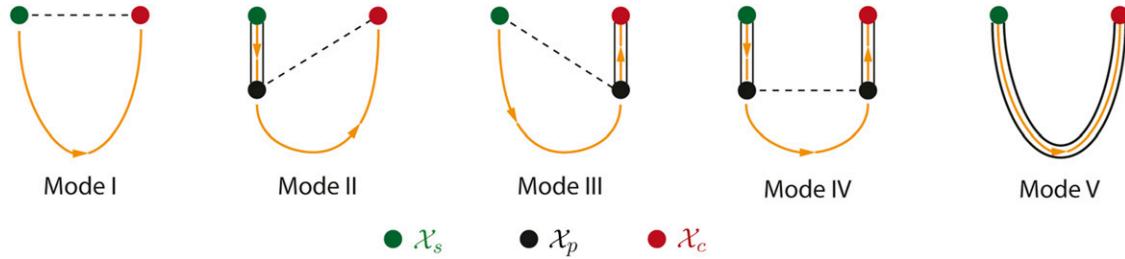


Figure 7. The five modes of behaviors (MoBs) proposed based on our planning and control framework to define individual motion steps in various tasks. Mode I: free reaching, Mode II: constrained retraction with free reaching, Mode III: constrained reaching, Mode IV: constrained retraction and reaching, Mode V: trajectory tracking.

Tasks	Sequence							
	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
Biopsy	Needle targeting to insertion pose ● \mathcal{X}_s	Insertion along navigated path ● \mathcal{X}_p	Retraction for next insertion pose ● \mathcal{X}_c					REPEAT Step 2 - Step 3 ● \mathcal{X}_p → ● \mathcal{X}_c
Palpation	Probe contact to tissue surface ● \mathcal{X}_s	Scanning along planned path ● \mathcal{X}_p	Retraction for next scanning pose ● \mathcal{X}_c					REPEAT Step 2 - Step 3 ● \mathcal{X}_p → ● \mathcal{X}_c
Debridement	Tool targeting to tissue ● \mathcal{X}_s	Tissue retraction & collection ● \mathcal{X}_c						REPEAT Step 1 - Step 2 ● \mathcal{X}_s → ● \mathcal{X}_c
Dissection	Tool targeting and dissection ● \mathcal{X}_s	Tool moving to next pose ● \mathcal{X}_c						REPEAT Step 1 - Step 2 ● \mathcal{X}_s → ● \mathcal{X}_c
Knot Tying	Suture grasping ● \mathcal{X}_s 1	Suture looping ● \mathcal{X}_p 1 2	Suture grasping ● \mathcal{X}_c 2	Knot tightening ● \mathcal{X}_s 1 2				REPEAT Step 1 - Step 4 ● \mathcal{X}_s → ● \mathcal{X}_c
Suturing	Needle targeting to insertion pose ● \mathcal{X}_s 1	Needle Insertion into tissue ● \mathcal{X}_p 1	Needle regrasping ● \mathcal{X}_c 2	Needle exiting from tissue ● \mathcal{X}_p 2	Suture pulling ● \mathcal{X}_c 2	Needle Hand-off ● \mathcal{X}_s 1 2	Tool Resetting ● \mathcal{X}_s 1 2	REPEAT Step 1 - Step 7 ● \mathcal{X}_s → ● \mathcal{X}_c



Figure 8. Demonstration of complete pipelines of various independent surgical tasks which could be fully formulated by our framework using the five MoBs for task automation.

surgeons. The initial/final state of the pegs is usually selected that corresponds to awkward robot configurations (i.e., either near-singularity or near-limit joint states). The robot trajectory needs to experience large-range orientation adjustment in order to align the ring to the peg to finish the manipulation properly, which might be hard to plan within confined space. To verify the validity of our framework to tackle such extreme scenarios, we assign three typical cases, as shown in Figure 9:

- Case I: Transfer the ring from a horizontal peg to an adjacent vertical peg.
- Case II: Transfer the ring from a horizontal peg to another horizontal peg by twisting the shaft through 180 degrees.
- Case III: Transfer the ring from a vertical peg to another vertical peg by twisting the shaft through 180 degrees.

The target states of the robot are known and remain static at this stage of simulation. The process is actuated using Mode III for our proposed SMP. Particularly, the differences between initial and final positions of the distal joints q_{w_1} , q_{w_2} , and q_l in three cases are $[1.16, 1.487, 2.1]^\top$ rad, $[1.672, 2.294, 1.296]^\top$ rad, and $[-2.853, 0.732, 2.2002]^\top$ rad, respectively, which indicate an average of $>100^\circ$ motion for each distal joint in each case. The results show that the DS-based framework could simultaneous plan and control the instrument's motion through twisted motions and reach the target configuration in all cases. The guaranty of global unique solution in Proposition 2 in our new robot model, the instrument smoothly reaches the target pose without being trapped in false equilibriums, which might appear in optimization-based approaches due to the nonconvex workspace manifold. Meanwhile, the trajectory remains

Table 2. Parameter setting of our algorithm in simulations.

Parameter	Value
Γ	$\begin{bmatrix} 0.1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$
\mathbf{l}_h	$[-1 \ 0 \ 0]^\top$
\mathbf{l}_v	$[0 \ 1 \ 0]^\top$
\mathbf{M}	$[0 \ 0 \ 1]^\top$
k_1	\times (random)
k_2	0.2
k_3	1
k_4	1
Γ	0.05
K	0.5
ϵ	0.01

feasible despite the target joint positions of specific joints being adjacent to its limits. We compare the resultant joint actions with and without deploying the SMA (as shown in [Figure 10](#)). Without SMA, the distal joint motions tend to overshoot from its final positions before convergence. This might result in exceeding joint limits ((e.g., q_t in Case I and II, and q_{w2} in Case III, note that the joint limit of q_{w2} and q_t are both $\pm \pi/2$ rad, illustrated by gray shaded areas). The convergent performance is also not affected by the sequential actions.

Then, we verify the capability of GVM to inherently avoid unnecessary contact to the target during manipulation. The main symbol of inadvertent tool-target collision is the time instant(s) that the end-effector position goes behind the target contact point, which is mathematically interpreted by the following metric:

$$\phi := \langle {}^g \mathbf{x}_t, \vec{z}_g \rangle < 0 \quad (37)$$

Guiding the end-effector through \S_p and \S_c by implementing GVM avoids $\phi < 0$ throughout three cases (refer to the dark green curves in [Figure 11](#)). Without GVM, the end-effector will intrude the target area (shown in [Figure 11](#) the red curve segments) during manipulation which could damage the tissue in real surgical procedures. GVM allows the twisted motions to be properly adjusted first safely away from the target before reaching it. Note that such performance is not particular to the selected cases, but could be easily guaranteed upon proper setting of \mathcal{X}_p according to \mathcal{X}_c in (27) could guarantee Note also from [Figure 11](#) that a reactive \mathcal{X}_g does not affect the convergent performance of instrument manipulation.

7.3. Scene II: Dual-arm collision avoidance

Then, we investigate the applicability of our algorithm to deal with collision avoidance. Aiming for a common but challenging scenario in robotic surgery, we consider a tool-tool collision situation where two robotic instruments are

manipulated to their respective targets, respectively, and have to move close to each other within shared workspace. The initial and final states of two instruments are defined such that their trajectories will make the robot skeleton collide to each other if no avoidance is actuated.

[Figure 12](#) shows the setup of this scene and the simulation snapshots for the dual-arm manipulation considering tool-tool collision avoidance. Our framework allows simultaneously dual-arm collision avoidance as the planning strategy is decentralized. The minimum tolerant distance is set to 10 mm as the tool shaft size is 8 mm, leaving the minimum tool-tool distance to be 2 mm. Both two instruments start collision avoidance as soon as the process begins and consider the other one as the moving obstacle. By inspecting the instant nearest obstacle point on the skeleton, two shafts move around each other and retract the distal part to prevent the whole robot from collision. After the other instrument stays clear from its target, the instrument proceeds on the tool-centric action and target-centric action.

[Figure 13](#) shows the resultant dual-arm trajectories. For comparison, the trajectories for direct target reaching are shown as well. It is clear that to avoid collision, the trajectories are more complex, but remain smooth and do not affect the subsequent target-reaching process where a constrained path is required. The collision avoidance input only applies on demand when the obstacle obstructs target reaching, and will smoothly fade once the obstacle is clear.

7.4. Scene III: Perturbed interaction

In this subsection, we validate the motion performance of the instrument when reacting to different target situations led by the GVM. A typical action during surgery is to manipulate the instrument's distal tool from a safe idle configuration to contacting a target object, usually for grasping purposes. The contact must be made via a proper reaching pose to ensure safe tool-target interaction. We use a half-circle (or 1/2) surgical needle as a target, where the robot aims to grasp the needle body at its two-thirds curvature point as a standard needle pick-up procedure. The needle is attached to a support to visually emphasize its pose change during simulation, which is also known to the robot's base. Again, we apply Mode III motion behavior to the instrument, where \mathcal{X}_g moves between \mathcal{X}_p and \mathcal{X}_c subject to on-the-fly regulation of τ . \mathcal{X}_p is assigned with the same orientation as of \mathcal{X}_c but with a negative z-axis offset position from the grasping point. [Figure 14](#) illustrates the tool reaching process of the instrument from its idle position $[-116.04, 16.85, -56.13]^\top$ mm to the ideal needle grasping point $[-135.68, 18.60, -61.17]^\top$ mm. The trajectory of the end-effector indicates that the robot performs tool-centric action that guides the tool to the pre-grasp pose first with τ maintaining close to 0. While \mathcal{X}_g is nearly aligned with \mathcal{X}_p , the GVP τ adaptively transits from 0 to 1 to further allow the instrument to eliminate the remaining distance for contact, that is, proceeding to target-centric action. Note that during such action, the orientation of the

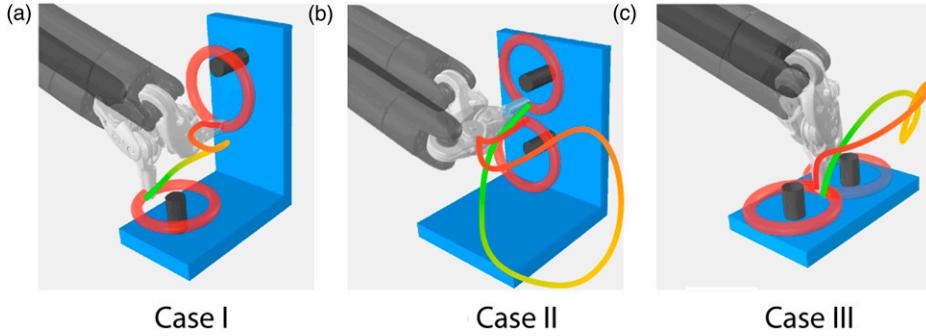


Figure 9. The three cases of pose-to-pose manipulation for simulated peg transfer tasks where the instrument needs to drastically twist the wrist. The end-effector trajectories are visualized from green to red along the evolution of time.

