

# A probabilistic framework for tracking uncertainties in robotic manipulation

Huy Nguyen and Quang-Cuong Pham

**Abstract**—Precisely tracking uncertainties is crucial for robots to successfully and safely operate in unstructured and dynamic environments. We present a probabilistic framework to precisely keep track of uncertainties throughout the entire manipulation process. In agreement with common manipulation pipelines, we decompose the process into two subsequent stages, namely perception and physical interaction. Each stage is associated with different sources and types of uncertainties, requiring different techniques. We discuss which representation of uncertainties is the most appropriate for each stage (e.g. as probability distributions in  $SE(3)$  during perception, as weighted particles during physical interactions), how to convert from one representation to another, and how to initialize or update the uncertainties at each step of the process (camera calibration, image processing, pushing, grasping, etc.). Finally, we demonstrate the benefit of this fine-grained knowledge of uncertainties in an actual assembly task.

**Index Terms**—manipulation, assembly, uncertainty

## I. INTRODUCTION

While tracking uncertainties has long been a central theme in mobile robotics [1], it has received comparatively less attention in industrial robotics in general, and robotic manipulation in particular. Yet, precisely tracking uncertainties is crucial for robots to successfully and safely operate in unstructured and dynamic environments which, according to many reports<sup>1</sup>, are becoming more common in the industry.

To illustrate, consider the assembly task depicted in Fig. 1. For the robot to successfully assemble the two objects, the poses of the objects relative to the robot must be known to the system. However, uncertainties are introduced into the object pose estimation during camera calibration and image processing. Next, these uncertainties evolve during the physical interactions between the robot and the objects (pushing, grasping, etc.). Thus, the actual object poses before the last step of the assembly might significantly differ from the assumed poses, which in turn may cause assembly failures.

While it is not possible to totally eliminate all uncertainties – because of the inherent noise introduced during the various perception and action stages – we argue in this paper that precisely keeping track of the uncertainties can significantly improve the speed and success rate of manipulation.

The authors are with the School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore. Corresponding author: Huy Nguyen (email: huy.nguyendinh09@gmail.com).

<sup>1</sup>See for instance “The Robotics Revolution: The Next Great Leap in Manufacturing” by the Boston Consulting Group <https://www.bcg.com/publications/2015/lean-manufacturing-innovation-robotics-revolution-next-great-leap-manufacturing.aspx>.

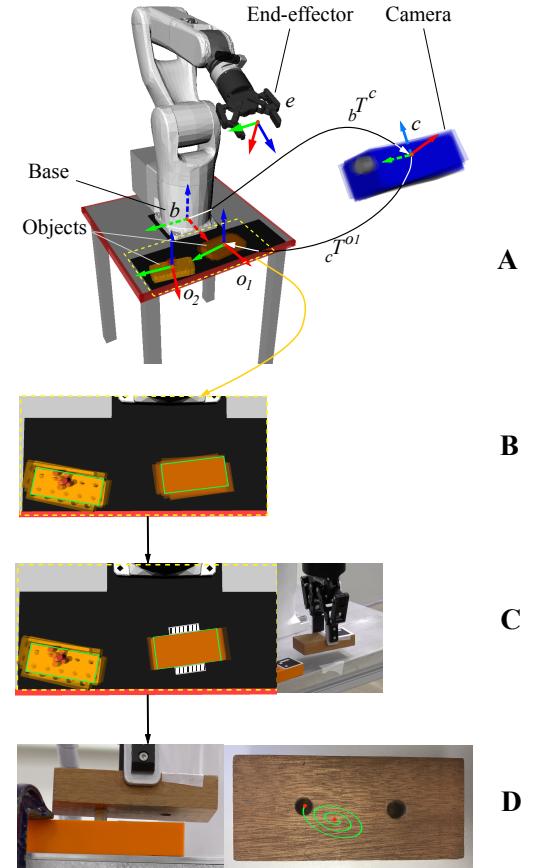


Figure 1. A typical assembly task. **A:** Setup, comprising a 6-DOF robot equipped with parallel-jaw gripper, a 3D camera, and objects to be assembled. **B:** Uncertainty in object pose estimation. **C:** The uncertainty on the pose of the brown object has evolved under the grasping action. **D:** The elliptical search pattern for assembly is optimized based on the fine-grained knowledge of the uncertainties.

More specifically, we present a probabilistic framework to precisely keep track of uncertainties throughout the entire manipulation process. In agreement with common manipulation pipelines [2], we decompose the process into two subsequent stages, namely perception and physical interaction. Each stage is associated with different sources and types of uncertainties, requiring different techniques. We discuss which representation of uncertainties is the most appropriate for each stage (e.g. as probability distributions in  $SE(3)$  during perception, as weighted particles during physical interactions), how to convert from one representation to another, and how to initialize or update the uncertainties at each step of the process (camera calibration, image processing, pushing, grasping, etc.). Finally,

we demonstrate the benefit of this fine-grained knowledge of uncertainties in an actual assembly task.

