CLOSED-LOOP PLANNING AND CONTROL OF STEERABLE MEDICAL NEEDLES

Sachin Patil

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Computer Science.

Chapel Hill
2013

Approved by:
Ron Alterovitz
Jur van den Berg
Ming C. Lin
Dinesh Manocha
Marc Niethammer
ABSTRACT

SACHIN PATIL: Closed-Loop Planning and Control of Steerable Medical Needles
(Under the direction of Ron Alterovitz)

Steerable needles have the potential to increase the effectiveness of needle-based clinical procedures such as biopsy, drug delivery, and radioactive seed implantation for cancer treatment. These needles can trace curved paths when inserted into tissue, thereby increasing maneuverability and targeting accuracy while reaching previously inaccessible targets that are behind sensitive or impenetrable anatomical regions. Guiding these flexible needles along an intended path requires continuously inserting and twisting the needle at its base, which is not intuitive for a human operator. In addition, the needle often deviates from its intended trajectory due to factors such as tissue deformation, needle-tissue interaction, noisy actuation and sensing, modeling errors, and involuntary patient motions. These challenges can be addressed with the assistance of robotic systems that automatically compensate for these perturbations by performing motion planning and feedback control of the needle in a closed-loop fashion under sensory feedback.

We present two approaches for efficient closed-loop guidance of steerable needles to targets within clinically acceptable accuracy while safely avoiding sensitive or impenetrable anatomical structures. The first approach uses a fast motion planning algorithm that unifies planning and control by continuously replanning, enabling correction for perturbations as they occur. We evaluate our method using a needle steering system in phantom and ex vivo animal tissues. The second approach integrates motion planning and feedback control of steerable needles in highly deformable environments. We demonstrate that this approach significantly improves the probability of success compared to prior approaches that either consider uncertainty or deformations but not both simultaneously. We also propose a data-driven method to estimate parameters of stochastic models of steerable needle motion. These models can be used to create realistic medical simulators for clinicians wanting to train for steerable needle procedures and to improve the effectiveness of existing planning and control methods.
This dissertation advances the state of the art in planning and control of steerable needles and is an important step towards realizing needle steering in clinical practice. The methods developed in this dissertation also generalize to important applications beyond medical needle steering, such as manipulating deformable objects and control of mobile robots.
To my parents – Chandrashekhar and Shubhangi Patil
ACKNOWLEDGEMENTS

I would like to first thank Ron Alterovitz for having me as his student and providing me the opportunity to pursue the research that led to this dissertation. It has been a privilege to have him as my advisor. His enthusiasm and love for his research area motivated me to pursue research as a career. He has been extremely helpful and always gave me insightful feedback on various problems. I thank him for being patient and having faith in my abilities, and providing me the freedom to explore things at my own pace, while always coming to the rescue whenever I got stuck or bogged down.

I would like to thank Jur van den Berg for being an excellent mentor to me throughout my PhD, and extensively collaborating with me on this work. He significantly enhanced my understanding of fundamental concepts involved in this work and provided a sounding board for my research ideas. We continue to have very exciting and fruitful discussions.

I would also like to thank Dinesh Manocha and Ming Lin for being on my defense committee and I am extremely grateful to them for giving me an opportunity to pursue research in graduate school when I first came to UNC Chapel Hill. This was instrumental in getting me excited about doing research in the area of motion planning. I would also like to thank Marc Niethammer for being on my defense committee and for this support and suggestions on how to improve this dissertation.

I thank my colleagues in the Computational Robotics group at UNC Chapel Hill for their invaluable feedback on this work and presentations over the years: Luis Torres, Wen Sun, Chris Bowen, and Edgar Lobaton. I would like to thank several close collaborators and co-authors, which include Robert Webster and Jessica Burgner at Vanderbilt University, Momen Abayazid, Guus Vrooijink, and Sarthak Misra at University of Twente, and Pieter Abbeel and Ken Goldberg at University of California at Berkeley. I am also grateful to the other faculty members at UNC Chapel Hill for their guidance and teaching me the nuts and bolts of doing research. I thank the staff in the Department of Computer Science, Janet Jones and Jodie Turnbull in particular, for shielding me from all administrative issues and making my graduate school experience hassle-free.
I am also indebted to the many close friends and colleagues for their company, advice, and constant encouragement throughout my stay in Chapel Hill – Vishal Verma, Vishnu Reddy, Ram Kumar, Stephen Guy, Luv Kohli, Peter Lincoln, Ryan Schubert, Sean Curtis, Anish Chandak, Jason Sewall, Nikunj Raghuvanshi. I would also like to thank Anya Derbakova for her support and encouragement, particularly in the months leading up to my defense. Many thanks to my friends from California, who had nothing to do with any of the work in this dissertation, but who always cheered me on and warmly welcomed me back whenever I was visiting – Devsmita Das and Suddha Basu, Nidhi Parashar and Prashant Baronia, and Priti Kadam and Aditya Prakash.

Finally, and most importantly, I thank my parents, Chandrashekhar and Shubhangi Patil, for enabling me over all these years, for giving me the wisdom and freedom to choose the path I wanted to take in life, and for their unconditional love and support through my moments of doubts.
# TABLE OF CONTENTS

LIST OF TABLES ........................................................................................................ xi

LIST OF FIGURES .................................................................................................... xii

1 Introduction ........................................................................................................ 1
   1.1 Steerable Needle Model ............................................................................. 2
   1.2 Robotic Needle Steering ........................................................................... 4
   1.3 Challenges for Closed-Loop Planning and Control ................................... 5
   1.4 Thesis Statement ......................................................................................... 7
   1.5 Main Results ............................................................................................... 7
      1.5.1 Needle Steering in 3D via Rapid Replanning .................................... 7
      1.5.2 Data-Driven Stochastic Models for Simulating Steerable Needle Procedures ... 8
      1.5.3 Unified Framework for Planning and Control in Deformable Environments ... 9
   1.6 Organization ............................................................................................... 10

2 Needle Steering in 3D via Rapid Replanning ....................................................... 11
   2.1 Rapid Replanning ....................................................................................... 11
   2.2 Related Work ............................................................................................ 13
   2.3 Objective .................................................................................................. 17
   2.4 Variable Curvature Needle Kinematic Model ............................................. 18
   2.5 Rapid Replanning Approach ..................................................................... 21
      2.5.1 Motion Planning ................................................................................ 21
      2.5.2 Optimizing Clinical Metrics .............................................................. 25
      2.5.3 Fault Tolerance ............................................................................... 27
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.6 Experimental Setup</td>
<td>27</td>
</tr>
<tr>
<td>2.6.1 System Components</td>
<td>27</td>
</tr>
<tr>
<td>2.6.2 Tissue Sample Materials</td>
<td>29</td>
</tr>
<tr>
<td>2.6.3 Needle Characterization</td>
<td>29</td>
</tr>
<tr>
<td>2.7 Experimental Evaluation</td>
<td>31</td>
</tr>
<tr>
<td>2.7.1 Evaluation in Tissue Phantoms</td>
<td>32</td>
</tr>
<tr>
<td>2.7.2 Evaluation in Tissue Phantoms with Obstacles</td>
<td>33</td>
</tr>
<tr>
<td>2.7.3 Evaluation in Ex Vivo Porcine Loin Tissue</td>
<td>34</td>
</tr>
<tr>
<td>2.7.4 Evaluation in Anthropomorphic Liver Phantom</td>
<td>35</td>
</tr>
<tr>
<td>2.8 Discussion</td>
<td>37</td>
</tr>
<tr>
<td>3 Data-Driven Stochastic Models For Simulating Steerable Needle Procedures</td>
<td>40</td>
</tr>
<tr>
<td>3.1 Simulating Steerable Needle Procedures</td>
<td>40</td>
</tr>
<tr>
<td>3.2 Related Work</td>
<td>43</td>
</tr>
<tr>
<td>3.3 Preliminaries and Objective</td>
<td>45</td>
</tr>
<tr>
<td>3.3.1 State Representation</td>
<td>45</td>
</tr>
<tr>
<td>3.3.2 Stochastic Kinematic Model</td>
<td>46</td>
</tr>
<tr>
<td>3.3.3 Stochastic Measurement Model</td>
<td>47</td>
</tr>
<tr>
<td>3.3.4 Problem Definition</td>
<td>48</td>
</tr>
<tr>
<td>3.4 State Estimation</td>
<td>48</td>
</tr>
<tr>
<td>3.4.1 Extended Kalman Filter</td>
<td>49</td>
</tr>
<tr>
<td>3.4.2 Extended Kalman Smoother</td>
<td>50</td>
</tr>
<tr>
<td>3.5 Expectation Maximization for Parameter Estimation</td>
<td>50</td>
</tr>
<tr>
<td>3.6 Parameter Estimation</td>
<td>54</td>
</tr>
<tr>
<td>3.6.1 Data Collection</td>
<td>54</td>
</tr>
<tr>
<td>3.6.2 Estimated Error Parameters: Sim-Test</td>
<td>55</td>
</tr>
<tr>
<td>3.6.3 Estimated Error Parameters: Ex-vivo Porcine Loin Tissue</td>
<td>59</td>
</tr>
<tr>
<td>3.7 Evaluation</td>
<td>61</td>
</tr>
</tbody>
</table>
LIST OF TABLES

2.1 Comparing cost metrics for needle insertion in Sim-test phantom tissues and ex vivo porcine loin tissue with virtual obstacles ........................................... 36

4.1 Comparison of our method with prior methods over 100 plans in terms of mean absolute error (MAE) from ground truth probability ................................. 91
LIST OF FIGURES

1.1 Anatomical scenario modeling the human male pelvic region for simulating delivery of radioactive doses to targets within the prostate for cancer treatment ...... 2

1.2 Mechanics of insertion of steerable needle with a pre-bent bevel-tip ............... 3

1.3 Workflow for a clinically viable robotic system for steerable needle procedures ...... 4

2.1 Closed-loop needle steering via rapid replanning ........................................... 12

2.2 Overview of our rapid replanning paradigm which relies on a fast sampling-based motion planner for closed-loop steering of the needle to the desired target region while avoiding anatomical obstacles ............................................. 17

2.3 Circular arc connecting the needle tip to a point defined in the local coordinate frame of the needle tip. The needle plan is composed of circular arcs of bounded curvature ................................................................. 19

2.4 Three stages of duty cycled steerable needle insertion ...................................... 21

2.5 Our needle steering system consists of a needle steering robot, a pre-bent, bevel-tip steerable needle, and an electromagnetic tracking system ......................... 28

2.6 Experiments in Sim-Test tissue phantom and ex vivo porcine loin tissue ........... 29

2.7 Characterization of the relationship between the duty cycling parameter for a steerable needle in Sim-Test tissue phantom and ex vivo porcine loin tissue .... 30

2.8 Experiments in a cuboidal shaped Sim-Test tissue phantom environment containing virtual obstacles for planning purposes .............................. 32

2.9 Comparison of targeting errors using closed-loop steering and open-loop execution for two clinically motivated optimization criteria ................................. 34

2.10 Experiments in an ex vivo porcine tissue sample containing virtual obstacles for planning purposes ................................................................. 35

2.11 Comparison of targeting errors using closed-loop steering and open-loop execution for two clinically motivated optimization criteria ................................. 37

2.12 Our needle steering system with rapid replanning as applied to a scenario motivated by the clinical task of ablating a tumor in the liver while avoiding the hepatic veins ................................................................. 38

3.1 We propose a simplified, stochastic model of steerable needle insertion and estimate the parameters of this stochastic model from data obtained from prior experiments and procedures ................................................. 42
3.2 Experimental data in Sim-Test: Normalized histograms showing the distribution of position and orientation errors along each of the axes .................. 58

3.3 Experimental data in ex vivo porcine loin tissue: Normalized histograms showing the distribution of position and orientation errors along each of the axes .... 61

3.4 Evaluation of our method by simulating steerable needle procedures with artificial noise based on our estimated error parameters in a cuboidal shaped Sim-Test tissue phantom and in an ex vivo porcine loin tissue sample ........ 62

3.5 Comparison of the targeting error using closed-loop, rapid replanning steering and open-loop execution for two clinically motivated optimization criteria in simulation in Sim-Test material ................................. 64

3.6 Comparison of the targeting error using closed-loop, rapid replanning steering and open-loop execution for two clinically motivated optimization criteria in simulation in ex vivo porcine loin tissue ......................... 65

4.1 Motion planning for guiding steerable needles under 2D image guidance: Our unified framework accounts for uncertainty in deformation models, noisy sensing, and unpredictable actuation, results in a significantly higher probability of success in plan execution as compared to current motion planning solutions .......................................................... 70

4.2 Schematic overview of our unified framework for planning and control under uncertainty in deformable environments ................................. 76

4.3 Model linearization computed numerically using a deformable object simulator ..... 79

4.4 Estimating the probability of success for a motion plan based on a priori probability distributions of the robot state ................................. 82

4.5 Truncating the joint conditional distributions of the robot state with respect to constraints imposed by obstacles in the workspace ......................... 84

4.6 Local convexification of free space and transforming the environment for truncating the joint conditional distributions of the robot state ............... 89

4.7 Comparison of the probability of collision estimated using our method and other prior methods for a nonholonomic bevel-tip steerable needle ............... 90

4.8 Plan executions with and without feedback under the influence of actuation noise ... 93

4.9 Percentage of successful plan executions of plan chosen by our method consistently outperforms plans that do not simultaneously account for both uncertainty and deformations .................................................. 94

4.10 Simulation of needle steering in a plane in the human liver for biopsy or drug delivery ................................................................. 96
CHAPTER 1

Introduction

Needles are widely used in many minimally invasive clinical procedures for diagnosis and treatment, including retrieval of tissue samples for biopsies, and administering therapies such as delivering drugs, implanting radioactive seeds for cancer treatment, and facilitating thermal ablation of cancerous tissue. Once inserted, a needle can also serve as a guide for introduction of a sheath through which catheters and other medical devices can be introduced to reach clinical targets deep inside a patient’s anatomy. As sensors, manipulators, and other medical devices continue to become smaller, applications for needle-based interventional procedures will continue to expand.

Currently, needle-based procedures are performed by clinicians using stiff needles under image guidance provided by modalities such as ultrasound, computed tomography (CT), and magnetic resonance imaging (MRI). Performing these procedures using stiff needles is limited to straight line paths between the needle entry location and the desired target, which can cause puncturing of sensitive tissues leading to increased patient trauma and recuperation times (Fig. 1.1a). Moreover, the use of stiff needles can result in deviations from the straight line path due to factors such as needle/tissue deformation, involuntary patient motions, and noisy actuation and imaging feedback, with limited ability to correct for this error during insertion (Abolhassani et al., 2007).

As an alternative to stiff needles, we consider a new class of needles that can steer along curved paths through soft tissues (Cowan et al., 2011). These steerable needles offer improved maneuverability within tissue during insertion and greater targeting accuracy. They also facilitate access to previously inaccessible clinical targets while avoiding puncturing sensitive anatomical tissues such as vital organs and vessels and avoiding impenetrable structures such as bones (Fig. 1.1b).
Figure 1.1: Anatomical scenario modeling the human male pelvic region for simulating delivery of radioactive doses to targets within the prostate for cancer treatment [reproduced with permission from (Chentanez et al., 2009)]. (a) Performing needle-based procedures using stiff needles can cause puncturing of sensitive tissues (such as Cowper’s glands in this instance) leading to increased patient trauma and recuperation times. (b) Using a steerable needle offers increased maneuverability while enabling access to previously inaccessible clinical targets without puncturing sensitive tissues.

1.1 Steerable Needle Model

In this work, we consider steerable needles with a flexible needle shaft and a bevel-tip (see Fig. 1.2a). Even though our emphasis is on bevel-tip steerable needles, the contributions of this dissertation are also applicable to other asymmetric-tip steerable needles such as needles with stylet tips (Okazawa et al., 2005) or programmable bevel-tips (Ko et al., 2011).

The needle is externally controlled by inserting and twisting the needle at its base. The needle is maneuvered within tissue by reorienting the bevel-tip through twists applied to the needle base. The flexible needle shaft bends due to reaction forces from the tissue on the bevel-tip, causing the needle to follow constant curvature paths within tissue in the direction of the bevel (Webster III et al., 2006). The natural curvature is dependent on the material properties of both the needle and the tissue.

The state of the needle tip is defined by the position and orientation as defined in a world coordinate frame (see Fig. 1.2b). The forward motion of the needle is subject to nonholonomic constraints, i.e., similar to a car or an airplane, the needle cannot instantaneously move in arbitrary directions in its 6D state space. Since only two control inputs (insertions and twists) are available to control the needle in its 6D state space, it is also an under-actuated system. In fact, the needle is not
small-time locally controllable, i.e., any state close to the current state is not reachable in arbitrarily small amounts of time by paths close to the current state (Murray et al., 1994; Kallem, 2008).

As with all real world robotic systems, there is uncertainty involved with steerable needle procedures and it causes the needle to deviate from its intended trajectory. This uncertainty can arise from several factors such as unpredictable actuation, noisy and partial measurements from sensors, unpredictable needle-tissue interaction, and involuntary patient motions.

In addition, the mechanics of steerable needle insertion into soft tissue depends on the interaction between the needle and the tissue and the deformations in both the flexible needle and tissue resulting from these interactions (Chentanez et al., 2009). The coupling between the needle and tissue introduces new complications which cause noisy needle behavior. For instance, tissue inhomogeneities can cause the needle shaft to deflect during insertion, and friction between the needle and tissue could cause torsional buildup along the needle shaft when twists are applied to the base.

Analyzing all aspects of steerable needle insertion and sources of uncertainty remains a challenging problem. For these reasons, guiding a steerable needle around anatomical obstacles to a target region under image guidance is challenging for a human operator. Creating a needle steering robotic system that assists clinicians and addresses these challenges could enable new needle-based procedures and substantially improve the clinical outcomes of some existing needle-based procedures.
Figure 1.3: Workflow for a clinically viable robotic system for steerable needle procedures. This dissertation advances the state of the art in planning and control of steerable needles and is an important step towards realizing needle steering in clinical practice. The clinician, who is responsible for the procedure, receives feedback from all components of the system, although these edges are omitted for succinctness.

1.2 Robotic Needle Steering

Several image-guided needle placement and insertion systems have been developed for stiff needles, including commercial systems (Philips Healthcare, 2012; General Electric (GE) Healthcare, 2012; InnerOptic Inc., 2012) and clinically viable academic prototypes (Fichtinger et al., 2008; Long et al., 2012; Hungr et al., 2012; Seifabadi et al., 2012). However, these systems are currently limited to stiff needles and optimize straight line paths between the insertion location and the target, which can potentially result in damage to sensitive tissues. Robotic needle steering systems for performing steerable needle procedures can help overcome these deficiencies.

Fig. 1.3 shows the workflow for a clinically viable robotic needle steering system. In the preoperative stage, high resolution scans (e.g., preoperative sensing modalities such as CT or MRI) of the anatomical region of interest in the patient are acquired. In the modeling phase, the clinician identifies sensitive structures such as glands or blood vessels and other obstacles such as bones. The clinician also specifies the insertion location of the needle, the clinical target, a clinically motivated optimization criteria, and a characterization of the steerable needle’s curvature. In the intraoperative stage, the planning and control system plans and controls the motion of the needle in a closed-loop
fashion using measurements of the state of the system (both needle and tissue) obtained from low resolution sensing devices such as ultrasound or electromagnetic tracking. The system is either fully autonomous or semi-autonomous with a clinician in the loop to adjust execution as needed.

1.3 Challenges for Closed-Loop Planning and Control

The planning and control system is an important part of the robotic needle steering system. We discuss the challenges associated with closed-loop planning and control of steerable needles below.

As discussed in Sec. 1.1, the needle is an under-actuated nonholonomic system. Motion planning and control for such systems is challenging. Since deformable objects have possibly infinite degrees of freedom, planning and control of a very high-dimensional, coupled system comprising of the needle and tissue is a hard problem. In addition, steerable needle procedures are subject to considerable deformations and uncertainty which are responsible for large deviations in the motion of the needle, anatomical structures, and clinical targets within the tissue.

A highly accurate, deterministic model of steerable needle motion in human tissues would make motion planning and control less challenging. In spite of many prior efforts (Cowan et al., 2011), modeling all aspects of steerable needle insertion remains challenging to accomplish efficiently and accurately because of the complexity of modeling deformation behaviors, modeling the interaction of the needle with the deformable object, and incorporating uncertainty from several sources. As a consequence, the model is often deliberately chosen to be a simplification. The mismatch between modeling and reality is treated as an additional source of uncertainty and is addressed using closed-loop planning and control techniques.

For computational tractability, we begin by using a simplified model for planning and control that assumes that the needle bends to follow the needle tip exactly, i.e., the tissue does not deform the needle shaft. We also assume that the insertions and twists applied to the needle base are directly and exactly transmitted to the needle tip, i.e., there is no buckling or torsion along the needle shaft. This implies that the motion of the needle is fully determined by the motion of the needle tip. Instead of dealing with a high-dimensional coupled system, we then consider the planning and control problem in a reduced 6D state space of the needle tip. Our objective is to address the uncertainty in steerable needle motion in soft tissue, which includes accounting for mismatch between modeling and reality.
It is also important to evaluate how good any planning and control approach is for steerable needle procedures because some executions are significantly better than others. There are three key clinically motivated criteria for evaluating an approach:

1. **Accuracy**: The success of any needle-based procedure depends on the targeting accuracy. Large errors between the final position of the needle tip and the desired clinical target can have undesirable consequences. For example, poor placement during biopsies leads to false negatives and inaccurate placement of radioactive seeds for cancer treatment destroys healthy instead of cancerous tissue (Bogdanich, 2009). Accurate needle insertion is sufficiently difficult that poor accuracy is common in practice. For example, experienced clinicians inserting radioactive seeds into the prostate gland for brachytherapy prostate cancer treatment experience average placement errors of 6.3 mm, about 15% of the prostate’s diameter (Taschereau et al., 2000). Other studies have found that the targeting errors for biopsies performed by experienced clinicians averages between 5.5–6.5 mm when performing procedures using rigid needles (Blumenfeld et al., 2007; Schouten et al., 2012). Robotic systems for placement and insertion of stiff needles achieve targeting errors to the order of $\approx 1–4$ mm (Fichtinger et al., 2008; Xu et al., 2010; Long et al., 2012; Seifabadi et al., 2012; Hungr et al., 2012).

2. **Safety**: It is crucial that the needle avoids impenetrable anatomical structures such as bones and sensitive structures such as glands and vessels. For instance, damage to glands during prostate biopsies can cause problems such as incontinence (Wilt et al., 2008). Another aspect of safety deals with minimizing the amount of tissue cut during a procedure. This might be important for procedures being performed in critical organs such as the brain (Field et al., 2001). Trajectories that maximize safety are less likely to cause collisions with anatomical structures when deviations occur, but these trajectories tend to be longer, thereby increasing the amount of tissue cut during the procedure. It is important to consider the trade-off between avoiding specified structures and tissue cut during a procedure.

3. **Efficiency**: It is also important to consider the total procedure time for steerable needle procedures. A 2005 study of U.S. hospitals found that operating room charges averaged $62 per minute (range: $22 to $133 per minute) (Shippert, 2005). For a needle steering robotic
system to be clinically viable, it would need to efficiently perform procedures within a few minutes to minimize risk to the patient and operating costs.

Considering these statistics, a robotic needle steering system that achieves mean targeting errors that are lower than current clinical practice, is capable of safely avoiding sensitive anatomical structures, and can facilitate efficient execution of procedures could improve patient care. It would have the added advantage of enabling clinicians to avoid anatomical obstacles and minimize tissue damage and patient trauma as compared to traditional stiff needles.

1.4 Thesis Statement

In this dissertation, we present approaches for closed-loop motion planning and control (or steering) of steerable needles under sensory feedback. The objective of planning and control can be stated as follows. Given the initial position and orientation of the needle tip, the target region, and a specification of the anatomy and characterization of the steerable needle’s properties, the objective is to steer the needle to clinical targets under sensory feedback while avoiding anatomical structures and optimizing clinically relevant criteria.

My thesis is as follows:

_Efficient motion planning and control techniques that consider uncertainty and deformations can facilitate closed-loop guidance of steerable needles to targets within clinically acceptable targeting accuracy while safely avoiding clinician-specified anatomical structures._

1.5 Main Results

In support of my thesis, I will present several results relating to efficient closed-loop planning and control of steerable needles.

1.5.1 Needle Steering in 3D via Rapid Replanning

In this work, we present a new approach for automatically guiding the steerable needle to targets in 3D environments while avoiding obstacles and compensating for real-world uncertainties. We unify planning and control using a new, fast algorithm that continuously replans the needle motion to
handle unexpected perturbations. In contrast to the standard practice of planning a feasible trajectory and then using a feedback controller for correcting uncertain perturbations, our motion planner is sufficiently fast enough to correct for perturbations in needle, obstacle, or target motion as they occur. Our rapid replanning approach is enabled by an efficient sampling-based rapidly exploring random tree (RRT) planner that achieves orders of magnitude reduction in computation time compared to prior 3D approaches by incorporating variable curvature kinematics and a novel distance metric for planning. We also consider two clinically motivated criteria for planning: minimizing insertion length (minimizing tissue cut) and maximizing clearance (safety) from critical anatomical structures.

We integrated our algorithm with a needle steering system consisting of a bevel-tip steerable needle, a needle steering robot, and an electromagnetic tracker for estimating the needle tip pose in tissue. Given preoperative medical images, the clinician specifies the insertion location and target region as well as obstacles comprising of sensitive structures such as glands or blood vessels and impenetrable structures such as bones. Our rapid replanner then automatically guides the needle around anatomical obstacles to the target region with high accuracy. We provide experimental results to demonstrate that the approach can work in a practical clinically-inspired scenario. In our experiments, the system guided the needle tip in 3D to targets in phantom and ex vivo animal tissues with obstacles and achieved targeting errors averaging below 3 mm. In achieving these low targeting errors, our rapid re-planning approach overcame substantial uncertainty: open-loop needle steering resulted in errors exceeding 1 cm. Our experiments demonstrate that our system can achieve targeting accuracy that exceeds current clinical practice while simultaneously enabling avoidance of obstacles. A preliminary version of this work was published in (Patil and Alterovitz, 2010a).

1.5.2 Data-Driven Stochastic Models for Simulating Steerable Needle Procedures

The goal of this work is to create a stochastic model that captures needle behavior. We present a data-driven method for estimating the parameters of a stochastic model of the needle motion. This stochastic model can be used to create realistic medical simulators for clinicians to train for steerable needle procedures and to improve the effectiveness of existing planning and control methods.

Since it is difficult to efficiently and accurately model the needle motion and its behavior within tissue and the model would vary on a per-patient basis, we instead use a simplified model of the needle tip motion and add stochasticity to the model to account for real-world perturbations occurring
during procedures. We also consider stochastic models to account for noisy and partial measurements obtained during intraoperative sensing. We use the expectation-maximization (EM) algorithm to estimate the parameters of this stochastic model using data gathered from prior experiments and procedures. We present simulation results to validate our stochastic model.

1.5.3 Unified Framework for Planning and Control in Deformable Environments

In this work, we present a unified framework for planning and control of steerable needles in deformable environments that simultaneously considers large deformations and substantial uncertainty. Prior motion planning methods for deformable environments either assume deterministic deformations, which may result in unsafe paths, or compute plans in a static world and consider deformations as a type of perturbation, which neglects the large time-dependent motions of the obstacles and target. Our unified framework, which simultaneously accounts for both uncertainty and deformations, results in a significantly higher probability of success in plan execution.