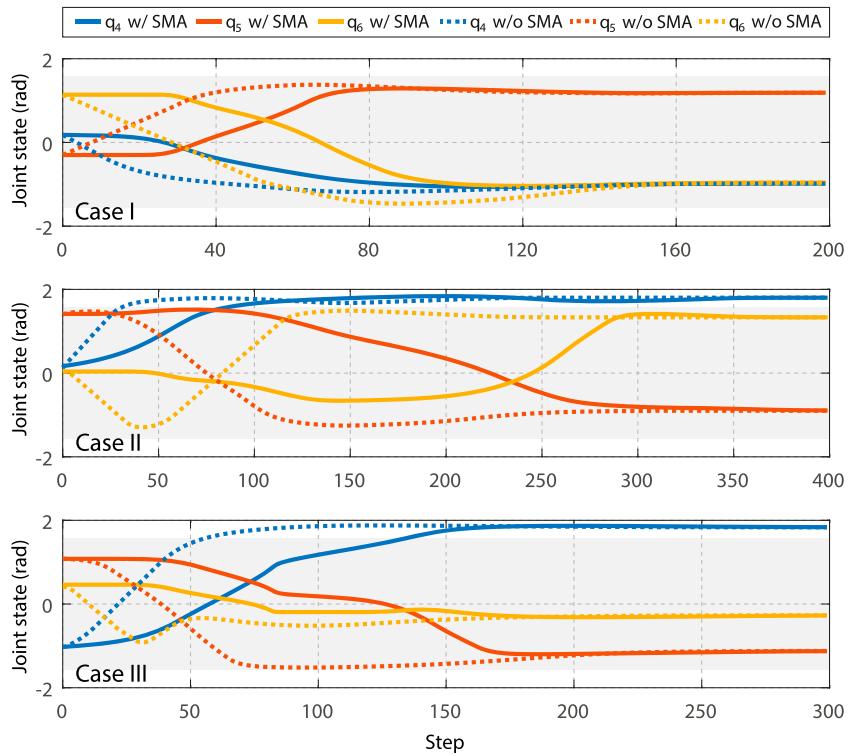


Figure 10. Evolution of distal joint positions for three cases of complex manipulation. The gray areas denote the $-\pi/2, \pi/2$ feasible joint range available to both q_5 and q_6 , where motion control without SMA tends to overshoot the joint positions before convergence, which might enter unreachable regions under near-limit configurations.

distal tool remains unchanged despite the varying \mathcal{X}_g at the final phase of contact motion, which helps achieve a proper tool-target contact configuration. It is also shown in Figure 15 that the regulation of the end-effector's orientation is already settling (with the robot approaching \mathcal{X}_p) before the target-centric action begins (with the robot approaching \mathcal{X}_c). This allows the tool to be “well-prepared” in advance to delicately reach the object, which will be shown later in our experiments to improve task-level reliability.

We next study how GVM will react to a target under different situations. The target pose of the robot might be perturbed, where the disturbances could be from sensoring

noises or physiological motions during surgical procedures. We add uniform random noises to the target 3D position with the magnitude of 0.003 mm for all axial directions and is deployed throughout the process. The orientation remains unchanged at this stage. The comparison of the target-reaching motion performance with and without applying noises is shown in Figure 15. It is clear that the perturbed setting results in not only longer convergent time of the tool’s position e_t , but also remaining a stabilized position difference (denoted by $e_{t,s}$) from the ideal contact point. This results from the untrackable noise which makes the robot controller hard to converge η sufficiently to 0, and thus prevent τ from

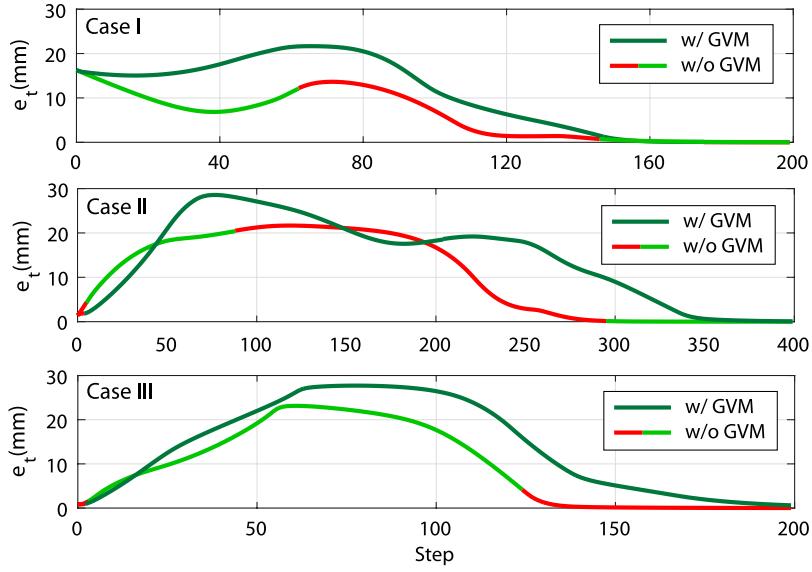


Figure 11. Instrument end-effector position error during target-reaching manipulation in the three cases with and without using GVM-based guidance. The red parts indicate the time instants where target intrusion happens (i.e., $\phi < 0$).

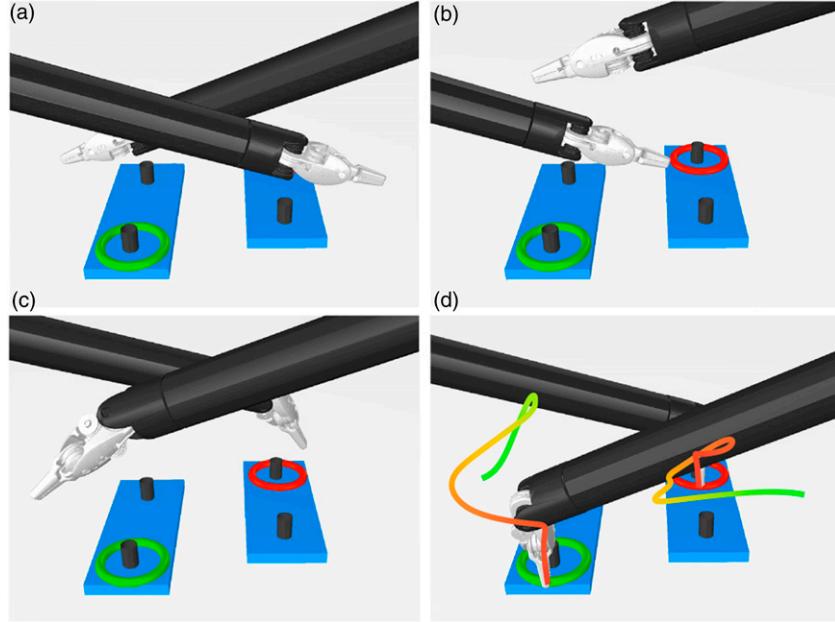


Figure 12. Simulation snapshots of dual-arm manipulation for ring grasping with tool-tool collision avoidance that evolves as a-d). In d), the back-projected instrument end-effector positions are shown that move from green to red.

settling to 1, although the orientation convergence is not affected. This indicates the noisy target could lead to an “inconfident” targeting attempt by our planning and control framework.

To further study the robot motion behavior using our system to deal with target disturbances, we apply different noises to the identical setup and further obtain the relationship between the e_t and the noise configurations, shown in Figure 16. Obviously, larger noises indicate larger stabilized position error $e_{t,s}$ from which the relationship is

relatively linear. Meanwhile, selecting smaller gain κ that tunes the reaction of $\dot{\tau}$ to η will lead to decreased $e_{t,s}$. This is because a smaller κ relaxes the condition of position alignment between \mathcal{X} to \mathcal{X}_g and ease τ to move to 1. When the noise magnitude is elevated to a certain turning point, the stabilized position rapidly increases but remain linear afterward. This is caused by the τ rewinding from nearly 1 to $\ll 1$ due to the noise-induced misalignment of the tool to the target (or the unsettled η). \mathcal{X}_g then returns to \mathcal{X}_p to make easier for the robot to reach and thus automatically creating

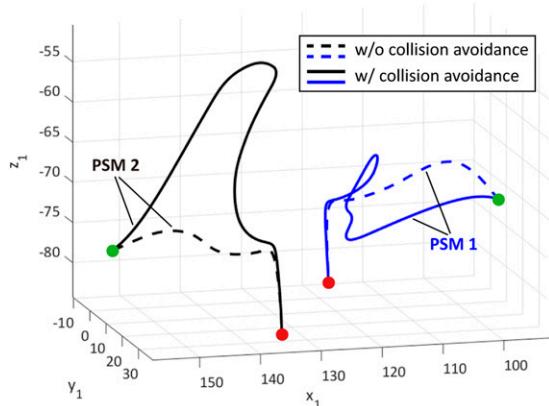


Figure 13. Motion trajectories of two robotic instruments during simultaneous ring grasping in Scene II (described by the base of PSM1 in mm) with and without applying collision avoidance. The green/red dots indicate initial/final positions of two robot end-effectors, respectively.

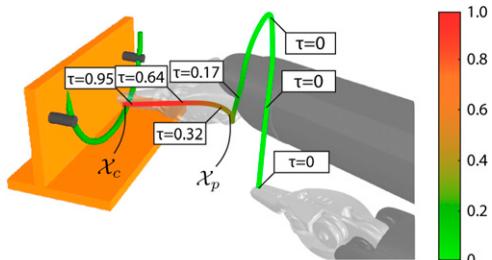


Figure 14. Illustration of pre-grasp manipulation and final constrained target contact planning and control upon evolution of τ based on (30).

a new balance. As the change of \mathcal{X}_g 's orientation provides extra disturbance to the position of \mathcal{X}_p , the turning point of $e_{t,s}$ upon increasing κ comes earlier, while the level-off $e_{t,s}$ is also higher than that without orientation noise. However, the maximum $e_{t,s}$ under different configuration will not exceed d in (35), as in the worst case, τ will only drop to 0. The above performance reflects that the disturbance of the target state could significantly change the motion behavior of the instrument when attempting to reach the target. Target perturbation will make the strategy cautious when performing the target-centric action, and might even abandon it if the disturbance is severe enough due to the impossibility to settle η to a. The instrument will still settle to \mathcal{X}_g with a contact-ready orientation upon tool-centric action, but then keeps a (preset) safe distance from the target. If the noise magnitude becomes small, the robot controller is still able to converge η close to 0 such that τ maintains positive as in (30). The strategy is then “confident” enough to guide the end-effector to the \mathcal{X}_c , or the final contact pose (e.g., $d < 2$ mm or so). Note that among all the above cases, a stabilized \mathcal{X} and \mathcal{X}_g is always achieved. The performance of the GVM using the proposed framework remains stable regardless of the noises being added.

7.5. Scene IV: Target-varying interaction

Finally, we conduct simulation based on Scene II to study how our framework could deal with a moving target. The instrument might reach a target which is being moved by another instrument (e.g., needle hand-off in wound suturing) or need to follow specific trajectories, such as needle insertion into tissue and palpation through tissue surfaces. The target is either moving or to be moved after being grasped. Here, we propose the following three typical cases that cover different target motion types to further evaluate the applicability of our framework to different manipulation strategies:

- Reaching an intermittently moving target along a linear trajectory with unchanged orientation.
- Reaching a continuously moving target along a pose-varying trajectory (moving position and orientation).
- Tracking an arc-shape trajectory (moving position and orientation).

The first and second cases that correspond to decentralized dual-arm coordination for target hand-off are set. In the first case, the final contact position \mathbf{x}_c (or the needle) moves from $[-102.59 \ -86.77 \ -62.46]^\top$ mm to $[-163.49 \ -85.99 \ -58.49]^\top$ mm through a straight line but halts at the midpoint for 100 steps. The movement starts at $t = 180$ and the overall travel distance is around 60 mm. The target orientation is $[2.70130.5169 \ -1.8761]$ and remains unchanged. The translating velocity of the needle is set to 0.1 mm per step. In Figure 17(a), the end-effector trajectory is first adjusted for tool-centric action for target orientation alignment, and then proceed to target-centric action while τ is moving to 1. However, after the needle starts to move (in orange), the misalignment between the tool and the target makes contact process immediately suspended, where the instrument position rapidly moves back to \mathcal{X}_p and track the needle's pose with a stabilized distance d . While the needle temporarily stops, the instrument attempts to reach the target again, and is suspended again as the needle moves. Finally, after the needle comes to its final stop, the instrument manages to reach the final contact point as τ eventually settles at 1.