The remainder of the paper is organized as follows. Section II summarizes related works. In Section III, we present our framework for precisely tracking uncertainties in manipulation. In Section IV, we demonstrate the benefit of the framework in several physical experiments. Finally, in Section V, we conclude and sketch some directions for future research.

## II. RELATED WORKS

### A. Uncertainty in robotics

Some fundamental works on modelling spatial uncertainties were studied in [3], [4], [5], [6]. The above developments have been utilized to establish a theoretical basis for a popular subfield of the *simultaneous localization and mapping* (SLAM) problem [1], [7]. However, most existing researches have only focused on *mobile robotics*. Not much attention was paid explicitly to the modelling, estimating and tracking the uncertainties in *manipulation tasks*. In fact, most manipulation tasks are sequences of primitive phases consisting of calibration, perception and manipulating actions. Existing works, however, have only focused on estimating uncertainties in individual phases, not throughout the whole process, where multiple sources of uncertainties interact and compound [8].

The works closest to our paper are [9], [10], in which authors presented a methodology for manipulating and propagating spatial uncertainties in a robotic assembly system for generating executable actions for accomplishing a desired task. Although the works proposed a novel framework to work with uncertainties, there was no discussion on how these uncertainties are estimated. Based on those works, [8] proposed methods for estimating the geometrical relationships between coordinate frames and the spatial uncertainties in estimated model parameters. However, later phases (i.e. manipulating actions) where contacts between robot(s) and the environment occur have not been investigated. On another note, their estimation model were based on an Euler angles parameterization, as opposed to the  $SE(3)$  formulation, is well-known to involve singularities.

### B. Representation and propagation of uncertainty

In early works by [11] and [12], the uncertainty of a pose is simply represented by worst-case bounds which include all possible errors (min-max approach). This simple approach usually results in conservative estimates which make it difficult to apply in decision-making process.

To address this limitation, the probabilistic approach which uses the calculus of probability theory and assign probabilities to all potential positions of the object has been proposed. The pose of an object is now represented by a probability distribution over the space [9], [13]. As a result, the probabilistic representation can make use of probability theory and, thus, provide more uncertainty manipulation methodologies (e.g. propagation, fusion, etc.) [3], [4], [14], [15]. Recently, this probabilistic approach has been further investigated in [5], [6], [16], [17] where they provided a rigorous treatment

of representing and propagating uncertainty on  $SO(3)$  and  $SE(3)$ .

### C. Alternative approaches to deal with uncertainty in manipulation

A common approach to deal with uncertainties while performing manipulation tasks is to plan the geometric motions of the objects and then execute these motions on a compliant robot. This simple approach has proved to be a key component for the success of many manipulation tasks [2], [18], [19], [20]. To compute robust motions under the presence of motion and sensing uncertainties, many works have incorporated such information into their planners [21], [22], [23]. This line of research are generally referred to as belief space planning [24], [25], [26]. Even though such approaches have enabled many works to reliably accomplish complex manipulation tasks in realistic, uncertain environment, they commonly requires a set of hypotheses of pose for every manipulated object in the environment. In fact, these sets of hypotheses are often assumed to be much larger than the real ones to account for ambiguous estimations. Consequently, robots have to perform many more redundant motions and spend more time to reason the effects of these uncertainties. This lack of access to explicit uncertainty information motivates us to study and proposed a new probabilistic framework for tracking uncertainties in manipulation tasks. In turn, the capabilities to obtain such fine-gain information enabled by our framework can also be utilized in these planning approaches to achieve a higher level of performance when dealing with uncertainties.

## III. A PROBABILISTIC FRAMEWORK FOR TRACKING UNCERTAINTY

In general, most typical manipulation tasks can be decomposed into two subsequent stages, namely perception and physical interactions, see Figure 1. Each stage is associated with different sources and types of uncertainties, hence requiring different techniques to deal with. The rest of this section will discuss which representation of uncertainties is the most appropriate for each stage, and how to initialize or update the uncertainties at each step of both stages.

### A. Choices of uncertainty representations

1) *During perception stage*: During this stage, we choose to represent uncertainties of 3D poses as probability distribution in  $SE(3)$ . This is accomplished by storing the means as a (singularity-free) transformation matrix and using a (constraint-sensitive) perturbation of the pose (with associated covariance matrices). In particular, we model the uncertainties of the rotation and the translation parts separately as follows:

We assume that the rotation part of a pose is corrupted as below

$$\mathbf{R} = \exp([\boldsymbol{\xi}_{\mathbf{R}}]) \bar{\mathbf{R}}, \quad (1)$$

where  $\bar{\mathbf{R}} \in SO(3)$  is the mean of  $\mathbf{R}$ , and the small perturbation variable  $\boldsymbol{\xi}_{\mathbf{R}} \in \mathbb{R}^3$  is zero-mean Gaussian with covariance matrice  $\boldsymbol{\Sigma}_{\mathbf{R}}$ .

The translation part of a pose is corrupted as below

$$\mathbf{t} = \boldsymbol{\xi}_t + \bar{\mathbf{t}}, \quad (2)$$

where  $\bar{\mathbf{t}} \in \mathbb{R}^3$  is the mean of  $\mathbf{t}$ , and  $\boldsymbol{\xi}_t \in \mathbb{R}^3$  is the zero-mean Gaussian perturbation with covariance matrix  $\Sigma_t$ .