Our method requires a simulator of steerable needle procedures. We use a sampling-based motion planner based on the simulator to generate a set of candidate motion plans that assume expected deformations. We then use the simulator and optimal control to numerically estimate time-dependent state distributions based on uncertain parameters (e.g. deformable material properties or actuation and sensing errors). We use this information to generate an optimal linear-quadratic (LQG) feedback controller for each candidate plan to mitigate any uncertainty in the expected deformations that occur during the actual execution of the plan. Since computing an optimal controller using the full deformable system state is computationally prohibitive, we observe that it is possible to formulate the optimal control problem using a subset of the full state space. We then present an efficient method to select the plan with the highest estimated probability of successfully avoiding obstacles and reaching the goal region. Using FEM-based simulation of deformable tissues, we demonstrate that our method can generate high quality plans for guiding steerable needles around obstacles to clinical targets under 2D image guidance in the presence of considerable deformations and uncertainty. We demonstrate that our approach significantly improves the probability of success compared to prior approaches based on standard feedback controllers or motion planners that do not simultaneously consider deformations and uncertainty. Results from this work were published in (Patil et al., 2011) and (Patil et al., 2012).
1.6 Organization

The rest of this dissertation is organized as follows. Chapter 2 presents the rapid replanning algorithm for closed-loop needle steering in 3D environments with obstacles. It also provides details about our needle steering system and presents experimental results for needle steering in phantom and ex vivo tissues. Chapter 3 describes a data-driven method for estimating parameters of a stochastic model of the needle behavior in tissue. Chapter 4 presents a unified planning framework that takes into account both uncertainty and deformations, by combining sampling-based motion planning in deformable environments with feedback control techniques. Finally, Chapter 5 presents a summary of our contributions and presents avenues for future work.
CHAPTER 2

Needle Steering in 3D via Rapid Replanning

In this chapter, we present a new approach for automatically guiding steerable needles to targets in 3D environments while avoiding obstacles and compensating for real-world perturbations. We unify planning and control using a new, fast algorithm that continuously replans the needle motion while optimizing clinically motivated criteria. Our rapid replanning approach is enabled by an efficient sampling-based rapidly exploring random tree (RRT) planner that achieves orders of magnitude reduction in computation time compared to prior 3D approaches by incorporating variable curvature kinematics and a novel distance metric for planning. We experimentally evaluate our approach for needle steering using tissue phantoms and ex vivo animal tissue. We demonstrate that our rapid replanning strategy successfully guides the needle around obstacles to desired 3D targets with an average error of less than 3 mm, which exceeds the current standard for clinical care.

2.1 Rapid Replanning

We present a novel approach for closed-loop needle steering in 3D environments with obstacles that unifies motion and planning and control. In contrast to the standard practice of planning a feasible trajectory and then using a feedback controller for correcting uncertain perturbations, our motion planner is fast enough to correct for perturbations in needle, obstacle, or target motion as they occur. This eliminates the need for a separate feedback controller, which can be difficult to create and tune for nonholonomic robots like steerable needles. Moreover, such controllers do not provide obstacle avoidance guarantees, which is critical for many steerable needle procedures. Another advantage of our approach is that the planner considers the global state of the system at each time step, rather than just locally correcting for perturbations. This could account for displacements of the target region and obstacles due to tissue deformation during the procedure.
Figure 2.1: Closed-loop needle steering via rapid replanning. Given the current needle tip state, target region, a specification of the anatomy, and characterization of the steerable needle’s properties, our approach uses a fast, randomized motion planner to compute in the available time many feasible motion plans across homotopy classes (top left). The method selects the best plan based on optimization criteria such as minimizing path length or maximizing clearance from obstacles (top right). We execute the first control input of the plan and measure the state of the needle tip (bottom). The actual state of the needle tip deviates from the model predicted state because of uncertainty. We repeat the planning process, hence replanning, starting from the actual needle tip state. This approach is made possible by a new, fast planner capable of computing hundreds of feasible plans per second.

Our new rapid replanning approach uses a customized, sampling-based motion planner that speeds up motion planning for steerable needles to the point that it can be done in real time with typical needle insertion velocities. To enable sub-second planning times, we leverage several observations and algorithmic advances. First, in contrast to prior motion planning approaches for needle steering, we relax the constant curvature path assumption by planning variable curvature paths and using duty cycled spinning during insertion (Minhas et al., 2007; Engh et al., 2010) to adjust the needle’s net curvature. Second, we propose a new distance metric for incremental expansion of the rapidly-exploring search tree to significantly improve planner performance. These help us achieve orders of magnitude reduction in computation time compared to prior sampling-based planners (Xu et al., 2009) and make the planner suitable for closed-loop needle steering. We also use the fast performance to enable us to consider two clinically motivated optimization criteria: minimizing insertion length and maximizing clearance from critical anatomical structures.

We present a complete needle steering system capable of automatically reaching targets in 3D environments while avoiding obstacles and compensating for real-world uncertainties. Our system...
consists of a bevel-tip steerable needle, a needle steering robot, and an electromagnetic tracker for estimating the needle tip pose in tissue. Given preoperative medical images, the clinician can specify the insertion location and target region as well as sensitive structures such as glands or blood vessels and other obstacles such as bones. Our rapid replanner then automatically guides the needle around anatomical obstacles to the target region with high accuracy.

We provide experimental results to demonstrate that the approach can work in a practical clinical scenario. In our experiments, the system guided the needle tip in 3D to targets in phantom and ex vivo animal tissues with obstacles and achieved targeting errors averaging below 3 mm. In achieving these low targeting errors, our rapid re-planning approach overcame substantial uncertainty: open-loop needle steering resulted in errors exceeding 1 cm. In comparison to current standards for clinical needle procedures, experienced physicians achieved targeting errors averaging 5.5–6.5 mm when performing procedures using stiff needles (Taschereau et al., 2000; Blumenfeld et al., 2007; Schouten et al., 2012). Our experiments demonstrate that our system can achieve targeting accuracy that exceeds current clinical practice while simultaneously enabling avoidance of obstacles.

The rest of this chapter is organized as follows: Sec. 2.2 provides a survey of previous work on planning and control for steerable needles. Sec. 2.3 summarizes the objective of our rapid replanning framework and Sec. 2.4 presents details of the variable curvature kinematic model of the motion of the needle tip used in this work. We provide details of our approach in Sec. 2.5 and describe our needle steering system used in our experiments in Sec. 2.6. Finally, Sec. 2.7 presents experimental results of needle steering in 3D environments with obstacles in phantom and ex vivo animal tissues and we present a discussion of our approach and directions for future work in this regard in Sec. 2.8.

### 2.2 Related Work

Several needle steering techniques have been developed that allow clinicians to adjust the needle path within tissue to improve targeting accuracy. These include bevel-tip flexible needles (Webster III et al., 2006), symmetric-tip needles that can be steered by applying forces at the base (DiMaio and Salcudean, 2003; Glozman and Shoham, 2007), stylet tips (Okazawa et al., 2005), programmable bevel-tip needles (Ko et al., 2011), and pre-bent concentric tubes (Webster III and Jones, 2010). Our
emphasis is on bevel-tip flexible needles, but our approach is also applicable to planning and control of needles with stylets and programmable bevel-tip needles.

Significant advancements have been made in modeling, motion planning, and controlling bevel-tip steerable needles (Cowan et al., 2011). A kinematic model generalizing a unicycle was experimentally validated by Webster et al. (Webster III et al., 2006). This model was augmented to include the effects of torsional friction along the needle shaft (Reed et al., 2009; Swensen and Cowan, 2012). Minhas et al. (Minhas et al., 2007; Engh et al., 2010) showed that in addition to insertion speed and rotation speed as control inputs, the curvature of the needle path can be controlled through duty cycled spinning of the needle during insertion. The mechanics and characteristics of steerable needles have been modeled for ex vivo tissue (Okamura et al., 2004; Misra et al., 2010), and in vivo tissue (Majewicz et al., 2012). Chentanez et al. (Chentanez et al., 2009) created a physically-based 3D simulation that uses nonlinear FEM for modeling needle and tissue deformations and accounts for topological changes occurring in the tissue mesh during needle insertion.

Motion planning for steerable needles in a plane (2D) has been extensively studied. Alterovitz et al. (Alterovitz et al., 2007) developed a planner for planar needle steering that addresses the issue of motion uncertainty by solving a Markov decision process (MDP) over a discretized state space and assuming full state observations. Kallem et al. (Kallem and Cowan, 2009; Kallem et al., 2010) developed a controller that stabilizes the needle to a desired plane. Reed et al. (Reed et al., 2011) coupled the controller to the planar planner (Alterovitz et al., 2007) to create an image-guided needle steering planning/control system and validated the approach in phantom tissues. These planning and control strategies do not generalize to 6D state spaces encountered in needle steering in 3D environments. Alterovitz et al. (Alterovitz et al., 2005) also take into account tissue deformations during needle insertion but this work does not address the issue of uncertainty due to noisy actuation, sensing, or errors in deformation modeling and simulation. Recently, Bernardes et al. (Bernardes et al., 2012) presented a semi-automated planar needle steering system that uses an adaptive path planning strategy to compensate for system uncertainties.

Motion planners for needle steering in 3D environments with obstacles have been proposed by Duindam et al. based on optimization (Duindam et al., 2008) and inverse kinematics (Duindam et al., 2010). Rapidly-exploring random trees (RRT) have been used in (Xu et al., 2009) to explore the full 6D state space of the needle tip to search for a feasible motion plan. Lobaton et al. (Lobaton et al., 2014)
compute motion plans for the steerable needle by using an approach based on sampling spheres of constant radii in the environment and concatenating trajectory segments on the surfaces of these spheres. Sampling-based motion planners for needle steering in 3D environments are computationally very expensive and not suitable for real-time operation. These planners also do not address the issue of uncertainty. A path-of-probability algorithm based on diffusion-based error propagation was developed by Park et al. (Park et al., 2010b), but this work does not take into account obstacles or noisy sensory feedback. Open-loop execution of motion plans result in large targeting errors due to deformation of the workspace and system uncertainties.

There is also prior work on controlling steerable needles in 3D environments to compensate for errors during insertion. Hauser et al. (Hauser et al., 2009) propose a real-time feedback controller which plans helical paths for 3D needle steering and experimentally demonstrate in simulation that the controller is robust to motion deviation and noisy sensing, even for a greedy control selection policy. Seiler et al. (Seiler et al., 2012) propose a fast trajectory correction method to compensate for uncertainty and deformations in the environment during needle insertion. These controllers either do not address the issue of obstacle avoidance or only propose a provisional solution and do not provide any guarantees on safety and performance in the presence of obstacles. Van den Berg et al. (van den Berg et al., 2010) propose a framework for planning and LQG-based feedback control of a steerable needle under motion and sensing uncertainty. This framework was extended by Patil et al. (Patil et al., 2011) for deformable environments. The LQG framework for estimation and control used in this approach does not address the issue of control saturation, which is a practical concern for needle steering. Prior LQG-based methods may fail due to control input saturation, which is a practical concern for needle steering, and cannot respond in real-time to significant perturbations not accounted for by the a priori model.

A core concept of this work involves the use of fast motion planning for obstacle avoidance and correcting the needle trajectory to account for uncertainties. There is considerable prior work in real-time replanning for dynamic environments using sampling-based planners and some of these techniques can also be applied to robots navigating in deformable environments under uncertainty (Frazzoli et al., 2002; Zucker et al., 2007). Some replanning approaches deal with discretized state spaces and propose efficient implementations of replanning algorithms that include D* and anytime A* algorithms based on classical heuristic search (van den Berg et al., 2006). Other methods
have been proposed to address the issue of safety in replanning in dynamic environments (Frazzoli et al., 2002). Model predictive control (MPC), also known as receding horizon control, is a form of replanning at each time step that formulates an optimal control problem truncated at some predetermined finite time horizon. Such techniques have been successfully applied to robot navigation (du Toit and Burdick, 2010). These methods address the issue of optimality of the computed motion plans but are difficult to scale to higher dimensional configuration spaces and continuous dynamical systems. Many approaches have sought to reduce the planning time to facilitate real-time replanning by reusing information from prior plans (Kuwata et al., 2009), using pre-computed coarse global plans to reduce planning workload (Zucker et al., 2007), or harness the computational power of modern hardware for efficient parallel planner implementations (Pan et al., 2010). Sampling-based planners can be practically applied to higher dimensional state spaces but the completeness of such methods has still not been fully explored (Hauser, 2012).

Several robot-assisted needle placement and insertion systems have been proposed for stiff needles, including systems that use either ultrasound or magnetic resonance tracking for biopsies (Hata et al., 2006; Boctor et al., 2008; Xu et al., 2010; Long et al., 2012; Seifabadi et al., 2012) and prostate high-dose rate brachytherapy (Susil et al., 2004; Fichtinger et al., 2008; Mozer et al., 2009; Hungr et al., 2012). These systems achieve targeting errors that are in the range of ≈ 1–4 mm. Experimental studies that compare the targeting error achieved by robotic needle insertion systems versus trained clinicians (Maier-Hein et al., 2009; Schouten et al., 2012) also report targeting errors of the same magnitude. The targeting accuracy of our system is within clinically acceptable thresholds with the added advantage of using steerable needles which allows us to avoid anatomical obstacles and minimize tissue damage and patient trauma.

Our sampling-based planner is customized for steerable needles (Patil and Alterovitz, 2010a; Patil et al., 2013) and we experimentally demonstrate in phantom and ex vivo animal tissues that our rapid replanning approach can be used for closed-loop needle steering in 3D environments. We have also integrated the planning approach with a robotic needle steering system to create the first fully integrated system that is capable of avoiding obstacles in 3D environments.
2.3 Objective

We specifically focus on bevel-tip steerable needles (Cowan et al., 2011), which move along an approximately circular arc of constant curvature $\kappa_0$ in the direction of the bevel when inserted into a tissue medium. The needle is controlled by two control inputs: insertion speed $v$ used to insert the needle and twist speed $\omega$ applied at the needle base used to reorient the bevel-tip.

To enable automatic needle steering, our system requires as input a specification of the anatomy. Given registered preoperative volumetric medical images that are standard in clinical care (e.g., CT scans or MRI), the clinician can specify the initial state of the needle tip $X_0 \in SE(3)$, a volumetric target region $\mathcal{P}_{\text{goal}} \subset \mathbb{R}^3$, and obstacles $\alpha_i \in \mathcal{O}$ that include sensitive structures such as glands or blood vessels and other obstacles such as bones that are represented as segmented volumes in the images. Our system also requires characterization of the steerable needle, including the natural maximum curvature of the needle $\kappa_0$ and the empirical relationship $\alpha = h[\kappa], 0 \leq \kappa \leq \kappa_0$ that relates the needle’s curvature to the duty cycling factor $\alpha$ (defined in Sec. 2.4).

The objective is to automatically steer the needle around clinician-specified anatomical obstacles while optimizing a clinician-specified clinical criteria. The criteria can include metrics such as minimizing insertion length (i.e., minimizing tissue damage) or maximizing clearance from obstacles (i.e., maximizing safety). Our approach, shown in Fig. 2.2, uses rapid replanning. Given the inputs specified above, the fast planner computes a large number of randomized plans, each defined as
a sequence of discrete controls that will steer the needle tip to the target region while avoiding anatomical obstacles. From the computed set of plans, the planner selects the plan that optimizes the clinical criteria. At the core of our rapid replanning approach is a fast motion planner for needle steering based on a customized rapidly exploring random tree (RRT) planner that incorporates variable curvature kinematics and a novel distance measure for planning (described in Sec. 2.5). To compensate for uncertain perturbations that occur during needle steering in tissues, the planning process is repeated at frequent intervals in a closed loop fashion using feedback from the electromagnetic tracking system to sense the needle tip pose at the beginning of each interval and replan a control sequence to reach the target region.

2.4 Variable Curvature Needle Kinematic Model

Our planner uses a variable curvature kinematic model of the motion of the needle tip’s trajectory as the needle is inserted in tissue. The kinematic model is deterministic and does not explicitly consider errors arising from factors such as tissue deformations, actuation errors, and noisy sensing. Our rapid replanning approach allows us to correct for these errors as they occur during the procedure.

We assume that the needle is flexurally flexible and torsionally stiff, i.e., the shaft exactly follows the needle tip, and the insertions and twists applied to the needle base are directly transmitted to the tip. The motion of the needle is then fully determined by the motion of the needle tip. The state of the entire needle is then described by the needle tip pose, represented as a $4 \times 4$ matrix $X = [R \ p] \in SE(3)$, where $p \in \mathbb{R}^3$ is the position of the needle tip and $R \in SO(3)$ is the rotation matrix that encodes the needle tip orientation relative to a world coordinate frame.

We extend the constant curvature unicycle kinematic model of the needle tip proposed by Webster et al. (Webster III et al., 2006) to consider the curvature $\kappa$ ($0 \leq \kappa \leq \kappa_0$) to be an additional control input parameter. Let $v$ be the insertion speed and $\dot{\omega}$ be the twist speed of the needle. Physically realizing the variable curvature model requires that the curvature $\kappa$ be realizable in terms of insertion speed $v$ and twist speed $\omega$, which are the only two physical control inputs to the system. We later show how the twist speed $\dot{\omega}$ can be converted to the physical twist speed $\omega$ using duty cycled spinning of the needle during insertion (Minhas et al., 2007).
Figure 2.3: (a) Local coordinate frame $X_t$ attached to the needle tip and a point in $\mathbb{R}^3: [x, y, z]^T$ defined in the local coordinate frame. The needle is inserted along the $z$-axis, and the needle rotates around a line parallel to the $x$-axis and passing through the point $[0, -r, 0]^T$. The circular arc traced out by the needle (shown in orange) is parameterized as a triplet $[l, \phi, r]$. (b) Needle plan composed of circular arcs of bounded curvature ($\kappa = 1/r$, $0 \leq \kappa \leq \kappa_0$) connects initial state $X_0$ and target region $P_{\text{goal}}$ while avoiding obstacles $O$.

Given the control input vector $u = [v, \hat{\omega}, \kappa]^T \in \mathbb{R}^3$, it is convenient to describe the kinematics in terms of the instantaneous twist $U \in se(3)$ expressed in the local coordinate frame attached to the needle tip (Fig. 2.3a), given by (Webster III et al., 2006; van den Berg et al., 2010):

$$U = \begin{bmatrix} \hat{\omega} \\ v \end{bmatrix}, \quad \hat{\omega} = \begin{bmatrix} 0 \\ 0 \\ \kappa \end{bmatrix}^T, \quad v = \begin{bmatrix} 0 \\ 0 \\ v \end{bmatrix}^T. \quad (2.1)$$

where the notation $[s]$ for a vector $s \in \mathbb{R}^3$ refers to the $3 \times 3$ skew-symmetric cross-product matrix. The discrete-time kinematics evolves over time interval $t$ as:

$$X_{t+1} = f[X_t, u, t] = X_t \exp(Ut). \quad (2.2)$$

where $\exp(\cdot)$ denotes the matrix exponential operator. Note that for the special case of $\kappa = \kappa_0$, $\hat{\omega} = \omega$ and Eqn. (2.2) reduces to the constant curvature kinematic model (Webster III et al., 2006).
Prior work on motion planning for steerable needles in 3D (Xu et al., 2009; Duindam et al., 2008, 2010; Park et al., 2010b) assumes $\kappa$ is a constant, which severely restricts the range of motion of the needle tip. This makes it difficult for planners to compute a feasible motion plan in 3D environments with obstacles, thus sacrificing optimality or completeness. In contrast, our motion planning method assumes a variable curvature kinematic model that allows us to compute trajectories composed of circular arcs of bounded curvature ($0 \leq \kappa \leq \kappa_0$) (as shown in Fig. 2.3b). This helps us to compute feasible motion plans with sub-second computation time.

We use results from Minhas et al. (Minhas et al., 2007), who demonstrated that any curvature $0 \leq \kappa \leq \kappa_0$ can be approximated by duty cycling the rotation of the needle, i.e., by alternating between (i) insertion without rotation, in which the needle follows a path of maximum curvature ($\kappa = \kappa_0$), and (ii) insertion with rotation, in which the needle moves straight ($\kappa = 0$) by spinning at a constant rate and stopping the spinning such that the tip is at the same axial angle every time. Duty cycling of steerable needles was successfully demonstrated in cadaver brains for neurosurgical procedures (Engh et al., 2010).

Let the control input $\mathbf{u} = [v, \dot{\omega}, \kappa]^T$ be applied over a time duration $\Delta$. Let $\delta$ be the duration of each duty cycling interval, which is composed of an insertion interval of duration $\delta_{\text{ins}}$ and a spin interval of duration $\delta_{\text{spin}}$, as illustrated in Fig. 2.4. Let $\alpha (0 \leq \alpha \leq 1)$ be the proportion of the time spent in spin intervals, i.e., $\alpha = \delta_{\text{spin}}/\delta$, where $\delta = \delta_{\text{ins}} + \delta_{\text{spin}}$. The empirical relationship between $\kappa$ and $\alpha$ is expressed as:

$$\alpha = h[\kappa], \ 0 \leq \kappa \leq \kappa_0,$$

(2.3)

where $h[\kappa]$ is dependent on the mechanical properties of the needle and tissue and is determined by fitting a polynomial function to the empirical data gathered during preoperative characterization experiments (as described in Sec. 2.4).

Duty cycling is implemented for needle steering by moving a fixed distance each cycle and spinning with a fixed twist speed $\omega_{\text{spin}}$. Given $\kappa$, we use Eqn. (2.3) to determine $\alpha$. Since the needle tip arrives at the same axial angle at the end of each spin interval, the duration of the spin interval $\delta_{\text{spin}} = (2k\pi/\omega_{\text{spin}}), k \in \mathbb{Z}$. We then compute the quantities $\delta = (\delta_{\text{spin}}/\alpha)$ and $\delta_{\text{ins}} = (\delta - \delta_{\text{spin}})$. 

20
\[ \Delta = \alpha \Delta \]

\[ \delta = \frac{2k\pi}{\omega_{\text{spin}}} \]

\[ \delta_{\text{spin}} = 2k\pi/\omega_{\text{spin}}, \quad k \in \mathbb{Z} \]

Figure 2.4: The time duration \( \Delta \) is split into three intervals of duration \( \delta \) each for \( \alpha = h[\kappa] = 1/3 \). Each interval is then composed of two intervals: (i) a spin interval of duration \( \delta_{\text{spin}} = (2k\pi/\omega_{\text{spin}}) \), \( k \in \mathbb{Z} \) in which the needle is both inserted and rotated, and (ii) an insertion interval of duration \( \delta_{\text{ins}} \) in which the needle is only inserted without any rotation.

The low level control inputs during a duty cycle interval are given by:

\[ v(t) = v, \quad 0 \leq t \leq \Delta / \delta \]

\[ \omega(t) = \begin{cases} \dot{\omega} + \omega_{\text{spin}} & \text{if } j \delta < t \leq j \delta + \delta_{\text{spin}} \\ \dot{\omega} & \text{if } j \delta + \delta_{\text{spin}} < t \leq (j + 1) \delta \end{cases} \]

where \( j \in \{0, 1, \ldots, \Delta / \delta\} \) and \( \Delta / \delta \) is the total number of duty cycle intervals required to span the duration \( \Delta \).

2.5 Rapid Replanning Approach

In this section, we present details of the individual components involved in our rapid replanning approach (Fig. 2.2) for closed-loop needle steering in 3D environments with obstacles.

2.5.1 Motion Planning

To enable motion planning for a rapid replanning approach, we create a fast motion planner for the steerable needle. We based our planner on a sampling-based rapidly exploring random tree (RRT) (LaValle, 2006), which is well suited for the under-actuated, nonholonomic steerable needle and also provides completeness guarantees, i.e., the probability of finding a solution converges to one, if a solution exists, as the number of samples approaches infinity.

The input to the planner is an initial state \( X_0 \), a target region \( \mathcal{P}_{\text{goal}} \), and the computation time available for planning \( \Gamma \). Our algorithm is based on the classic RRT, which proceeds as follows. The planner incrementally builds a tree \( \mathcal{T} \) over the state space, while satisfying nonholonomic motion
constraints of the system and avoiding obstacles in the environment. To expand the tree $T$, a random state $X_{\text{rand}}$ is sampled from the state space. The algorithm identifies a node in the tree $X_{\text{near}}$, that is closest to the sample $X_{\text{rand}}$, as defined by a specified distance metric $\rho[\cdot]$. The algorithm attempts to expand $T$ towards $X_{\text{rand}}$ based upon the best control input $u$ and the resulting state $X_{\text{new}}$ is added to the tree. This process is repeated until either the tree $T$ connects $X_0$ and $P_{\text{goal}}$ or the available computation time is exceeded, in which case the algorithm reports that a solution cannot be found. A feasible plan $\Psi$ is extracted from the tree by traversing it backwards from the goal node to the root.

For a nonholonomic system like the steerable needle, finding the best control input to a sampled state requires solving a difficult two-point boundary value problem of connecting two states in $SE(3)$. Prior RRT-based needle steering planners (Xu et al., 2009) avoid this by performing deterministic or uniform random sampling of control inputs to determine the best control input $[v, \omega]^T \in \mathbb{R}^2$ that leads the needle tip to a new state $X_{\text{new}}$ closest to $X_{\text{rand}}$. Since these methods also assume the constant curvature kinematic model, the limited range of motion of the needle tip requires a large number of control samples to make progress towards the sampled state. This results in wasted computational effort and is a major computational bottleneck.

To enable efficient planning, we customize the classic RRT algorithm for steerable needles by leveraging several observations and algorithmic improvements. We consider variable curvature kinematics (Sec. 2.4) and introduce a new distance metric $\rho[\cdot]$. We present each step of the algorithm (outlined in Alg. 1) in detail below. In the available computation time, we compute many feasible bounded curvature ($0 \leq \kappa \leq \kappa_0$) trajectories through 3D environments with obstacles.