Case II is similar to Case I but with 6-DoF pose change over time. By reserving the translation motion of the needle, we add rotary motion of the needle with respect to its own frame, that is, $0.15 * \sin 0.025t$ rad along y-axis and $0.25 * \sin 0.025t$ along z-axis. The motion starts at $t = 150$ and continuously lasts for $\delta t = 600$. As shown in Figure 17(b), the end-effector manages to track the moving \mathcal{X}_g induced by \mathcal{X}_c without contacting the needle. When the motion disappears, the tool is finally guided to the contact point where the needle could be directly grasped.

In Case III, we show how GVM contribute to stable trajectory tracking of the instrument. We first define an arc-shaped trajectory whose distance to the rotation center is 15 mm, and the total equivalent rotation range is $2/3\pi$. Such

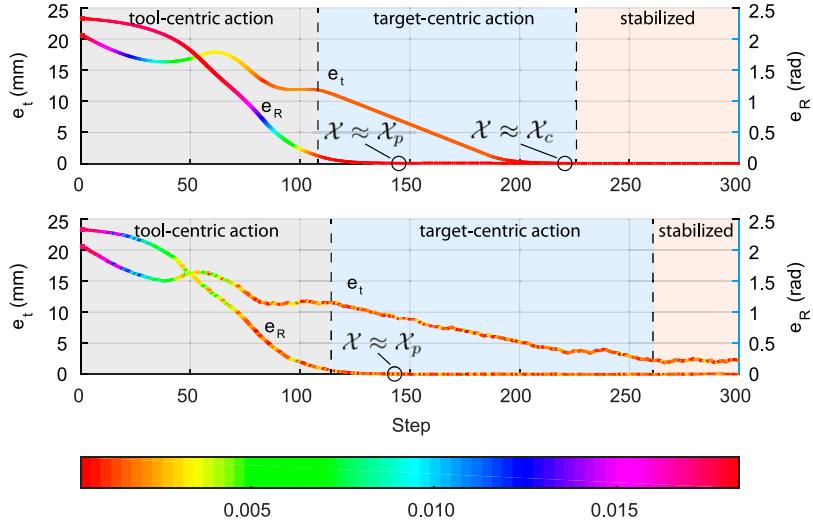


Figure 15. Evolution of two examples of target reaching shows that orientation is alignment in tool-centric action phase and then comes to translation to eliminate the remaining distance to avoid premature collision.

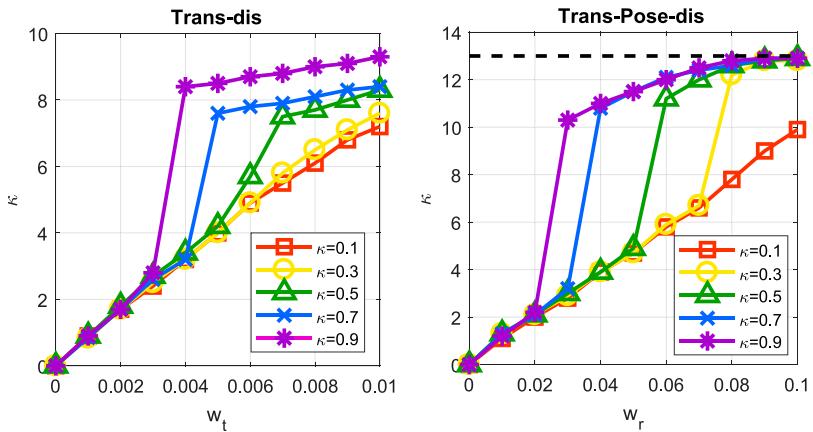


Figure 16. Stabilized end-effector distance from the final contact state \mathcal{X}_c under different tuning gain κ and different types of noises.

needle insertion path is commonly used in MIS to minimize tissue trauma during penetration. The GVP τ becomes the parametric value that guides \mathcal{X}_g along the trajectory. The robot will move to the starting point of the trajectory first and then follow the trajectory as τ elevates to 1. While the robot output meets disturbances, that is, deliberate deviation of the tool from the prescribed path, the value of τ will not move until the controller guides the tool back on where it deviates from (illustrated in Figure 17(c)).

7.6. Summary of simulations

The simulation results demonstrate the motion-level behavior of our robot planning and control framework in different scenes. Overall, the GVM method is validated to be capable to dealing with the following three scenarios:

- Pose transfer through extreme configurations: The planning and control motions remain feasible and

efficient to large-range pose adjustment (>100 average distal joint position differences) near joint limits.

- Stable and reactive target contact: The robot motion is guided via a bidirectional manner, that is, robot adaptively decides whether to reach or to stay clear from the target by evaluating the on-the-fly tool-target configuration.
- Smooth trajectory tracking: The robot is capable of following a predefined path with controller disturbances, where the tracking process automatically halts due to path deviation and will proceed after the deviation is cleared.

Basically, the framework provides a reactive planning and control process which could be used to characterized different types of elementary steps in surgical procedures, and could lead to a “fail-in-safety” performance which highly reacts to the on-the-fly information. This is essential to further improve task-level reliability when automating a

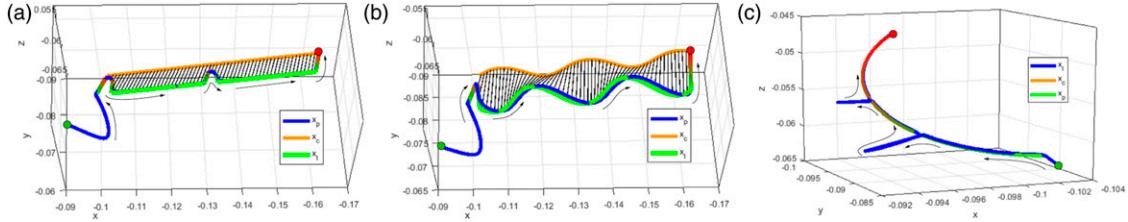


Figure 17. Resultant trajectories of the instrument in all three cases in Simulation Scene III. The orange and blue curve represents the target and the instrument’s end-effector position, respectively. The curve with changing color between green and red is the trajectory of the pre-grasp point, or the position evolution corresponding to \mathcal{X}_p .

sequence of motion steps in order to complete a task successfully. The quantified results will be later addressed in experiments.

8. Experiments and Results

8.1. System setup

In this section, we present results of on-site experiments for the robot to automatically perform simulated surgical tasks. We use a dVRK with two PSMs as the robot platform. Both PSMs are equipped with LNDs, which is a common type of instrument adopted when conducting tissue arrangement and suturing in RAMIS. Both instruments have a 7-DoF joint set with the first four accounting for the RCM motions for minimally invasive setup, the middle two joints for providing wristed motions, and the last DoF for controlling the tool actuation (or the jaw opening angle). The physical setup is shown in [Figure 18](#). We use a Core i7 3.4 GHz user-side PC controller (with 156 GB RAM and without GPU acceleration) which connects to the robot controller with TCP/TP. The cisst/SAW software environment is used and the algorithms run in MATLAB 2020b under the Robot Operating System (ROS). A pair of industrial CMOS cameras is used for online visual feedback data acquisition with 640×480 resolution and 30 frames per second. The base position of the two robots (i.e., the RCM) is set such that their distances to the target to be operated are both around 150 mm, which is commonly adopted in the clinical practice of RAMIS. The camera is fixed between two PSMs, which is a typical setup scale for multi-port RAMIS ([Escobar and Falcone, 2014](#)). The positions of the left camera relative to PSM1 and PSM2 are $[34.5, 70.2, 39.1]^\top$ mm and $[66.9, 104.6, 51.60]^\top$ mm, respectively. The overall control loop of the system is around 20 Hz without software acceleration. The robot end-effector motions are saturated to 0.8 mm per step, which indicates a maximum 16 mm/s linear velocity for the tool.

We define the jaw angle to stay 0 during tool-centric actions because the 7th is not considered in tool-centric actions. While the instrument enters target-centric actions (with $\tau > 0$ and $\dot{\tau} > 0$), the jaw angle smoothly opens to a

predefined angle (30°) and will only close when $\tau \rightarrow 1$. When inverting the trajectory, the above sequence will be toggled accordingly. This applies to any elementary motions which involve Modes II, III, and IV.

Several sensing algorithms are required for providing accurate 3D information to facilitate task completion is introduced in this subsection. The focal lengths of the two cameras are tuned to 30 mm.⁵ The image resolution of both cameras are set to 640×480 which acquire video streams both at 30 frames per second. The stereo camera has been calibrated in advance ([Zhang, 2000](#)) with its backprojection being 0.1582 pixel upon data acquisition of 30 pairs of sample images. The camera-robot transformation has been calibrated using the method based on our previous work ([Zhong et al., 2020](#)) to all the engaged PSMs. The performance of our framework is systematically assessed by respectively performing single-arm tasks like debridement and membrane dissection, a dual-arm task like suturing. Task effectiveness in terms of efficiency, success rate and accuracy will be validated.

8.2. Task I: Debridement

The first scenario simulates the surgical tissue debridement procedure, which is designed to illustrate the performance of our framework to automate repetitive reach-and-grasp motions of the robotic instrument to different individual targets. We fabricate a group of L-shape silicone gel blocks to simulate tissue debris. Each block has a bounded 12 mm cube of size which simulates a plumped piece of tissue with a graspable edge, and is attached by a unique fiducial marker of AprilTag ([Olson, 2011](#)) with 25h9 tag family for 6-DoF target localization and multi-target identification. The target is online detected and tracked during the task. We apply the MoB to be Mode II (refer to [Figure 8](#)) for automating the block reaching with grasping and Mode III for retracting with tissue releasing for debris collection to a plastic plate.

We design a multi-block debridement task to be completed by PSM 1. The blocks are random placed within a rough volume of $50 \times 50 \text{ mm}^2$ table surface, whose orientation are randomly settled as well. As all the blocks are identical in sizes, their poses are online tracked by

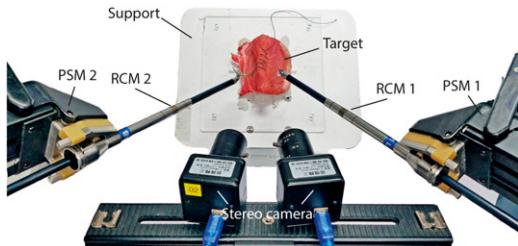


Figure 18. The experimental setup.

using the fiducial markers to generate grasping configurations. Their pick-up sequence is then randomly assigned to avoid adding bias to statistical performance analysis. The releasing pose \mathbf{T}_{rel} of the robot's end-effector for debris collection is assumed constant with

$$\mathbf{T}_{rel} = \begin{bmatrix} -0.924 & -0.222 & 0.310 & 0.117 \\ -0.176 & 0.970 & 0.170 & -0.050 \\ -0.338 & 0.103 & -0.936 & -0.0527 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (38)$$

that is, 138.4 mm distance from the robot base. The tool's opening DoF is controlled for after the current manipulation step is completed.

We have conducted ten trials of six-block debridement and ten trials for four blocks. As the main goal of the task is to clear the blocks from their initial positions, we evaluate the quality of this task by assessing the overall success rate of the whole task and the time distribution for each pick-up step. Figure 19 illustrates the snapshots of two trials of six-block trials. The overall success rate for the six-block and the four-block cases are all 90%, that is, both 9 out of 10. Note that for each case, we rearrange the blocks by random placement. Despite different poses, the instrument manages to collect all blocks smoothly. The algorithm could guide the instrument adjacent to the block first and then finalize the contact reaching to achieve a safe grasp thanks to the GVM scheme. It avoids premature tool-target collision during tool-centric action phase which can prevent the tissue from being dislocated or being damaged unexpectedly. This is important as it is normally difficult to track the instant change of tissue's position shortly before contact due to visual occlusion. We also show in Figure 20 that during the grasping attempt, inaccurate instrument localization might unexpectedly deviate the target from its initial position. Our framework could rapidly react by stating clear from the target first and initiates further adjustment. As long as the target is detectable, the instrument will start another attempt once the tool-target configuration is desirable again. Figure 21 shows the average execution time of each motion step during the task. The median duration for collecting each block is 12.8/11.9 s for four/six block

trials. The four-block cases take longer time to complete one debris collection, as the more scattered placement of the blocks leads to longer instrument travel distances. The random block setting does not significant affect the completion time for each attempt, which shows good consistency.