This representation is free of singularities and avoids the need to enforce constraints when solving optimal estimation problems. Moreover, such representation also allows us to utilize the Lie groups and their associated mathematical machinery to provide a rigorous treatment to address many essential operations, i.e. propagation and fusion of uncertainties (see later in III-B). Thanks to these advancement, we are able to store uncertainty information as analytical form and evaluate the uncertainties of object pose estimation in timely fashion.

In fact, the representation also implies that rotation and translation noises are independent<sup>2</sup>. Because of this assumption, the space we consider is, strictly speaking,  $SO(3) \times \mathbb{R}^3$ , rather than  $SE(3)$ .

Under this representation, belief states are represented by multivariate normal distributions (Gaussian distributions). Although it can bring us a fast, analytical method to estimate uncertainties in the perception stage, there are a number of shortcomings. Most importantly, Gaussian distributions are unimodal (probability density exhibits single peak). However, in other cases where one must handle multimodal distributions, i.e. during the physical interaction stage, particles are more preferable.

*2) During physical interaction stage:* Unlike the perception stage where camera measurements are easily used to infer the poses of the objects, physical interactions only provide local information about those poses. Consequently, motion and observation models during this stage are often highly non-linear and lack of analytic derivatives. Moreover, the resulting belief state is usually non-Gaussian and may be multi-modal. This, therefore, naturally leads to the use of particle filter and its variants to track the belief state during the physical interaction stage.

### B. Perception

Regarding the perception stage, we are interested in estimating the poses of the objects and their associated uncertainties with respect to the robot base frame. As illustrated in Figure 1, such uncertainties are definitely originated from three main sources which include: (i) the uncertainty of the object pose estimation in the camera frame, (ii) the uncertainty of the camera position (commonly known as the hand-eye transformation), and (iii) the uncertainty of the robot end-effector positioning.

$${}_b\mathbf{T}^{o1} = {}_b\mathbf{T}^c {}_c\mathbf{T}^{o1} \quad (3)$$

We note that in this setting the camera is mounted at a fixed position in the environment (not on the robot end-effector).

<sup>2</sup>Since there is in general no bi-invariant distance on  $SE(3)$  [27], when solving optimal estimation problems, finding the rotation and translation components simultaneously would require a non-trivial rotation/translation weighting in any cases. As a result, we choose to model the uncertainties of the rotation and the translation parts separately instead of working directly on  $SE(3)$ .

Hence, the uncertainty of the robot end-effector positioning does not contribute directly to the uncertainties of the object poses. Instead, it affects the uncertainty of the camera position via the hand-eye calibration process [28].

In this section, we shall provide a methodology to estimate explicitly these sources of uncertainties. We will also give a discussion on how the uncertainties can be propagated so that the poses of the objects and their associated uncertainties with respect to the robot base frame can be eventually obtained.

#### 1) Estimating uncertainties:

- *Uncertainty on the robot positioning  ${}_b\mathbf{T}^e$ :* Thanks to the advancement of calibration methods and measurement equipment, robots are now better built with higher accuracy. After calibration, a common positioned control industrial robot can achieve sub-millimeter in the mean position errors of the end-effector. For example, our experiments show that the mean of the position errors of our calibrated manipulator is about 0.3 mm. This can be considered negligible compared to other sources of uncertainties in our system, e.g. the uncertainty of the pose estimation using our camera system. Because of this result, we find it usually safe to assume accurate joint measurements and robot model.
- *Uncertainty on the hand-eye calibration  ${}_b\mathbf{T}^c$ :* Commonly, this problem can be formulated as: solve for  $\mathbf{X}$  in  $\mathbf{AX} = \mathbf{XB}$ , where  $\mathbf{X}$  is the unknown  $4 \times 4$  hand-eye transformation matrix and  $\mathbf{A}$  and  $\mathbf{B}$  are known  $4 \times 4$  transformation matrices.

Here, we are interested, not merely in solving for the hand-eye transformation, but more comprehensively, in evaluating its covariance. To address this problem, we follow [28] where they propose a novel algorithm to rigorously work out such a derivation. In particular, the technique exploits the benefits of applying optimization techniques directly on  $SE(3)$  to obtain the derivation of the covariance of the hand-eye transformation.

- *Uncertainty on the camera-based pose estimation  ${}_c\mathbf{T}^o$ :* Regarding the camera-based pose estimation problem, its uncertainty mainly depends on the intrinsic calibration of the camera. However, a model to capture the uncertainties of the camera parameters is highly non-linear and difficult to be evaluated. Hence, we decide to estimate such uncertainty by drawing a number of samples and performing Monte Carlo estimation. Despite the fact that this approach is slow, it is proven to bring out a more accurate estimation. A more detailed clarification of how this Monte Carlo estimation performed in the real systems can be found in the Experimental Section of [28].

- 2) Propagating uncertainties:* After obtaining all sources of uncertainties in the perception stage, we are now in a position to estimate the uncertainty of the object pose with respect to the robot base frame. In this work, we chose to follow the method proposed in [28] (Appendix A), where the covariance propagation method are derived for the case where rotation and translation are decoupled.