The individual function definitions in Alg. 1 are as follows:

*random_point_in$\mathbb{R}^3()$: To avoid solving the SE(3) two-point boundary value problem or performing random sampling of control inputs, we sample a random point $p_{\text{rand}} \in \mathbb{R}^3$ in the workspace as opposed to sampling a random state in $SE(3)$. The sampled point can then be connected to a given state $X_{\text{near}} = \begin{bmatrix} R_{\text{near}} & p_{\text{near}} \end{bmatrix}$ directly using a circular arc parameterized by $[l, \phi, r]^T$, where $l$ is the arc length, $\phi$ is the change in orientation of the needle tip coordinate frame $X_{\text{near}}$ around the $z_{\text{near}}$-axis, and $r$ is the arc radius (Fig. 2.3a). Let $[x, y, z]^T = R_{\text{near}}^T(p_{\text{rand}} - p_{\text{near}})$ be the coordinates...
Algorithm 1: $\Psi \leftarrow \text{needle}_RRT\_\text{planner}(X_0, P_{\text{goal}}, \Gamma)$

1: $\mathcal{T} \leftarrow \text{initialize}_\text{tree}(X_0)$
2: $\tau \leftarrow 0$
3: while ($\mathcal{T} \cap P_{\text{goal}} = \emptyset \land \tau < \Gamma$) do
4:     $p_{\text{rand}} \leftarrow \text{random\_point\_in\_R}^3()$
5:     $X_{\text{near}} \leftarrow \text{nearest\_neighbor}(p_{\text{rand}}, \mathcal{T})$
6:     $u \leftarrow \text{control\_inputs}(X_{\text{near}}, p_{\text{rand}}, \Delta)$
7:     $X_{\text{new}} \leftarrow f[X_{\text{near}}, u, \Delta]$
8:     if collision\_free($X_{\text{near}}, X_{\text{new}}, u, \Delta$) then
9:         $\mathcal{T} \leftarrow \text{add\_vertex}(X_{\text{new}})$
10:        $\mathcal{T} \leftarrow \text{add\_edge}(X_{\text{near}}, X_{\text{new}}, u, \Delta)$
11:     end if
12:     if $p_{\text{new}} \in P_{\text{goal}}$ then
13:        $\Psi \leftarrow \text{extract\_plan}(\mathcal{T}, X_{\text{new}})$
14:     end if
15:     $\tau \leftarrow \text{update\_time}()$
16: end while
17: return $\Psi$

of $p_{\text{rand}}$ in the local coordinate frame of $X_{\text{near}}$. The parameters of the circular arc are then given by:

$$r = \frac{x^2 + y^2 + z^2}{2 \sqrt{x^2 + y^2}}$$  \hspace{1cm} (2.6)

$$\phi = \arctan(x, -y)$$  \hspace{1cm} (2.7)

$$l = r \theta = r \arctan(z, r - \sqrt{x^2 + y^2}).$$  \hspace{1cm} (2.8)

To accelerate motion planning for steerable needles, we incorporate two forms of biasing that empirically result in significant performance gains. First, we bias the growth of the tree $\mathcal{T}$ towards the target region $P_{\text{goal}}$ by sampling from $P_{\text{goal}}$ with a higher probability than the rest of the workspace. Second, whenever a new node $X_{\text{new}}$ is added to the tree, the planner attempts to connect $X_{\text{new}}$ to a randomly sampled point in $P_{\text{goal}}$.

control\_inputs($\cdot$): Given a circular arc parameterized as $[l, \phi, r]$ and a given time interval $\Delta$, we derive the augmented control input vector required to compute the new state of the needle tip $X_{\text{new}}$. First, we reorient the needle tip by $\phi$ radians such that the circular arc is contained in the plane defined by the $y$-$z$ axes in the reoriented local coordinate frame $X_{\text{near}}^r$, which is obtained by applying a rotation of $\phi$ radians around the $z$-axis to the current state $X_{\text{near}}$. We then compute the augmented control input $u = [v, \dot{\omega}, \kappa]^T$ that steers the needle tip along a circular arc of length $l$ and
radius \( r \) using the relations: \( v = l / \Delta \), \( \dot{\omega} = 0 \), and \( \kappa = 1 / r \). We compute \( X_{\text{new}} \) by applying \( u \) to the reoriented frame \( X_{\text{near}}^r \) for a time duration \( \Delta \) according to Eqns. (2.1) and (2.2).

nearest_neighbor(·): The efficiency with which the RRT algorithm is able to explore the state space is highly sensitive to the distance metric \( \rho[·] \) used to compute the nearest node in the tree. In the presence of nonholonomic constraints, widely used metrics like the Euclidean distance are a very poor approximation of the true distance between points in the constrained state space. The performance of the RRT planner degrades as a result of repeated attempts at extending the same nodes in the tree without making sufficient progress (Shkolnik et al., 2009).

We introduce a new distance metric customized for steerable needles that accounts for the needle’s nonholonomic constraint as well as the buckling of the needle in soft tissue. Since the needle has a maximum curvature \( \kappa_0 \), not all sampled points will be reachable from a given state because of the nonholonomic constraints of the needle. The reachable set from a state \( X_{\text{near}} = \begin{bmatrix} R_{\text{near}} & p_{\text{near}} & 0 \\ 0 & 1 \end{bmatrix} \) consists of all points that can be connected to \( p_{\text{near}} \) by a circular arc that has a radius \( r \geq 1 / \kappa_0 \) and is tangent to the \( z_{\text{near}} \)-axis of the local coordinate frame. This definition of the reachable set also directly relates to the distance metric \( \rho[·] \) that is used to select the tree node that is nearest to the sampled point \( p_{\text{rand}} \). Accordingly, we define the distance metric \( \rho[X_{\text{rand}}, p_{\text{rand}}] \) as the length of such a circular arc connecting \( p_{\text{rand}} \) and \( X_{\text{near}} \) if \( p_{\text{rand}} \) is in the reachable set of \( X_{\text{near}} \), and infinity otherwise, i.e.,

\[
\rho[X_{\text{rand}}, p_{\text{rand}}] = \begin{cases} 
  l(= r\theta) & \text{if } r \geq 1 / \kappa_0 \land \theta \geq 0 \\
  \infty & \text{otherwise}
\end{cases} \quad (2.9)
\]

This strategy restricts the search domain to only those nodes that are within the reachable set of the nearest node \( X_{\text{near}} \), thus increasing the likelihood of state space coverage (Shkolnik et al., 2009).

It is important to prevent buckling of the needle shaft, which may occur during insertion because of reaction forces from the tissue. This implies that not all points in the reachable set can be physically accessed by the steerable needle from some poses. In our experiments, we have observed that the needle starts to buckle roughly when the needle tip heading is greater than \( \pi / 2 \) radians from its initial orientation. Formally stating, given the current state \( X_{\text{near}} \) and a sampled point \( p_{\text{rand}} \) and \( z_{\text{rand}} \) is the insertion \( z \)-axis tangent to the circular arc connecting \( X_{\text{near}} \) and \( p_{\text{rand}} \) at \( p_{\text{rand}} \), the needle would buckle if \( R_{\text{near}} z_{\text{rand}} \cdot z_0 \leq 0 \), where \( z_0 \) is the insertion \( z \)-axis at the initial state \( X_0 \) in the world.
coordinate frame. We preclude such points from being added to the tree by setting the distance to these points to infinity.

To efficiently search for the nearest node in the tree to a sampled point \( p_{\text{rand}} \) according to this distance measure, we use the fact that the needle will not traverse a circular arc spanning more than \( \pi/2 \) radians due to buckling considerations and it follows that if the point is unreachable from a state \( X_{\text{near}} \), i.e., \( \rho[X_{\text{near}}, p_{\text{rand}}] = \infty \), then \( p_{\text{rand}} \) is unreachable from all descendant nodes in the sub-tree of the node corresponding to state \( X_{\text{near}} \). Hence, the nearest node can be found using a depth-first traversal of the tree, pruning parts of the tree if a node is encountered from which \( p_{\text{rand}} \) is not reachable. Even though the worst case complexity of the nearest neighbor search is still linear in the number of nodes in the tree, pruning away parts of the tree in practice results in a significant performance improvement, especially as the tree grows larger.

\texttt{collision\_free(\cdot)}: To enable obstacle avoidance, only collision free arcs are added to the tree. We check if the circular arc connecting \( X_{\text{near}} \) and \( p_{\text{rand}} \) is collision free by approximating it as a sequence of line segments and checking if all the segments are collision free. Since the obstacle definitions are obtained from segmentation of 3D scans, the obstacle meshes are likely to be non-manifold. We use the SOLID library (van den Bergen, 2004) for detecting collisions with arbitrary, polyhedral obstacles at interactive rates.

\texttt{extract\_plan(\cdot)}: When the position \( p_{\text{new}} \) of a newly added state \( X_{\text{new}} \) is found to lie in the target region \( P_{\text{goal}} \), the planner terminates. By traversing the tree \( T \) backwards from the goal state to the root, we obtain a trajectory composed of piecewise circular arcs of bounded curvature \( (0 \leq \kappa \leq \kappa_0) \). We extract a motion plan \( \Psi \) comprised of a discrete sequence of control inputs, in terms of the insertion speed \( v(t) \) and twist speed \( \omega(t) \), that guide the needle to the target along the computed trajectory. For each circular arc parameterized by a triplet \([l, \phi, r]\) in the trajectory, we first reorient the needle tip by \( \phi \) radians by applying a control input \( \omega = \omega_{\text{spin}} \) for a duration of \( \phi/\omega_{\text{spin}} \). We then compute the factor \( \alpha \) based on the curvature \( \kappa = 1/r \) using Eqn. (2.3). Given the control input \( u = [l/\Delta, 0, 1/r]^T \), we compute the controls \([v(t), \omega(t)]^T\) for traversing the circular arc in a plane using Eqns. (2.4) and (2.5).

### 2.5.2 Optimizing Clinical Metrics

We consider the following clinically motivated criteria \( c[\Psi] \) for quantifying plan optimality:
• Minimizing the total needle insertion length (shortest path), i.e., minimizing $c[\Psi] = \int_{0}^{T} v(t)\,dt$. This metric is relevant to procedures in vital organs such as the brain where limiting tissue damage is important (Field et al., 2001). Shortest trajectories, however, often pass in close proximity to obstacles, thereby increasing the likelihood of collisions. Since it is also important to avoid critical structures, we artificially enlarge all obstacles by a predefined safety buffer $\epsilon$ using Minkowski sums (van den Bergen, 2004).

• Maximizing the minimum clearance from obstacles (maximum clearance), i.e., maximizing $c[\Psi] = \max_{0<t\leq T} \min_{i \in O} d[p_t, o_i]$, where $d[p_t, o_i]$ is the distance of the needle tip $p_t$ from obstacle $o_i \in O$. Trajectories that have a greater minimum clearance from obstacles are safer because they are less likely to collide with anatomical obstacles when deviations occur. Such trajectories, however, tend to be longer, thereby increasing the amount of tissue cut during the procedure. This metric could be useful when obstacle avoidance is critical but other tissue damage is manageable, e.g., in liver or muscular tissue.

The correct choice of the optimization criterion will vary by specific procedure, and we will assume that the clinician will select $c[\Psi]$ based on the requirements of the procedure. To compute a plan that optimizes $c[\Psi]$ as best as possible in the allowable computation time, we use our fast, randomized planning algorithm to compute hundreds of different feasible motion plans in a second and then select the plan that performs best under the selected criterion.

We note that any sampling-based motion planner, including our method, cannot guarantee that a globally optimal solution will be found in a finite time interval. Methods like RRT* (Karaman and Frazzoli, 2010) can compute optimal motion plans as computation time is allowed to increase, but cannot guarantee optimality in finite time and will not be efficient for needle steering due to their requirement of a solver for two-point boundary value problems. Our method will explore the steerable needle’s state space and repeatedly generate independent paths in search of a higher quality solution, and the best found path will progressively improve over the duration of the time interval. A further advantage of our approach is that it is trivially parallelizable, allowing for plan quality to improve as the number of cores in modern multi-core architectures increases.
2.5.3 Fault Tolerance

We empirically observed that for larger insertion lengths, the cumulative uncertainty can cause the target region to fall outside the reachable set of the steerable needle. We handle such situations gracefully by using fault tolerant heuristics to improve targeting accuracy.

When the target region is unreachable, we artificially enlarge the target region to include points in the workspace that are still reachable from the current state of the needle tip. This allows us to use the planner to compute a feasible plan. If the target is still unreachable, we modify the existing plan by optimizing over a discretized set of possible twists $[0 \leq \phi < 2\pi]$ to select the best twist that would guide the needle tip closest to the target region. We continue execution until we either reach the target region or insert the needle tip past the target region, in which case we terminate execution.

We plan to investigate the use of advanced trajectory optimization approaches (Seiler et al., 2012) in future work to further improve targeting accuracy when the target falls outside the reachable set.

2.6 Experimental Setup

We describe our needle steering system, shown in Fig. 2.5, and our experimental setup.

2.6.1 System Components

Bevel-tip steerable needle: We use needles fabricated from nitinol. In our experiments, we used two needles with tube outer diameters of 0.92 mm and 0.88 mm, henceforth referred to as Needle 1 and Needle 2, respectively. To enable steering at tight curvatures, the needles (1) incorporate a hand-machined bevel tip, and (2) are pre-bent just behind the bevel tip (Reed et al., 2011).

Needle steering robot: The robotic actuation unit contains a single carriage actuated by a lead screw. The needle is gripped by a custom brass collet, which is housed in a rotary bearing attached to the carriage and is fixed to a toothed pulley via two set screws. The pulley is actuated via a belt drive by motors attached to the carriage. The needle is inserted through a hole in the front plate of the robot. Buckling of the needle during insertion is prevented using an external telescoping sheath (Webster III et al., 2006). The robot controls both needle insertion and axial twists via Maxon DC motors, with low-level PID control implemented using a Galil DMC 4080 Motion Control unit. Additional details
Figure 2.5: Our needle steering system consists of a needle steering robot, a pre-bent, bevel-tip steerable needle, and an electromagnetic tracking system. We present details of the hardware system in Sec. 2.6. We performed experiments using phantom tissue (shown here) and ex vivo porcine tissue.

of the needle steering system are available in (Das et al., 2010). Our planner, implemented in C++, runs on a PC and sends the control inputs to the robot controller via PCI bus.

**Electromagnetic tracking system**: Accurate needle steering requires sensing the state of the needle tip position and orientation. Highly accurate approaches for state estimation include using stereo cameras (Webster III et al., 2006; Reed et al., 2011) or fluoroscopic images (Majewicz et al., 2012), but these approaches either cannot be used in opaque media such as the human body or can result in high radiation exposure to the patient for longer procedures.

We use an electromagnetic tracking system (Aurora® v1, Northern Digital Inc., Canada) (Northern Digital Inc., 2012) for tracking the needle tip pose. Embedded within the tip of the tube is a 5-DOF magnetic tracking coil, the position and orientation of which (other than the roll about the needle axis) can be measured by the system. We estimate the roll of the needle using encoders on the servo motor that applies axial twists at the needle base. Electromagnetic tracking is a cost effective and non-invasive method for reliably sensing the state of the needle tip in opaque tissue. The manufacturer specifications for the standard deviation of the error in sensing the position along any given axis is 0.7 mm and in sensing an angle is 0.2° (Northern Digital Inc., 2012).
2.6.2 Tissue Sample Materials

Tissue phantom: We first evaluate our system using a tissue phantom composed of an animal-protein-based gel marketed as the Simulated Muscle Tissue Ballistic Test Media (Sim-Test) from Corbin™, Inc (Corbin Manufacturing and Supply, Inc., 2012). This material, which was used in prior needle steering experiments (Webster III et al., 2006), is a close match to muscle tissue in terms of density and elasticity. This material can be used consistently without refrigeration, unlike water-based or gelatin-based gels. It is also water-soluble and can be diluted and cast into desired shapes for experiments. We cast the Sim-Test material into a cuboidal block of approximate dimensions 11 cm × 7 cm × 15 cm for our experiments (Fig. 2.6).

Ex vivo porcine loin tissue: We also evaluate our system in fresh ex vivo porcine loin tissue. This particular tissue is from the central spine of the animal and is tender because it is primarily comprised of muscles that are used for posture rather than locomotion. The portion of porcine loin used in our experiments had approximate dimensions of 10 cm × 5 cm × 19 cm. It was inhomogeneous and comprised of both muscular and fatty tissue types (Fig. 2.6).

2.6.3 Needle Characterization

The approach presented in Sec. 2.5 requires that we characterize the maximum curvature of the needle $\kappa_0$ and the empirical relationship $h[\kappa]$ between the curvature $\kappa$ and the duty cycling factor $\alpha$. We empirically determined that $h[\kappa]$ is dependent on the mechanical properties of the needle and the tissue and is not necessarily linear as demonstrated by prior work with duty cycled needle steering in a gelatinous phantom (Minhas et al., 2007).

To construct the relationship $h[\kappa]$, we varied the value of $\alpha$ between 0 and 1 in increments of 0.1. We then computed the duration of the duty cycling interval $\delta$ for a time interval $\Delta = 1 \text{ sec}$ (Sec. 2.4).
Given a fixed insertion speed $v_{ins}$ and twist speed $\omega_{spin}$, we commanded the actuators during each duty cycling interval with control inputs computed by substituting $v = v_{ins}$ in Eqn. (2.4) and $\dot{\omega} = 0$ in Eqn. (2.5).

The application of these controls causes the needle tip to traverse a circular arc of variable curvature $\kappa$ in a plane. We performed repeated insertions of both needles for up to 10 cm in both the Sim-Test tissue phantom and ex vivo porcine loin tissue. We computed a best-fit polynomial curve with a fixed maximum degree ($= 3$) that minimized the sum of the squared errors of the data points from the curve. This curve defines the relationship $\alpha = h[\kappa]$. An important point to note is that the smaller the distance $v_{ins}\delta$ traveled by the needle tip in every duty cycling interval, the better the approximation of $\kappa$. However, we empirically observed that for an insertion distance per duty cycling interval of less than 0.5 cm, the effect of inserting the needle without spinning was negligible, i.e., the effective curvature was close to 0. This is important because it physically limited the interval lengths at which we could replan during closed-loop steering to at least 0.5 cm.

Figure 2.7: Characterization of the relationship $\alpha = h[\kappa]$ (Eqn. (2.3)) for Needle1 and Needle 2 in Sim-Test tissue phantom and ex vivo porcine loin tissue.
To determine the effective curvature $\kappa$ of the planar arc, we recorded the state of the needle tip $X_t = \begin{bmatrix} R_t & p_t \end{bmatrix}$ after the end of each duty cycling interval for $N$ such intervals. We observed that the needle tip deviated from the plane because of initialization errors and other sources of uncertainty. To robustly estimate $\kappa$, we fit a circle to the set of 3D points given by $p_t \in \mathbb{R}^3, t = 0, \ldots, N$. We accomplished this by first computing a best-fit plane that minimized the sum of the squared orthogonal distances from each point to the plane by performing principal component analysis (PCA) on the set of points. We then projected the points onto the first two principal components that span the plane and then fitted a circle to the set of projected 2D points using a robust circle fitting algorithm (Taubin, 1991). The curvature $\kappa$ was obtained by taking the reciprocal of the radius of this fitted circle. Fig. 2.7 shows the relationship $\alpha = h[\kappa]$ for Needle 1 and Needle 2 in Sim-Test tissue phantom and ex vivo porcine loin tissue. Needle 1 achieved a maximum curvature $\kappa_0 = 0.11 \text{ cm}^{-1}$ in Sim-Test (Fig. 2.7a). Needle 2 had a lesser outer diameter (0.88 mm) and achieved a maximum curvature of $\kappa_0 = 0.15 \text{ cm}^{-1}$ in Sim-Test and a maximum curvature of $\kappa_0 = 0.073 \text{ cm}^{-1}$ in porcine loin tissue (Figs. 2.7b and 2.7c). Fig. 2.7 also shows the best-fit curves for $h[\kappa]$ for each of the needle-tissue combinations considered. In particular, we found that any value of $\alpha > 0.5$ for duty cycled insertion in ex vivo porcine loin tissue resulted in a 0 effective curvature, which explains the lack of empirical data points in Fig. 2.7c.

2.7 Experimental Evaluation

We evaluate our new needle steering system in tissue phantoms and ex vivo porcine loin tissue to demonstrate the ability to steer needles to targets with clinically acceptable accuracy while avoiding obstacles. For all the experiments described below, we consider a spherical target region of 1 mm and measure the targeting accuracy of the needle tip by computing the distance between the center of this spherical target region and the final needle tip position after insertion. We executed the motion planner on an Intel® i7 3.33 Ghz PC. We set the replanning interval $\Delta$ to 1 sec and allocated 1 sec of computation time per replanning step, which is a sufficiently short time interval for clinical applications that require needle insertion depths of $\approx 10$–15 cm.
Figure 2.8: We perform experiments in a cuboidal shaped Sim-Test tissue phantom (see Fig. 2.6). We assume that the workspace contains obstacles (shown in yellow) for planning purposes. We selected 10 randomly chosen targets (shown in red) in the workspace that are located at insertion depths ranging from 9 cm to 11.5 cm from the face of the cuboidal block. The insertion location of the needle is marked in green and the needle is inserted into a face of the cuboidal block.

2.7.1 Evaluation in Tissue Phantoms

We first evaluated our needle steering system in the Sim-Test tissue phantom described in Sec. 2.6.2. We chose 10 random target regions in the workspace at distances ranging from 9 cm to 11.5 cm from the face of the cuboidal block through which the needle is inserted as shown in Fig. 2.8. To evaluate the accuracy of the proposed system, we performed 3 insertions for each of the 10 targets under closed-loop rapid replanning using Needle 1 (0.92 mm diameter) and the shortest path metric. We achieved a mean targeting error of 1.07 mm ($\pm$ 0.59 mm).

To assess the impact of uncertainty, we also performed an open-loop execution for each target and achieved an average error of 9.57 mm ($\pm$ 2.95 mm). The open-loop execution results show that, even for homogeneous tissue phantoms, perturbations due to uncertainty can lead to large errors if not corrected. Our rapid replanning
approach significantly improves the targeting accuracy by accounting for errors and perturbations as they occur.

2.7.2 Evaluation in Tissue Phantoms with Obstacles

We next evaluated the needle steering system in the Sim-Test tissue phantom with virtual (not physically embedded) obstacles. We created four scenes, shown in Fig. 2.8, with virtual obstacles. Scenes #1 and #2 contain spherical obstacles which obstruct the path to some of the considered targets. Scene #3 contains two box-like obstacles that create a narrow passage that the needle must go through before reaching the targets. Scene #4 is challenging since the obstacles create a narrow passage that force the needle to traverse two-bend trajectories around obstacles to reach the targets.

We first evaluated our approach using 3 insertions for each of the 10 targets in each scene. We used Needle 1 and the maximum clearance metric for these insertions. The mean targeting error for each of the scenes was 1.24 mm (± 0.71 mm), 1.29 mm (± 0.79 mm), 1.12 mm (± 0.9 mm), and 1.25 mm (± 0.84 mm), respectively. Even with obstacles that restrict the navigable space in the environment, our approach successfully steered the needle to the target without collisions in any of the insertions.

We also evaluated the impact of optimization criteria (i.e., shortest path or maximum clearance) on target accuracy and obstacle avoidance for each scene. We used Needle 2 and chose 3 out of the 10 target regions in the workspace. For the shortest path criterion, we enlarged all obstacles by a safety buffer of 5 mm. We performed 3 insertions for each of the two criteria using our rapid replanning approach. Fig. 2.9 shows the means and standard deviations of the targeting error for each optimization criteria. The maximum mean closed-loop rapid replanning targeting error was 1.7 mm for the shortest path criterion and 1.66 mm for the clearance criterion. To illustrate the impact of uncertainty, we also ran the system using an open-loop plan for each target and scene. For the open-loop insertions, the mean targeting errors were as high as 10 mm for the shortest path criterion and 9.1 cm for the clearance criterion. Our closed-loop, rapid replanning approach significantly reduces targeting errors compared to open-loop execution of motion plans. In terms of the optimization criteria, we found that both the shortest path and maximum clearance criteria perform equally well. It is important to note that the shortest path criterion requires the clinician
to specify the safety buffer around obstacles whereas the clearance criterion does not require any specification of parameters. We used a safety buffer of 5 mm for the shortest path experiments.

### 2.7.3 Evaluation in Ex Vivo Porcine Loin Tissue

We also evaluate our rapid replanning approach in ex vivo porcine loin tissue samples as shown in Fig. 2.6. We created two scenes with virtual obstacles (shown in Fig. 2.10). The two scenes are similar to scenes constructed earlier (Fig. 2.8) and are modified to take into account the different dimensions of the workspace. We use two cylindrical obstacles in Scene #1 and two box-like obstacles in Scene #2, which create a narrow passage and require the needle to traverse two-bend trajectories around the obstacles to reach the target regions.

We used Needle 2 for this set of experiments and considered 3 randomly chosen targets in the workspace shown in Fig. 2.10. We evaluated the system for both the shortest path and maximum clearance criteria in each of these scenes using 3 insertions per target for each criterion. For the shortest path criterion, we enlarged all obstacles by a safety buffer of 5 mm. As before, we also execute the system using an open-loop motion plan for each target for comparison.

Fig. 2.11 shows the mean targeting error and standard deviations of the targeting error for each of the two criteria for steering using our closed-loop rapid replanner and using an open-loop plan. The mean targeting error for the shortest path criterion for both scenes was 3.6 mm (± 1.85 mm) for
Figure 2.10: We perform targeting experiments in an ex vivo porcine loin tissue sample. We assume that the approximately cuboidal workspace contains virtual obstacles (shown in yellow) for planning purposes. The insertion location of the needle is marked in green and the needle is inserted into the face of the tissue sample. We selected 3 randomly chosen targets (shown in red) in the workspace that are located at insertion depths ranging from 10 cm to 11 cm from the insertion face.

Our closed-loop rapid replanner and 10 mm (± 2.6 mm) for open-loop steering. The targeting errors are larger than in Sim-Test phantom tissue because of the anisotropic nature of interaction between needle and tissue and heterogeneity of the tissue sample. In spite of the slightly larger errors, the targeting errors using our approach are within clinically acceptable thresholds and are significantly smaller than open-loop steering. The mean targeting error for the clearance criterion for both scenes was 2.6 mm (± 1.2 mm) for our closed-loop rapid replanner and 15.6 mm (± 3 mm) for open-loop steering. Two of the open-loop insertions collided with the virtual obstacles. In contrast, our rapid replanning approach steered the needle safely to the target region. In terms of the optimization criteria, we found that the maximum clearance criterion worked better than the shortest path criterion because of the narrow passage in the environment, which is further constricted by imposing an artificial safety buffer in case of the shortest path criterion.

2.7.4 Evaluation in Anthropomorphic Liver Phantom

We apply our needle steering system to a scenario motivated by the clinical task of ablating a tumor in the liver while avoiding the hepatic veins. We built the anthropomorphic liver phantom that models the hepatic veins based on the hepatic vein anatomical model provided by Desser et al. (Desser et al., 2003). In this experiment, the obstacles are physically embedded in a tissue phantom. We modeled the major hepatic veins (middle, left, and right) and the inferior vena cava using hollow tubing (see Fig. 2.12) so that the veins would be visible on preoperative CT images. We constructed a tumor from plastic that was roughly spherical and 5 mm in diameter. The tumor model
was coated with calcium sulfate so that it would be visible in the CT images. We placed the model veins, model tumor, and fiducial markers for registration in a box which we filled with Sim-Test to create the anthropomorphic tissue phantom (see Fig. 2.12b).

After the phantom was constructed, we used a portable flat-panel CT scanner to obtain preoperative images of the environment (see Fig. 2.12c). We specified 5 insertion locations on the surface of the box and specified 5 different target sites on the tumor for ablation. We also segmented the major vessels and the tumor from the CT scans to obtain obstacle meshes for planning. We used Needle 2 for this set of experiments. For each pair of insertion location and target region, we performed closed-loop steering using our rapid replanning approach using the clearance optimization criterion. We did not perform open-loop steering in this experiment to avoid damaging the needle in case it collided with the model veins during the procedure.

<table>
<thead>
<tr>
<th>Scene</th>
<th>Target #</th>
<th>Closed-loop Steering</th>
<th></th>
<th>Open-loop Steering</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Insertion Length (cm)</td>
<td>Target Error (mm)</td>
<td>Insertion Length (cm)</td>
<td>Target Error (mm)</td>
</tr>
<tr>
<td>SimTest #1</td>
<td>1</td>
<td>10.66</td>
<td>0.54</td>
<td>10.83</td>
<td>1.82</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>11.49</td>
<td>0.70</td>
<td>11.61</td>
<td>1.61</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>11.43</td>
<td>1.67</td>
<td>11.65</td>
<td>1.35</td>
</tr>
<tr>
<td>SimTest #2</td>
<td>1</td>
<td>10.68</td>
<td>0.92</td>
<td>10.70</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>11.51</td>
<td>0.68</td>
<td>11.57</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>11.57</td>
<td>0.38</td>
<td>11.66</td>
<td>0.89</td>
</tr>
<tr>
<td>SimTest #3</td>
<td>1</td>
<td>10.36</td>
<td>0.31</td>
<td>10.70</td>
<td>0.83</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>11.59</td>
<td>2.60</td>
<td>11.54</td>
<td>1.26</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>11.64</td>
<td>2.39</td>
<td>11.66</td>
<td>2.72</td>
</tr>
<tr>
<td>SimTest #4</td>
<td>1</td>
<td>10.57</td>
<td>0.85</td>
<td>10.64</td>
<td>1.56</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>10.68</td>
<td>1.13</td>
<td>11.60</td>
<td>0.58</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>11.58</td>
<td>2.95</td>
<td>11.39</td>
<td>2.85</td>
</tr>
<tr>
<td>Porcine Loin #1</td>
<td>1</td>
<td>10.31(±0.11)</td>
<td>2.69(±1.95)</td>
<td>10.47(±0.24)</td>
<td>2.01(±1.64)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>11.51(±0.36)</td>
<td>4.09(±1.39)</td>
<td>11.58(±0.16)</td>
<td>1.70(±0.37)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>11.34(±0.03)</td>
<td>3.58(±1.04)</td>
<td>11.37(±0.04)</td>
<td>3.02(±1.16)</td>
</tr>
<tr>
<td>Porcine Loin #2</td>
<td>1</td>
<td>10.34(±0.04)</td>
<td>2.48(±2.06)</td>
<td>10.34(±0.07)</td>
<td>3.77(±0.47)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>11.69(±0.24)</td>
<td>6.07(±1.34)</td>
<td>11.47(±0.01)</td>
<td>1.78(±0.54)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>11.45(±0.05)</td>
<td>2.60(±1.49)</td>
<td>11.42(±0.08)</td>
<td>3.26(±1.13)</td>
</tr>
</tbody>
</table>

1 Obstacles enlarged by a safety buffer of 0.5 cm

Table 2.1: Comparing cost metrics for needle insertion in Sim-test phantom tissues and ex vivo porcine loin tissue with virtual obstacles. Target coordinates used for these experiments (in cm): 1) [2.48, −0.6, 10.0], 2) [0.14, −0.1, 11.45], and 3) [1.27, −2.21, 11.17].
In each instance, our rapid replanning approach successfully steered the needle to the target region on the tumor surface while avoiding the hepatic veins, with an average error of 2.38 mm (± 1.02 mm) over up to 15.5 cm insertion length. The targeting error was higher in this case as compared to the targeting experiments in Sim-Test phantom tissue with virtual obstacles (see Fig. 2.9). We suspect this is because the embedded physical obstacles, unlike the virtual obstacles, were rigid and resulted in larger errors in the motion of the needle tip. We note that the rigid nature of our model differs from actual anatomical vessels and organs, which would deform with the surrounding tissue due to interaction forces between the needle and the tissue.