8.3. Task II: Tissue membrane dissection

We next set up an experiment to simulate the dissection task to a tissue membrane, which is a typically demanded in surgeries involving organ removal like laparoscopic cholecystectomy (Reynolds Jr, 2001). The task includes highly repetitive cutting sub-steps in order to dissect the superficial layer of the tissue via a prescribed trajectory. We use a soft handkerchief paper with average thickness of 1 mm as a phantom tissue layer. The layer is fixed by a static support which keeps the tension of the surface to simulate the connectivity of the target to surrounding tissues and to reduce tissue deformation. PSM2 is used to perform the task using the Potts Scissors instrument to provide cutting shear force to the target. The tissue is of size 90 mm * 80 mm, and the dissection path on the tissue is marked by a red line that goes across the layer surface. We set two types of path, one starts from the short edge and ends at the opposite one (with total length of 90 mm), and the other spans diagonally to the midpoint of the long edge (with totally length of around 60 mm). The 3D positions of the path endpoints are detected first using the stereo camera and then are used to interpolate the path profile by comparing the backprojection error with respect to the observed one. Finally, and the interval distance d_c to between consecutive cutting steps is computed to determine the cutting points. For empirical experiences, we select such distance to be $d_c = 3$ mm, which indicates a 30-time cutting motion sequence for a 90 mm path and 20 cuts for 60 mm path.

We have conducted five consecutive trials for this task. Three trials are set We use Mode III to perform the dissection procedure (could be also solved by Mode IV but slightly slower) and the clamping motion of the scissors is done after settling the manipulation. The setting of d_c is set smaller than the length of the scissors (around 9 mm) allows a backward shifting before proceeding to the next step to avoid the tissue-tool adhesivity to after each cutting. During the procedure, the instrument automatically plans and dissects the phantom layer progressively through repetitive cutting steps and separates the layer in half via the labelled paths. Figure 22 shows the snapshots of the recorded clips of the real-time automated dissection process using a single instrument. The robot successfully completes four out of five trials with the overall success rate thus being 80%. The failure case is caused by tool-tissue entanglement with uncontrollable path deviation. The average duration for executing the whole task is 81.1 ±

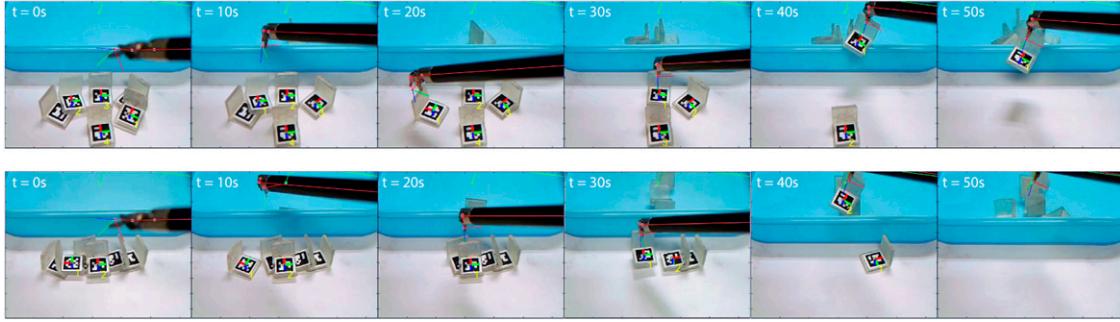


Figure 19. Frames of the simulated debridement task using a single robotic instrument. The red and green robot skeleton represents the backprojected layout of the instrument's current and tissue releasing configuration on the camera image. The Cartesian frame of the instrument's end-effector (i.e., the tool) and the blocks are also visualized.

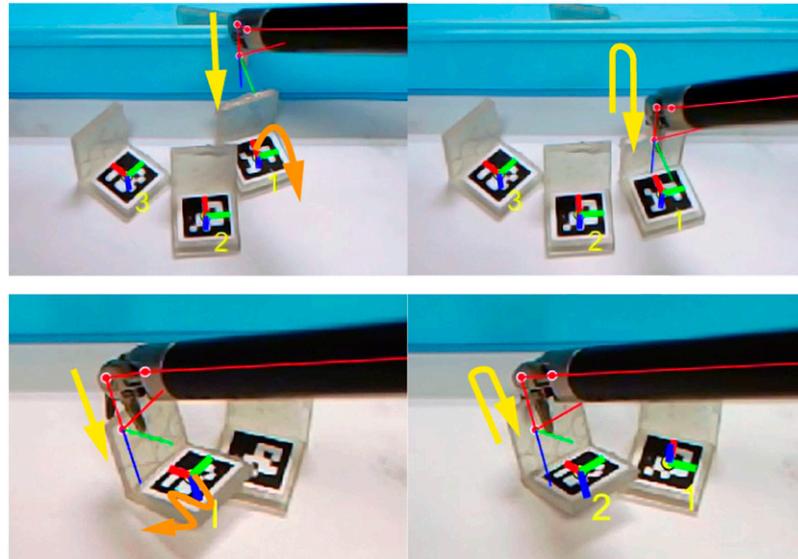


Figure 20. Target uncertainty during task execution, where the block is unexpectedly deviated to a new position (deviated motion marked in orange arrows) due to premature tool-target collisions. The planned/re-planned tool position is conceptually denoted as yellow arrows.

1.2 s and 54.3 ± 0.9 s for the 90 mm case and 50 mm case, respectively. The time cost for performing each single cutting step is 3.9 ± 0.3 s. The dissected layout of the phantom tissue layers are shown in Figure 23. Upon manual measurement, the maximum lateral deviation from the prescribed linear path is around 1.5 mm for both 50 mm and 90 mm case (we assume the highest resolution for measurement is 0.5 mm as the smallest scale of the used ruler is 1 mm).

8.4. Task III: Wound suturing

We finally conduct autonomous dual-arm multi-throw wound suturing as a comprehensive evaluation of our framework to perform complicated tasks. The task appears in most surgical interventions, where a suturing needle is guided through wound edges for tissue approximation and wound closure. The guideline is characterized by our

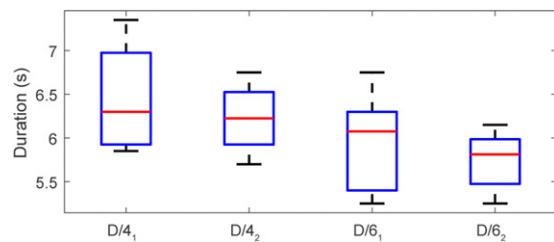


Figure 21. Motion step durations during debridement, where the number 4/6 indicates the number of fragments, and the subscript 1/2 denotes the reaching/collecting motion of the fragment, respectively.

framework in Figure 8 into seven individual steps for each throw.

8.4.1. Scenario setting. To construct the setup, we use two PSMs which are both mounted with LNDs as the surgical

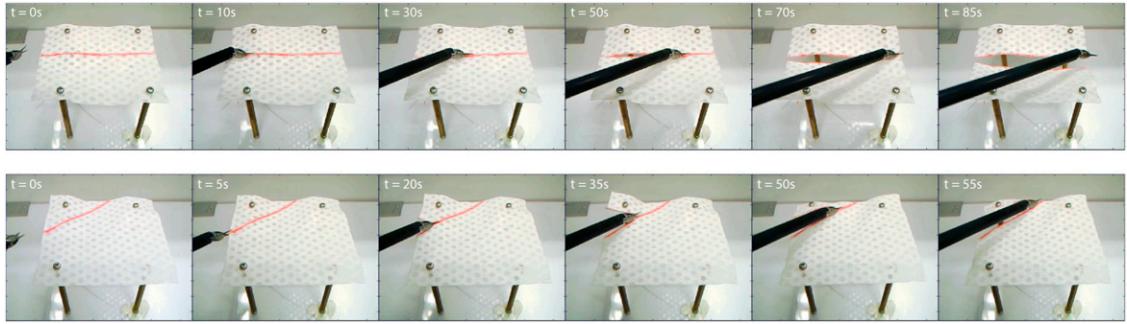


Figure 22. Frames of automated phantom layer dissection procedures. The first and second rows indicate the 90 mm case and 60 mm case, respectively, through paths with different directions.

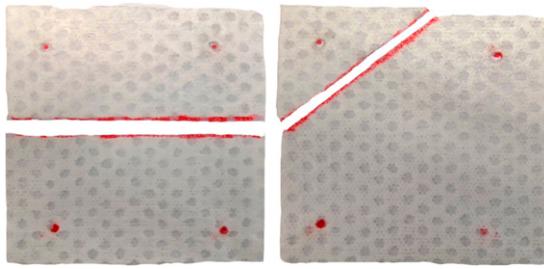


Figure 23. Demonstration so the dissected phantom layers placed on the table with the prescribed cutting paths in red lines, and the fixation points to the support marked in red dots.

tools. To regulate the in-hand needle grasping position, we adopt the work in Sen et al. (2016) that uses 3D-printed PVS needle holder mounted along the jaws to enhance needle manipulation accuracy. The stereo camera is placed between the two distal tools as a three-point invasion in typical laparoscopy (Horgan and Vanuno, 2001).

Two types of tissues are used in our experiments. The first type is the artificial soft tissue that consists of two layers, the outer layer to simulate tissue skin and the inner layer for the dermal structure. Both layers are made from synthetic gel. We also prepare a piece of porcine tissue by using two separate parts to create a lumped wound for ease of selecting needle insertion orientation. The tissues are both fixed on a support to prevent unnecessary movements. The positions of the wound edges are computed from the user-input wound endpoints via stereo images. In addition, we introduce two parameters: the stitching width d_w that denotes the distance between the needle's entry and exit point on the tissue, and the throw distance d_t that denotes the (ideal) parallel distance of the suture between consecutive throws. Without loss of generality, we set $d_w = 10$ mm and $d_t = 5$ mm as a resembling scale setting in standard robot-assisted surgical training procedures (Garcia-Ruiz et al., 1998).

8.4.2. Single-throw suturing. We first show the capability of our framework to complete a single-throw pipeline in

suturing. **Figure 24** shows the snapshots of individual motion steps of a single-throw pipeline on the artificial tissue. The needle's insertion and exiting point on the tissue is directly assigned by user input that yields the preset d_w and d_t . The instruments start from preset idle configurations which are cleared from the target tissue. The manipulation in each step is converged to the goal configuration with $\|\eta\|$ down to a scale ϵ_g before proceeding to the next step. To perform target-centric action to the needle during targeting, the tool-needle transformation is computed using image-based needle pose detection at the start of a throw. During needle hand-off, two instruments work in a decentralized manner, where PSM2 guides the needle to a precomputed hand-off pose on top of the wound, with PSM1 automatically track the motion of PSM2 to decide a proper hand-off configuration subject to GVM scheme. An example is particularly shown in **Figure 25** that demonstrates the decentralized needle hand-off process by human–robot collaboration. The complete parameter settings for this scenario are shown in **Table 3**. Note that the pipeline in a single throw covers target contact (e.g., grasping the needle), trajectory tracking (e.g., needle insertion/existing), and dual-arm coordination (e.g., needle hand-off). Due to the theoretical proofs of trajectory feasibility and motion stability, all the elementary motions are characterized and automated, involving target contact (including needle insertion, regrasping, and hand-off) while the path avoids premature collisions to the target tissue (refer to our [supplementary video](#) for demonstration).