### C. Physical interactions

In this work, we focus on manipulating tasks that usually perform on industrial assembly where common elements, i.e. force/torque sensors and parallel-jaw grippers, are ubiquitous due to their strength, robustness, cost-effectiveness, ease of integration, and many more. The work in this paper will only cover two manipulating actions commonly performed in manipulation, which are planar grasping actions and touch-based localization.

1) *Problem formulation:* Given the initial distribution of the objects obtained from the previous perception stage, the goal is to update the object distributions after the planar grasping action or/and the touch-based localization. As commonly known, the Bayesian filter is considered as the most general algorithm for filtering a belief state given initial knowledge and a sequence of actions and observations. Hence, the problem will be cast into the Bayesian framework and be addressed as a nonlinear filtering problem as shown below.

Let  $\mathbf{X}$  be the state of a dynamical system which evolves under actions  $\mathbf{u}$  and provides observations  $\mathbf{y}$ . Starting with  $P(\mathbf{X}_0)$  – the *prior* distribution over the state  $\mathbf{X}$  – the goal is to recursively update the following conditional probability

$$P(\mathbf{X}_{t+1}|\mathbf{y}) = \eta P(\mathbf{y}|\mathbf{X}_t)P(\mathbf{X}_t). \quad (4)$$

Here  $P(\mathbf{X}_{t+1}|\mathbf{y})$  is known as the *posterior*, which represent our uncertain belief about the state  $\mathbf{X}$  after having incorporated the measurement  $\mathbf{y}$ . On the right-hand side, the first factor  $P(\mathbf{y}|\mathbf{X}_t)$  is the *measurement probability*, which encodes the likelihood of the measurement given the state (*measurement model*). The second factor  $P(\mathbf{X}_t)$  is the *prior*, which represents our belief about  $\mathbf{X}$  before obtaining the measurements  $\mathbf{y}$ . The factor  $\eta$  is a normalizing factor independent of the state  $\mathbf{X}_t$  and needs not be computed.

Over the next sections, we will discuss how to apply this filter into two common contact actions, which are planar grasping actions and touch-based localization.

2) *Planar grasping:* Regarding planar grasping action, the state is the pose  $\mathbf{X} \in SE(2)$  of the manipulated object. Actions are motions of the hand, given by the velocity  $\mathbf{u}$ . During contact, the object moves with a velocity  $f_\phi(\mathbf{X}, \mathbf{u})$  where the function  $f$  encodes the physics of the object motion in response to the intended motion of the gripper. The parameter  $\phi$  includes environmental properties. In this work, we particularly build analytical state estimators to track the poses of the objects from the post-grasp gripper distances based on the works from [29], [30]. This motion model is constructed following the assumption of quasi-static rigid body mechanics with Coulomb friction. Such assumption not only allows us to attain more approachable and simpler models, but also well suit the scale and speed of our application.

Relating to the Bayesian updates, the particle filter first samples the particles  $\mathbf{X}^i$  from the prior distribution, then uses the motion model in order to simulate and obtain  $\mathbf{X}^{i,\text{sim}}$ . Once the final object poses are acquired, their associated finger widths  $d_{\mathbf{X}^{i,\text{sim}}}$  can be estimated. This information will be used to compute an important weight for each forward-simulated particle. During the weighting step, the particles which are

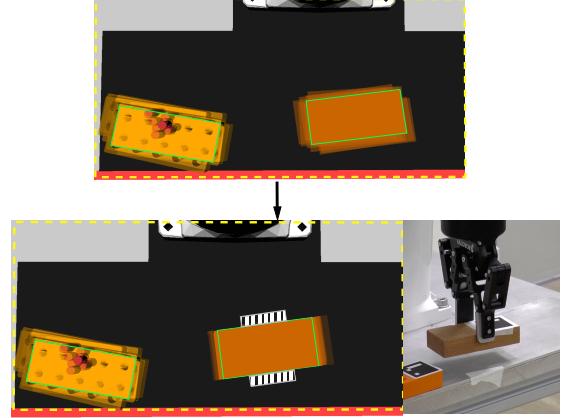


Figure 2. Once a planar grasping action is successfully performed, it sometimes can significantly improve the estimation of the objects in the closing direction of the gripper.

consistent with our measurement  $\mathbf{y}$  will be assigned with higher probability. In particular, the post grasp distance measurements are assumed to be corrupted by Gaussian noise with the variance  $\sigma_d$ . The measurement probability is computed as follows

$$P(\mathbf{y}|\mathbf{X}_t^i) = \eta_y \exp\left(-\frac{1}{2} \frac{(\mathbf{y} - d_{\mathbf{X}_t^i, \text{sim}})^2}{\sigma_d^2}\right), \quad (5)$$

where  $\eta_y$  is a constant and will be taken into account during the normalization.