2.8 Discussion

We have presented a novel approach for unifying planning and control of steerable needles in 3D environments with obstacles and real-world perturbations. Our approach relies on a fast RRT motion planner for steerable needles that uses variable curvature kinematics and a novel distance measure for planning that speeds up motion planning to the point that it can be done in real time with typical needle insertion velocities. We use the fast performance to enable us to consider two clinically motivated optimization criteria: minimizing insertion length and maximizing clearance from critical anatomical structures. Our approach accounts for perturbations as they occur, thus eliminating

![Targeting Error (Porcine Loin)](image)

Figure 2.11: We compare the targeting error using closed-loop, rapid replanning steering and open-loop execution for the two proposed optimization criteria in ex vivo porcine loin tissue. Our approach significantly outperforms open-loop execution. Error bars indicate one standard deviation of the targeting error over repeated trials. Detailed experimental results are available in Tab. 2.1.
Figure 2.12: We applied our needle steering system with rapid replanning to a scenario motivated by the clinical task of ablating a tumor in the liver while avoiding the hepatic veins. (a) We constructed an anthropomorphic liver phantom that includes the major hepatic veins in the liver (right) based on an anatomical model provided by Desser et al. (Fig. 1 in (Desser et al., 2003)). The model was built to scale to match human liver dimensions and is shown next to a geometrically correct human liver model manufactured based on segmented CT images of a human patient. (b) We placed the model in a container that was filled with Sim-Test material to create the liver phantom for experiments. The hepatic vein model placed in a box along with fiducial markers for registration before it was cast in Sim-Test phantom tissue material. (c) We used a portable flat-panel CT scanner to obtain preoperative images of the environment while the electromagnetic tracking system provided measurements of the position and orientation of the needle tip during the procedure. (d) We specified the insertion location and target region and annotated segmented structures such as veins that needed to be avoided. We illustrate feasible motion plans (shown in green) computed at time step 1. (e) Via rapid replanning, our planner successfully guided the needle (reconstructed from CT scans after the procedure) between the middle and left hepatic veins to reach the target on the surface of the tumor.

the need for modeling complex phenomena such as needle and tissue deformation, needle-tissue interaction, and torsional build-up along the needle shaft. Our approach also eliminates the need for designing complex feedback controllers that try to guide the needle tip along a pre-planned trajectory by compensating for errors as they occur.
We also presented the first fully integrated, automated needle steering system that is capable of avoiding obstacles in 3D environments with real-world perturbations. We experimentally evaluated our system by performing procedures in tissue phantoms and ex vivo porcine loin tissue. In our anatomic liver scenario, we took steps toward demonstrating how this system could be used for clinical steerable needle procedures. Our experimental results demonstrate that our rapid re-planning strategy successfully guides the needle to desired targets while avoiding obstacles with an average error of less than 3 mm, which is within clinically acceptable thresholds and better than the accuracy achieved by trained clinicians. In addition to accuracy, our system offers the added advantage of automatically avoiding sensitive structures.

This research is another step towards realizing needle steering in actual clinical practice. There are several avenues for improving this work. First, tissue damage due to duty cycled spinning of the needle is a clinically-relevant concern. Previous studies in phantom tissues (Reed et al., 2011) and in vivo tissues (Engh et al., 2010) showed that duty cycled spinning can leave corkscrew trails within the cut tissue. Recently, Swaney et al. (Swaney et al., 2012) showed that flexible joints can be used to connect the bevel tip to minimize tissue damage during duty cycled spinning of the needle. Second, it is possible to use ideas from Patil et al. (Patil et al., 2011) to plan in deformable environments instead of quasi-static environments to increase the probability of successful plan execution in scenarios with very large deformations. Third, we assume that the measurements of the state of the needle tip obtained from the magnetic tracking system are accurate. Since the measurement noise (Northern Digital Inc., 2012) is very low when compared to the errors in the needle tip motion, this is not a major concern. However, there is always a possibility of improving the quality of localization of the needle tip and improve targeting accuracy by using a Kalman filter for state estimation (Kallem et al., 2010; van den Berg et al., 2010).

Further, we plan to investigate methods to avoid detailed characterization of the needle in tissue by estimating the curvature of the needle and estimate points on the calibration curve as the procedure is being performed. Finally, we would like to evaluate our approach in in vivo tissues where accurate needle characterization cannot always be performed and there are other sources of uncertainty such as involuntary patient motions and unforeseen needle tip deflections due to tissue membranes.
CHAPTER 3

Data-Driven Stochastic Models For Simulating Steerable Needle Procedures

In this chapter, we present a data-driven stochastic model of steerable needle insertion for simulating steerable needle procedures. We describe the need for a stochastic model of steerable needle insertion and the challenges associated with constructing such a model. In this work, we consider a model that incorporates a stochastic motion model of the needle tip pose and a stochastic measurement model of the partial and possibly noisy measurements of the needle tip pose. We describe an expectation maximization (EM) algorithm for estimating the parameters of the stochastic model from data gathered from experiments and prior procedures. We validate the stochastic model by comparing the targeting error achieved in simulated steerable needle procedures using our stochastic model vis-a-vis targeting errors achieved in needle procedures performed in tissue phantoms and ex vivo porcine loin tissue.

3.1 Simulating Steerable Needle Procedures

Computer simulations of surgical procedures could help clinicians to train in simulated environments constructed from preoperative imaging data without risks to patient safety. Studies indicate that surgical skills learned using computational simulators directly improve operating room performance by significantly decreasing procedure time and reducing the frequency of medical errors by sixfold compared to traditional training (Seymour et al., 2002; Satava, 2005; Gallagher et al., 2005). Simulations of surgical procedures have also been used for preoperative planning and optimization (Alterovitz, 2006; Taylor, 2006).

In the context of steerable needle procedures, simulations could help clinicians to train on a per-patient basis for semi-autonomous needle insertion procedures. Simulations could also be used
for preoperative planning such as optimizing the entry position and orientation of the needle tip to maximize the chances of successful plan execution (Alterovitz, 2006; Dehghan and Salcudean, 2009), or for optimizing the placement of sensors in the environment (van den Berg et al., 2010). Also, the effectiveness of motion planning and control algorithms extensively relies on the underlying robot motion model, which can take the form of a simulation.

For a simulation to be useful for the above applications, the simulation must be based on a model of steerable needle motion in human tissues that is of sufficient accuracy. However, an accurate model for simulation is difficult to specify and accurate simulation is difficult to achieve for several reasons, as described in Chapter 1. Modeling and analyzing all aspects and sources of uncertainty of steerable needle insertion remains a challenging problem in spite of many prior efforts (Cowan et al., 2011; Okamura et al., 2010). A secondary consideration is computational complexity of the model, which should be low for real-time interactive simulators and for simulators used by motion planners. Prior work has investigated models of varying complexity, ranging from simple kinematic models (Webster III et al., 2006) to complex finite element models (Chentanez et al., 2009) but fail to account for uncertainty in needle motion.

Simulations could be made more useful if we had a motion model of the needle tip that not only predicts the expected (mean) evolution of the needle tip as a function of the control inputs, but also gives information about the distribution of the possible outcomes under perturbations. A priori information about the distribution of possible needle states due to perturbations is important for modeling variability and designing effective planning and control algorithms that take uncertainty into account without the need to model all the factors responsible for this uncertainty. Also, any motion model requires parameters, which can be estimated solely using means or by distributions.

To this end, we deliberately choose a simple kinematic model and use a data-driven approach to learn time-dependent Gaussian distributions of pose uncertainty, which enables very fast simulation and accurately models needle steering in representative tissues. The simplified kinematic model describes the expected needle tip motion and we add stochasticity to the model to model the mismatch between modeling and reality. Fig. 3.1 gives an overview of our proposed approach for estimating the parameters of this stochastic model based on data collected from prior experiments and procedures.

For computational tractability, our simulation of needle steering begins by using a simplified model of needle steering that assumes that the needle bends to follow the needle tip exactly, i.e., the
Figure 3.1: Developing models for fast, yet accurate, simulations of steerable needle procedures is challenging. We propose a simplified, stochastic model of steerable needle insertion and estimate the parameters of this stochastic model from data obtained from prior experiments and procedures.

tissue does not deform the needle shaft. We also assume that the insertions and twists applied to the needle base are directly and exactly transmitted to the needle tip, i.e., there is no buckling or torsion along the needle shaft. This implies that the motion of the needle is fully determined by the motion of the needle tip. Instead of dealing with a high-dimensional coupled system, we model needle motion in a reduced 6D state space of the needle tip. Our objective is to address the uncertainty in steerable needle motion in soft tissue, which includes accounting for mismatch between modeling and reality.

One of the challenges that is addressed in this work is to estimate the stochastic modeling parameters in the presence of sensing uncertainty. This fact is overlooked by prior work in estimating model parameters for steerable needles which assume perfect measurements. In reality, most imaging and sensing modalities such as computed tomography (CT) imaging, 3D ultrasound tracking, or electromagnetic tracking systems have errors associated with measurements. We use the expectation maximization (EM) algorithm to alternate till convergence between Bayesian estimates of the needle tip poses that best explain the observed data and maximum likelihood estimation of the model parameters. The goal of our work is to create a stochastic model that captures needle behavior to inform the design, motion planning, state estimation, and control for robotic needle steering systems.

The remainder of this chapter is organized as follows. We discuss previous work on modeling of needle behavior and motion in soft tissue in Sec. 3.2, and formally state our objective for estimating a stochastic model of needle motion in Sec. 3.3. Sec. 3.4 provides an overview of extended Kalman
filtering and smoothing for steerable needles. We describe an expectation-maximization (EM) algorithm for estimating the stochastic model parameters in Sec. 3.5 and present results of parameter estimation from data gathered from multiple procedures performed in animal-protein-based gel and ex vivo porcine loin tissue in Sec. 3.6. We present simulation results for validation of our method in Sec. 3.7 and analyze our method and discuss future work in Sec. 3.8.

3.2 Related Work

Significant advancements have been made in modeling bevel-tip steerable needles (Cowan et al., 2011). A kinematic model of the nonholonomic motion of the needle tip was proposed and experimentally validated by Webster et al. (Webster III et al., 2006). The model was augmented to include the effects of torsional friction along the needle shaft (Reed et al., 2009; Swensen and Cowan, 2012). Okamura et al. (Okamura et al., 2004) identified interaction forces due to puncture, cutting, and friction that develop during needle insertion through tissue. Minhas et al. showed that the curvature of the needle path can be controlled through duty cycled spinning of the needle during insertion (Minhas et al., 2007). The mechanics of steerable needles during insertion in ex vivo tissue and in vivo tissue have been characterized by Majewicz et al. (Majewicz et al., 2012).

Motion planning and control for steerable needles in a plane (2D) has been extensively studied (Alterovitz, 2006; Bernardes et al., 2012; Kallem et al., 2010; Ko et al., 2011). Kallem (Kallem and Cowan, 2009) developed a controller that stabilizes the needle to a desired plane. Reed et al. (Reed et al., 2011) coupled this controller to a planar planner (Alterovitz et al., 2007) to create an integrated needle steering system. Motion planners have also been developed for needle steering in 3D environments with obstacles (Hauser et al., 2009; Xu et al., 2009; Duindam et al., 2010; Patil and Alterovitz, 2010a; Seiler et al., 2012). These planners and controllers require a model of the kinematics of needle tip and all the above works are based on the deterministic kinematic model proposed by Webster et al. (Webster III et al., 2006). A planning and control framework based on LQG control has been proposed for planning and control of steerable needles that consider the effect of uncertainty during insertion in static environments (van den Berg et al., 2010) and deformable tissue (Patil et al., 2011). This approach assumes a stochastic model of needle insertion but do not address the issue of characterization of the stochastic modeling parameters. Park et al. (Park et al.,
2010b) proposed a path-of-probability algorithm based on diffusion-based error propagation that considers uncertainty characterized using data-driven models (Park et al., 2010a) but this work only considers uncertainty in the control inputs, only considers data based on inserting the needle without any twists, and does not consider sensing uncertainty or a variable curvature kinematic model.

The challenge in coming up with stochastic models for steerable needle procedures is that we cannot learn patient-specific models because of lack of availability of data till the procedure is actually performed. Webster et al. (Webster III et al., 2006) fit empirical model parameters to a deterministic kinematic model of the needle tip motion based on data gathered from repeated insertions within phantom tissue. Since this is not feasible for clinical procedures, Misra et al. (Misra et al., 2010) proposed a mechanics-based model parametric model of needle-tissue interaction for asymmetric-tip steering that captures the salient behavior of needle-tissue interaction but such a model is computationally complex and is not suitable for fast planning and control of needles. Chentanez et al. (Chentanez et al., 2009) created a physically-based 3D simulation that uses nonlinear FEM for modeling needle and tissue deformations but this model is also computationally expensive and numerical stability is an issue with full-blown physically-based simulation.

Since it is difficult to develop an accurate model of needle motion in soft tissue, we instead use a generalized kinematic motion model and add stochasticity to the model to subsume the mismatch between modeling and reality. We then refine the parameters of our stochastic model to incorporate data from experiments and actual procedures as it is made available. Our expectation-maximization (EM) (Dempster et al., 1977) approach is very similar for estimating parameters of a stochastic model for training Kalman filters (Abbeel et al., 2005), system identification of dynamic systems (Ghahramani and Hinton, 1996), and quantifying statistical similarity between simulation models and real-world data for complex, aggregate systems like human crowds (Guy et al., 2012). We propose a stochastic model specific to steerable needle insertion for the purposes of simulation for facilitating preoperative optimization of clinical procedures and motion planning and control of steerable needles in soft tissue without having to model all the complex phenomena that govern steerable needle insertion.
3.3 Preliminaries and Objective

We next formally define our stochastic kinematic model and establish the objective of our work.

3.3.1 State Representation

The complete state of the steerable needle during the procedure is completely characterized by the high-dimensional coupled system comprising of both the needle and the tissue. To make the problem tractable, we assume that the state of the needle is completely described by the needle tip pose (as described in Chapter 1). The pose can be represented as a $4 \times 4$ matrix $X = \begin{bmatrix} R & p \\ 0 & 1 \end{bmatrix} \in SE(3)$, where $p \in \mathbb{R}^3$ is the position of the needle tip and $R \in SO(3)$ is the rotation matrix that encodes the needle tip orientation relative to a world coordinate frame.

State estimation and control in the $SE(3)$ group is a hard problem (Park et al., 2008) because it is difficult to preserve the orthogonality constraint on the rotation matrix $R$. We choose an equivalent, alternate representation of the pose for the sake of convenience. The pose can also be represented by a vector $x = [p \; r] \in \mathbb{R}^6$, comprising of the position $p \in \mathbb{R}^3$ and the orientation is described as a rotation of angle $||r||$ about axis $r \in \mathbb{R}^3$.

The $X \in SE(3)$ representation is convenient for kinematic modeling and motion planning purposes but the $x \in \mathbb{R}^6$ is convenient for state estimation and control purposes. We can conveniently transform between the two state representations. Given a state $X = \begin{bmatrix} R & p \\ 0 & 1 \end{bmatrix}$, the equivalent state $x \in \mathbb{R}^6$ is given by

$$x = \begin{bmatrix} p \\ r \end{bmatrix}, \text{ where } \begin{bmatrix} r \\ p \\ 0 & 1 \end{bmatrix} = \log(X),$$

(3.1)

where $\log(\cdot)$ is the matrix logarithm that defines a mapping from $SE(3) \rightarrow se(3)$ (Murray et al., 1994) and the notation $[r]$ for a vector $r = [r_x, r_y, r_z]^T \in \mathbb{R}^3$ refers to the $3 \times 3$ skew-symmetric cross-product matrix

$$[r] = \begin{bmatrix} 0 & -r_z & r_y \\ r_z & 0 & -r_x \\ -r_y & r_x & 0 \end{bmatrix}.$$  

(3.2)
Conversely, given a state $x = [P]$, the equivalent state $X \in SE(3)$ is given by

$$X = \exp\left(\begin{bmatrix}[r] & P \\ 0 & 1 \end{bmatrix}\right),$$  

(3.3)

where $\exp(\cdot)$ is the matrix exponential that defines a mapping from $se(3) \to SE(3)$ and has a convenient closed-form analytical expression (Murray et al., 1994).

### 3.3.2 Stochastic Kinematic Model

As described in Chapter 2, using a variable curvature kinematic model for describing the needle tip motion has several advantages over prior approaches that used a constant curvature kinematic model that severely restricts the range of motion of the needle tip.

Let $v$ be the insertion speed and $\dot{\omega}$ be the twist speed of the needle. We extend the constant curvature model proposed by Webster et al. (Webster III et al., 2006) to consider the curvature $\kappa$ ($0 \leq \kappa \leq \kappa_0$) as an additional control input parameter. The control input vector to the variable curvature model is denoted as $u = [v, \dot{\omega}, \kappa]^T \in \mathbb{R}^3$. The variable curvature $\kappa$ can be realized by duty cycled rotation of the needle (as described in Chapter 2).

Since it is challenging to develop accurate models of the kinematics of the needle tip, we deliberately use a simplified model for the expected motion of the needle tip and represent the cumulative effect of all the sources of uncertainty, including the mismatch between modeling and reality, using an additive stochastic error term. The central limit theorem (CLT), loosely speaking, suggests that the resultant distribution arising from the combination of several independent sources of error can be well modeled as a Gaussian. We therefore represent the error term to be drawn from a Gaussian distribution with mean $\mu_t$ and variance $M_t$ at each time $t$.

Formally stating, we represent the discretized stochastic kinematics of the needle tip in terms of the needle tip pose $x_t \in \mathbb{R}^6$ and the control input $u_t \in \mathbb{R}^3$ applied at time $t$, as:

$$x_{t+1} = f[x_t, u_t] + m_t, \quad m_t \sim \mathcal{N}[\mu_t, M_t],$$  

(3.4)
where \( m_t \in \mathbb{R}^6 \) represents the cumulative error between the true state at the next time step \( x_{t+1} \) and the state estimated by the function \( f \) and is assumed to be drawn from a Gaussian distribution. Eqn. (3.4) can be equivalently expressed as

\[
x_{t+1} \sim \mathcal{N}[f[x_t, u_t] + \mu_t, M_t]
\] (3.5)

We base our function \( f \) on the idealized kinematic model proposed by Webster et al. (Webster III et al., 2006). This is the same variable curvature kinematic model considered in Chapter 2 and is provided here for the sake of completeness. Given the control input vector \( u = [v, \hat{\omega}, \kappa]^T \), it is convenient to describe the kinematics in terms of the instantaneous twist \( U \in se(3) \) expressed in a local coordinate frame attached to the needle tip, given by (van den Berg et al., 2010):

\[
U = \begin{bmatrix}
[\hat{\omega}] & v \\
0 & 0
\end{bmatrix}, \quad \hat{\omega} = \begin{bmatrix} v\kappa & 0 & \hat{\omega} \end{bmatrix}^T, \quad v = \begin{bmatrix} 0 & 0 & v \end{bmatrix}^T.
\] (3.6)

where the notation \([s]\) for a vector \( s \in \mathbb{R}^3 \) refers to the \( 3 \times 3 \) skew-symmetric cross-product matrix. The discrete-time kinematics evolves over time interval \( t \) as:

\[
X_{t+1} = f[X_t, u, t] = X_t \exp(U t).
\] (3.7)

where \( \exp(\cdot) \) denotes the matrix exponential operator. In this case, \( \exp : se(3) \rightarrow SE(3) \) denotes the exponential map (Murray et al., 1994).

### 3.3.3 Stochastic Measurement Model

As the steerable needle is inserted into tissue, noisy and possibly partial sensor measurements of the needle tip state are obtained according to a known stochastic measurement model:

\[
\tilde{z}_t = h[x_t] + n_t, \quad n_t \sim \mathcal{N}[0, N_t],
\] (3.8)

where \( \tilde{z}_t \) is the measurement obtained at time \( t \) and \( n_t \) is the zero-mean Gaussian noise with variance \( N \) that models the sensing uncertainty, which is usually made available from manufacturer
specifications or can be estimated using system identification. Eqn. (3.8) can also be expressed as:

\[ z_t \sim \mathcal{N}[h[x_t], N_t]. \quad (3.9) \]

For instance, if we use a 5DOF sensor of an electromagnetic tracking system (Northern Digital Inc., 2012), we might obtain measurements of 5 of the 6 degrees of freedom and the sixth degree of freedom, which is the needle twist, will have to be estimated. Similarly, if we use 3D ultrasound to track the needle tip, we can only measure the position of the needle tip and the orientation of the tip has to be estimated.

### 3.3.4 Problem Definition

The problem can now be formally defined as follows:

**Input:** We are given a set of \( K \) needle trajectories \( \Gamma = \{ \gamma^1, \ldots, \gamma^K \} \), obtained from executing steerable needle motion plans within tissue. Each trajectory \( \gamma^k \) comprises of a series of measurements \( z^k_{0:T_k} = \{ z^k_0, \ldots, z^k_{T_k} \} \) that provide a noisy (and potentially partial) measurement corresponding to the (unknown) true state of the needle tip \( x^k_{0:T_k} = \{ x^k_0, \ldots, x^k_{T_k} \} \) at the corresponding time step. We assume that the given trajectories are representative of the errors encountered during steerable needle procedures and span the space of all control input and state output responses. We are also given the sequence of control inputs \( u^k_{0:T_k-1} = \{ u^k_0, \ldots, u^k_{T_k-1} \} \) that are applied to the base of the steerable needle at the corresponding time steps.

**Output:** The objective is to estimate the means \( \mu_t \) and the variances \( M_t \) of the error distribution (Eqn. (3.4)) that best explain the series of measurements \( \{ z^1_{0:T_1}, \ldots, z^K_{0:T_K} \} \) obtained from the \( K \) trajectories. The goal is to estimate the parameters of the stochastic model for simulating steerable needle procedures for preoperative procedure optimization, motion planning, state estimation, and control of steerable needles.

### 3.4 State Estimation

We provide details of our state estimation framework for estimating the pose of the needle tip, which is later used for estimating the stochastic model parameters. We use an extended Kalman
smoother (Simon, 2006) for state estimation, which requires linear(ized) kinematics and measurement models. It is also important to account for the non-zero mean of the error term $\mu_t$. We use estimates of $\mu_t$ and $M_t$ for estimating the state distributions and we assume that the measurement noise variances $N_t$ are known. We will later show how the estimates of $\mu_t$ and $M_t$ are revised.

### 3.4.1 Extended Kalman Filter

Given a trajectory $\gamma$, we first need to estimate the most likely true states of the needle tip $x_{0:T}$ given the set of noisy measurements $z_{0:T}$. This estimation can be performed using Bayesian inference and we use the extended Kalman filter (EKF) for this purpose. The EKF keeps track of the estimate $\hat{x}_t$ of the state $x_t$ and variance $P_t$ of the true state $x_t$ (which is assumed to be Gaussian) during execution. It continually performs two steps; a control update to propagate the applied control input $u_t$, and a measurement update to incorporate the measurement $z_t$.

We will use the following notation (where $E[\cdot|\cdot]$ is the conditional expectation):

\[ \hat{x}_t = E[x_t|z_{0:t}], \quad P_t = E[(x_t - \hat{x}_t|t)(x_t - \hat{x}_t|t)^T|z_{0:t}], \]
\[ \hat{x}_{t+1} = E[x_{t+1}|z_{0:t}], \quad P_{t+1} = E[(x_{t+1} - \hat{x}_{t+1}|t)(x_{t+1} - \hat{x}_{t+1}|t)^T|z_{0:t}], \quad (3.10) \]

We assume that the distribution of the initial state is known, i.e., $x_0 \sim N[\hat{x}_{0|0}, P_{0|0}]$ is given. Given Eqns. (3.4) and (3.8), the control and measurement updates proceed as follows:

**Control update step:**

\[ \hat{x}_{t+1|t} = f[\hat{x}_t|t, u_t] + \mu_t, \quad (3.11) \]
\[ P_{t+1|t} = A_t P_t A_t^T + M_t, \quad (3.12) \]

where $A_t$ is the Jacobian matrix given by $\frac{\partial f}{\partial x}[\hat{x}_t|t, u_t]$.

**Measurement update step:**

\[ K_{t+1} = P_{t+1|t} H_{t+1}^T (H_{t+1} P_{t+1|t} H_{t+1}^T + N_t)^{-1}, \quad (3.13) \]
\[ \hat{x}_{t+1|t+1} = \hat{x}_{t+1|t} + K_{t+1}(z_{t+1} - h[\hat{x}_{t+1|t}]), \quad (3.14) \]
\[ P_{t+1|t+1} = (I - K_{t+1} H_{t+1}) P_{t+1|t}, \quad (3.15) \]
where $H_{t+1}$ is the Jacobian matrix given by $\frac{\partial h}{\partial \hat{x}[\hat{x}_{t+1}|t]}$.

The Jacobian matrices can either be computed by numerical finite differences or by using the analytical expressions for the Jacobians provided by (van den Berg et al., 2010), which are exact and do not suffer from numerical errors.

### 3.4.2 Extended Kalman Smoother

The Kalman filter iteratively estimates the state distributions $N[\hat{x}_t|t, P_t|t]$ based on measurements received in the past relative to time step $t$. However, once execution finishes, we can incorporate measurements received in the future relative to time $t$ to obtain a more refined estimate.

Estimators that take into account both past and future are often called *smoothers*. The Kalman smoother estimates the distribution $N[\hat{x}_t|T, P_t|T]$ by first performing a forward pass of the Kalman filter in time. That allows us to compute the distributions $N[\hat{x}_t+1|t, P_t+1|t]$ and $N[\hat{x}_t+1|t+1, P_t+1|t+1]$ for time $0 \leq t < T$. We then perform a backward pass in time that updates the estimates of the state distributions as follows:

**Kalman smoothing backward step:**

\[
L_t = P_t|t A^T_t P_{t+1}|t, \quad (3.16)
\]
\[
\hat{x}_t|T = \hat{x}_t|t + L_t (\hat{x}_{t+1}|T - \hat{x}_{t+1}|t), \quad (3.17)
\]
\[
P_t|T = P_t|t + L_t (P_{t+1}|T - P_{t+1}|t) L^T_t. \quad (3.18)
\]

Note that $(P_{t+1}|T - P_{t+1}|t) < 0$ as the uncertainty over $\hat{x}_{t+1}$ is smaller when conditioned on all observations, than when only conditioned only on past observations. The Kalman smoother is used in the post-processing step to yield $\hat{x}_t|T$ as the optimal estimate of the state at time $t$ and the variance $P_t|T$ serves as a measure of uncertainty around the state estimate.