8.4.3. Multi-throw suturing. We define a five-throw suturing task for artificial tissue and a four-throw suturing task on porcine tissue to comprehensively evaluate the task-level applicability of our framework. The suturing will be performed using single-continuous suture pattern, which is a common and efficient type of suturing technique in RAMIS. We do not include knot tying at the end of suturing due to the complexity of detecting the suture's 3D topology, which is not the main goal of this work. The needle's insertion/exiting positions are

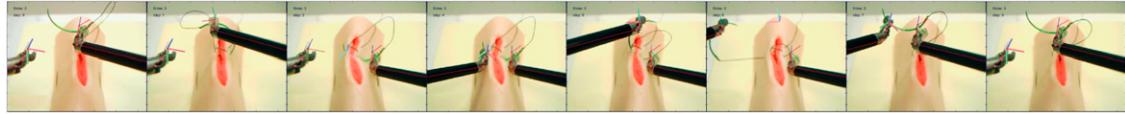


Figure 24. Step illustration of a single-throw suturing during task automation. From left to right: 1) idle (the start of a throw); 2) needle (pre-entry) targeting; 3) needle insertion; 4) needle regrasping; 5) needle exiting; 6) suture pulling; 7) needle hand-off; and 8) resetting (for starting the next throw).

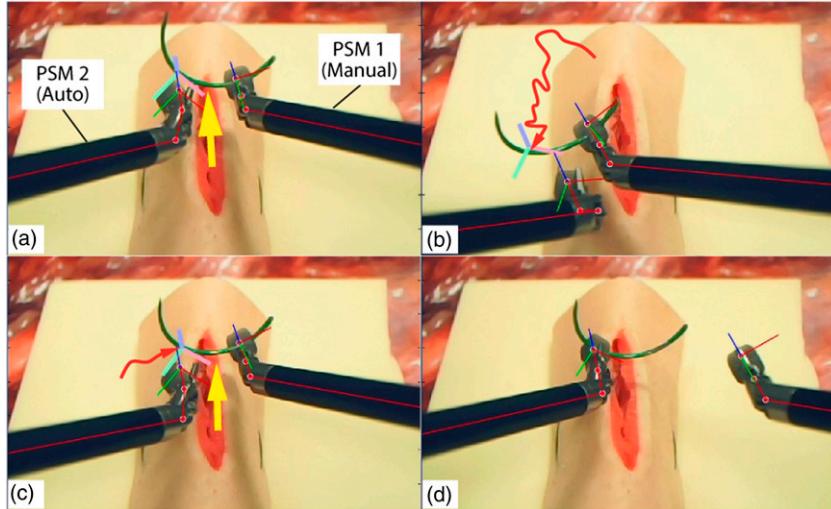


Figure 25. When PSM1 is manually controlled by a user (with the shaky trajectory in red arrows), PSM2 stably tracks and makes contact attempt (in yellow arrow) when the tool-target alignment becomes desirable.

Table 3. Task-relevant parameters of our framework for multi-throw suturing.

	Pos. of idled PSM1/2	Pos. of first insertion/exiting point	Pos. of last insertion/exiting point	Pos. of wound endpoints
Artificial tissue	$[0.131 \ -0.067 \ -0.054]^T$ $[-0.091 \ -0.012 \ -0.045]^T$	$[0.117 \ -0.066 \ -0.096]^T$ $[-0.133 \ -0.044 \ -0.091]^T$	$[0.111 \ -0.053 \ -0.090]^T$ $[-0.122 \ -0.032 \ -0.095]^T$	$[0.130 \ -0.066 \ -0.093]^T$ $[0.116 \ -0.007 \ -0.060]^T$
Porcine tissue	$[0.116 \ -0.051 \ -0.055]^T$ $[-0.099 \ -0.044 \ -0.091]^T$	$[0.117 \ -0.074 \ -0.086]^T$ $[-0.140 \ -0.042 \ -0.079]^T$	$[0.103 \ -0.053 \ -0.082]^T$ $[-0.123 \ -0.020 \ -0.078]^T$	$[0.131 \ -0.071 \ -0.083]^T$ $[0.119 \ -0.045 \ -0.081]^T$

*All positions (unit: m) are calculated relative to the respective robot base except the wound endpoints positions (from the camera).

computed based on two virtual 3D markers (named as the “wound positioners”) manually input by the users via stereo images prior to the start of the task. The needle insertion orientation is computable once the insertion and exiting point is known (Pedram et al., 2017). Note that we select d_s as 8 mm for the porcine tissue to evaluate task performance under different settings. The moving distance of the suture pulling step is computed based on its consumed length of suture in each throw, which is generically given as 14 mm measured from teleoperation based on the assigned d_w and d_t . Trials on each type of tissue are conducted continuously to maintain identical setting throughout the experiments.

We perform ten trials for each type of tissue, respectively, under identical setting. The task-level assessment include task success rate to evaluate reliability, duration of a single throw and the entire task for in-process performance, and resultant suture pattern accuracy for task quality. Figure 26 illustrates the frames of task execution process on both the artificial and porcine tissue viewing from the sensing (left) camera. The task is defined as “successful” only if the suture passes through both edges of the wound for enough times without user interference. The overall success rate of performing five-throw suturing is 80% on the artificial tissue (i.e., 8/10) and 70% on the porcine tissue (i.e., 7/10). The failure cases are either subject to needle-suture

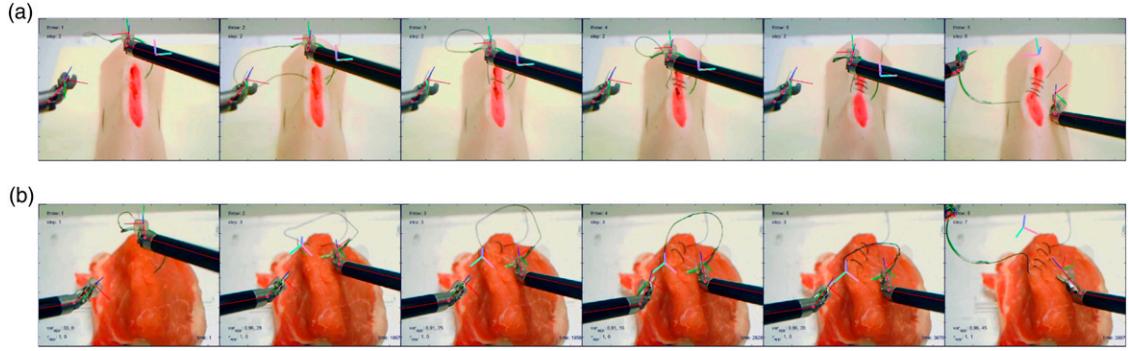


Figure 26. Frames of dual-arm wound suturing using our framework on the artificial tissue (first row) and the porcine tissue (second row). Each image is captured at the same moment in the step (during needle targeting in the first row and during needle insertion in the next row).

entanglement or missing the exiting point during insertion. It could be further elevated to 90% (i.e., 9/10) for both the artificial tissue and porcine tissue suturing, respectively, if manual suture arrangement is provided during the task. Note that the above data is calculated by directly regarding a task as a failed trial once any step failure occurs during the process. For step-wise failure, we manually halt the task once a failed step occurs and rearrange the needle/suture accordingly, and then let the robot proceed to the rest of the task to compute the accumulative number of step failure. In this case, a total of 5 motion failures during 10 trials of five-throw suturing (or among the totally 640 motion steps), which indicates 0.8% of motion failure.

Suturing the porcine tissue owns lower success rate, as the tissue surface might encounter irreversible deformation after many stitches. The trajectories of the two instruments' distal tool for completing such task are shown in Figure 27. It can be seen that the in-process trajectories maintain good consistency among individual throws without unpredictable movements. Figure 28 demonstrates the suturing accuracy on the artificial tissue.⁶ Figure 29 shows the time cost for completing individual throws in the task. The average time of performing one throw is 48.3 ± 1.4 s and 45.1 ± 1.0 s for artificial/porcine tissue, with the time variation percentage for being only 2.9% and 2.2%, respectively, which shows good step-level consistency. The average time cost for completing each throw gradually decreases (shown in Figure 29), as the decreased residual suture length lowers the travel distance for suture pulling. Furthermore, we characterize the outcome quality by manually measuring the deviation of the insertion/exiting position accuracy. The total average accuracy is 1.7 ± 1.1 mm, with the maximum 2.0 mm of wound deviation as the number of throws increases due to the irreversible deformation after several trials. The deformation remains limited influence on suturing accuracy, and is regarded tolerable unless it causes misplacement of the needle during insertion (which is directly regarded as failure).

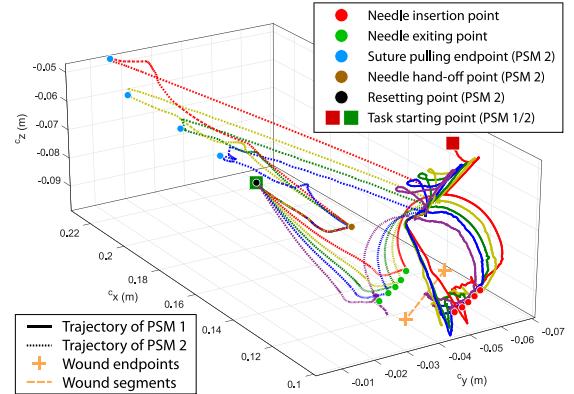


Figure 27. Trajectories of the two instruments during the five-throw suturing task, where the red/yellow/green/blue/purple indicates the trajectories from 1st to 5th throw, respectively.

8.5. Overall assessment

In this subsection, we validate the task-level performance of our framework by comparing with existing works reporting either manual task operations by novice surgeons, or state-of-the-art automation approaches. We focus on aspects including the time cost, the step/task failure, and number of individual trials in each task which are all metrics that reflect the task execution quality and consistency during performance assessment (Martin et al., 1997). The existing works which aim to automate the identical tasks (e.g., continuous debridement and suturing) using the standard guidelines with articulated surgical robots will be especially selected and compared to ours as side-by-side analysis. Note that we are unable to provide identical task-relevant parameter settings for performance comparison, as many works do not provide detailed parameter settings. The comparison will be conducted based on their best reported results as well.

We first summarize the duration performance from the above experiments into time cost for completing each

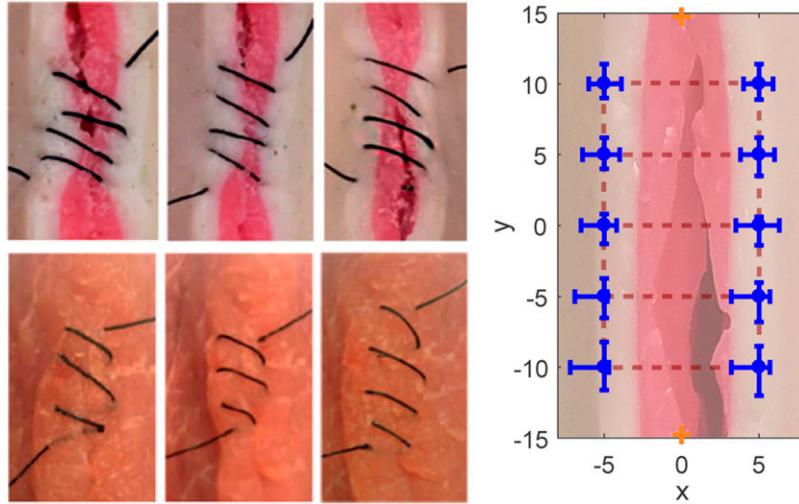


Figure 28. The suture pattern on along the tissue wound edges after five-throw suturing (left-upper three: results on artificial tissue; left-bottom three: results on porcine tissue), and the manually measured suturing accuracy on the insertion/exiting positions (the right figure).

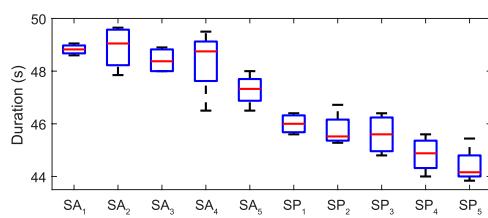


Figure 29. Duration required for each throw for suturing an artificial tissue and a porcine tissue, respectively. The gradual descent appears via consecutive throws in attributed to the shorter residual length of suture and thus shorter suture pulling distance under saturated tool velocity.