It is also worth-noting that once a planar grasping action is successfully performed, it can significantly improve the estimation of the objects in the closing direction of the gripper. Especially, in the cases where the geometries of the grasped objects are simple, e.g. rectangular boxes, the uncertainties of the objects could shrink to an one-dimensional distribution, see Figure 2. Besides, in this work, we also assume that the objects do not end up slipping out or being jammed at an undesired position. This assumption is considered reasonable in this case because the object geometries are simple and the slow movement of the fingers is always under control.

3) *Touch-based localization:* With regards to touch-based localization, this type of interaction is often employed to further improve the pose estimation of the objects. As mentioned earlier, the objects to be located are, in general, assumed to be static during the measurement collection. This commonly-chosen assumption is realistic in such cases of handling very slight contact and the objects that are heavy or mounted on a support fixture to prevent their possible movements.

In this problem, the state is the pose  $\mathbf{X} \in SE(3)$  of an object  $\mathcal{O}$  with a known shape. The measurements  $\mathbf{y} = \mathbf{y}_0, \dots, \mathbf{y}_n$  are obtained by touching the object with the end-effector of the robot (as shown in Figure 3). Each measurement  $\mathbf{y}_k := (\mathbf{y}_k^{\text{pos}}, \mathbf{y}_k^{\text{nor}})$  consists of an contact position  $\mathbf{y}_k^{\text{pos}}$  and a contact normal  $\mathbf{y}_k^{\text{nor}}$ .

In fact, many works have been proposed and are able to solve the 6-DOF localization problem efficiently and reliably. In this work, we use the *proximity measurement model* and *Scaling Series* method in [31], [32] to perform the touch-based



Figure 3. The object localization problem via touch. The photo shows the robot interacting with a wooden stick using a pin and its force/torque sensor.

localization owing to its computational efficiency.

#### IV. APPLICATION TO PART ASSEMBLY

In this section, we discuss an application of our framework to parts assembly tasks. The robotic platform used in this work is characterized by cost-efficient, off-the-shelf components combined with a classical position-control industrial manipulator. In particular, the main components of our platform are:

- 1 Denso VS060: Six-axis industrial manipulator.
- 1 Robotiq Gripper 2-Finger 85: Parallel adaptive gripper designed for industrial applications. Closure position, velocity and force can be controlled. The gripper opening goes from 0 to 85 mm. The grip force ranges from 30 to 100 N.
- 1 ATI Gamma Force-Torque (F/T) Sensor: It measures all six components of force and torque. They are calibrated with the following sensing ranges:  $f = [32, 32, 100]$  N and  $\tau = [2.5, 2.5, 2.5]$  Nm.
- 1 Ensenso 3D camera N35-802-16-BL

##### A. Single pin insertion

To illustrate the effectiveness of the framework, we first consider a dexterous task: a cylindrical peg-in-hole task. In this task, two rectangular boxes ( $20 \times 50 \times 110$ ) mm are placed randomly on a known table. The first box  $\text{Obj}_1$ , also the static one, is clamped down on the table with a cylindrical pin ( $r = 4$  mm,  $l = 30$  mm) pre-inserted. The second box  $\text{Obj}_2$ , considered as the movable one, will be grasped by the robot. Besides, the camera mounted on a fixed tri-pod will be used to localize the two above boxes (see Figure 1). As shown in the Figure, to perform its task, the robot arm needs to grasp  $\text{Obj}_2$  and insert it into  $\text{Obj}_1$ . After it completes the insertion, it will then release  $\text{Obj}_2$ . Note that the geometries of all objects are assumed to be known.

This is a good example for our framework illustration since most of the capabilities we have developed are required to perform this task. The task also constitutes one of the key steps in many assembly processes. In addition, it yields a fully quantifiable way to measure the performance of a robotic manipulation system. For these reasons, it is considered as a good test for the generalizability and simplicity of our approach.

Regarding the execution of the task, the process is naturally divided into three sequent sub-tasks as below:

- locate the two objects by using the camera;

- pick  $\text{Obj}_2$ ;
- compliant insertion of the two objects.

1) *Perception*: A camera mounted on a fixed tri-pod is used to localize the two boxes. In this setting, we note that the transformation of an object relative to the robot base is given by

$$_b\mathbf{T}^o = {}_b\mathbf{T}^c {}_c\mathbf{T}^o. \quad (6)$$

This implies that the uncertainty of the robot positioning does not directly affect the uncertainty of the object pose with respect to the robot base. However, as mentioned in Section III-B, the uncertainty of the robot positioning (if existing) also affects the estimation of the  $_b\mathbf{T}^c$  via the hand-eye calibration. Again, it is noted that we assume the uncertainty of our manipulator is negligible compared to the uncertainty of the pose estimation using our camera system. We follow the proposed framework in Section III-B to estimate the poses of the two objects and their associated uncertainties.

2) *Pick Obj<sub>2</sub>*: Once the two objects are located, the robot arm needs to perform the picking task. In fact, this task resembles the planar grasping action which has been discussed earlier in Section III-C2.