### 3.5 Expectation Maximization for Parameter Estimation

For the sake of brevity, we only consider a single trajectory $\gamma$ composed of $T$ time steps and we will later generalize to a set of trajectories $\Gamma = \{\gamma^1, \ldots, \gamma^K\}$. Given a set of measurements $z_{0:T}$, the objective is to estimate the parameters of the stochastic model $\mu_t$ and $M_t$, $t \in \{0, \ldots, T\}$, such
that the measurements become as likely as possible. This can be formally stated as follows:

\[
\{\mu_{0:T}, M_{0:T}\} = \arg\max_{\mu_t \in \mathbb{R}^6, M_t \in \mathcal{M}^6} L[\mu_{0:T}, M_{0:T}|z_{0:T}] = \arg\max_{\mu_t \in \mathbb{R}^6, M_t \in \mathcal{M}^6} p[z_{0:T}|\mu_{0:T}, M_{0:T}], \tag{3.19}
\]

where the likelihood \(L[\cdot]\) is a function of the parameters of a statistical model and the likelihood of a set of parameter values \(\mu_{0:T}, M_{0:T}\), given measurements \(z_{0:T}\), is equal to the probability of those measurements given those parameter values. Also, \(\mathcal{M}^6\) is the set of all \(6 \times 6\) symmetric positive definite matrices.

We use the expectation-maximization (EM) algorithm (Dempster et al., 1977) to find the maximum likelihood estimates of the parameters \(\mu_{0:T}\) and \(M_{0:T}\). The principal idea of the EM algorithm is to find the maximum likelihood estimates of parameters of a model, where the model depends on unobserved latent variables in addition to unknown parameters and known data observations. It is mathematically convenient to maximize the log-likelihood instead of the likelihood since the logarithm cancels against the exponent in the probability density function of a Gaussian distribution, which is equivalent since the logarithm is a monotonic function. The EM iteration alternates between performing an expectation (E) step, which creates a function for the expectation of the log-likelihood evaluated using the current estimate for the parameters, and a maximization (M) step, which computes parameters maximizing the expected log-likelihood found on the E step. These parameter-estimates are then used to determine the distribution of the latent variables in the next E step.

Instead of directly maximizing the joint log-likelihood of the measurements, we first infer the distributions of true states \(N[\hat{x}_{t+1|T}, P_{t+1|T}]\) using our best guesses of \(\mu_t\) and \(M_t\), and then use these distributions to estimate \(\mu_{0:T}\) and \(M_{0:T}\) by maximizing the joint log-likelihood. This iterative approach to estimation is guaranteed to converge in a coordinate-ascent manner to a locally optimal estimate of the distributions of the true states and the stochastic model parameters \(\mu_{0:T}\) and \(M_{0:T}\).

The objective now is to maximize the joint log-likelihood \(LL[\mu_{0:T}, M_{0:T}|x_{0:T}, z_{0:T}]\) given by:

\[
\log p[x_{0:T}, z_{0:T}|\mu_{0:T}, M_{0:T}] = \log \left( \prod_{t=0}^{T} p(x_t|x_{t-1})p(z_t|x_t) \right)
\]

\[
= \sum_{t=0}^{T} \log p[x_t|x_{t-1}] + \sum_{t=0}^{T} \log p[z_t|x_t]. \tag{3.20}
\]
**E step:** Using the relation \( x_t \sim \mathcal{N}[f[x_{t-1}, u_{t-1}] + \mu_t, M_t] \) from Eqn. (3.5), the first term in Eqn. (3.20) evaluates to:

\[
\sum_{t=0}^{T} \log p[x_t|x_{t-1}] = \sum_{t=0}^{T} \frac{e^{-\frac{1}{2}(x_t-f[x_{t-1}, u_{t-1}]-\mu_t)^T M_t^{-1}(x_t-f[x_{t-1}, u_{t-1}]-\mu_t)} (2\pi)^{\dim[z]/2} |M_t|^{1/2}}{(2\pi)^{\dim[z]/2} |M_t|^{1/2}}
\]

\[
= \sum_{t=0}^{T} \left( -3 \log(2\pi) + \frac{1}{2} \log |M_t| - \frac{1}{2} (x_t - f[x_{t-1}, u_{t-1}] - \mu_t)^T M_t^{-1}(x_t - f[x_{t-1}, u_{t-1}] - \mu_t) \right).
\]

Similarly, using \( z_t \sim \mathcal{N}[h[x_t], N] \) from Eqn. (3.9), the second term in Eqn. (3.20) evaluates to:

\[
\sum_{t=0}^{T} \log p[z_t|x_t] = \sum_{t=0}^{T} \frac{e^{-\frac{1}{2}(z_t-h[x_t])^T N_t^{-1}(z_t-h[x_t])}}{(2\pi)^{|z|/2} |N_t|^{1/2}}
\]

\[
= \sum_{t=0}^{T} \left( -\frac{|z|}{2} \log(2\pi) + \frac{1}{2} \log |N_t| - \frac{1}{2} (z_t - h[x_t])^T N_t^{-1}(z_t - h[x_t]) \right)
\]

Since Eqn. (3.22) does not contain \( \mu_{0:T} \) or \( M_{0:T} \), we ignore this term henceforth because it does not contribute to the maximization of the log-likelihood. Hence, the log-likelihood is now given by:

\[
LL[\mu_{0:T}, M_{0:T}|x_{0:T}, z_{0:T}]
\]

\[
= \sum_{t=0}^{T} \left( \frac{\log |M_t|}{2} - \frac{1}{2} (x_t - f[x_{t-1}, u_{t-1}] - \mu_t)^T M_t^{-1}(x_t - f[x_{t-1}, u_{t-1}] - \mu_t) \right) + \Phi
\]

\[
= \sum_{t=0}^{T} \frac{\log |M_t|}{2} - \frac{\text{Tr}}{2} \left( \sum_{t=0}^{T} M_t^{-1}(x_t - f[x_{t-1}, u_{t-1}] - \mu_t)(x_t - f[x_{t-1}, u_{t-1}] - \mu_t)^T \right) + \Phi.
\]

(3.23)

where \( \Phi \) is an expression independent of both \( \mu_{0:T} \) and \( M_{0:T} \), and we use the identities \( a = \text{Tr}(a) \) if \( a \) is a scalar, \( \text{Tr}(A + B) = \text{Tr}(A) + \text{Tr}(B) \) and \( \text{Tr}(AB) = \text{Tr}(BA) \) for matrices \( A \) and \( B \).

Since this expression cannot be directly evaluated since the true states \( x_{0:T} \) are unknown, we compute the expectation of the joint log-likelihood conditioned on the set of measurements \( z_{0:T} \), i.e., \( E[LL[\mu_{0:T}, M_{0:T}|x_{0:T}, z_{0:T}]]_{z_{0:T}} \) using an extended Kalman smoother (Sec. 3.4.2).
expectation of the log-likelihood is given by:

\[
E[LL(\mu_{0:T}, M_{0:T}|x_{0:T}, z_{0:T})|z_{0:T}]
\]

\[
= \sum_{t=0}^{T} \log |M_t^{-1}| \frac{1}{2} - \frac{1}{2} \text{Tr} \left( \sum_{t=0}^{T} M_t^{-1} E[(x_t - f[x_{t-1}, u_{t-1}] - \mu_t)(x_t - f[x_{t-1}, u_{t-1}] - \mu_t)^T|z_{0:T}] \right). \tag{3.24}
\]

We use the first order Taylor series approximation for \( f[x_{t-1}, u_{t-1}] \) by linearizing around \( \hat{x}_{t-1}|T \) to obtain the approximation \( f[x_{t-1}, u_{t-1}] \approx f[\hat{x}_{t-1}|T, u_{t-1}] + A_{t-1}(x_{t-1} - \hat{x}_{t-1}|T) \), where \( A_{t-1} \) is the Jacobian matrix given by \( \frac{\partial f}{\partial x}[x_{t-1}|T, u_{t-1}] \). Substituting these terms in the expectation term of Eqn. (3.24), the expectation term simplifies to (van den Berg, 2010):

\[
E[(x_t - f[x_{t-1}, u_{t-1}] - \mu_t)(x_t - f[x_{t-1}, u_{t-1}] - \mu_t)^T|z_{0:T}]
\]

\[
= P_{t|T} + A_{t-1}P_{t-1}|T A_{t-1}^T - P_{t|T}L_{t-1}^T A_{t-1}^T - A_{t-1}L_{t-1}P_{t|T}
\]

\[
+ (\hat{x}_{t|T} - f[\hat{x}_{t-1}|T, u_{t-1}] - \mu_t)(\hat{x}_{t|T} - f[\hat{x}_{t-1}|T, u_{t-1}] - \mu_t)^T. \tag{3.25}
\]

The expectation of the log-likelihood \( E[LL(\mu_{0:T}, M_{0:T}|x_{0:T}, z_{0:T})|z_{0:T}] \) is now completely given by Eqns. (3.24) and (3.25).

**M step:** We then determine \( \mu_{0:T} \) that maximizes the expectation of the log-likelihood by individually taking the derivatives of the expression in Eqn. (3.24) with respect to \( \mu_t, t \in \{0, \ldots, T\} \) and equating it to zero to give:

\[
\mu_t = (\hat{x}_{t|T} - f[\hat{x}_{t-1}|T, u_{t-1}]). \tag{3.26}
\]

Similarly, we can find \( M_{0:T} \) that maximizes the log-likelihood by taking the derivative of the expression in Eqn. (3.24) individually with respect to \( M_t^{-1} \) (which is equivalent to maximizing with respect to \( M_t \)) and equating it to zero to give:

\[
M_t = E[(x_t - f[x_{t-1}, u_{t-1}] - \mu_t)(x_t - f[x_{t-1}, u_{t-1}] - \mu_t)^T|z_{0:T}], \tag{3.27}
\]

where the expectation term is evaluated using Eqn. (3.25).
We generalize this to estimate $\mu_{0:T}$ and $M_{0:T}$ across the set of trajectories $\Gamma = \{\gamma^1, \ldots, \gamma^K\}$ by summing the expectation terms over all trajectories to compute the following solution:

$$
\mu_t = \frac{1}{K} \sum_{k=1}^{K} (\hat{x}_{t|T_k}^k - f(\hat{x}_{t-1|T_k}^k, u_{t-1}^k)),
$$

(3.28)

$$
M_t = \frac{1}{K} \sum_{k=1}^{K} \mathbb{E}[(x_t^k - f(x_{t-1}^k, u_{t-1}^k) - \mu_t)(x_t^k - f(x_{t-1}^k, u_{t-1}^k) - \mu_t)^T | z_{0:T_k}].
$$

(3.29)

The EM algorithm can be summarized as follows. We select initial values of $\mu_t = \mu_{t}^0$ and $M_t = M_{t}^0$ to begin the iterative process. At each iteration $j$, we use estimates of $\mu_t$ and $M_t$ from the previous iteration, $\mu_t^{j-1}$ and $M_t^{j-1}$, to compute the expectation of the log-likelihood (E step) and then compute revised estimates of the parameters. This iterative process is repeated till the convergence criterion is met.

We consider the means $\mu_t$ to have converged when the difference in the norm of each of the individual $\mu_t^j$ and the previous estimate $\mu_t^{j-1}$ is below a user defined threshold $\epsilon_\mu$. As for $M$, there are several choices for the convergence criterion, including the determinant $|M|$, the maximum eigenvalue, or the trace $\text{Tr}(M)$. The main disadvantage of the determinant is that a small determinant can correspond to a very elongated ellipse. As opposed to the maximum of the eigenvalues, the trace (which equals the sum of all eigenvalues) represents the uncertainty in all directions equally. We consider the covariances $M_t$ to have converged when the difference in the Frobenius norm (based on the trace criterion) of each of the individual $M_t^j$ and the previous estimate $M_t^{j-1}$ is below a user defined threshold $\epsilon_M$. We use a threshold value of $\epsilon_\mu = \epsilon_M = 1e^{-4}$ in our experiments.

### 3.6 Parameter Estimation

#### 3.6.1 Data Collection

We gathered data from multiple procedures performed in two kinds of materials (as outlined in Sec. 2.6): (i) animal-protein-based gel marketed as Simulated Muscle Tissue Ballistic Test Media (Sim-Test) (Corbin Manufacturing and Supply, Inc., 2012), and (ii) ex vivo porcine loin tissue.

We recorded measurements of the state of the needle tip using an electromagnetic tracking system (Northern Digital Inc., 2012) (Sec. 2.6) at 1 cm intervals ranging up to 10 cm of needle
insertion length for 200 trajectories in Sim-Test material and 50 trajectories in porcine loin tissue. We obtained measurements of 5 of the 6 degrees of freedom from the tracking system and the sixth degree of freedom, which is the needle twist, was measured using the encoders on the motor used for applying twists to the needle base. Since the state of the needle tip was recorded as a $4 \times 4$ matrix $X_t = \begin{bmatrix} R_t & p_t \\ 0 & 1 \end{bmatrix} \in SE(3)$ at each time step $t$, we used Eqn. 3.1 to convert to the state representation $x_t = \begin{bmatrix} p_t \\ r_t \end{bmatrix} \in \mathbb{R}^6$ required for our method. This gives the following stochastic measurement model using Eqn. (3.8):

$$z_t = x_t + n_t, \quad n_t \sim \mathcal{N}[0, N_t].$$

For the electromagnetic tracking system, the manufacturer specifications for the standard deviation of the error in estimating the position along any given axis is 0.07 cm and the angle is $0.2^\circ$ (or equivalently 0.0035 radians) (Northern Digital Inc., 2012). We assume the rotary encoders on the motor used to determine the twist angle around the insertion axis to be accurate, and assume the standard deviation to be 0.0035 radians. This gives us the following constant covariance matrix $N_t$:

$$N_t = \begin{bmatrix} (0.07)^2 I_3 \times 3 & 0 \\ 0 & (0.0035)^2 I_3 \times 3 \end{bmatrix},$$

where $I_3 \times 3$ is the $3 \times 3$ identity matrix. We also recorded the control inputs applied for actuation of the needle at each time step $t$ in terms of the insertion speed $v_t$ and twist speed $\omega_t$. It is important to note here that we have assumed the torsion along the needle shaft to be negligible but studies have shown that this might not always be case (Reed et al., 2009; Swensen and Cowan, 2012). We could also use data from prior papers on torsion compensation to compute more accurate results and we leave this as a topic for future work.

### 3.6.2 Estimated Error Parameters: Sim-Test

We considered 200 trajectories performed in the Sim-Test medium of up to 10 cm of insertion length, with measurements taken at 1 cm intervals. We initialized the expectation-maximization (EM) algorithm with the following parameters:

$$\mu_t^0 = [0, 0, 0, 0, 0]^T, \quad M_t^0 = 10^{-4} I_{6 \times 6},$$

55
where $I_{6\times6}$ is the $6 \times 6$ identity matrix.

We then applied the EM algorithm to the empirical data by alternating between Bayesian estimates of the needle tip state obtained using the extended Kalman smoother (Sec. 3.4) and maximum likelihood estimation of the parameters $\mu_{0,9}$ and $M_{0,9}$. The parameters obtained after convergence of the EM algorithm are given below:

$$
\begin{align*}
\mu_0 &= 10^{-3}, \\
&= \begin{bmatrix}
14.64 \\
-56.24 \\
-124.48 \\
-1.92 \\
2.00 \\
0.05
\end{bmatrix}, \\
M_0 &= 10^{-4}, \\
&= \begin{bmatrix}
28.55 & 3.15 & 5.89 & -0.09 & 4.44 & -0.15 \\
3.15 & 31.05 & 5.18 & -4.60 & -3.87 & 0.22 \\
5.89 & 5.18 & 9.33 & -0.42 & -0.48 & 0.04 \\
-0.09 & -4.60 & -0.42 & 5.87 & -1.81 & 0.10 \\
4.44 & -3.87 & -0.48 & -1.81 & 11.97 & -0.56 \\
-0.15 & 0.22 & 0.04 & 0.10 & -0.56 & 0.03
\end{bmatrix}, \\
\mu_1 &= 10^{-3}, \\
&= \begin{bmatrix}
6.72 \\
-52.49 \\
-69.64 \\
-3.39 \\
8.95 \\
-0.56
\end{bmatrix}, \\
M_1 &= 10^{-4}, \\
&= \begin{bmatrix}
20.13 & 1.20 & -1.60 & 0.17 & -1.06 & -0.28 \\
1.20 & 23.96 & 0.40 & -2.70 & -2.07 & -0.03 \\
-1.60 & 0.40 & 8.81 & -1.02 & 1.01 & 0.03 \\
0.17 & -2.70 & -1.02 & 4.64 & -1.49 & 0.09 \\
-1.06 & -2.07 & 1.01 & -1.49 & 9.64 & -0.67 \\
-0.28 & -0.03 & 0.03 & 0.09 & -0.67 & 0.18
\end{bmatrix}, \\
\mu_2 &= 10^{-3}, \\
&= \begin{bmatrix}
4.65 \\
-44.52 \\
-43.79 \\
-3.69 \\
5.75 \\
-0.53
\end{bmatrix}, \\
M_2 &= 10^{-4}, \\
&= \begin{bmatrix}
11.35 & 3.01 & 0.51 & -0.15 & 2.77 & -0.81 \\
3.01 & 11.15 & -0.43 & -2.27 & 0.96 & -0.05 \\
0.51 & -0.43 & 7.33 & -0.67 & 2.58 & 0.06 \\
-0.15 & -2.27 & -0.67 & 4.49 & -1.74 & 0.18 \\
2.77 & 0.96 & 2.58 & -1.74 & 8.14 & -0.46 \\
-0.81 & -0.05 & 0.06 & 0.18 & -0.46 & 0.30
\end{bmatrix}, \\
\mu_3 &= 10^{-3}, \\
&= \begin{bmatrix}
6.12 \\
-42.29 \\
-26.20 \\
-7.35 \\
7.15 \\
-0.14
\end{bmatrix}, \\
M_3 &= 10^{-4}, \\
&= \begin{bmatrix}
9.73 & 0.64 & -1.37 & 0.37 & 1.86 & -0.84 \\
0.64 & 14.48 & -0.32 & -3.06 & 0.35 & -0.02 \\
-1.37 & -0.32 & 6.53 & -0.53 & 1.38 & 0.15 \\
0.37 & -3.06 & -0.53 & 4.41 & -0.45 & 0.13 \\
1.86 & 0.35 & 1.38 & -0.45 & 6.24 & -0.45 \\
-0.84 & -0.02 & 0.15 & 0.13 & -0.45 & 0.50
\end{bmatrix}, \\
\mu_4 &= 10^{-3}, \\
&= \begin{bmatrix}
5.87 \\
-44.26 \\
-17.39 \\
-7.75 \\
7.05 \\
-0.20
\end{bmatrix}, \\
M_4 &= 10^{-4}, \\
&= \begin{bmatrix}
9.82 & -0.56 & -1.26 & 0.01 & 1.61 & -0.85 \\
-0.56 & 14.97 & 2.61 & -3.62 & -0.52 & -0.18 \\
-1.26 & 2.61 & 6.24 & -0.82 & 0.53 & 0.30 \\
0.01 & -3.62 & -0.82 & 4.10 & -0.23 & 0.10 \\
1.61 & -0.52 & 0.53 & -0.23 & 5.36 & -0.36 \\
-0.85 & -0.18 & 0.30 & 0.10 & -0.36 & 0.62
\end{bmatrix}, \\
\mu_5 &= 10^{-3}, \\
&= \begin{bmatrix}
10.97 \\
-44.74 \\
-12.62 \\
-7.75 \\
9.38 \\
0.27
\end{bmatrix}, \\
M_5 &= 10^{-4}, \\
&= \begin{bmatrix}
12.29 & 1.02 & -1.37 & 0.43 & 2.91 & -0.86 \\
1.02 & 16.12 & 2.48 & -3.84 & -0.48 & -0.09 \\
-1.37 & 2.48 & 6.45 & -0.96 & 0.54 & 0.28 \\
0.43 & -3.84 & -0.96 & 3.81 & -0.21 & 0.13 \\
2.91 & -0.48 & 0.54 & -0.21 & 6.64 & -0.31 \\
-0.86 & -0.09 & 0.28 & 0.13 & -0.31 & 0.89
\end{bmatrix}, \\
\mu_6 &= 10^{-3}, \\
&= \begin{bmatrix}
17.65 \\
-48.16 \\
-8.46 \\
-9.29 \\
10.28 \\
-0.88
\end{bmatrix}, \\
M_6 &= 10^{-4}, \\
&= \begin{bmatrix}
19.73 & -0.88 & -1.73 & 0.56 & 3.29 & 0.19 \\
-0.88 & 18.10 & 0.81 & -5.13 & 1.05 & 0.70 \\
-1.73 & 0.81 & 4.65 & -0.16 & -0.27 & -0.14 \\
0.56 & -5.13 & -0.16 & 4.25 & -0.28 & -0.22 \\
3.29 & 1.05 & -0.27 & -0.28 & 5.75 & -0.24 \\
0.19 & 0.70 & -0.14 & -0.22 & -0.24 & 1.19
\end{bmatrix}, \\
\mu_7 &= 10^{-3}, \\
&= \begin{bmatrix}
18.56 \\
-47.77 \\
-11.95 \\
11.54 \\
0.18
\end{bmatrix}, \\
M_7 &= 10^{-4}, \\
&= \begin{bmatrix}
19.67 & -0.77 & -1.80 & 0.47 & 4.66 & -0.07 \\
-0.77 & 22.09 & -0.39 & -3.76 & 0.91 & 0.88 \\
-1.80 & -0.39 & 5.34 & 0.38 & -0.88 & 0.26 \\
0.47 & -3.76 & 0.38 & 3.88 & -0.16 & -0.10 \\
4.66 & 0.91 & -0.88 & -0.16 & 6.05 & 0.21 \\
-0.07 & 0.88 & 0.26 & -0.10 & 0.21 & 1.21
\end{bmatrix}
\end{align*}

where the units of the dimensions of the state pertaining to the position of the needle tip are in cm and the units pertaining to the orientation are in radians.

In contrast to prior work that characterizes noise models for needle insertion in terms of stochastic control inputs (Park et al., 2010a), our data-driven noise model directly operates on the needle tip pose. It is important to note that the error parameters estimated using our method also have a non-zero mean (Eqns. (3.33) - (3.42)). The mean of the deviation in position along the insertion axis (z-axis) is negative – in particular, the value is -0.125 cm for the first time step (Eqn. (3.33), and then it consistently decreases to near zero (Eqn. (3.42)). This indicates that the predicted position of the needle tip along the insertion axis overshoots the actual needle tip position, which can be attributed to the fact that the needle buckles slightly outside the tissue inside the robot’s telescoping shaft when initially inserted into tissue and the needle tip does not move as far as predicted by the kinematic model. After the needle is within the tissue, the deterministic kinematic model does a better job of predicting the position of the needle tip along the insertion axis.

We compared the computed distribution of position and orientation errors along each of the x, y, and z—axes to the distribution of the errors as computed by the deterministic motion model given by Eqn. (2.2) as applied to the smoothed state estimates \( \hat{x}_i^j | T_i \). In particular, we considered the distribution of the error \( \hat{x}_i^j | T_i \) between the smoothed state estimate \( \hat{x}_i^j | T_i \) and the state predicted by the deterministic kinematic model based on the smoothed state estimate \( \hat{x}_{t-1}^j | T_i \) and control input \( u_{t-1}^j \) at each time step \( t \) for each \( i = 1, \ldots, K \) trajectories as given by:

\[
\hat{x}_i^j | T_i = (\hat{x}_i^j | T_i - f(\hat{x}_{t-1}^j | T_i, u_{t-1}^j)), \quad 1 \leq i \leq K
\]  

\[
\mu_s = 10^{-3}, \quad M_s = 10^{-4}.
\]

\[
\begin{bmatrix}
16.23 \\
-45.17 \\
-1.53 \\
-9.65 \\
7.98 \\
2.20
\end{bmatrix}
\]

\[
\begin{bmatrix}
26.94 \\
-0.74 \\
-0.83 \\
1.41 \\
3.86 \\
0.75
\end{bmatrix}
\]
Figure 3.2: Experimental data in Sim-Test: Normalized histograms showing the distribution of position and orientation errors along each of the x, y, z-axes obtained by computing $\bar{x}_i^1 = (\bar{x}_i^1|9 - f(\bar{x}_0^i|9, u_0^i))$, $1 \leq i \leq K$ using the smoothed estimates of the needle tip state for each of the K trajectories. The normal distribution $\mathcal{N}[\mu_0, M_0]$ based on the estimated parameters of our stochastic model and marginalized over each of the individual dimensions of the state is shown in red.

According to the stochastic kinematic model in Eqn. (3.4), the residual error $\bar{x}_{t|T_i}$ at each time step $t$ for the $i^{th}$ trajectory is a sample from the distribution $\mathcal{N}[\mu_t, M_t]$. Fig. 3.2 shows the normal error distribution $\mathcal{N}[\mu_0, M_0]$ for the first time step estimated using our method and the normalized histogram of the distribution of position and orientation errors along each of the axes obtained by computing $\bar{x}_i^1 = (\bar{x}_i^1|9 - f(\bar{x}_0^i|9, u_0^i))$, $1 \leq i \leq K$. The number of histogram bins is computed according to the Freedman-Diaconis rule (Freedman and Diaconis, 1981) for computing histograms for density estimation. Notice that the mean of the position error along the insertion axis (z-axis) is negative, which implies that the deterministic kinematic model overpredicts the actual needle tip position along the insertion axis.