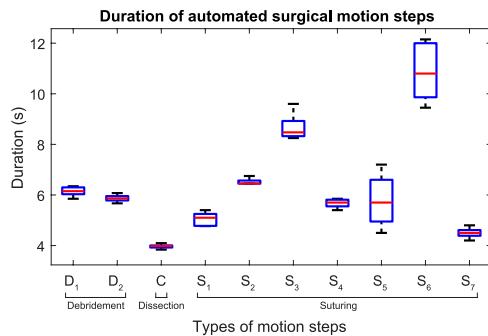


Figure 30. Comparison of duration of individual motion steps in different surgical tasks (D indicates debridement task, C indicates dissection task, S indicates suturing task).

task-relevant motion step (results shown in Figure 30). The time cost for completing the different task-relevant motions under different setups are shown in Figure 31. We calculate the variation percentage of the duration for each step in terms of the average time cost is 3.3%/4.5%

for the 4-/6-block debridement, 3.7%/3.7% for the 5/8 mm dissection procedure, and 1.8%/1.4% for the four-/five-throw suturing on the artificial/porcine tissue, respectively. They indicate good consistency of robot's planning and control performance under different setups. As there are current no existing works that automate different surgical tasks, the performance could be used as a benchmark for future works to refer to for multi-task automation.

We have also listed results among different works in terms of setups, the time cost, and the accuracy (if available), where the results are shown in Table 4. For comparing task duration, we post-process the data of each work by computing the average time to execute one single step for each task (as shown in the data in blue font). Our framework takes averagely 6.2 ± 0.3 s to finish collecting a single fragment in debridement, which is slightly faster than many of the results using manual operations by novices, but a lot greater in time consistency. Regarding suturing, our framework takes averagely 46.7 ± 1.3 s for a single-throw suturing throw on the assigned tissue surface (maximum tool velocity currently set to 16 mm/s), which is significantly more efficient (mostly 1.5-2 times faster) than manual operations, and the time variation is more than an order of magnitude smaller. Note that we are also the first to conduct five-throw suturing on phantom tissue and also the first to use wristed robotic instruments for multi-throw suturing on porcine tissue. The performance is mainly attributed to a close-form solution for integrated motion planning and control framework, which owns looping time of < 1 ms without software acceleration, and does not involve data-driven knowledge or iterative computations. Although the works appeared in Table 4 own different setups and the performance might not be compared with identical scenarios, we show

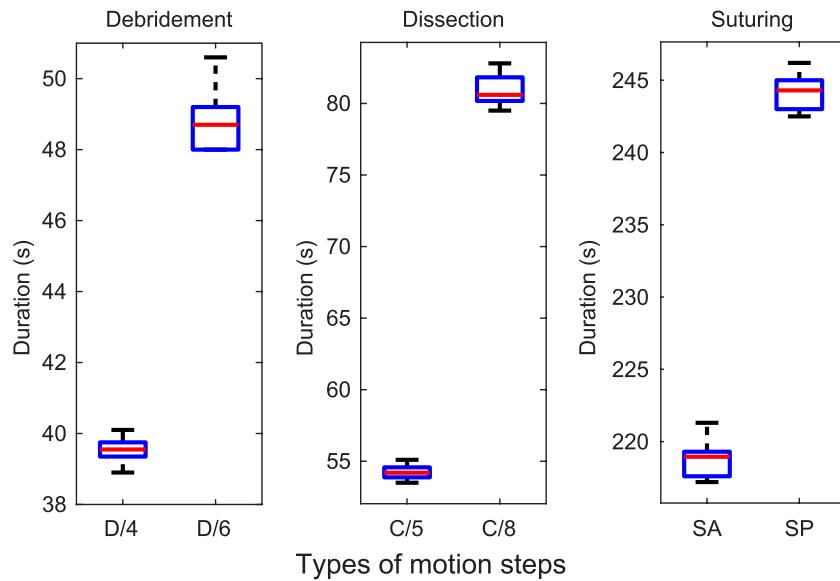


Figure 31. Time consistency of completing task-relevant motions with different settings (D/4 and D/6 denote 4-/6-block debridement, C/5 and C/5 denote 50-/80-mm dissection path length, SA and SP denote suturing on artificial/porcine tissue, respectively).

Table 4. Comparison of manual/robot-assisted/autonomous tasks execution in physical environments. The time cost in blue font indicates the computed equivalent time cost for each elementary procedure in a task. Failure rate (column "Fail") is computed with average number of failures per task and/or total accumulative failure percentage (N.A. indicates that the data is not given or uninterpretable).

Works	Tasks							
	Target	Debridement			Suturing			
		Qty.	Time cost (s)	Fail	Setting	Qty.	Time cost (s)	
Garcia-Ruiz et al. (1998)	Bead	10	M 68±32 (6.8) RZ 183±42 (18.3)	5% 3%	Needle 2-O Width 10 mm	4	M 154±40 (38.5) RZ 605±57 (151.25)	8%
Dakin and Gagner (2003)	Peanut	10	M 53.5 (5.35) RZ 128.5 (12.85) RD 61.0 (6.1)	2.22 (22.2%) 3.11 (31.1%) 2.44 (24.4%)	Needle 4-O Width 5 mm	4	M 172.0 (43.0) RZ 426.1 (106.5) RD 236.2 (59.1)	3 2.56 1.78
Hubens et al. (2003)	Ring	1	M 19.0±5.7 RD 9.5±8.7	2.25 0 *median	Needle 4-O Width 5 mm	4	M 356.4±112.6 (89.1) RD 60.5±39.2 (15.2)	38 2
Shah et al. (2009)			N.A.		Needle 20 mm Width 5 mm	4	RZ 258±93 (64.5)	53
Fard et al. (2018)			N.A.		Needle 4-O Width 5 mm	4	RD 126.84±55.86 (31.7)	N.A.
Kehoe et al. (2014)	Foam	1	M 29.0 AR 91.8	5.0% 8.7%			N.A.	
Mahler et al. (2014)	Foam	1	AR 15.8	N.A.			N.A.	
Seita et al. (2018)	Seed	8	AD 57.62 (7.20)	10 (8.3%)			N.A.	
Leonard et al. (2014)			N.A.		Needle 3-O Width 10 mm	9	M 560.4±358.6 (62.3) RD 342.6±226.2 (38.1) A* 64.51±0.81 (7.2)	N.A.
Sen et al. (2016)			N.A.		Needle 39 mm Width 5 mm	4	M 136.85 (34.2) AD 383.00 (95.8)	50%
Pedram et al. (2020)			N.A.		Needle 30.55 mm Width 16 mm	1	AR ~200	N.A.
Ours	Syn-Gel	6	AD 48.8±1.3 (6.2±0.3)	0.1 1.7%	Needle 34 mm Width 10 mm	5	AD 243.5±1.9 (46.7±1.3)	0.5 1.4%

Notation of task setting in the “time cost” columns.

First letter: 1) M: manual execution with regular instruments; 2) R: robot-assisted surgeon-centered execution; 3) A: autonomy.
Second letter: 1) Z: Zeus robot; 2) D: dVSS/dVRK; 3) R: Raven Robot; 4) *: other types of robots.

the capability of our work to be a baseline for generic surgery autonomy.

9. Discussions and Conclusions

In this paper, we have presented a generic planning and control framework to automate different types of surgical tasks. The framework systematically addressed both surgery-specific motion safety and task-level characterization, which has been weakly explored. The DS-based controller with SMA globally guarantees the step-level motion stability and trajectory reachability via Lyapunov stability using NSS as system states. The model is efficient, differentiable without iterations, and can solve constrained trajectories without high-dimensional workspace analysis. Meanwhile, the GVM provides adaptive re-planning that guides the robot through constrained path to contact target properly based on on-the-fly situations. The framework was then extended to construct complete pipelines of various surgical tasks including single-instrument tasks like debridement, tissue dissection, and dual-instrument tasks like suturing.

The proposed framework has been validated through simulations for performance study, and then through experimental study on three different surgical tasks (but not limited to). The performance comparison shows that our results own good task efficiency, reliability, and performance consistency that outperform manual operations by novice surgeons and state-of-the-art automation algorithms. Despite the absence of detailed parameter settings provided by existing works, our still represents the farthest step of our framework to generically handle continuous multi-step suturing and debridement. The framework could be used when the camera is moving, which commonly appears in surgery, as long as the RCMs of the endoscopic camera and the instrument is not changed. It could potentially be adopted to larger domain of robot-enabled surgical tasks that require more complicated end-effector motions with wristed profiles.

We also address several limitations in this work. We do not actively regulate the online tissue deformation induced by the task motions. This is reported in our suturing experiments that might result in failure of the task due to the inaccurate needle insertion. This could be solved by integrating tension information to formulate the tissue's dynamic change. One could either directly measure the tool-tissue interaction force (requiring force sensors which could not be achieved by dVRK), or to adopt vision-based tracking of tissue deformation using trackable surface features (mostly individual points). The deformation relative to its resting state could be online monitored to estimate the (relative) tension. Excessive deviations could be modelled as repellent factors and avoided by intermediate poses in goal-varying manipulation. Meanwhile, the camera-robot transformation is pre-calculated using efficient autonomous calibration approach in [Zhong et al. \(2020\)](#). It is not updated during the task, where the

leading positioning error of the instrument (even though not significant) is not online refined, but currently being capable of obtaining good task success rate. Once it could be updated to minimize the residual positioning error, the motion accuracy apart from the planning and control framework could be further improved, which is currently our ongoing work.

There are several directions extendable from this work. The task-level performance and assessment results could be used as a benchmark to be addressed by future approaches under such setting. We are also developing learning-based perception algorithms to localize target surgical areas such that our framework could be applied to complex surgical environment that targets clinical applications. Moreover, we would like to explore active tissue deformation regulation during task automation to further improve task-level reliability under dynamic/uncertain physical situations.

Acknowledgements

We thank Mr. Dickson Chun Fung Li and Ms. Yaqing Wang for their assistance in the setup of physical experiments.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work is supported in part by the Shenzhen Portion of Shenzhen-Hong Kong Science and Technology Innovation Cooperation Zone under HZQB-KCZYB-20200089, in part of the HK RGC under T42-409/18-R and 14202918, in part by the Multi-Scale Medical Robotics Centre, InnoHK, and in part by the VC Fund 4930745 of the CUHK T Stone Robotics Institute.

ORCID iDs

Fangxun Zhong  <https://orcid.org/0000-0002-1151-1995>

Yun-Hui Liu  <https://orcid.org/0000-0002-3625-6679>

Supplemental Material

Supplemental material for this article is available online.

Notes

1. Examples of robotized surgical instruments that adopt the design of the wristed structure include da Vinci Xi (Intuitive Surgical), Versius (CMR Surgical), Micro Hand S (Weigao Group Medical), and RAVEN (Applied Dexterity).
2. While satisfying (7), for sake of simplicity, we define the node vectors and the end-effector frame axes (using DHc) such that $\mathbf{H} = \mathbf{I}_{3 \times 3}$ and $k_s = k_w = k_t = 1$.
3. The approximation is ensured by an additional assumption that s must be continuously and stably guided to \mathbf{x}_{s_g} . We will shortly provide a DS that regulates s to satisfy such performance.

4. Such setting of λ could guarantee a sequenced actuation process using SMA especially when the instrument experiences weak manipulability and/or extreme joint positions which could result in drastic joint control input.
5. The Karl Storz 1 S stereo laparoscope owns a focal length of 15–200 mm
6. We only complete the suturing accuracy measurement to the trials targeting artificial tissue, as for the porcine tissue, it becomes difficult to find visual reference on its surface for measurement after the task.
7. This assumption is made by assuming that the lateral moving range of $\mathbf{x}_s(\mathbf{q})$ is much smaller than $\|\mathbf{x}_s(\mathbf{q})\|$, which is reasonable in the confined intra-corporeal surgical workspace, as it tolerates much smaller lateral end-effector motions compared to longitudinal motions.