To be specific, before grasping  $\text{Obj}_2$ , the robot arm moves its gripper to a pre-grasp position which is above the object. The gripper, then, gradually close its fingers until it exceeds the force limit. After the execution of the grasping action, we read the distance between the fingers from the Robotiq gripper encoder, and update the distribution of the object. In this step, the previous information obtained from the perception sub-task is used as the initial uncertainty region of  $\text{Obj}_2$ . Note that since we assume the table plane to be known, we then project  $\text{Obj}_2$  uncertainty region onto this surface. Moreover, the pose of  $\text{Obj}_2$  and its associated uncertainty will be presented in the gripper local frame as  $\text{Obj}_2$  is now rigidly grasped by the gripper.

3) *Compliant pin insertion*: Next, pin insertion is the final sub-task to be executed. Due to the uncertainties on the position of the objects ( $\text{Obj}_1$  and  $\text{Obj}_2$ ), the exact position of the hole/pin is unknown. Moreover, given the parameters of the peg-in-hole set-up, we observe that the insertion will fail in case the position errors are more than 0.5 mm. To cope with such uncertainty, many researchers have proposed to apply force-controlled exploration. For example, [2] used a super-imposed spiral search pattern. In their setting, the hole position is uncertain. The robot is controlled to follow the spiral pattern while maintaining the contact between the pin and the hole. Once the hole is found, the exploration will be immediately terminated. Even though this strategy is able to precisely and consistently detect holes and perform tight pin insertions (100% success across the 28 insertions), it is indeed a “blind searching” algorithm since there is no information about the uncertainties of the objects to be used. This limitation motivates us to study how such uncertainties can be employed to design a better compliant insertion strategy as below.

After the gripper grasps  $\text{Obj}_2$ , the robot arm moves the  $\text{Obj}_2$  to the pre-inserting position. It, then, moves down until it touches the surface of  $\text{Obj}_1$ . Once they are in contact, the

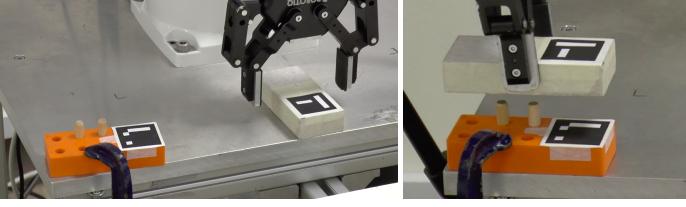


Figure 4. We exploit the benefits of having access to uncertainties information in a double pins insertion task.

relative transformation between the two objects is as follows:

$$\text{\scriptsize{\textit{o1}}} \mathbf{T}^{\text{\scriptsize{\textit{o2}}}} = \text{\scriptsize{\textit{o1}}} \mathbf{T}^b {}_b \mathbf{T}^e {}_e \mathbf{T}^{\text{\scriptsize{\textit{o2}}}}, \quad (7)$$

where  $\text{\scriptsize{\textit{o1}}} \mathbf{T}^b$  is the transformation of the base with respect to  $\text{\scriptsize{\textit{Obj1}}}$  frame,  ${}_e \mathbf{T}^{\text{\scriptsize{\textit{o2}}}}$  is the relative transformation between  $\text{\scriptsize{\textit{Obj2}}}$  and the end-effector (gripper). As the uncertainty of  ${}_e \mathbf{T}^{\text{\scriptsize{\textit{o2}}}}$  is represented by particles, we assume it to be a single peak Gaussian distribution and perform an empirical estimation to obtain its parameters<sup>3</sup>. The uncertainty of  $\text{\scriptsize{\textit{o1}}} \mathbf{T}^b$  is, then, obtained by the forward propagation method. The search pattern is now optimized based on such fine-grained knowledge of the uncertainty. Instead of using the “circular” spiral trajectory, we change the spiral pattern which follows the elliptical shape of the one-standard-deviation covariance ellipsoid of the  $\text{\scriptsize{\textit{o1}}} \mathbf{T}^b$  distribution (see Figure 1).

Compared to the traditional “circular” spiral search strategy, our new proposed spiral pattern can successfully perform the insertion at a faster speed. The Table I shows the experimental results of over 50 insertions for each method. As expected, we achieve 100% success across all the insertions. In addition, our method requires only  $11.2 \pm 4.5$  seconds to insert successfully, which is about two times faster compared to the traditional search strategy.

	Circular spiral search	Elliptical spiral search
Time(s)	$22.8 \pm 5.7$	$11.2 \pm 4.5$

Table I

WE COMPARE OUR ELLIPTICAL SPIRAL SEARCH WITH THE TRADITIONAL “CIRCULAR” SPIRAL SEARCH STRATEGY.

### B. Double pin insertion

In previous experiment, we have studied *single* pin insertion task in which only the uncertainties in the translation parts of the object poses were employed to derive the new search strategy. We now study another example where both uncertainties of rotation and translation parts are used in planning the insertion strategy. In particular, we keep the experiment setting similar to last task except that both objects now have two pins/holes (see Figure 4).

In this experiment, we assume  $\text{\scriptsize{\textit{Obj2}}}$  is fully constrained and its pose is determined before the execution of the compliant double pin insertion. Note that this can be done by simply

<sup>3</sup>Such assumption is suitable for our experiment where object geometry is simple. For better capture of the multi-modal of the distribution, the Gaussian mixture model is recommended.

bringing the object to contact with a known surface in order to further reduce the uncertainty after the grasping action.