We also estimated a single mean $\mu$ and covariance $M$ for the error distribution based on the data collected in Sim-Test by summing the relevant expectation terms over all time steps $\{0 : T\}$ and
across the set of trajectories $\Gamma$ given by the following:

$$
\mu = \frac{1}{K} \sum_{k=1}^{K} \left( \frac{1}{T_k} \sum_{t=0}^{T_k} (\hat{x}_t^k - f[\hat{x}_{t-1}^k, u_{t-1}^k]) \right),
$$

$$
M = \frac{1}{K} \sum_{k=1}^{K} \left( \frac{1}{T_k} \sum_{t=0}^{T_k} E[(x_t^k - f[x_{t-1}^k, u_{t-1}^k] - \mu_t)(x_t^k - f[x_{t-1}^k, u_{t-1}^k] - \mu_t)^T | z_{0:T_k}] \right). \tag{3.45}
$$

The parameters $\mu$ and $M$ estimated using our method are given below:

$$
\begin{bmatrix}
11.56 \\
-46.26 \\
-32.18 \\
-7.43 \\
7.83 \\
0.32
\end{bmatrix}
\Rightarrow
\begin{bmatrix}
17.70 & 0.47 & 0.25 & 0.27 & 2.68 & -0.20 \\
0.47 & 20.07 & 2.77 & -3.77 & -0.28 & 0.30 \\
0.25 & 2.77 & 21.24 & -1.68 & 1.28 & 0.35 \\
0.27 & -3.77 & -1.68 & 4.44 & -0.68 & -0.04 \\
2.68 & -0.28 & 1.28 & -0.68 & 7.21 & -0.22 \\
-0.20 & 0.30 & 0.35 & -0.04 & -0.22 & 0.80
\end{bmatrix}
\Rightarrow
\begin{bmatrix}
M = 10^{-4} \\
\mu = 10^{-3}
\end{bmatrix}
$$

### 3.6.3 Estimated Error Parameters: Ex-vivo Porcine Loin Tissue

We considered 50 trajectories performed in ex-vivo porcine loin tissue of up to 10 cm of insertion length, with measurements taken at 1 cm intervals. We initialized the expectation-maximization (EM) algorithm with the following parameters:

$$
\mu_t^0 = [0, 0, 0, 0, 0]^T, \quad M_t^0 = 10^{-4}I_{6 \times 6},
$$

where $I_{6 \times 6}$ is the $6 \times 6$ identity matrix. The estimated means $\mu_{0:9}$ and covariances $M_{0:9}$ of the error distributions are given below:

$$
\begin{bmatrix}
-28.96 \\
-26.92 \\
-152.02 \\
-26.21 \\
-0.16 \\
0.22
\end{bmatrix}
\Rightarrow
\begin{bmatrix}
-19.24 \\
3.54 \times 10^{-1} \\
7.82 \times 10^{-1} \\
-0.22 \\
-1.84 \times 10^{-1} \\
0.09 \times 10^{-1}
\end{bmatrix}
\Rightarrow
\begin{bmatrix}
\mu_0 = 10^{-3} \\
M_0 = 10^{-4}
\end{bmatrix}
$$

$$
\begin{bmatrix}
-15.14 \\
-20.59 \\
-102.82 \\
-8.34 \\
19.92 \\
-0.68
\end{bmatrix}
\Rightarrow
\begin{bmatrix}
42.80 \\
4.07 \\
97.26 \\
6.32 \\
-18.40 \\
-0.58 \times 10^{-1}
\end{bmatrix}
\Rightarrow
\begin{bmatrix}
\mu_1 = 10^{-3} \\
M_1 = 10^{-4}
\end{bmatrix}
$$

$$
\begin{bmatrix}
3.52 \\
-23.24 \\
-65.64 \\
-3.88 \\
28.73 \\
-0.86
\end{bmatrix}
\Rightarrow
\begin{bmatrix}
43.94 \\
-2.24 \times 10^{-1} \\
-4.35 \times 10^{-1} \\
3.99 \times 10^{-1} \\
8.97 \times 10^{-1} \\
-0.63 \times 10^{-1}
\end{bmatrix}
\Rightarrow
\begin{bmatrix}
\mu_2 = 10^{-3} \\
M_2 = 10^{-4}
\end{bmatrix}
$$

59
It is important to note that the error parameters estimated using our method also have a non-zero mean in ex vivo porcine loin tissue (Eqns. (3.48) - (3.57)). The mean of the deviation in position along the insertion axis (z-axis) is negative – in particular, the value is -0.152 cm for the first time step (Eqn. (3.48), and then it decreases to \(\approx -0.03\) cm towards the last time step (Eqn. (3.57)). This implies that the deterministic kinematic model overpredicts the actual needle tip position along the insertion axis in porcine tissue.

Fig. 3.3 shows the normal error distribution \(\mathcal{N}(\mu_0, M_0)\) for the first time step estimated using our method and the normalized histogram of the distribution of position and orientation errors along
Figure 3.3: Experimental data in ex vivo porcine loin tissue: Normalized histograms showing the distribution of position and orientation errors along each of the $x$, $y$, $z$-axes obtained by computing $\bar{x}^i_{1;1|9} = (\hat{x}^i_{1|9} - f(\hat{x}^i_{0|9}, u^i_0))$, $1 \leq i \leq K$ using the smoothed estimates of the needle tip state for each of the $K$ trajectories. The normal distribution $N[\mu, M]$ based on the estimated parameters of our stochastic model and marginalized over each of the individual state dimensions is shown in red.

each of the axes obtained by computing $\bar{x}^i_{1} = (\hat{x}^i_{1|9} - f(\hat{x}^i_{0|9}, u^i_0))$, $1 \leq i \leq K$. Notice that the mean of the position error along the insertion axis ($z$-axis) is negative, which implies that the deterministic kinematic model overpredicts the actual needle tip position along the insertion axis.

We also estimated a single mean $\mu$ and covariance $M$ for the error distribution based on the data collected in Sim-Test by summing the relevant expectation terms over all time steps $\{0 : T\}$ and across the set of trajectories $\Gamma$ using Eqns. (3.44) and (3.45). The parameters $\mu$ and $M$ estimated using our method are given below:

$$
\mu = 10^{-3} \cdot \begin{bmatrix} -6.40 \\ -14.58 \\ -53.78 \\ -15.39 \\ 16.08 \\ -0.14 \end{bmatrix}, \quad M = 10^{-4} \cdot \begin{bmatrix} 49.50 & 5.38 & 4.08 & -1.38 & 9.71 & -0.35 \\ 5.38 & 51.84 & -0.07 & -11.87 & 2.86 & -0.15 \\ 4.08 & -0.07 & 31.44 & -0.05 & 1.45 & -0.22 \\ -1.38 & -11.87 & -0.05 & 11.25 & -1.03 & 0.21 \\ 9.71 & 2.86 & 1.45 & -1.03 & 10.21 & -0.40 \\ -0.35 & -0.15 & -0.22 & 0.21 & -0.40 & 0.18 \end{bmatrix}
$$

(3.58)

3.7 Evaluation

Given this stochastic model of needle insertion, we can simulate perturbations during steerable needle procedures by adding simulated noise to the needle tip pose at each time step based on the
Figure 3.4: We simulated steerable needle procedures with simulated noise based on our estimated error parameters in a cuboidal shaped Sim-Test tissue phantom [(a) - (d)] and in an ex vivo porcine loin tissue sample [(e) - (f)]. These are the same environments used to gather experimental data for estimating the parameters of the error distribution (Sec. 2.7). The workspace contains obstacles (shown in yellow) in each of the cases. The insertion location of the needle is marked in green and the needle is inserted into a face of the cuboidal block.

error parameters estimated using our method. Given the state of the needle tip $x_t$ at time $t$ and applied control input $u_t$, the state is propagated forward in simulation according to Eqn. (3.5) as $x_{t+1} \sim N[f(x_t, u_t) + \mu_t, M_t]$. Since the kinematic model of the motion of the needle tip (Eqn. (2.2)) uses the $SE(3)$ representation of the state, we use Eqns. (3.3) and (3.1) to transform between the two equivalent state representations. Since the real state of the needle tip is not known, we used an extended Kalman filter (Sec. 3.4) to estimate the current state $\hat{x}_{t|t}$ at time $t$.

To validate our stochastic model, we simulated steerable needle insertions in the same environments used to gather empirical data provided as input to our method (see Fig. 3.4). We used a curvature of $\kappa_0 = 0.15 \text{ cm}^{-1}$ in Sim-Test and a curvature of $\kappa_0 = 0.073 \text{ cm}^{-1}$, obtained by characterizing the maximum curvature of the needle in these media (Sec. 2.6.3).
3.7.1 Simulation Results: Sim-Test

We performed 10 simulation runs with simulated perturbations for each target location in the environment starting from the same initial insertion location. We performed both open-loop executions of feasible motion plans and closed-loop steering using our rapid replanning approach (Chapter. 2). We also evaluated the impact of optimization criteria (i.e., shortest path or maximum clearance). This is similar to the experiments conducted in the actual Sim-Test material (Sec. 2.7.1). For the shortest path criterion, we enlarged all obstacles by a safety buffer of 5 mm.

Fig. 3.5b shows the mean targeting error and standard deviations across multiple simulated insertions using the means $\mu_t$ and covariances $M_t$ estimated using our method (Eqns. (3.33) - (3.42)). For the sake of comparison, we also include the mean targeting error and corresponding standard deviations for the actual experiments in Sim-Test as described in Sec. 2.7 (Fig. 3.5a) and when we used a single estimated mean $\mu$ and covariance $M$ as given by Eqn. (3.46) (Fig. 3.5c).

The mean targeting error of $\approx 8$ mm for open-loop execution of motion plans for both the shortest path criterion and maximum clearance for all the test scenes (Scene #1 - #4) are in close agreement with the experimental data gathered from physical experiments in Sim-Test material (Fig. 3.5a). We only have limited statistics for the physical experiments from 3 runs for each target within the Sim-Test material. The mean targeting error for closed-loop execution of motion plans for both criteria are slightly smaller than the errors encountered in practice. This can be attributed to the fact that we have limited experimental data for each scenario for the physical experiments.

3.7.2 Simulation Results: ex vivo Porcine Loin Tissue

We performed 10 simulation runs with simulated perturbations for each target location in the environment starting from the same initial insertion location. Similar to the Sim-Test material, we performed both open-loop executions of feasible motion plans and closed-loop steering using our rapid replanning approach (Chapter. 2) for both optimization criteria. This is similar to the experiments conducted in the ex vivo porcine loin tissue sample (Sec. 2.7.3). For the shortest path criterion, we enlarged all obstacles by a safety buffer of 5 mm.

Fig. 3.6b shows the mean targeting error and standard deviations across multiple simulated insertions using the means $\mu_t$ and covariances $M_t$ estimated using our method (Eqns. (3.48) - (3.57)).
Figure 3.5: We compare the targeting error using closed-loop, rapid replanning steering and open-loop execution for the two proposed optimization criteria in simulation in Sim-Test material ((b)-(c)). Error bars indicate one standard deviation of the targeting error over repeated trials. The errors, both for open-loop execution and closed-loop steering using our rapid replanning approach are comparable to the errors encountered in steerable needle insertions in actual Sim-Test material (a).
Figure 3.6: We compare the targeting error using closed-loop, rapid replanning steering and open-loop execution for the two proposed optimization criteria in simulation in ex vivo porcine loin tissue. Error bars indicate one standard deviation of the targeting error over repeated trials. The errors, both for open-loop execution and closed-loop steering using our rapid replanning approach are comparable to the errors encountered in steerable needle insertions in actual ex vivo porcine tissue.
For comparison, we also include the mean targeting error and corresponding standard deviations for the actual experiments in ex vivo porcine loin tissue sample as described in Sec. 2.7.3 (Fig. 3.6a) and when we used a single estimated mean $\mu$ and covariance $M$ as given by Eqn. (3.58) (Fig. 3.6c).

The mean targeting error for open-loop execution of motion plans for the maximum clearance criterion for the two test scenes (Scene #1 and #2) are in close agreement. However, the mean targeting error and standard deviation error for the shortest path criterion is overestimated in simulation. This can be attributed to the fact that we only have limited statistics for the physical experiments from 3 runs for each target within the ex vivo porcine loin tissue sample. The mean targeting error for closed-loop execution of motion plans for both criteria are comparable to the errors encountered in practice (Fig. 3.6a).

3.8 Discussion

We presented a data-driven stochastic model of the motion of the needle tip. We use an expectation maximization (EM) algorithm for estimating the parameters of this stochastic model from data gathered from experiments and prior procedures. Since modeling all sources of error and uncertainty during steerable needle procedures is challenging, our data-driven method provides an alternate means of creating stochastic models that capture needle behavior to serve as a basis for algorithms for preoperative procedure optimization, motion planning, state estimation, and control of steerable needles. Since the objective is to capture the cumulative effects of all sources of error and other unmodeled effects, we assume that the available data is representative of the errors encountered during steerable needle procedures. The hypothesis is that the greater the amount of available data, the better our understanding will be of uncertainty in steerable needle procedures. Our method estimates different parameters $\mu_t$ and $M_t$ for each time step $t$ to better capture uncertainty due to local effects by explicitly considering the impact of insertion depth on uncertainty.

Our approach has a few limitations. We are restricted by our assumption of a discrete time kinematic motion model. Since the model parameters are estimated based on measurements that are obtained at discrete time intervals and the number of time intervals in the procedures, our stochastic model cannot be used to simulate needle procedures that might vary greatly in the time step or length of the needle shaft inserted at every time step. Finally, we do not estimate the curvature of the needle,
which is important from the perspective of state estimation, planning, and control. The dependence of our method on the curvature occurs through the computation of the physical control inputs that are applied to the needle during duty cycled rotation of the needle. However, our method is general enough to work for needles with similar curvature characteristics.
CHAPTER 4

Unified Framework for Planning and Control in Deformable Environments

In this chapter, we present a unified framework for motion planning and feedback control for closed-loop steering of steerable needles in deformable tissue. We use a sampling-based motion planner based on a physically-based simulator of the deformable environment to generate a set of candidate plans based on expected deformations. We use the simulator and optimal control to numerically estimate time-dependent state distributions based on uncertain parameters (e.g. deformable material properties, actuation errors, and noisy sensing) and then select the plan with the highest estimated probability of successfully avoiding obstacles and reaching the target region. Using FEM-based simulation of deformable tissues, we demonstrate that our method can generate high quality plans for guiding steerable needles around obstacles to the desired target region under considerable deformations and uncertainty under 2D image guidance.

4.1 Planning and Control in Deformable Environments

Motion planning for steerable needles in highly deformable environments is challenging because it requires anticipation of deformations in the environment while simultaneously considering uncertainty in those deformations and uncertainty in the sensing of the system state. Prior work in motion planning has considered the effect of predictable deformations with no uncertainty. These methods often use a physically-based simulation of deforming objects to generate feasible motion plans (Lamiraux and Kavraki, 2001; Bayazit et al., 2002; Gayle et al., 2005; Rodriguez et al., 2006; Frank et al., 2008). These planners assume that the simulator accurately predicts deformations of the objects in the environment, which is seldom the case. Other prior work in motion planning has focused on uncertainty and modeled deformations as a type of uncertain perturbation, enabling the use of
standard feedback controllers (Alterovitz et al., 2007; van den Berg et al., 2011). However, these approaches will not work effectively for problems involving large, history-dependent deformations that fall outside the realm of small perturbations. For example, steerable needle procedures involve interaction between the needle and soft tissues that deform significantly and in a history-dependent manner, and these large deformations cannot be predicted with high accuracy due to uncertainty in the underlying tissue properties.

Our goal is to compute motion plans that maximize the probability of success in challenging environments with uncertain deformations. In this work, the probability of success is defined as the probability of successfully avoiding obstacles and reaching the desired target region. We present a unified approach that combines motion planning with sensing and feedback control for generation and execution of robust motion plans in deformable environments. We consider uncertainty due to noise in actuation and sensor measurements, as well as uncertainty in deformations arising from erroneous material model assumptions and inaccurate needle/tissue interaction models. To the best of our knowledge, our method is the first framework that enables effective computation of motion plans in environments with both large deformations and substantial uncertainty.

Our method requires a simulator of deformable objects. We use the finite element method (FEM) for simulating deformations in our examples, but our approach is generally applicable to other simulation techniques as well. We use a sampling-based motion planner in conjunction with the simulator to generate a set of candidate motion plans that assume expected deformations. Our method then uses the simulator and optimal control to numerically estimate time-dependent state distributions based on uncertain parameters (e.g. deformable material properties or actuation errors). We use this information to generate an optimal linear-quadratic (LQG) feedback controller (Bertsekas, 2007) for each candidate plan to mitigate any uncertainty in the expected deformations that occur during the actual execution of the plan. Since computing an optimal controller using the full deformable system state is computationally prohibitive, we observe that it is possible to formulate the optimal control problem using a subset of the full state space. We then use an extension of the LQG-MP framework (van den Berg et al., 2011) to select the plan with the highest estimated probability of successfully avoiding obstacles and reaching the target region.

We demonstrate the ability of our method to generate high quality plans for guiding steerable needles in deformable tissue under 2D image guidance (as shown in Fig. 4.1). We show that our
Figure 4.1: We illustrate plans for guiding steerable needles under 2D image guidance. Current motion planning solutions for deformable environments either assume deterministic deformations (b), which may result in paths through narrow passageways that are highly likely to result in obstacle collision, or compute plans in a static world and consider deformations as a type of perturbation (c), which neglects the large time-dependent motions of the obstacles and target. Our unified framework (d), which accounts for uncertainty in deformation models, noisy sensing, and unpredictable actuation, results in a significantly higher probability of success ($p_s$) in plan execution.

The rest of this chapter is organized as follows: Sec. 4.2 provides a survey of related work in the relevant areas. Sec. 4.3 summarizes the objective and Sec. 4.4 presents details of our unified framework for planning and control in deformable environments. Sec. 4.5 describes a fast, analytical method to estimate the probability of success for a mobile robot such as a steerable needle executing a given motion plan under Gaussian models of uncertainty. Finally, Sec. 4.6 presents simulation results that demonstrate that our approach significantly improves the probability of success of procedures as compared to prior work under a wide variety of simulated scenarios.

### 4.2 Related Work

This section surveys prior work in related areas, including deformable modeling and simulation, motion planning in deformable environments, and motion planning under uncertainty.
4.2.1 Deformation Modeling and Simulation

Physically-based simulation of deformable objects is a well studied area in solid mechanics (Zienkiewicz et al., 2005) and computer graphics (Nealen et al., 2006). Deformation simulation techniques can be broadly classified into two categories: (i) mesh-based methods, and (ii) mesh-free methods. Mass-spring systems, boundary element methods, and finite element methods (FEM) are a few of the popular simulation techniques available. The choice of the simulation technique is application-specific and influences the accuracy of the estimated deformations. Continuum mesh-based methods such as finite-element method (FEM) simulate deformations accurately at increased computational cost, and the model parameters for these methods are based on actual material properties such as Young’s modulus and Poisson’s ratio. Nonlinear finite element methods (FEM) are preferred when it is important to simulate deformations accurately and have been successfully used for simulating realistic deformations arising in surgical tissue simulations (Nienhuys, 2003; Chentanez et al., 2009). We use a geometrically nonlinear FEM model for simulating object deformations in this work, but the proposed approach is equally applicable to other simulation techniques as well.

4.2.2 Motion Planning in Deformable Environments

Robot motion planning is an active area of research and has historically focused on planning in static environments with rigid objects (LaValle, 2006). Recent work has begun to explore motion planning for deformable robots operating in possibly deforming environments. Bayazit et al. (Bayazit et al., 2002) propose a two-tier approach based on probabilistic roadmaps (PRM) combined with free-form deformations to plan paths for deformable robots. Rodriguez et al. (Rodriguez et al., 2006) use rapidly exploring random trees (RRT) to plan for robots in completely deformable environments. Gayle et al. (Gayle et al., 2005) and Moss et al. (Moss et al., 2008) propose a constraint based motion planning method for deformable robots modeled as mass-spring systems and using finite element methods, respectively. Frank et al. (Frank et al., 2009) use co-rotational FEM with a PRM planner to achieve significant speedups for planning in deformable environments. Motion planning algorithms have also been developed for clinical applications including deformable catheters traveling through body cavities (Gayle et al., 2005), deformable linear objects such as sutures (Moll and Kavraki, 2006), flexible needle devices traveling through deformable tissue (Alterovitz et al., 2005), and
automating atomic subtasks such as tissue retraction for robot-assisted surgery (Patil and Alterovitz, 2010b). All prior work on motion planning in deformable environments assume that deformations are deterministic and do not consider uncertainty due to noisy actuation, noisy and partial sensing, or deformation modeling and simulation errors.

A significant body of work exists on motion planning and control for bevel-tip steerable needles. Tissue deformation (without uncertainty) has been taken into account using 2D FEM simulation of soft tissue (Alterovitz et al., 2005). Motion uncertainty has been considered in 2D for cases with negligible deformations (Alterovitz et al., 2007). Planners for static 3D environments with obstacles have been proposed (Duindam et al., 2008; Patil and Alterovitz, 2010a). Hauser et al. (Hauser et al., 2009) use a fast feedback controller for guiding steerable needles in deformable tissue but do not address obstacle avoidance. Seiler et al. (Seiler et al., 2012) proposed a fast trajectory correction method to compensate for uncertainty during needle insertion. These controllers either do not consider obstacle avoidance or do not provide any guarantees on performance in the presence of obstacles. Recently, Van den Berg et al. (van den Berg et al., 2010) proposed an LQG feedback controller for addressing motion and sensing uncertainty in steerable needle insertion, but this work does not take into account displacement of the target and obstacles due to tissue deformation and uncertainty resulting from deformation. No prior work has successfully computed motion plans in highly deformable environments with uncertainty and obstacles.

4.2.3 Motion Planning under Uncertainty

Motion planning under uncertainty has received considerable attention in the past decade because uncertainty is an important concern when dealing with real-world robotic systems and environments. The uncertainty typically originates from noisy actuation, noisy and partial sensor measurements, and uncertainty about the environment. Sampling-based planners have been proposed to explicitly account for motion uncertainty (Kewlani et al., 2009; Alterovitz et al., 2007). Another class of planners has explicitly focused on uncertainty in the environment and obstacles contained therein (Burns and Brock, 2007; Guibas et al., 2008). Motion planning under both motion and sensing uncertainty is actually an instance of a partially observable Markov decision process (POMDP) (Roy et al., 1999), which suffer from the curse of dimensionality. Efficient point-based value iteration methods have proposed recently to compute approximate solutions to POMDP problems (Kurniawati et al., 2008).
but these approaches are still computationally expensive and cannot be used for steerable needle planning and control because of the discretization of the action and observation space.

A large body of work has appeared recently that assumes Gaussian models for motion and sensing uncertainty. These approaches use a combination of a Kalman filter (Simon, 2006) for state estimation and standard feedback control techniques (Bertsekas, 2007) to compensate for uncertainty. A few approaches explicitly incorporate uncertainty in sampling-based motion planning to compute high quality motion plans in terms of safety (Prentice and Roy, 2009; van den Berg et al., 2011; Bry and Roy, 2011). (Platt et al., 2010; van den Berg et al., 2012) use iterative trajectory optimization methods to efficiently solve for a locally optimal plan and an associated control policy. There have been attempts to treat deformations as a sort of uncertain perturbation and compensating for deformations using feedback controllers (van den Berg et al., 2010), but these methods do not explicitly account for large deformations and displacement of the obstacles and the target configuration. To our best knowledge, there is no prior work that offers a principled solution for the motion planning under uncertainty problem in deformable environments with obstacles.

In the presence of uncertainty in planning and execution, it is important to compute motion plans that optimize a user-specified objective to evaluate the quality of the motion plans. Prentice and Roy (Prentice and Roy, 2009) compute plans that minimize the maximum uncertainty along the motion plan. Van den Berg et al. (van den Berg et al., 2011) propose metrics to evaluate motion plans based on the probability of collision and uncertainty at the target configuration. (Bry and Roy, 2011; Vitus and Tomlin, 2011) also aim to minimize the probability of collision of the robot during runtime execution. All these approaches characterize uncertainty in terms of a priori Gaussian distributions of the robot state, but assume for the sake of simplicity that the distributions along the plan are independent. This is an incorrect assumption because it does not account for possible collisions with obstacles during plan execution (Greytak, 2009). It is important to be able to accurately compute these metrics without relying on computationally expensive Monte-Carlo simulations to obtain a good estimate (Lambert et al., 2008; du Toit and Burdick, 2011).
4.3 Objective

We consider a steerable needle being inserted into deformable tissue. We assume that time is discretized into stages of equal duration and a control \( u_t \) from its control input space \( U = \mathbb{R}^{n_u} \) is executed at each time step \( t \). Our objective, formalized below, is to compute a motion plan and associated feedback controller to maximize the probability of success.

Our planner requires as input a simulator of the coupled deformable system that models the motion of the needle and deformations in the surrounding tissue (Sec. 1.3). Given a model of anatomical structures such as glands, vessels, and bones in the environment, a set of simulation parameters \( s \), and a control input \( u_t \) for the robot, the simulator \( g \) computes the expected motions of the robot and deformations in the environment. We define the state space of the deformable system to be \( \mathcal{Y} \). A deformable system state \( y_t \in \mathcal{Y} \) at time \( t \) encodes all necessary information to save the state of the simulator and be able to restart it, including deformations, dynamics parameters, friction states, configuration of the needle tip etc. Formally, the simulator \( g \) evolves as:

\[
y_{t+1} = g[y_t, u_t, s].
\] (4.1)

The simulator \( g \) acts as a deterministic function, but the deformations and motions of the robot and environment may be highly uncertain. The material properties of deformable objects (e.g. Young’s modulus and Poisson’s ratio (Nealen et al., 2006)), interaction parameters such as friction, and deviation from commanded actuation are all uncertain and can significantly affect the motion of the needle tip and deformations in the environment. Due to this uncertainty, we assume \( s \) comes from a known distribution \( S \) and we assume that the expected value of \( s \) is 0 for brevity. Selecting different values of \( s \sim S \) allows us to use the deterministic simulator to consider variations in outcome.

The dimension of the deformable system state space \( \mathcal{Y} \) is very high, possibly containing thousands of elements for large, complex meshes of deformable objects (Chentanez et al., 2009). Computing optimal control policies directly in this high-dimensional state space would be computationally prohibitive. Furthermore, while the initial model of the environment may be known, it is unlikely that it will be possible to sense and track the entire deformation over time when executing a plan. Instead of working with the entire deformable state \( \mathcal{Y} \), we pose the control problem by considering
a reduced-dimensional space $X \subset \mathbb{R}^n$ that only contains the attributes that define the state of the needle (the position and orientation of the needle tip) and/or points in the deformable environment that are necessary for collision detection.

We assume that $X \subset Y$. Given a state $x^*_t \in X$ at time $t$ and the applied control input $u^*_t \in U$, the expected state at time $t + 1$, $x^*_{t+1} \in X$, evolves as:

$$x^*_{t+1} = f[x^*_t, u^*_t, s],$$

(4.2)

where $f$ is based on the simulator $g$ (Eqn. (4.1)). It is important to note that during the actual execution of a given plan, the true (unknown) state $x_t$ departs from the expected output of the deterministic simulation because of uncertainty arising from noisy control inputs and uncertainty in the deformation models. The uncertainty in the true state $x_t$ is modeled by adding a stochastic noise term $m_t$, which is assumed to be zero-mean Gaussian with variance $M_t$. It is important to note that in contrast to the kinematic models considered in Chapters 2 and 3, the function $f$ is based on a physically-based simulation and does not assume a base kinematic motion model for the needle tip.

During the actual plan execution, we also assume that sensors provide us with partial and noisy information about the state according to a given stochastic observation model:

$$z_t = h[x_t, n_t], \quad n_t \sim \mathcal{N}[0, N_t],$$

(4.3)

where $z_t$ is the measurement obtained at time $t$ that relates to the true state $x_t$ through function $h$, and $n_t$ is the measurement noise that we assume is drawn from a zero-mean Gaussian distribution with known variance $N_t$.

To account for uncertainty during the execution of a plan, we assume that the needle is controlled in a closed-loop fashion using the linear quadratic Gaussian (LQG) feedback controller and state estimator framework (Stengel, 1994). The LQG controller uses a linear quadratic regulator (LQR) control law that operates on the estimate of the robot state and aims to keep the robot close to the nominal plan and uses a Kalman filter for state estimation.

Our objective of planning and control for the needle in deformable environments can now be formally stated as follows:
Objective: Given a start state $x^{\text{start}} \in \mathcal{X}$ and a target region $\mathcal{X}^{\text{target}} \subset \mathcal{X}$, generate a motion plan and associated feedback controller that maximizes the probability of successfully avoiding obstacles and reaching the target region.

Input: Deformable system simulator $g$, deformable system model parameters $s$ and their distribution $S$, sensor observation model, start state $x^{\text{start}}$, and target region $\mathcal{X}^{\text{target}} \subset \mathcal{X}$.

Output: A motion plan composed of a series of states and corresponding control inputs, $\pi : (x^*_0, u^*_0, \ldots, x^*_t, u^*_t, \ldots, x^*_\ell), 0 \leq t < \ell$ where $\ell$ is the number of discrete stages in the plan and $x^*_{t+1} = f[x^*_t, u^*_t, 0]$. Also, $x^*_0 = x^{\text{start}}, x^*_\ell \in \mathcal{X}^{\text{target}}$, and the associated feedback controller for handling uncertainty arising from simulation errors, noisy sensing, and unpredictable actuation.