References

- Adhami L and Coste-Maniere E (2003) Optimal planning for minimally invasive surgical robots. *IEEE Transactions on Robotics and Automation* 19(5): 854–863.
- Agustinos A, Wolf R, Long J-A, et al. (2014) Visual servoing of a robotic endoscope holder based on surgical instrument tracking. In: *5th IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechatronics*. Piscataway, NJ: IEEE, pp. 13–18.
- Amato NM and Wu Y (1996) A randomized roadmap method for path and manipulation planning. *Proceedings of IEEE International Conference on Robotics and Automation* 1: 113–120.
- Berenson D, Srinivasa SS, Ferguson D, et al. (2009) Manipulation planning on constraint manifolds. In: *2009 IEEE International Conference on Robotics and Automation*. Piscataway, NJ: IEEE, pp. 625–632.
- Billard A and Kragic D (2019) Trends and challenges in robot manipulation. *Science* 364(6446): eaat8414.
- Cao C, MacKenzie CL and Payandeh S (1996) Task and motion analyses in endoscopic surgery. In: *Proceedings ASME Dynamic Systems and Control Division*. New York: ASME, pp. 583–590.
- Chiu Z-Y, Richter F, Funk EK, et al. (2021) Bimanual regrasping for suture needles using reinforcement learning for rapid motion planning. In: *2021 IEEE International Conference on Robotics and Automation (ICRA)*. Piscataway, NJ: IEEE, pp. 7737–7743.
- Dakin G and Gagner M (2003) Comparison of laparoscopic skills performance between standard instruments and two surgical robotic systems. *Surgical Endoscopy* 17(4): 574–579.
- Davies BL, Hibberd R, Ng W, et al. (1991) A surgeon robot for prostatectomies. In: *Fifth International Conference on Advanced Robotics' Robots in Unstructured Environments*. Piscataway, NJ: IEEE, pp. 871–875.
- D'Ettorre C, Dwyer G, Du X, et al. (2018) Automated pick-up of suturing needles for robotic surgical assistance. In: *2018 IEEE International Conference on Robotics and Automation (ICRA)*. Piscataway, NJ: IEEE, pp. 1370–1377.
- Escobar P and Falcone T (2014) *Atlas of Single-Port, Laparoscopic, and Robotic Surgery*. Nueva York: Springer.
- Fang K, Zhu Y, Garg A, et al. (2020) Learning task-oriented grasping for tool manipulation from simulated self-supervision. *The International Journal of Robotics Research* 39(2–3): 202–216.
- Fard MJ, Ameri S, Darin Ellis R, et al. (2018) Automated robot-assisted surgical skill evaluation: Predictive analytics approach. *The International Journal of Medical Robotics and Computer Assisted Surgery* 14(1): e1850.
- Ferbach P and Barraquand J (1997) A method of progressive constraints for manipulation planning. *IEEE Transactions on Robotics and Automation* 13(4): 473–485.
- Fontanelli GA, Selvaggio M, Ferro M, et al. (2018) A v-rep simulator for the da vinci research kit robotic platform. In: *2018 7th IEEE International Conference on Biomedical Robotics and Biomechatronics (Biorob)*. Piscataway, NJ: IEEE, pp. 1056–1061.
- From PJ (2013) On the Kinematics of Robotic-assisted Minimally Invasive Surgery. *Modeling, Identification and Control: A Norwegian Research Bulletin* 34(2): 69–82.
- Funda J, Taylor RH, Eldridge B, et al. (1996) Constrained cartesian motion control for teleoperated surgical robots. *IEEE Transactions on Robotics and Automation* 12(3): 453–465.
- Garcia-Ruiz A, Gagner M, Miller JH, et al. (1998) Manual vs robotically assisted laparoscopic surgery in the performance of basic manipulation and suturing tasks. *Archives of Surgery* 133(9): 957–961.
- Haidegger T (2019) Autonomy for surgical robots: Concepts and paradigms. *IEEE Transactions on Medical Robotics and Bionics* 1(2): 65–76.
- Horgan S and Vanuno D (2001) Robots in laparoscopic surgery. *Journal of Laparoendoscopic & Advanced Surgical Techniques. Part A* 11(6): 415–419.
- Hu D, Gong Y, Hannaford B, et al. (2015) Semi-autonomous simulated brain tumor ablation with ravenii surgical robot using behavior tree. In: *2015 IEEE International Conference on Robotics and Automation (ICRA)*. Piscataway, NJ: IEEE, pp. 3868–3875.
- Hubens G, Coveliers H, Balliu L, et al. (2003) A performance study comparing manual and robotically assisted laparoscopic surgery using the da vinci system. *Surgical Endoscopy* 17(10): 1595–1599.
- Hwang M, Seita D, Thananjeyan B, et al. (2020) Applying depth-sensing to automated surgical manipulation with a da vinci robot. In: *2020 International Symposium on Medical Robotics (ISMR)*. Piscataway, NJ: IEEE, pp. 22–29.
- Hwang YK and Ahuja N (1992) A potential field approach to path planning. *IEEE Transactions on Robotics and Automation* 8(1): 23–32.
- Jun S-K, Narayanan MS, Agarwal P, et al. (2012) Robotic minimally invasive surgical skill assessment based on automated video-analysis motion studies. In: *2012 4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)*. Piscataway, NJ: IEEE, pp. 25–31.
- Kapoor A and Taylor RH (2008) A constrained optimization approach to virtual fixtures for multi-handed tasks. In:

- 2008 IEEE International Conference on Robotics and Automation. Piscataway, NJ: IEEE, pp. 3401–3406.
- Kehoe B, Kahn G, Mahler J, et al. (2014) Autonomous multilateral debridement with the raven surgical robot. In: 2014 IEEE International Conference on Robotics and Automation (ICRA). Piscataway, NJ: IEEE, pp. 1432–1439.
- Krupa A, Gangloff J, Doignon C, et al. (2003) Autonomous 3-d positioning of surgical instruments in robotized laparoscopic surgery using visual servoing. *IEEE Transactions on Robotics and Automation* 19(5): 842–853.
- Kwoh YS, Hou J, Jonckheere EA, et al. (1988) A robot with improved absolute positioning accuracy for ct guided stereotactic brain surgery. *IEEE Transactions on Bio-Medical Engineering* 35(2): 153–160.
- Leonard S, Wu KL, Kim Y, et al. (2014) Smart tissue anastomosis robot (star): A vision-guided robotics system for laparoscopic suturing. *IEEE Transactions on Bio-Medical Engineering* 61(4): 1305–1317.
- Li T-Y and Latombe J-C (1997) On-line manipulation planning for two robot arms in a dynamic environment. *The International Journal of Robotics Research* 16(2): 144–167.
- Li Y and Chen Y (2015) The ultimate hyper redundant robotic arm based on omnidirectional joints. In: 2015 IEEE International Conference on Mechatronics and Automation (ICMA). Piscataway, NJ: IEEE, pp. 1840–1845.
- Li Y, Richter F, Lu J, et al. (2020) Super: A surgical perception framework for endoscopic tissue manipulation with surgical robotics. *IEEE Robotics and Automation Letters* 5(2): 2294–2301.
- Loukas C (2018) Video content analysis of surgical procedures. *Surgical Endoscopy* 32(2): 553–568.
- Mahler J, Krishnan S, Laskey M, et al. (2014) Learning accurate kinematic control of cable-driven surgical robots using data cleaning and gaussian process regression. In: 2014 IEEE international conference on automation science and engineering (CASE). Piscataway, NJ: IEEE, pp. 532–539.
- Marinho MM, Adorno BV, Harada K, et al. (2019) Dynamic active constraints for surgical robots using vector-field inequalities. *IEEE Transactions on Robotics* 35(5): 1166–1185.
- Martin J, Regehr G, Reznick R, et al. (1997) Objective structured assessment of technical skill (osats) for surgical residents. *The British Journal of Surgery* 84(2): 273–278.
- McKinley S, Sen S, Garg A, et al. (2015) Autonomous tumor localization and extraction: Palpation, incision, debridement and adhesive closure with the da vinci research kit. In: Hamlyn Surgical Robotics Conference, London. London: Imperial College, p. 27.
- Mikada T, Kanno T, Kawase T, et al. (2020) Suturing support by human cooperative robot control using deep learning. *IEEE Access* 8: 167739–167746.
- Mirrazavi Salehian SS, Figueroa N and Billard A (2018) A unified framework for coordinated multi-arm motion planning. *The International Journal of Robotics Research* 37(10): 1205–1232.
- Murali A, Sen S, Kehoe B, et al. (2015) Learning by observation for surgical subtasks: Multilateral cutting of 3d viscoelastic and 2d orthotropic tissue phantoms. In: 2015 IEEE International Conference on Robotics and Automation (ICRA). Piscataway, NJ: IEEE, pp. 1202–1209.
- Nageotte F, Zanne P, De Mathelin M, et al. (2005) A circular needle path planning method for suturing in laparoscopic surgery. In: Proceedings of the 2005 IEEE International Conference on Robotics and Automation. Piscataway, NJ: IEEE, pp. 514–519.
- Nageotte F, Zanne P, Doignon C, et al. (2006) Visual servoing-based endoscopic path following for robot-assisted laparoscopic surgery. In: 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway, NJ: IEEE, pp. 2364–2369.
- Nageotte F, Zanne P, Doignon C, et al. (2009) Stitching planning in laparoscopic surgery: Towards robot-assisted suturing. *The International Journal of Robotics Research* 28(10): 1303–1321.
- Nagy DÁ, Nagy TD, Elek R, et al. (2018) Ontology-based surgical subtask automation, automating blunt dissection. *Journal of Medical Robotics Research* 03(03n04): 1841005.
- Narazaki K, Oleynikov D and Stergiou N (2006) Robotic surgery training and performance: identifying objective variables for quantifying the extent of proficiency. *Surgical Endoscopy* 20(1): 96–103.
- Nichols KA and Okamura AM (2015) Methods to segment hard inclusions in soft tissue during autonomous robotic palpation. *IEEE Transactions on Robotics* 31(2): 344–354.
- Olson E (2011) Apriltag: A robust and flexible visual fiducial system. In: 2011 IEEE international conference on robotics and automation. Piscataway, NJ: IEEE, pp. 3400–3407.
- Osa T, Staub C and Knoll A (2010) Framework of automatic robot surgery system using visual servoing. In: 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway, NJ: IEEE, pp. 1837–1842.
- Osa T, Sugita N and Mitsuishi M (2018) Online trajectory planning and force control for automation of surgical tasks. *IEEE Transactions on Automation Science and Engineering* 15(2): 675–691.
- Patil S and Alterovitz R (2010) Toward automated tissue retraction in robot-assisted surgery. In: 2010 IEEE International Conference on Robotics and Automation. Piscataway, NJ: IEEE, pp. 2088–2094.
- Pedram SA, Ferguson P, Ma J, et al. (2017) Autonomous suturing via surgical robot: An algorithm for optimal selection of needle diameter, shape, and path. In: 2017 IEEE International conference on robotics and automation (ICRA). Piscataway, NJ: IEEE, pp. 2391–2398.
- Pedram SA, Shin C, Ferguson PW, et al. (2021) Autonomous suturing framework and quantification using a cable-driven surgical robot. *IEEE Transactions on Robotics* 37(2): 404–417.
- Preda N, Manurung A, Lambercy O, et al. (2015) Motion planning for a multi-arm surgical robot using both sampling-based algorithms and motion primitives. In: 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Piscataway, NJ: IEEE, pp. 1422–1427.