To perform the double pin insertion, we decouple the task into two force-controlled explorations. First, the robot arm moves  $\text{\scriptsize{\textit{Obj2}}}$  to the pre-inserting position, slightly tilts  $\text{\scriptsize{\textit{Obj2}}}$ , then starts moving down until contact is detected. We then perform the elliptical spiral search to find the precise position of the first pin/hole using the information of the uncertainty of the translation part. Second, we find the next pin/hole by simply rotating  $\text{\scriptsize{\textit{Obj2}}}$  around the first pin/hole axis while maintaining contact with the surface. In this case, the information of the uncertainty of the rotation part are employed to determine the bound angles of the exploration. As expected, we also achieve 100% success across 25 attempts. The average running time of the force-controlled exploration is  $28.3 \pm 7.1$  seconds.

The video of the two experiments can be found at <https://youtu.be/rbo4QL40fqU>.

### V. CONCLUSION

In this paper, we have presented a probabilistic framework to precisely keep track of the uncertainties throughout the entire manipulation process. In order to do so, we decompose the manipulation task into two subsequent stages, namely perception and physical interactions. Each stage is associated with different sources and types of uncertainties, requiring different techniques. We discuss which representation of uncertainties is the most appropriate for each stage (e.g. as probability distributions in  $SE(3)$  during perception, as weighted particles during physical interactions), how to convert from one representation to another, and how to initialize or update the uncertainties at each step of the process (camera calibration, image processing, pushing, grasping, etc.).

We have shown that precisely keeping track of the uncertainties in the system can significantly improve the speed and success rate of the manipulation. For example, as shown in our experiment on the single cylindrical peg-in-hole task, the spiral search operation was accelerated by two times. We also demonstrated how the information of uncertainties in the system can be applied into more complex task, i.e. double pin insertion. In fact, without such information, it would be extremely challenging to perform the mentioned tasks successfully.

In addition to the mentioned complex insertion tasks, we believe these techniques could find application in various manipulation tasks that require the knowledge of the uncertainties of the object poses. One application is to estimate the success probability of a particular action. This estimation plays an important role as it enables the robot to improve the overall success rate of the task by not performing unnecessary motions. Another application is to plan the action that can actively reduce the uncertainties in the assembly process and simultaneously decrease the number of actions required. Moreover, future work can also inherit the proposed techniques to employ in bimanual manipulation tasks with higher complexity.

It is also worth-noting that the benefits of the precise information about the uncertainties inevitably comes with an

expense. The reason is that it takes more time for the robot to reason and obtain such information from the measurements after each physical interaction. In fact, both planar grasping action and touch-based localization need to be taken into account when dealing with this challenge. Regarding the planar grasping action, [33] have proposed a solution in which the forward motion model is approximated by kernel conditional density estimation (KCDE). Nevertheless, in order to properly address this problem, further investigation to improve the overall speed of the estimation process is required in future works.

#### ACKNOWLEDGMENT

This work was supported in part by NTUitive Gap Fund NGF-2016-01-028 and SMART Innovation Grant NG000074-ENG.