4.4 Approach

We provide a schematic overview of our method in Fig. 4.2. We use a simulation-based RRT motion planner (LaValle, 2006) to generate a set of candidate plans based on expected deformations computed using the simulator. We then use the simulator and numerical methods to linearize the model around each computed plan to compute a linear-quadratic Gaussian (LQG) controller (Stengel, 1994). We use an extension of the LQG-MP framework (van den Berg et al., 2011; Patil et al., 2012) to estimate time-dependent state distributions, which are used to select the plan with the highest estimated probability of successfully avoiding obstacles and reaching the target region.

4.4.1 Motion Planning

Given a start state $x^{\text{start}} \in \mathcal{X}$ and a target region $\mathcal{X}^{\text{target}} \subset \mathcal{X}$, we use the RRT motion planner (LaValle, 2006) to generate a set of feasible motion plans $\Pi$ that connect the start state and the
target region and avoid obstacles in the environment. The plans are generated assuming expected deformations (i.e. there is no uncertainty in the deformations).

The RRT algorithm incrementally builds a tree-like structure over the state space \( \mathcal{X} \). Each node in the RRT tree stores the state \( x \) as well as the full deformable system state \( y \in \mathcal{Y} \) to enable continuation of the simulation if the node is expanded. At each iteration of the algorithm, we generate a random state \( \text{sample} \in \mathcal{X} \) and use a nearest-neighbor algorithm to identify the node in the tree closest to \( \text{sample} \). We then attempt to expand the tree towards \( \text{sample} \) by choosing the best known control input \( u \in \mathcal{U} \) obtained by sampling. For each node expansion, we simulate the deformations in the environment starting from the deformed system state stored at the node.

The RRT tree is grown until the maximum number of permissible iterations is exceeded. A nominal plan composed of a sequence of states and corresponding control inputs \( \pi : (x^*_0, u^*_0, \ldots, x^*_\ell, u^*_\ell) \) can be extracted by traversing the tree from a node in the target region \( \mathcal{X}_{\text{target}} \) to the root of the tree containing the start state \( x^{\text{start}} \). Multiple such plans are extracted from the tree to generate a set of candidate motion plans \( \Pi \). It is important to note that the motion plan selected for execution is selected from this set \( \Pi \) and may not necessarily be globally optimal.

### 4.4.2 Model Linearization

Executing a plan computed in the previous section in an open-loop manner is unlikely to reach the target region in practice because of uncertainty due to several factors such as unpredictable needle and tissue interaction, unpredictable actuation, noisy (and possibly) sensor measurements, and uncertainty arising from the deformable object simulation itself. We create a feedback controller to mitigate these uncertainties.

In general, physically-based deformable object simulations can be highly nonlinear, and system identification for nonlinear control for such simulators can be difficult. However, since the deformable system is controlled to stay close to the computed plan, we approximate these nonlinear models by locally linearizing the model around the plan \( \pi \). In essence, we consider uncertainty about the expected deformation rather than taking the prior approach of considering the deformation itself to be a type of uncertainty (as is described in Chapter 3).

It is convenient to express the control problem in terms of the deviation from the plan. By defining the deviation in the state as \( \bar{x}_t = (x_t - x^*_t) \), deviation in control input as \( \bar{u}_t = (u_t - u^*_t) \),
and deviation from the actual measurement as 
\[ \bar{z}_t = (z_t - h(x^*_t, 0)), \]
the dynamics and observation models given by Eqns. (4.2) and (4.3) can be linearized as:

\[
\begin{align*}
\bar{x}_{t+1} &= A_t \bar{x}_t + B_t \bar{u}_t + m_t, \quad m_t \sim N(0, M_t], \\
\bar{z}_t &= H_t \bar{x}_t + W_t n_t, \quad n_t \sim N(0, N_t],
\end{align*}
\]

where

\[
\begin{align*}
A_t &= \frac{\partial f}{\partial x}[x^*_t, u^*_t, 0], \quad B_t = \frac{\partial f}{\partial u}[x^*_t, u^*_t, 0], \\
H_t &= \frac{\partial h}{\partial x}[x^*_t, 0], \quad W_t = \frac{\partial h}{\partial n}[x^*_t, 0]
\end{align*}
\]

are the Jacobian matrices of \( f \) and \( h \) along a given plan \( \pi \).

If the dynamics model \( f \) is known, then the Jacobian matrices can be computed analytically. In the case of physically-based simulators for deformable objects, it is typically difficult or impossible to compute these Jacobian matrices analytically because the simulations in general cannot be written using closed-form formulas and are not directly differentiable. Given a nominal plan \( \pi : (\ldots, x^*_{t-1}, u^*_{t-1}, x^*_t, u^*_t, x^*_{t+1}, u^*_{t+1}, \ldots) \), we numerically estimate the Jacobian matrices \( A_t \) and \( B_t \) at each time-step along the plan. The Jacobian matrices of the given observation model \( h, H_t \) and \( W_t \), are computed analytically. The stochastic noise term \( m_t \) models the uncertainty due to simulation parameters \( s \sim S \) and is assumed to be drawn from a zero-mean Gaussian distribution with variance \( M_t \), which is also numerically estimated as described below.

**Numerical estimation of matrix** \( A_t \): In the absence of any deviation from the control input \( u^*_t \), the Jacobian matrix \( A_t \in \mathbb{R}^{n_x \times n_x} \) describes how the deviation in state evolves from \( \bar{x}_t \) to \( \bar{x}_{t+1} \) when the corresponding nominal control input \( u^*_t \) is applied to the perturbed state \( x_t = (x^*_t + \bar{x}_t) \). This is given by the relation \( \bar{x}_{t+1} = A_t \bar{x}_t \). By performing \( K = n_x \) independent simulation runs from states \((x^*_t + \bar{x}_t^k), k \in \{1, 2, \ldots, K \}\) to states \((x^*_{t+1} + \bar{x}_{t+1}^k)\), we can appropriately assemble the deviations in the consecutive states into matrices \( \bar{X}_t \) and \( \bar{X}_{t+1} \) and estimate \( A_t \) numerically as \( A_t = \bar{X}_{t+1} \bar{X}_t^{-1} \).

It is not possible to perturb the nominal state \( x^*_t \) to \((x^*_t + \bar{x}_t)\) within a physically-based deformable simulator. For instance, in the case of a needle procedure, this would correspond to the displacement and re-orientation of the tip, which is not possible without affecting the entire needle trajectory and
adjusting needle/tissue interaction parameters. So instead we perform $K$ independent simulation runs from the previous state $x^*_t$, where each run $k \in \{1, 2, \ldots, K\}$ involves the application of control input $u^*_t$ perturbed by noise $w_k$, where we assume that $w_k$ is drawn from a user-specified zero-mean Gaussian with variance $W$. This generates a set of states $(x^*_t + x^*_k), k \in \{1, 2, \ldots, K\}$ within the simulator, as shown in Fig. 4.3b. We then perform $K$ independent simulation runs from the set of states $(x^*_t + x^*_k)$ by applying the nominal control input $u^*_t$ to the robot to generate a set of states $(x^*_t + x^*_k)$. We then assemble the deviations from the nominal state $x^*_t$ into a matrix $X_t$. The matrix $A_t$ can then be estimated as described above. In practice, we found that performing $K (> n_x)$ simulation runs and solving the resultant over-determined least-squares problem $A_tX_t = \tilde{X}_{t+1}$ by taking the Moore-Penrose pseudo-inverse of matrix $X_t$ yielded better results at the expense of computational overhead. We used $K = 2n_x$ simulation runs for all our experiments.

**Numerical estimation of matrix $B_t$:** In the absence of any deviation from the nominal state $x^*_t$, the matrix $B_t \in \mathbb{R}^{p \times q}$ describes the relationship between deviations in control input $\bar{u}_t$ and the deviation in the subsequent state $x_{t+1}$. This is given by the relation $\bar{x}_{t+1} = B_t\bar{u}_t$. We perform $K = n_u$ independent simulation runs from the nominal state $x^*_t$ by perturbing individual elements of the control input $u^*_t$ by $e_k \cdot \epsilon, k \in \{1, 2, \ldots, K\}$, where $e_k$ is a unit vector given by $(\ldots, 0, 0, 1_k, 0, 0, \ldots)$.
and $\epsilon$ is a user-defined perturbation constant. This generates a set of states $(x_{t+1}^* + \bar{x}_{t+1}^k)$, as shown in Fig. 4.3c. The deviations $\bar{x}_{t+1}^k$, scaled by the $\epsilon$, comprise the individual columns of the matrix $B_t$.

**Numerical estimation of matrix $M_t$:** The stochastic noise term $m_t$ models uncertainty in the simulation parameters $s \sim S$. It is reasonable to assume that the uncertainty arising from a large number of sources can be modeled as a Gaussian distribution. Here, $m_t$ is assumed to be drawn from a zero-mean Gaussian distribution with variance $M_t$. Since the uncertainty distributions modeling the noise in control inputs and variance in simulation parameters $S$ are supplied by the user, the variance $M_t$ can be estimated a priori by perturbing the simulation parameters $s$ independently and estimating the parameters of the Gaussian distribution that models the resulting uncertainty.

### 4.4.3 LQG Control

Given linear(ized) dynamics and observation models and a quadratic cost function, the optimal approach for executing a plan is to use a linear-quadratic regulator (LQR) feedback controller in combination with a Kalman filter for state estimation. This ensemble is called linear-quadratic Gaussian (LQG) control (Stengel, 1994) and is provably optimal for state estimation and control for linear systems. We use an extended Kalman filter (EKF) (Simon, 2006) for optimal state estimation during actual execution of the plan. The Kalman filter keeps track of the estimate $\hat{x}_t = E[\bar{x}_t]$ and variance of the deviation in true state during control. The estimate $\hat{x}_t$ evolves according to:

$$\hat{x}_t = K_t \bar{z}_t + (I - K_t H_t)(A_t - 1 \hat{x}_{t-1} + B_t - 1 \bar{u}_{t-1}),$$  \hspace{1cm} (4.8)

where $K_t$, $0 \leq t \leq \ell$ are the Kalman gain matrices (Simon, 2006). Note that the Kalman-gain matrices $K_t$ can be computed in advance (i.e. before execution).

We compensate for uncertainty during plan execution by using an LQR feedback controller that aims to keep the true state close to the corresponding nominal state in the plan. The LQR formulation seeks the optimal control inputs by minimizing a quadratic cost function that seeks to simultaneously minimize deviations from the plan and deviations from the control input. Solving the cost function gives the control policy related to the state deviation estimate as:

$$\bar{u}_t = L_t \hat{x}_t,$$ \hspace{1cm} (4.9)
where $L_t, 0 \leq t < \ell$ are the LQR control gains that are pre-computed using a standard recursive procedure. We refer the reader to (Stengel, 1994; van den Berg et al., 2011) for additional details.

### 4.4.4 Selecting a High Quality Plan

We use an extension of the LQG-MP framework (van den Berg et al., 2011) to select a plan with the highest estimated probability of successfully avoiding obstacles and reaching the target region. Given the LQG controller for a plan $\pi$, we compute the a priori distributions of the deviation in the true state $\bar{x}_t$ and the estimated deviation $\hat{x}_t$. These distributions are used to estimate the probability of success, which is given by the product of the probability of colliding with obstacles during plan execution and the probability of reaching the target region, which is obtained by sampling the a priori distribution at the final time-step and determining how many of those samples lie within the target region. It is important to take into account the deformed configurations of the target and obstacles in the environment at each time-step along the plan. The computation of the probability of success is described in detail in the next section.

### 4.5 Estimating Probability of Success

For steerable needle procedures, the motion plan chosen for execution should be as safe as possible such that there is minimal risk that the robot will collide with obstacles in the environment and should lead the needle to the desired target region. The probability of success is computed as the product of the probability of collision along a given plan and the probability of successfully reaching the target region. Estimating the probability of success of a motion plan before actual execution is a critical step in many motion planning algorithms that consider and compensate for uncertainty.

The challenge lies in effectively computing the probability of collision along a plan. Prior work on motion planning under uncertainty has used both sampling-based and analytical approaches to estimating probability of collision. Naïve Monte Carlo sampling strategies can estimate the probability of collision by computing the ratio of the number of simulated executions that are collision free (Lambert et al., 2008; du Toit and Burdick, 2011). This approach requires a large number of simulations to obtain a reliable estimate, which requires more computation time than analytical approaches. Monte Carlo sampling also offers no guarantee that it will not underestimate
Figure 4.4: We estimate the probability of success for a motion plan based on a priori probability distributions of the robot state. The probability of success is given by the product of the probability of collision along the plan and the probability of reaching the target region. The probability of collision at each stage of the plan is conditioned on the previous stages being collision free. We compute truncated a priori distributions that discount plan executions (black dots) that collide with obstacles. Propagating the truncated distributions (black ellipses) accounts for only the collision free samples (red dots), resulting in accurate estimation of the probability of collision. Prior methods that use the unconditional distributions (gray ellipses) to estimate the collision probability result in an overly conservative estimate.

the probability of collision, resulting in violation of safety requirements. Under the assumption of Gaussian motion and sensing uncertainty, probability of collision can be estimated quickly based on a priori probability distributions of the robot state (van den Berg et al., 2011; Bry and Roy, 2011; Vitus and Tomlin, 2011). However, prior methods typically “approximate” the collision probability of a plan by assuming the probabilities of collision at stages along the plan are independent. Formally speaking, let $x_t \in \mathcal{X}$ denote the state of the robot at stage $t$ along the plan, and $\mathcal{X}_F \subset \mathcal{X}$ denote the feasible space not occupied by obstacles. Prior methods assume that the probability that a plan consisting of $\ell$ stages is collision free is given by $p(\bigwedge_{t=0}^{\ell} x_t \in \mathcal{X}_F) \approx \prod_{t=0}^{\ell} p(x_t \in \mathcal{X}_F)$. This yields an overly conservative estimate of the probability of collision (see Fig. 4.4), which might result in overly conservative motion plans and, depending on the safety required by the motion planner, may result in failure to find a feasible plan even if one exists.

In this work, we present a fast, analytical method to estimate the probability of collision, and consequently, the probability of success for a mobile robot such as a steerable needle executing a given motion plan under Gaussian models of motion and sensing uncertainty. Our approach for estimating the probability of collision accounts for the fact that the distribution of the state at each
stage along the plan is conditioned on the previous stages being collision free, i.e., the probability that a plan is collision free is given by

\[ p(\bigcap_{t=0}^{\ell} x_t \in X_F) = \prod_{t=0}^{\ell} p(x_t \in X_F \mid \bigcap_{i=0}^{t-1} x_i \in X_F). \]

This amounts to propagating the a priori distributions forward in time in such a way that instances that collide with obstacles are discounted from the propagation (Fig. 4.4). For this we propose a novel method to truncate the a priori distributions with respect to obstacles, approximate the truncated distributions by Gaussians, and propagate the truncated distributions forward in time. This results in an accurate estimate of the conditional distributions, and consequently, enables accurate estimation of the collision probability. Our algorithm also computes an estimate that is conservative; our goal is to not underestimate the probability of success in order to ensure that safety requirements are satisfied.

The use of truncated Gaussian distributions (Johnson et al., 1994) has been previously explored in the context of optimal state estimation with state constraints (Simon, 2006), but this work does not consider motion uncertainty. Greytak (Greytak, 2009) provides an analytical method to compute the probability of collision using truncated Gaussians but does not consider sensing uncertainty. Toussaint (Toussaint, 2009b) uses truncated Gaussians in an expectation-propagation framework for Bayesian inference, but the truncation result is dependent on the order in which constraints are processed, which leads to problems with convergence of the algorithm (Toussaint, 2009a). In contrast, we propose a novel order-independent algorithm for truncating Gaussian distributions with respect to hard state constraints. Our method can be used to quantify the safety of a plan (van den Berg et al., 2011; Patil et al., 2011), to improve quality of estimation of collision chance constraints (Bry and Roy, 2011; Vitus and Tomlin, 2011), or to elegantly account for hard state constraints imposed by obstacles in optimization based (Erez and Smart, 2010) or inference based (Toussaint, 2009b) planning methods. Our truncation approach is also directly applicable to the important problem of optimal state estimation with hard state constraints (Simon, 2006).

### 4.5.1 A Priori State Distributions

The objective of our method can be formally stated as follows. Given Gaussian models of motion and sensing uncertainty, a description of the obstacles in the environment, a nominal motion plan, and associated feedback controller and state estimator, the objective is to compute the probability of collision, and consequently, the probability of success of a given plan.
Under the given assumptions, the probability distributions of the robot state can be characterized a priori, i.e. before execution. Combining Eqns. (4.4), (4.5), (4.8), and (4.9), the true state deviation \( \bar{x}_t \), and the estimate \( \hat{x}_t \), jointly evolve as (van den Berg et al., 2011):

\[
\begin{bmatrix}
\bar{x}_t \\
\hat{x}_t 
\end{bmatrix} =
\begin{bmatrix}
A_{t-1} & B_{t-1}L_{t-1} \\
K_t H_t A_{t-1} & A_{t-1} + B_{t-1}L_{t-1} - K_t H_t A_{t-1}
\end{bmatrix}
\begin{bmatrix}
\bar{x}_{t-1} \\
\hat{x}_{t-1}
\end{bmatrix} +
\begin{bmatrix}
I & 0 \\
K_t H_t & K_t W_t
\end{bmatrix}
\begin{bmatrix}
m_{t-1} \\
n_t
\end{bmatrix}
\sim \mathcal{N}\left(0, \begin{bmatrix} M_{t-1} & 0 \\ 0 & N_t \end{bmatrix} \right).
\tag{4.10}
\]

We can write this equation in shorthand (for appropriate definitions of \( y_t, q_t, F_t, G_t, \), and \( Q_t \)) as:

\[
y_t = F_t y_{t-1} + G_t q_t, \quad q_t \sim \mathcal{N}\left(0, Q_t \right).
\tag{4.11}
\]
The mean $\hat{y}_t \in \mathbb{R}^{2nx}$ and associated variance $R_t = \text{Var}[y_t]$, propagate according to:

$$
\hat{y}_t = F_t \hat{y}_{t-1}, \quad \hat{y}_0 = 0,
$$

$$
R_t = F_t R_{t-1} F_t^T + G_t Q_t G_t^T, \quad R_0 = \begin{bmatrix} \text{Var}[\tilde{x}_0] & 0 \\ 0 & 0 \end{bmatrix}.
$$

The unconditional a priori distribution of the state $x_t$ at stage $t$ is then given by the marginal $x_t \sim \mathcal{N}[x^*_t + \Lambda \hat{y}_t, \Lambda R_t \Lambda^T]$, where $\Lambda = [I \ 0]$. To accurately estimate the probability of collision, we need to estimate the a priori state distributions at each stage along the plan that are conditioned on the previous stages being collision free, i.e. the distributions $(x_t \mid \bigwedge_{i=0}^{t-1} x_i \in \mathcal{X}_F)$. To this end, we pursue a recursive approach similar as above to propagate the conditional distributions.

Let $y_{t|s}$ denote the joint distribution of the true state deviation and its estimate at time $t$ conditioned on the state being collision free for all stages $0, \ldots, s$:

$$
y_{t|s} = \left( \begin{array}{c} \hat{x}_t \\ \hat{\tilde{x}}_t \end{array} \right) \mid \bigwedge_{i=0}^{s} x_i \in \mathcal{X}_F).
$$

We then repeatedly, for each stage $t$ of the plan, carry out the following steps. Assume we are given the joint conditional distribution $y_{t|t-1}$ as approximated by a Gaussian distribution $\mathcal{N}[\hat{y}_{t|t-1}, R_{t|t-1}]$. We then approximate the distribution $y_{t|t} \sim \mathcal{N}[\hat{y}_{t|t}, R_{t|t}]$ of all collision-free states at stage $t$ by truncating the distribution $y_{t|t-1}$ against the obstacles in the environment. Truncating the distribution effectively discounts all colliding states from the distribution (Fig. 4.4), and results in a shift of the mean and variance by $\Delta y_t$ and $\Delta R_t$ (as described in Sec. 4.5.2), respectively:

$$
\hat{y}_{t|t} = \hat{y}_{t|t-1} - \Delta y_t
$$

$$
R_{t|t} = R_{t|t-1} - \Delta R_t
$$
Using Eqns. (4.12) and (4.13), the conditional mean and variance are then propagated according to:
\[
\hat{y}_{t+1|t} = F_{t+1} \hat{y}_t|t, \tag{4.17}
\]
\[
R_{t+1|t} = F_{t+1} R_t|t F_{t+1}^T + G_{t+1} Q_{t+1} G_{t+1}^T. \tag{4.18}
\]

The recursion then continues. The initial conditions are set by defining \( \hat{y}_{0|-1} = \hat{y}_0 = 0 \) and \( R_{0|-1} = R_0 = \begin{bmatrix} \text{Var}[\bar{x}_0] & 0 \\ 0 & 0 \end{bmatrix} \).

At each stage of the recursion, the marginal \( x_{t|t-1} \sim \mathcal{N}((x_t^* + \Lambda \hat{y}_{t|t-1}), \Lambda R_{t|t-1} \Lambda^T) \) of the joint distribution \( y_{t|t-1} \) gives the a priori distribution of the robot state \( x_t \) given that all the previous states \( [x_0, \ldots, x_{t-1}] \) are collision free.

### 4.5.2 Truncating A Priori Distributions

At each stage \( t \) of the plan, we approximate the distribution of the feasible robot states with a truncated Gaussian distribution (Johnson et al., 1994). For the sake of brevity, we assume that the feasible region containing the state at each stage \( t \) is convex and is described by the conjunction of \( k \) linear inequality constraints as \( \bigcap_{i=0}^k \mathbf{a}_i x_i \leq b_i \). We later extend this analysis in Sec. 4.5.3 to non-convex regions by constructing a locally convex feasible region around the robot state.

Since the true state deviation and its estimate are correlated (Eqn. (4.10)), it is important to truncate the joint conditional distribution \( \mathcal{N}[\hat{y}_{t|t-1}, R_{t|t-1}] \) in \( \mathbb{R}^{2n_x} \), with respect to the \( k \) constraints. The \( i \)th linear constraint is then represented in \( \mathbb{R}^{2n_x} \) as \( \tilde{\mathbf{a}}_i^T \hat{y}_{t|t-1} \leq \tilde{b}_i \), where \( \tilde{\mathbf{a}}_i = [\mathbf{a}_i^T \ 0] \), and \( \tilde{b}_i = (b_i - \mathbf{a}_i^T x_t^*) \). We truncate the joint conditional distribution with respect to each constraint in a sequential manner and then accumulate the effect of truncation over all the constraints. We propose a novel truncation method that does not depend on the order in which the constraints are processed.

Given the \( i \)th constraint \( \tilde{\mathbf{a}}_i^T y_{t|t-1} \leq \tilde{b}_i \), we apply an affine transformation \( y_{t|t-1}^i = \tilde{\mathbf{a}}_i^T y_{t|t-1} \) to transform the conditional distribution \( \mathcal{N}[\hat{y}_{t|t-1}, R_{t|t-1}] \), to a 1D Gaussian \( \mathcal{N}[\tilde{\mathbf{a}}_i^T \hat{y}_{t|t-1}, \tilde{\mathbf{a}}_i^T R_{t|t-1} \tilde{\mathbf{a}}_i] \) along an axis normal to the constraint (as shown in Fig. 4.5). The problem now simplifies to truncating the 1D Gaussian distribution at a specified upper bound given by \( y_{t|t-1}^i = \tilde{b}_i \), which is well-known from standard statistical literature (Johnson et al., 1994). The mean, \( \mu_i \) and variance, \( \sigma_i^2 \)
of the truncated 1D Gaussian $y_{t|t}^i$ is given by:

$$
\mu_i = \hat{a}_i^T \hat{y}_{t|t-1} + \lambda(\alpha_i) \sqrt{\hat{a}_i^T R_{t|t-1} \hat{a}_i}, \quad (4.19)
$$

$$
\sigma_i^2 = \hat{a}_i^T R_{t|t-1} \hat{a}_i (1 - \lambda(\alpha_i)^2 + \alpha_i \lambda(\alpha_i)), \quad (4.20)
$$

where

$$
\alpha_i = \frac{(\tilde{b}_i - \hat{a}_i^T \hat{y}_{t|t-1})}{\sqrt{\hat{a}_i^T R_{t|t-1} \hat{a}_i}}, \quad \lambda(\alpha_i) = \frac{pdf(\alpha_i)}{cdf(\alpha_i)}. \quad (4.21)
$$

Here, $\lambda(\alpha_i)$ is the ratio of the standard Gaussian (mean 0 and variance 1) probability distribution function and the standard Gaussian cumulative distribution function evaluated at $\alpha_i$. Note that $(1 - cdf(\alpha_i))$ is the area under the Gaussian that lies beyond the constraint (shaded in black in Fig. 4.5), and is the probability that the robot lies in the infeasible region of the $i$th constraint.

The mean and variance of the truncated distribution $y_{t|t}$ are found by conditioning the joint distribution $(y_{t|t-1}, y_{t|t-1}^i)$, on the truncated 1D distribution $y_{t|t}^i$: $y_{t|t} = (y_{t|t-1}|y_{t|t-1}^i = y_{t|t}^i)$ as follows. We can construct the joint distribution of the conditional distribution $y_{t|t-1} \sim \mathcal{N}[\hat{y}_{t|t-1}, R_{t|t-1}]$, and the transformed 1D distribution $y_{t|t-1}^i \sim \mathcal{N}[\hat{a}_i^T \hat{y}_{t|t-1}, \hat{a}_i^T R_{t|t-1} \hat{a}_i]$ as:

$$
(y_{t|t-1}, y_{t|t-1}^i) \sim \mathcal{N}\left[\begin{bmatrix} \hat{y}_{t|t-1} \\ \hat{a}_i^T \hat{y}_{t|t-1} \end{bmatrix}, \begin{bmatrix} R_{t|t-1} & R_{t|t-1} \hat{a}_i \\ \hat{a}_i^T R_{t|t-1} & \hat{a}_i^T R_{t|t-1} \hat{a}_i \end{bmatrix}\right]. \quad (4.22)
$$

Using the law of iterated expectations and the law of total variances (Movellan, 2011), we reconstruct the truncated mean and variance of the joint distribution by conditioning on the truncated 1D distribution $y_{t|t}^i \sim \mathcal{N}[\mu_i, \sigma_i^2]$ (Eqn. 4.19, 4.20), according to:

$$
(y_{t|t-1}|y_{t|t}^i = y_{t|t}^i) \sim \mathcal{N}[\hat{y}_{t|t-1} - L(\hat{a}_i^T \hat{y}_{t|t-1} - \mu_i), R_{t|t-1} - L(\hat{a}_i^T R_{t|t-1} \hat{a}_i - \sigma_i^2) L^T], \quad (4.23)
$$

where $L = \frac{R_{t|t-1} \hat{a}_i}{\hat{a}_i^T R_{t|t-1} \hat{a}_i}$. The shift in the mean due to truncation due to the $i$th constraint is given by:

$$
\Delta y_{t}^i = \frac{R_{t|t-1} \hat{a}_i}{\hat{a}_i^T R_{t|t-1} \hat{a}_i} (\hat{a}_i^T \hat{y}_{t|t-1} - \mu_i), \quad (4.24)
$$
and the shift in variance is given by:

\[
\Delta R_t^i = \frac{R_t^{t-1} \tilde{a}_t}{\tilde{a}_t^T R_t^{t-1} \tilde{a}_t} \left( \tilde{a}_t^T R_t^{t-1} \tilde{a}_t - \sigma_t^2 \right) \frac{\tilde{a}_t^T R_t^{t-1} \tilde{a}_t}{\tilde{a}_t^T R_t^{t-1} \tilde{a}_t}.
\]

(4.25)

The cumulative shift in the mean due to truncation due to \( k \) constraints is then \( \Delta y_t = \sum_{i=0}^{k} \Delta y_t^i \), and the cumulative change in variance is \( \Delta R_t = \sum_{i=0}^{k} \Delta R_t^i \). The mean and variance of the truncated conditional distributions are then propagated recursively using Eqns (4.17) and (4.18).