- Reynolds W, Jr (2001) The first laparoscopic cholecystectomy. *JSLS : Journal of the Society of Laparoendoscopic Surgeons* 5(1): 89–94.
- Richter F, Shen S, Liu F, et al. (2021) Autonomous robotic suction to clear the surgical field for hemostasis using image-based blood flow detection. *IEEE Robotics and Automation Letters* 6(2): 1383–1390.
- Salehian SSM, Khoramshahi M and Billard A (2016) A Dynamical system approach for softly catching a flying object: Theory and Experiment. *IEEE Transactions on Robotics* 32(ARTICLE): 462–471.
- Schulman J, Gupta A, Venkatesan S, et al. (2013) A case study of trajectory transfer through non-rigid registration for a simplified suturing scenario. In: *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*. Piscataway, NJ: IEEE, pp. 4111–4117.
- Schwaner KL, Dall'Alba D, Jensen PT, et al. (2021) Autonomous needle manipulation for robotic surgical suturing based on skills learned from demonstration. In: *2021 IEEE 17th International Conference on Automation Science and Engineering (CASE)*. Piscataway, NJ: IEEE, pp. 235–241.
- Schweikard A, Adler JR and Latombe J-C (1993) Motion planning in stereotaxic radiosurgery. *IEEE Transactions on Robotics and Automation* 9(6): 764–774.
- Seita D, Krishnan S, Fox R, et al. (2018) Fast and reliable autonomous surgical debridement with cable-driven robots using a two-phase calibration procedure. In: *2018 IEEE International Conference on Robotics and Automation (ICRA)*. Piscataway, NJ: IEEE, pp. 6651–6658.
- Sen S, Garg A, Gealy DV, et al. (2016) Automating multi-throw multilateral surgical suturing with a mechanical needle guide and sequential convex optimization. In: *2016 IEEE international conference on robotics and automation (ICRA)*. Piscataway, NJ: IEEE, pp. 4178–4185.
- Shademan A, Decker RS, Opfermann JD, et al. (2016) Supervised autonomous robotic soft tissue surgery. *Science Translational Medicine* 8(337): 337ra64.
- Shah RD, Cao A, Golenberg L, et al. (2009) Performance of basic manipulation and intracorporeal suturing tasks in a robotic surgical system: single-versus dual-monitor views. *Surgical Endoscopy* 23(4): 727–733.
- Siméon T, Laumond J-P, Cortés J, et al. (2004) Manipulation planning with probabilistic roadmaps. *The International Journal of Robotics Research* 23(7–8): 729–746.
- Slotine J-JE and Li W (1991) *Applied Nonlinear Control*. Englewood Cliffs, NJ: Prentice hall, Volume 199.
- Sozzi A, Bonfè M, Farsoni S, et al. (2019) Dynamic motion planning for autonomous assistive surgical robots. *Electronics* 8(9): 957.
- Stilman M, Schamburek J-U, Kuffner J, et al. (2007) Manipulation planning among movable obstacles. In: *Proceedings 2007 IEEE international conference on robotics and automation*. Piscataway, NJ: IEEE, pp. 3327–3332.
- Van Den Berg J, Miller S, Duckworth D, et al. (2010) Superhuman performance of surgical tasks by robots using iterative learning from human-guided demonstrations. In: *2010 IEEE International Conference on Robotics and Automation*. Piscataway, NJ: IEEE, pp. 2074–2081.
- Varier VM, Rajamani DK, Goldfarb N, et al. (2020) Collaborative suturing: A reinforcement learning approach to automate hand-off task in suturing for surgical robots. In: *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. Piscataway, NJ: IEEE, pp. 1380–1386.
- Voros S, Long J-A and Cinquin P (2006) Automatic localization of laparoscopic instruments for the visual servoing of an endoscopic camera holder. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 535–542.
- Watanabe K, Kanno T, Ito K, et al. (2016) Human-integrated automation of suturing task with one-master two-slave system for laparoscopic surgery. In: *2016 IEEE International Conference on Advanced Intelligent Mechatronics (AIM)*. Piscataway, NJ: IEEE, pp. 1180–1185.
- Watanabe K, Kanno T, Ito K, et al. (2018) Single-master dual-slave surgical robot with automated relay of suture needle. *IEEE Transactions on Industrial Electronics* 65(8): 6343–6351.
- Wei G-Q, Arbter K and Hirzinger G (1997) Real-time visual servoing for laparoscopic surgery. controlling robot motion with color image segmentation. *IEEE Engineering in Medicine and Biology Magazine: The Quarterly Magazine of the Engineering in Medicine & Biology Society* 16(1): 40–45.
- Yang B, Chen W, Wang Z, et al. (2019) Adaptive fov control of laparoscopes with programmable composed constraints. *IEEE Transactions on Medical Robotics and Bionics* 1(4): 206–217.
- Zhang Z (2000) A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22(11): 1330–1334.
- Zhong F and Liu Y (2018) Image-based 3d pose reconstruction of surgical needle for robot-assisted laparoscopic suturing. *Chinese Journal of Electronics* 27(3): 476–482.
- Zhong F, Navarro-Alarcon D, Wang Z, et al. (2016) Adaptive 3d pose computation of suturing needle using constraints from static monocular image feedback. In: *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Piscataway, NJ: IEEE, pp. 5521–5526.
- Zhong F, Wang Y, Wang Z, et al. (2019) Dual-arm robotic needle insertion with active tissue deformation for autonomous suturing. *IEEE Robotics and Automation Letters* 4(3): 2669–2676.
- Zhong F, Wang Z, Chen W, et al. (2020) Hand-eye calibration of surgical instrument for robotic surgery using interactive manipulation. *IEEE Robotics and Automation Letters* 5(2): 1540–1547.
- Zucker M, Ratliff N, Dragan AD, et al. (2013) Chomp: Covariant hamiltonian optimization for motion planning. *The International Journal of Robotics Research* 32(9–10): 1164–1193.

Appendix

Appendix A. Anatomy of “Weak Coupling” on a Wristed Robot

We take the EndoWrist robotic surgical instrument structure as an example. Recalling the robot kinematics in (1) and (2), we use the skeleton nodes in $\mathcal{N} = \{n_1, n_2, n_3\}$ to describe the 3D position of the end-effector as follows:

$$\mathbf{x}_t = \overrightarrow{\mathcal{O}n_1} + \overrightarrow{n_1n_2} + \overrightarrow{n_2n_3} \quad (39a)$$

$$= \begin{bmatrix} q_3 \\ l_w \\ l_t \end{bmatrix}^\top \begin{bmatrix} \mathbf{R}_s(\mathbf{q}_s)\bar{\mathbf{r}}_s \\ \mathbf{R}_w(\mathbf{q}_s, \mathbf{q}_w)\bar{\mathbf{r}}_w \\ \mathbf{R}_t(\mathbf{q})\bar{\mathbf{r}}_t \end{bmatrix} \quad (39b)$$

where $l_{w/t}$ denote the link lengths of the instrument and q_3 is length of the shaft that passes the RCM, $\bar{\mathbf{r}}_{s/w/t}$ compute the directional vectors of the link’s centerline fixed in local frame, \mathbf{x}_t is coincided with \mathcal{O}_e due to standard DHc, and $\mathbf{R}(\cdot)$ denote the rotation matrices. Differentiating (39) yields:

$$\dot{\mathbf{x}}_t = \begin{bmatrix} q_3 \mathbf{S}_s(\mathbf{q}_s) + \mathbf{R}_s(\mathbf{q}_s)\bar{\mathbf{r}}_s \\ l_w \mathbf{R}_s^s \mathbf{S}_w(\mathbf{q}_w)\bar{\mathbf{r}}_w \\ l_t \mathbf{R}_w(\mathbf{q}_s, \mathbf{q}_w)^w \mathbf{S}_t(q_t)\bar{\mathbf{r}}_t \end{bmatrix}^\top \begin{bmatrix} \dot{\mathbf{q}}_s \\ \dot{\mathbf{q}}_w \\ \dot{\mathbf{q}}_t \end{bmatrix} \quad (40)$$

where $\mathbf{S}(\cdot)$ are the local Jacobian matrices; rearranging (39b) further leads to the following factored form:

$$\dot{\mathbf{x}}_t = \mathbf{L}_1 \mathbf{A}_1(\mathbf{q}_s) \dot{\mathbf{q}}_s + \mathbf{L}_2 \mathbf{A}_2(\mathbf{q}_w) \dot{\mathbf{q}}_w + \mathbf{L}_3 \mathbf{A}_3(q_t) \dot{q}_t \quad (41)$$

where $\mathbf{A}_{1/2/3}$ are the interaction matrices, with $\mathbf{L}(\cdot)$ being

$$\mathbf{L}_1 = \begin{bmatrix} l_s & 0 & 0 \\ 0 & l_s & 0 \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{L}_2 = \begin{bmatrix} l_w & 0 \\ 0 & l_w \end{bmatrix}, \mathbf{L}_3 = l_t \quad (42)$$

which act as coefficient matrices during regulation of individual terms in (41). In robotic MIS, the recommended insertion length of the instrument shaft that passes the RCM (or the trocar entry), that is, q_3 , is around 150 mm to 200 mm (Escobar and Falcone, 2014). The links of the wristed

structure (taking the Large Needle Driver provided by the DVSS) are $[l_w, l_t] = [9.1, 9.5]$ mm, respectively. This indicates that practically, we can assume $q_3 \gg l_{w,t}$, which will arise the following motion properties of the instrument upon such kinematic constraint:

- **Property 1.** The position of the end-effector $\mathbf{x}_t(t)$ (i.e., n_1) is roughly estimated as $\mathbf{x}_s(t)$ (the position of n_1) under any feasible $\mathbf{q}(t)$:

$$\mathbf{x}_t(\mathbf{q}(t)) \approx \mathbf{x}_s(\mathbf{q}_s) \forall \mathbf{q}(t) \text{ s.t. } q_3 \gg l_{w,t} \quad (43)$$

which could be easily derived from (39) as the norm of $\overrightarrow{n_1n_2}$ and $\overrightarrow{n_2n_3}$ could be neglected. This also indicates that the adjustment of the end-effector position is dominated by \mathbf{q}_s , as $\mathbf{x}_s(\mathbf{q}_s)$.

- **Property 2.** If the end-effector motion space $\mathbf{x}_t(\mathbf{q})$ is highly restricted,⁷ the rotation matrix $\mathbf{R}_s(\mathbf{q}_s)$ could be considered unchanged, that is,

$$\mathbf{R}_s(\mathbf{q}_s) \approx \mathbf{R}_s(\mathbf{q}_{s_0}) \forall \mathbf{q} \text{ s.t. } q_3 \gg l_{w,t} \quad (44)$$

as $\mathbf{x}_s(\mathbf{q}_s) = \mathbf{R}_s(\mathbf{q}_s)\bar{\mathbf{r}}_s$ (43) is also restricted to a small domain, leading $\mathbf{R}_s(\mathbf{q}_s) \rightarrow \mathbf{I}_{3 \times 3}$. Here, \mathbf{q}_{s_0} denotes the static initial configuration of the robot prior to the control. This also indicates that the adjustment of the end-effector rotation is dominated by \mathbf{q}_w and q_t , as $\mathbf{R}_t(\mathbf{q}) \approx \mathbf{R}_s(\mathbf{q}_{s_0})\mathbf{R}_w^w(\mathbf{q}_w)\mathbf{R}_t(\mathbf{q}_t)$.

To summarize, the adjustment of the instrument’s end-effector position and orientation are dominated by different (unique) set of robot joints in a relatively decoupled manner. We name the them as the “weak coupling” effect during robot motion control and is naturally available to any serial robot manipulators with long proximal links and a wristed end-effector with short distal links (e.g., yielding $q_3 \gg l_{w,t}$ in this case). However, it should be noted that we only utilize these approximations to design a well-performed robot controller instead of direct analytical computation, as $q_3 < 100l_{w,t}$.