#### REFERENCES

- [1] S. Thrun, W. Burgard, and D. Fox, *Probabilistic robotics*. MIT press, 2005.
- [2] F. Suárez-Ruiz, X. Zhou, and Q.-C. Pham, “Can robots assemble an ikea chair?” *Science Robotics*, vol. 3, no. 17, p. eaat6385, 2018.
- [3] R. Smith, M. Self, and P. Cheeseman, “Estimating uncertain spatial relationships in robotics,” in *Autonomous robot vehicles*. Springer, 1990, pp. 167–193.
- [4] H. F. Durrant-Whyte, “Uncertainty geometry in robotics,” *IEEE Journal of Robotics and Automation*, vol. 4, no. 1, pp. 23–31, 1988.
- [5] Y. Wang and G. S. Chirikjian, “Nonparametric second-order theory of error propagation on motion groups,” *International Journal of Robotics Research*, vol. 27, no. 11-12, pp. 1258–1273, 2008.
- [6] T. D. Barfoot and P. T. Furgale, “Associating uncertainty with three-dimensional poses for use in estimation problems,” *IEEE Transactions on Robotics*, vol. 30, no. 3, pp. 679 – 693, June 2014.
- [7] H. Durrant-Whyte and T. Bailey, “Simultaneous localization and mapping: part i,” *Robotics & Automation Magazine, IEEE*, vol. 13, no. 2, pp. 99–110, 2006.
- [8] M. Sallinen, “Modelling and estimation of spatial relationships in sensor-based robot workcells,” Ph.D. dissertation, University of Oulu (Oulu, Finland), 2003.
- [9] S. Su and C. Lee, “Uncertainty manipulation and propagation and verification of applicability of actions in assembly tasks,” in *IEEE International Conference on Robotics and Automation*, vol. 3, 1991, pp. 2471–2476.
- [10] S.-F. Su and C. G. Lee, “Manipulation and propagation of uncertainty and verification of applicability of actions in assembly tasks,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 22, no. 6, pp. 1376–1389, 1992.
- [11] R. H. Taylor, “The synthesis of manipulator control programs from task-level specifications.” Ph.D. dissertation, Stanford University, Stanford, CA, USA, 1976, aAI7707174.
- [12] R. A. Brooks, “Symbolic error analysis and robot planning,” *The International Journal of Robotics Research*, vol. 1, no. 4, pp. 29–78, 1982.
- [13] S.-F. Su and C. G. Lee, “Manipulation and propagation of uncertainty and verification of applicability of actions in assembly tasks,” *Systems Science and Cybernetics, IEEE Transactions on*, vol. 22, no. 6, pp. 1376–1389, 1992.
- [14] R. A. Brooks, “Visual map making for a mobile robot,” in *Robotics and Automation, 1985. Proceedings. IEEE International Conference on*, vol. 2. IEEE, 1985, pp. 824–829.
- [15] R. C. Smith and P. Cheeseman, “On the representation and estimation of spatial uncertainty,” *IEEE Journal of Robotics and Automation*, vol. 5, no. 4, pp. 56–68, 1986.
- [16] G. Chirikjian, *Stochastic Models, Information Theory, and Lie Groups, Volume 1: Classical Results and Geometric Methods*. Springer Science & Business Media, 2009, vol. 1.
- [17] ———, *Stochastic Models, Information Theory, and Lie Groups, Volume 2: Analytic Methods and Modern Applications*. Springer Science & Business Media, 2011, vol. 2.
- [18] R. A. Knepper, T. Layton, J. Romanishin, and D. Rus, “Ikeabot: An autonomous multi-robot coordinated furniture assembly system,” in *Robotics and Automation (ICRA), 2013 IEEE International Conference on*. IEEE, 2013, pp. 855–862.
- [19] A. Wahrburg, S. Zeiss, B. Matthias, and H. Ding, “Contact force estimation for robotic assembly using motor torques,” in *Automation Science and Engineering (CASE), 2014 IEEE International Conference on*. IEEE, 2014, pp. 1252–1257.
- [20] K. Van Wyk, M. Culleton, J. Falco, and K. Kelly, “Comparative peg-in-hole testing of a force-based manipulation controlled robotic hand,” *IEEE Transactions on Robotics*, vol. 34, no. 2, pp. 542–549, 2018.
- [21] C. Phillips-Grafflin and D. Berenson, “Planning and resilient execution of policies for manipulation in contact with actuation uncertainty,” *arXiv preprint arXiv:1703.10261*, 2017.
- [22] A. Sieverling, C. Eppner, F. Wolff, and O. Brock, “Interleaving motion in contact and in free space for planning under uncertainty,” in *Intelligent Robots and Systems (IROS), 2017 IEEE/RSJ International Conference on*. IEEE, 2017, pp. 4011–4073.
- [23] F. Wirnshofer, P. S. Schmitt, W. Feiten, G. v. Wichert, and W. Burgard, “Robust, compliant assembly via optimal belief space planning,” in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 1–5.
- [24] R. Platt Jr, R. Tedrake, L. Kaelbling, and T. Lozano-Perez, “Belief space planning assuming maximum likelihood observations,” in *Proceedings of the Robotics: Science and Systems Conference, 6th*, 2010.
- [25] N. A. Melchior and R. Simmons, “Particle rrt for path planning with uncertainty,” in *Robotics and Automation, 2007 IEEE International Conference on*. IEEE, 2007, pp. 1617–1624.
- [26] K. Hauser, “Randomized belief-space replanning in partially-observable continuous spaces,” in *Algorithmic Foundations of Robotics IX*. Springer, 2010, pp. 193–209.
- [27] F. C. Park and B. Ravani, “Smooth invariant interpolation of rotations,” *ACM Transactions on Graphics (TOG)*, vol. 16, no. 3, pp. 277–295, 1997.
- [28] H. Nguyen and Q. Pham, “On the covariance of  $x$  in  $ax=xb$ ,” *IEEE Transactions on Robotics*, vol. 34, no. 6, pp. 1651–1658, 2018.
- [29] M. T. Mason, “Mechanics and planning of manipulator pushing operations,” *The International Journal of Robotics Research*, vol. 5, no. 3, pp. 53–71, 1986.
- [30] S. Goyal, A. Ruina, and J. Papadopoulos, “Planar sliding with dry friction part 1. limit surface and moment function,” *Wear (Amsterdam, Netherlands)*, vol. 143, no. 2, pp. 307–330, 1991.
- [31] A. Petrovskaya and O. Khatib, “Global localization of objects via touch,” *IEEE Transactions on Robotics*, vol. 27, no. 3, pp. 569–585, 2011.
- [32] H. Nguyen and Q.-C. Pham, “Touch-based object localization in cluttered environments,” *arXiv preprint arXiv:1709.09317*, 2017.
- [33] J. Zhou, R. Paolini, A. M. Johnson, J. A. Bagnell, and M. T. Mason, “A probabilistic planning framework for planar grasping under uncertainty,” *IEEE Robotics and Automation Letters*, vol. 2, no. 4, pp. 2111–2118, 2017.