### 4.5.3 Estimating the Probability of Collision

We use the truncated conditional distributions to estimate the overall probability of collision of the given plan, based on the conditional probabilities of collisions at each stage along the plan. Given the joint conditional distribution at stage \( t \), \( \mathcal{N}[\tilde{y}_t|t-1, R_t|t-1] \), and the set of \( k \) linear constraints that define the locally convex region of free space containing the robot, we compute a lower bound for the probability of the robot being collision free using Boole’s inequality, as (Vitus and Tomlin, 2011):

\[
p(\tilde{x}_t|t-1 \in \mathcal{X}_F^c) \geq 1 - p(\bigvee_{i=0}^{k} \tilde{a}_t^T \tilde{y}_t|t-1 > \tilde{b}_i)
\geq 1 - \sum_{i=0}^{k} (1 - \text{cdf}(\alpha_i)).
\]

(4.26)

The overall probability that the robot does not collide with any obstacle for the duration \( \ell \) of the plan, is given by:

\[
p(\bigwedge_{t=0}^{\ell} \tilde{x}_t \in \mathcal{X}_F^c) = \prod_{t=0}^{\ell} p(\tilde{x}_t|t-1 \in \mathcal{X}_F^c),
\]

(4.27)

and the overall probability of collision is provided by the complement \((1 - p(\bigwedge_{t=0}^{\ell} \tilde{x}_t \in \mathcal{X}_F^c))\).

We extend our analysis to non-convex regions by truncating the joint conditional distributions with respect to linear constraints that define a locally convex region of free space containing the robot. For the sake of simplicity, we assume that only the robot position is relevant for collision detection. At each stage \( t \), we compute the marginal distribution \( \mathcal{N}[\tilde{p}_t|t-1, \Sigma_t|t-1] \) of the conditional distribution \( \mathcal{N}[\tilde{y}_t|t-1 + [\tilde{x}_t^T \Sigma_t^{-1}], R_t|t-1] \) over the dimensions of the robot state that describe the robot position \( \tilde{p}_t|t-1 \). We outline a greedy method that computes a locally convex region of free space such that the probability that the distribution \( \mathcal{N}[\tilde{p}_t|t-1, \Sigma_t|t-1] \) lies beyond the convex region is minimal.

88
Figure 4.6: We transform the environment such that the distribution of the robot position (left) is converted to a unit sphere (right). We then sequentially process the obstacle geometry in increasing order of distance from the origin. The linear constraints that define a locally convex region of the free space are determined by the normal to the vector of closest approach (shown in red). The locally convex region constructed using our approach for this example is defined by three constraints determined in order of their indices: $C_1$, $C_2$, and $C_3$.

Adopting the approach suggested in (van den Berg et al., 2011), we linearly transform the environment geometry by applying the transform $U_{t-1}$, where $\Sigma_t |_{t-1} = U_t U^T_t$ is the Cholesky decomposition. This transforms the uncertainty distribution of the robot position to a Gaussian distribution with zero mean and unit variance, which is a unit sphere in Euclidean space centered at the origin. The spherical symmetry simplifies the task of constructing a nonconservative convex region of free space around the distribution of the position of the robot (Fig. 4.6).

We construct the convex region using a sequential process. We consider the closest point on the obstacle geometry from the origin. The linear truncation constraint $a^T_i p_{t|t-1} \leq b_i$, is defined by the normal to the vector of closest approach to the obstacle. We then prune away all geometry that lies in the infeasible half space $a^T_i p_{t|t-1} > b_i$ of the constraint, and continue the process by considering the closest point on the remaining obstacle geometry to the origin. This procedure is repeated until all geometry has been pruned away. It is important to note that our convexification method works in a greedy fashion and is not guaranteed to find the least conservative convex bounding region.

4.5.4 Validation

We also apply our method for computing the probability of collision for a steerable needle navigating in a 3D environment with obstacles (Fig. 4.7a) and stochastic dynamics and partial and noisy sensing feedback. We initialize our method with a nominal plan computed using an RRT planner (Sec. 4.4). We tested our C++ implementation on a 3.33 GHz Intel® i7™ PC.
Figure 4.7: Nonholonomic bevel-tip steerable needle: (a) Unconditional distributions (solid gray ellipsoids corresponding to 3 standard deviations) provide an overly conservative approximation of the uncertainty. Our method computes conditional distributions (black wireframe ellipsoids), which provide an accurate estimate of the probability distributions of the feasible robot states (shown in red). The collision probability estimated by our method is 54.5%, while the ground truth probability is 52.4%. (b) Zoomed in view of the conditional distributions in the narrow corridor. (c) Comparison of our method to Monte Carlo simulations. (d) The probability of collision estimated by our method for a second plan is 43.9%, while the ground truth probability is 42.2%.

We follow the variable curvature stochastic kinematic model (Chapter 3). We also assume that we receive partial, noisy feedback on only the position of the needle tip $p$, and not its orientation. This is a reasonable assumption since current medical imaging technologies such as ultrasound do not allow for measuring the full state of the needle tip (as the imaging resolution is often too low to infer its orientation). The noise in the sensor measurement is modeled as $n \sim \mathcal{N}[0, N]$. This gives the stochastic measurement model, $h[x, n] = p + n$. We follow the approach in (van den Berg et al., 2010) to approximate the given nonlinear dynamics and measurement models with local linearizations around the nominal plan.

We validate our method by comparing the estimated collision probability with the ground truth probability computed using a million Monte Carlo simulations (considered as ground truth) of the given motion plan and counting the ratio of collision free simulations. Each execution is simulated in a closed-loop fashion using the given linear feedback controller and a Kalman filter, and with artificially generated motion and measurement noise.

Fig. 4.7b shows the discrepancy between the unconditional and conditional distributions in the presence of obstacles. The conditional distributions computed using our method provide an accurate estimate of the distribution of the collision free robot states along the plan, thus providing an accurate estimate of the probability of collision, and consequently, the probability of success.
Table 4.1 compares the probability of collision estimated by our method against the ground truth probability computed using Monte Carlo simulations for the scenarios considered above. Our estimate lies within 5% of the ground truth value. It is important to note that Monte Carlo simulations provide an unbiased estimate of the probability of collision, and can underestimate the probability if a sufficiently large number of samples are not considered. In contrast, our method provides a conservative estimate of the probability. Each Monte Carlo simulation takes 0.69 milliseconds while our method takes 7.4 milliseconds. It takes 2000 simulations to arrive within the accuracy bounds of our method (Fig. 4.7c), which corresponds to 1.4 seconds of computation time. Even neglecting the fact that Monte Carlo simulations underestimate the collision probability, it still takes 3000 simulations to arrive within the accuracy bounds of our method. This corresponds to over a second of computation time just to estimate the collision probability, which is undesirable for real-time motion planning under uncertainty. Our method provides accurate, yet conservative, estimates of the collision probability while incurring negligible computational overhead. This makes it especially suitable for online planning algorithms that explicitly consider uncertainty.

We compare our method to prior methods that rely on a priori state distributions to estimate the collision probability. We generated a set of 100 plans using the RRT planner using randomly initialized start states. For each plan, we estimated the collision probability using our method, applying Boole’s inequality to the unconditional distributions (Vitus and Tomlin, 2011), and LQG-MP (van den Berg et al., 2011). We use the mean error as a metric to compare the probability estimates to the ground truth probability. As summarized in Table 4.1, the estimate computing using our method reduces the estimation error by more than 25% as compared to the collision quality metric provided by LQG-MP (van den Berg et al., 2011) and the collision probability computed using the unconditional distributions directly (Vitus and Tomlin, 2011). It is important to note that all these estimation methods, including ours, provide a conservative bound for the collision probability.

<table>
<thead>
<tr>
<th></th>
<th>Our method</th>
<th>Unconditional (Vitus and Tomlin, 2011)</th>
<th>LQG-MP (van den Berg et al., 2011)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE (%)</td>
<td>5.0 (± 3)</td>
<td>20.7 (± 7)</td>
<td>61.7 (± 12)</td>
</tr>
<tr>
<td>Avg. Time (ms)</td>
<td>14</td>
<td>12</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 4.1: Comparison of our method with prior methods over 100 plans in terms of mean absolute error (MAE) from ground truth probability. Standard deviations provided in parentheses.
4.5.5 Discussion

Our fast, analytical method is used to efficiently estimate the a priori probability of collision, and consequently, the probability of success for a mobile robot such as a steerable needle operating under Gaussian models of uncertainty. We consider the correlations between the a priori probability distributions of the robot state, to accurately estimate the true distributions and consequently, the probability of success. Our method is computationally fast, enabling its use in online motion planning, and computes conservative estimates of the probability of success. We extend this framework to motion planning under uncertainty in deformable environments by estimating the probability of success based on the (expected) deformed configurations of the obstacles at each time instant along the plan. This is used to select a high quality plan and its associated feedback controller for execution.

4.6 Results

We present simulation results for our approach to guide a steerable needle through a planar deformable environment with obstacles to reach a desired target region. We use a kinematic model for a simple car without reverse (LaValle, 2006) that traverses a planar deformable environment and pushes the surface forward as it advances. This model is equivalent to a steerable needle moving through deformable tissue under 2D image guidance. Inserting needles into tissues causes the surrounding tissue to deform, thereby introducing a challenging planning problem. In addition, factors such as tissue deformation, needle-tissue interaction uncertainty, actuation and sensing uncertainty, and involuntary patient movements further complicate the task (Chap. 1).

We model the deformable environment as a unit square that is fixed at the four corners, as shown in Fig. 4.1a. The state of the needle $x_t = [x_t, y_t, \theta_t]^T \in \mathbb{R}^3$ is a vector consisting of its position $[x_t, y_t]^T$ and its orientation $\theta_t$ at time $t$. We do not include any points from the deformable environment in the definition of the state $x_t$. The control input $u_t = [v_t, \phi_t]^T \in \mathbb{R}^2$ consists of the insertion speed of the needle $v_t \in [0, v_{max}]$ and the steering angle $\phi_t \in [-\phi_{max}, \phi_{max}]$. For a bevel-tip steerable needle, the continuous steering angle can be transformed into duty-cycling parameters for
Figure 4.8: (a) Actuation noise, in the absence of feedback, causes the needle to veer away from the target region in the deformable environment. (b) Our feedback controller compensates for this noise to guide the robot to the target region. (c) A much stiffer than expected material composing the deformable environment causes the needle to miss the target region in the absence of feedback. (d) Our controller guides the robot back to the target region even with high variance in material parameters.

The state of the needle then evolves as:

$$\mathbf{x}_{t+1} = \begin{bmatrix} x_t + tv_t \cos \theta_t \\ y_t + tv_t \sin \theta_t \\ \theta_t + tv_t \tan(\phi_t)/(r_{\min} \tan(\phi_{\max})) \end{bmatrix}$$

where $t$ is the duration of a time-step and $r_{\min}$ is the minimum turning radius of the nonholonomic robot. In our experiments, we used $v_{\max} = 0.5$, $\phi_{\max} = \pi/3$, $\tau = 0.1$, and $r_{\min} = 0.25$ where all quantities are expressed in appropriate units. We also assume that we only receive feedback on the position of the needle and not its orientation, i.e. $\mathbf{h}[\mathbf{x}_t] = [x_t, y_t]$. This is a reasonable assumption since current medical imaging technologies such as ultrasound do not allow for measuring the full state of the needle tip (as the imaging resolution is often too low to infer its orientation). The noise in the sensor measurement is modeled as $\mathbf{n}_t \sim \mathcal{N}(0, \Sigma)$, where the variance in sensing noise $\Sigma$ is known.

The motion of the needle exerts an interaction force of magnitude $f_{\text{int}} = 1$ in the direction of its movement, thus deforming the environment. We model the environment as an isotropic, linearly elastic material and use a linear FEM simulator to compute the deformations in the environment as a result of the applied forces.

Given the initial state $\mathbf{x}_0$ of the needle and the target region $\mathcal{P}_\text{target} \subset \mathbb{R}^2$, the planning objective for the robot is to move to the target region without colliding with any obstacles (Fig. 4.1a). We used
the RRT algorithm as described in Sec. 4.4.1 to generate a set of 100 candidate motion plans, which took 81 seconds. Fig. 4.1b shows one such plan. We executed the remainder of our method to select a high quality plan, as shown in Fig. 4.1d. This process took 57 seconds. The output of our method is the selected motion plan and the corresponding LQG controller.

To demonstrate the effectiveness of the controller computed by our method, we show actual plan executions under actuation noise and varying material properties (Fig. 4.8). The controller computed by our method is successfully able to correct for uncertainty in actuation and deformation modeling.

To evaluate our method, we simulated 1000 executions of the planner solution to evaluate the percentage of successful plan executions. Execution of a plan is considered successful if the target region is reached and obstacles are avoided. It is important to note that each simulated plan execution assumes no knowledge of the internal deformable model being used. We considered the following scenarios to model uncertainty in actuation, sensing, and deformation modeling:

1. **Material properties (low variance):** We model the environment to have Young’s modulus \( E \sim \mathcal{N}(50.0, 5.0) \) and Poisson’s ratio \( \nu \sim \mathcal{N}(0.4, 1e^{-04}) \).

2. **Material properties (high variance):** We model the environment to have Young’s modulus \( E \sim \mathcal{N}(50.0, 75.0) \) and Poisson’s ratio \( \nu \sim \mathcal{N}(0.4, 4e^{-04}) \).
3. **Actuation noise**: The control inputs $\begin{bmatrix} \nu_t \\ \phi_t \end{bmatrix}$ are perturbed by noise drawn from a zero-mean Gaussian distribution $\mathcal{N}(0, \begin{bmatrix} 0.005 & 0.0 \\ 0.0 & 0.025 \end{bmatrix})$.

4. **Number of mesh elements**: The number of elements in the discretized mesh of the environment is uniformly distributed between 50 and 1000 elements, where the expected mesh contains 175 elements.

5. **Interaction force**: We model the interaction force $f^{\text{int}}$ to be uniformly distributed between 0 (no deformations) and 2 units.

We compared the results of our method to two existing approaches. First, we consider **RRT with deformations** in which we use a RRT plan as computed in Sec. 4.4.1 but do not explicitly consider uncertainty (i.e. no feedback control). Second, we consider **LQG-MP only** in which a plan is selected by LQG-MP (van den Berg et al., 2011) from a set of candidate plans computed in a static environment, which treats deformation as a source of uncertainty. For each approach, we simulate plan execution 1000 times using the scenarios above.

As shown in Fig. 4.9, our method consistently yields a significantly higher probability of success compared to existing approaches that plan in deterministic deformable environments without uncertainty or consider deformation as a type of uncertain perturbation. The poor rate of success of the RRT with deformations plan is due to the lack of any feedback control to compensate for uncertainty due to actuation noise and deformation modeling errors. On the other hand, the LQG-MP plan has a poor rate of success because it fails to account for large time-dependent motions of the target and obstacles due to deformations. We also evaluated the impact of the LQG-MP plan selection step of our method. Averaging across all the scenarios considered above, our method without the LQG-MP step performed 34% better than RRT with deformations and 17% worse than our complete method. The feedback controller significantly improves the rate of success by mitigating uncertainty encountered during execution of the plan. The use of the LQG-MP step further improves the rate of success at the expense of additional computation.

### 4.6.1 Clinical Scenario

We also evaluated our method in simulation in a clinical scenario involving planar needle steering within the human liver for biopsy or drug delivery (as shown in Fig. 4.10). We obtained planar
Figure 4.10: Simulation of needle steering in a plane in the human liver for biopsy or drug delivery. 
(a) Planar imaging slices from scans of a human liver from the U.S. National Library of Medicine’s 
Visible Human project. The initial state of the needle is specified (in red), the target region is marked 
in green, and the vessels inside the liver are segmented and marked as obstacles (shown in orange). (b) 
Ignoring uncertainty can lead to selection of a plan that goes through a very narrow passage between 
the vessels, thereby considerably decreasing the probability of success, even when using a feedback 
controller. (c) Consideration of uncertainty under the probability of success criterion efficiently 
computed using our method selects a high quality plan that has sufficient clearance from the vessels, 
thereby maximizing probability of successful plan execution. (d) Considerable uncertainty due to 
actuation and sensing errors, and errors in estimating modeling parameters causes the needle to hit a 
vessel in the absence of feedback. (e) Our deformation-aware controller guides the needle back to 
the target region even under considerable uncertainty.

imaging slides from computed tomography (CT) scans of a human liver from the U.S. National 
Library of Medicine’s Visible Human project database (U.S. National Library of Medicine, 2012). In 
a clinical procedure, high resolution scans would typically be obtained in the preoperative stage of 
the procedure. A clinician would specify the initial state of the needle (position and orientation), a 
desired target region, and specify major blood vessels that must be avoided during the procedure to 
prevent hemorrhaging during the procedure (Fig. 4.10a).

Figs. 4.10b and 4.10c demonstrate the importance of plan selection during preoperative opti-
mization (Sec. 4.5). Ignoring uncertainty can lead to selection of a plan that goes through a very
narrow passage between the vessels (Fig. 4.10b), thereby considerably decreasing the probability of success, even when using a feedback controller. On the other hand, if we choose a high quality plan based on our probability of success criterion, we can compute safe plans that have sufficient clearance from the blood vessels in case of unexpected deviation in the needle pose (Fig. 4.10c), thereby maximizing the probability of successful plan execution.

We simulate actual execution of the procedure with the chosen high quality plan and LQG feedback controller by considering uncertainty due to actuation and sensing errors, and modeling errors due to improper initialization of material parameters. The needle hits the vessel while executing the high quality plan without feedback in the presence of uncertainty (Fig. 4.10d). In contrast, our deformation-aware controller guides the needle back to the target while safely avoiding obstacles under considerable uncertainty (Fig. 4.10e).

4.7 Discussion

We have introduced a new, unified framework for planning and control under uncertainty in highly deformable environments that maximizes the probability of success by accounting for uncertainty in deformation models, noisy sensing, and unpredictable actuation. Unlike prior planners that assume deterministic deformations or treat deformations as a disturbance, our method explicitly considers uncertainty in large, time-dependent deformations. Although the method requires a simulator of the deformable environment, we place no significant restrictions on the simulator used. We have shown that our approach can generate high quality plans for guiding steerable needles through highly deformable tissue under 2D image guidance.

Our approach has a few limitations. First, we operate under the assumption of Gaussian models of uncertainty. This might not be an acceptable approximation in applications where multi-modal beliefs are expected to appear. However, preliminary results from Chapter 3 indicate that the Gaussian approximation is well founded for the problem of needle steering in soft tissue. Second, the feedback controller used in this work does not take bounds on the physical control inputs that can be applied to the system. This is a practical concern for needle steering, since curvature greater than the maximum curvature of the needle cannot be realized. We plan to combine our approach with model predictive control based methods that would use a fast planner (Chapter 2) to correct unexpected perturbations.
that cannot be effectively corrected by the LQG controller. Our analytical method for efficiently estimating the probability of success works well for steerable needles but we plan to extend this method to be applicable to non-point robots.

In future work, we plan to investigate improvements to each component in Fig. 4.2. Replacing the standard LQR control framework with integrated approaches for planning and control that compute an approximate solution to the POMDP problem (Platt et al., 2010; van den Berg et al., 2012) may improve controller performance. Similarly, replacing the standard Kalman filter with variants such as the unscented Kalman filter or a particle filter (Simon, 2006) may improve the quality of state estimation during plan execution. Our approach also assumes that a simulator that computes the expected deformations in the environment is available, but such simulators are difficult to construct for general applications. We envision that advances in computational modeling and simulation will further increase the applicability of our method. We also plan to investigate parallelizing the model linearization, which involves multiple, independent simulation runs, to reduce computation times.
CHAPTER 5

Conclusion and Future Work

Steerable medical needles have the potential to improve health care by improving the effectiveness of needle-based clinical procedures such as biopsy, drug delivery, neurosurgery, and radioactive seed implantation for cancer treatment. However, several hurdles need to be overcome before needle steering can be realized in practice. A big part of the challenge stems from the difficulty involved in accurately guiding these needles to clinical targets while safely avoiding sensitive and impenetrable anatomical structures. Creating a needle steering robotic system that assists clinicians and addresses these challenges could enable new needle-based procedures and substantially improve the clinical outcomes of some existing needle-based procedures.

In this dissertation, we have addressed a number of issues related to planning and control of steerable medical needles. We have demonstrated that efficient motion planning and control approaches can facilitate closed-loop guidance of steerable needles to clinical targets within clinically acceptable accuracy while avoiding sensitive and impenetrable anatomical structures. We have proposed two approaches for closed-loop planning and control of steerable needles, overcoming substantial deformations and uncertainty in the process. We have also proposed a data-driven method for creating stochastic models of steerable needle insertion, which could be used to create realistic medical training simulators of steerable needle procedures and to improve the effectiveness of existing planning and control techniques.

Our results, albeit presented in the context of medical needle steering, could be adapted to a number of applications, including manipulation of deformable objects and planning and control of mobile robots.

The main results from this dissertation are summarized below.
5.1 Summary of Results

We revisit the thesis from Chapter 1:

*Efficient motion planning and control techniques that consider uncertainty and deformations can facilitate closed-loop guidance of steerable needles to targets within clinically acceptable targeting accuracy while safely avoiding clinician-specified anatomical structures.*

In support of this thesis, we presented two approaches for closed-loop planning and control of steerable needles in soft tissue for automatically guiding the steerable needle to targets in 3D environments while avoiding obstacles and compensating for real-world uncertainties. The first approach uses a fast planner that can be used as a controller (Chapter 2). The second approach uses a physically-based simulator to model expected deformations to compute a motion plan and correct perturbations using a feedback controller (Chapter 4). We showed that the rapid replanning strategy was able to efficiently guide the needle safely to targets in phantom and ex vivo animal tissue with accuracy exceeding the accuracy of current clinical care. The unified planning and control framework for steerable needle insertion in soft tissue was shown to be superior to prior approaches that consider either uncertainty or deformations but not both simultaneously.

We also presented a data-driven method for creating stochastic models of needle insertion in soft tissue (Chapter 3). This method relies on data gathered from prior experiments and procedures and can be constantly refined as the availability of data from steerable needle procedures increases. We envision that this method could be an integral component for modeling real-world perturbations in simulations designed to help clinicians to train for steerable needle insertion procedures or for designing efficient planning and control algorithms that maximize the chances of successful plan execution. We created stochastic models for needle insertion for tissue phantoms and ex vivo porcine tissue based on real data gathered from needle insertions in these tissue sample materials. We also evaluated the accuracy of these models by simulating steerable needle insertions in these materials.

5.2 Limitations

For each of the techniques presented in this dissertation, we have discussed several limitations in the corresponding chapters. Here, we summarize the key limitations of each.
The rapid replanning approach described in Chapter 2 requires a detailed characterization of the steerable needle’s curvature for best results. Since we cannot obtain a detailed characterization on a per-patient basis before the procedure is performed, we could treat this as an additional source of uncertainty and account for erroneous characterization during closed-loop planning and control. The optimality and completeness of the rapid replanning approach is also an open area of research.

The data-driven method for estimating the parameters of a stochastic model of steerable needle insertion was presented in Chapter 3. The method assumes a discrete time kinematic motion model. Since the model parameters are estimated based on measurements that are obtained at discrete intervals of needle insertion length, our stochastic model cannot be used to simulate needle procedures that might vary greatly in the length of the needle shaft inserted at every insertion step. Finally, we do not estimate the curvature of the needle, which is important from the perspective of state estimation, planning, and control. The dependence of our method on the curvature occurs through the computation of the physical control inputs that are applied to the needle during duty cycled rotation of the needle. However, our method will likely work well for needles with similar curvature characteristics.

The unified framework of planning and control in uncertain, deformable environments presented in Chapter 3 constructs a feedback controller based on linearization of the physically-based deformable simulator. The linearization is only valid in regions close to the nominal trajectory of the needle. In the case of large perturbations, this might lead to sub-optimal performance in terms of the targeting accuracy and avoidance of critical anatomical structures. We plan to investigate refining the controller as the procedure is being performed for improving the probability of success.

5.3 Future Work

There are many exciting areas for further work regarding medical needle steering, in terms of developing its further applications and realizing it in actual clinical practice, and in formalizing its theoretical basis. We have described in each chapter some of the avenues of future work with regards to each specific application. Here, we discuss some of the broader possibilities and open problems.

**Teleoperation of steerable needle procedures:** A major challenge in using steerable needles for clinical procedures is to involve a clinician in the loop for reasons of accountability and patient
safety (Taylor, 2006; Okamura et al., 2010). Teleoperated robotic systems such as the da Vinci® surgical system manufactured by Intuitive Surgical (Intuitive Surgical da Vinci® Surgical System, 2012) are extensively used for minimally invasive surgical procedures worldwide. We plan to explore the use of our fast motion planning algorithm (Chapter 2) for clinician-teleoperated position control of the needle tip.

Optimality and completeness of replanning: An important issue pertaining to replanning is the completeness of the approach, i.e., to guarantee that a solution would be found if one exists. The completeness properties of sampling-based motion planners such as the RRT algorithm has been thoroughly investigated (LaValle, 2006). Since replanning involves using a sampling-based planning to find a solution at each replanning step, it remains to be investigated if replanning for steerable needles can be implemented in a manner that is fast in practice, is guaranteed to find a solution if one exists, and provides some level of optimally guarantee.

Maximize probability of success in replanning: A naïve strategy to ensure safe avoidance of critical anatomical structures is to maximize clearance from the obstacles, but a more principled approach would be to explicitly consider a stochastic model of the system dynamics (Chapter 3) and select motion plans that would minimize the probability of collisions with obstacles. This would involve extending the replanning technique presented in Chapter 2 to incorporate duty cycled spinning of the needle, and efficiently computing the probability of collision, and consequently the probability of success, using the method suggested in Chapter 4 (Patil et al., 2012). A motion planner based on a metric that quantifies probability of collision could enable a principled approach to maximizing probability of success.

Online estimation of model parameters: The unified planning and control framework presented in Chapter 4 deals with two fundamentally different kinds of uncertainty. The first kind deals with uncertainty in modeling and simulation, which is a function of the model parameters such as constituent material properties and friction and contact coefficients. Modeling errors can cause large systemic changes that cannot be corrected using a static control policy, which is computed a priori assuming a set of model parameter values. The second kind of uncertainty arises due to temporally-varying factors such as actuation and sensing, which is a continuous random process. The estimation process described in Chapter 3 can theoretically be used for online estimation of model
parameters as well. These model parameters could be used to revise the feedback control policy, resulting in a higher probability of success.

**Integrated planning and control using approximate POMDP solvers:** In the most general setting, the planning and control problem is a partially observable Markov decision process (POMDP) problem (Thrun et al., 2005). Recent results (Platt et al., 2010; van den Berg et al., 2012) have proposed approaches for efficiently computing locally optimal solutions to the POMDP problem by planing in an augmented state space called the belief space, which comprises of both the system state and a parameterized representation of the uncertainty associated with the state. These approaches yield control policies over the belief space, as opposed to control policies computed over the state space (Bertsekas, 2007). It remains to be investigated if these approaches can be scaled and extended to steerable needle insertion in soft tissue.

**Simultaneous optimization of motion and sensor plans:** An important aspect of motion planning under uncertainty problems is to consider the effect of sensing modalities on the planning and execution of motion plans. The type and placement of sensors in the environment can have a major impact on the quality of the computed plans in terms of minimizing uncertainty, and the ability of the robot to follow the computed motion plan. In recent work (van den Berg et al., 2010), we proposed a naïve sampling-based strategy to simultaneously compute motion and sensor plans. Our method showed that such a strategy can improve the probability of success when compared to prior approaches that either compute motion plans or sensor placements but not both simultaneously. Designing algorithms for simultaneously planning needle steering motions and sensor placements for a variety of clinically relevant sensors such as X-ray imaging devices or ultrasound probes is a promising topic for future research.
BIBLIOGRAPHY


107


