2013 IEEE/RSJ International Conference on
Intelligent Robots and Systems (IROS)
November 3-7, 2013. Tokyo, Japan

Sequential Trajectory Re-planning with Tactile Information Gain for Dexterous Grasping under Object-pose Uncertainty

Claudio Zito*, Marek S. Kopicki*, Rustam Stolkin*, Christoph Borst**, Florian Schmidt**,
Maximo A. Roa** and Jeremy L. Wyatt*

Abstract— **Dexterous grasping of objects with uncertain pose is a hard unsolved problem in robotics. This paper solves this problem using information gain re-planning. First we show how tactile information, acquired during a failed attempt to grasp an object can be used to refine the estimate of that object's pose. Second, we show how this information can be used to replan new reach to grasp trajectories for successive grasp attempts. Finally we show how reach-to-grasp trajectories can be modified, so that they maximise the expected tactile information gain, while simultaneously delivering the hand to the grasp configuration that is most likely to succeed. Our main novel outcome is thus to enable tactile information gain planning for Dexterous, high degree of freedom (DoFs) manipulators. We achieve this using a combination of information gain planning, hierarchical probabilistic roadmap planning, and belief updating from tactile sensors for objects with non-Gaussian pose uncertainty in 6 dimensions. The method is demonstrated in trials with simulated robots. Sequential re-planning is shown to achieve a greater success rate than single grasp attempts, and trajectories that maximise information gain require fewer re-planning iterations than conventional planning methods before a grasp is achieved.**

I. INTRODUCTION

In robot grasping, there is typically uncertainty associated with the location of the object to be grasped. However, if the object is not in its expected location, then a robot equipped with tactile sensors, or torque sensors at finger joints, may gain information to help refine localisation knowledge from tactile contacts (or lack of such contacts) during the execution of a reach to grasp trajectory. In this paper we i) describe how to iteratively update localisation knowledge using tactile observations from a previous grasp attempt, ii) show how successive grasp trajectories can be planned with respect to these iteratively refined object poses, and iii) show how each reach-to-grasp trajectory can be deliberately planned to maximise new tactile information gain, while also reaching towards the expected grasp location derived from previous information. This paper is an extension of our early work [18].

The main contributions of this work are i) using a hierarchical sample-based path planner, here a Probabilistic Roadmap (PRM) planner, which encodes expected information gain (in a low-dimensional belief space) in its trajectory segments to extend the work in [13], ii) refining the expected object pose by using an observation model for contact sensing by a multi-finger hand that palpates a 3D object to be grasped, and iii) showing that this approach enables planning for robot manipulators with 21 DoFs and non-Gaussian object pose uncertainty in 6 dimensions.

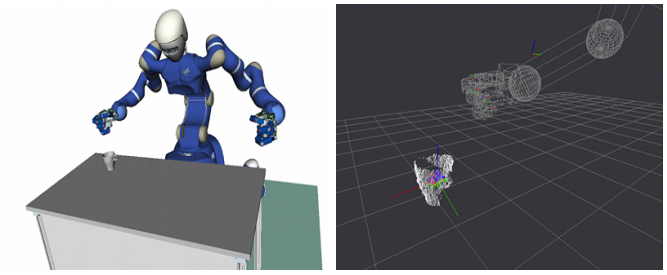

Fig. 1. Justin and a known mug to grasp in OpenRave simulator (left). Partial point cloud of the mug used for localisation (right).

We make several assumptions. First we assume that the object is of known shape, in the sense that a high density point cloud model or mesh model is available. Second we assume that a pre-computed grasp (i.e. a set of finger contacts) and its pre-grasp configuration are known a priori for this object. Third we assume the availability of an unreliable estimate of the object's pose.

In our scenario we employ a depth image obtained from an Asus Xtion Pro depth camera. This gives an incomplete point cloud of the object surface. Using a model fitting procedure similar to the sampling from surflet pairs method presented in [4], a probability density (or belief state) over the object pose is estimated, represented as a particle set. Given this distribution, a reach-to-grasp trajectory is planned. This trajectory has as its goal configuration the pre-computed grasp under the assumption that the object is at a pose corresponding to the mean position of the particle set. The path to this goal configuration is found using a stochastic motion planner. This planner works with a cost function that allows deviations from the shortest path that maximise the chance of gathering tactile observations that will reduce pose uncertainty in the object location. If unexpected observations occured during the execution of the planned trajectory, then such observations (both tactile contact and no-contact signals) collected at poses along the reach-to-grasp trajectory are used to perform a particle filter update. Replanning then occurs with the new pose distribution, and a new reach-to-grasp trajectory can be constructed. This process is then iterated until a successful grasp is achieved.

This work was supported by the EC FP7 ICT programme through the FP7 projects GeRT (248273) and PaCMan (600918).

*IRLab, School of Computer Science, University of Birmingham, United Kingdom, {cxz004,msk,stolkinr,jlw}@cs.bham.ac.uk

**German Aerospace Center, DLR, Wessling, Germany, firstname.lastname@dlr.de

The benefit of planning with beliefs is the ability to reason about the informational effects of sequences of actions. In typical belief space planning this means performing a kind of pre-posterior analysis, in which planned actions (here trajectories) cause imagined observations that are in turn used to perform a Bayes' update of the belief state. As the belief space grows exponentially in the length of the planned action-observation sequence these methods are exact but inefficient.

Our work instead builds on the approach of Platt et al. [13], [14]. That work plans a sequence of actions that will generate observations that distinguish a hypothesised state from competing hypotheses while also reaching a goal position. The informational value of a trajectory is the difference in the expected observations between the hypothesised position and each alternative. This approach allows us to track high-dimensional belief states using an accurate filter defined by the user, but reduces the complexity of planning in belief spaces by approximating the informational value of actions from a low-dimensional subspace of the belief state. Platt et al. applied this to planning for a two degree of freedom manipulator using a laser range finder for observations, and employed an optimisation framework for planning. The algorithm is proved to localise the true state of the system in 1 dimension and to reach a goal region with high probability. In contrast to [13], our approach encodes information gathering actions to localise an object to be grasped in 6 dimensions while simultaneously attempts to achieve the task. Similarly to [13], [14], our method is guaranteed to converge to the true state of the system in which a reach-to-grasp trajectory suceeds with high probability. Here we also show how this approach can be extended to planning the motion of a manipulator performing multi-finger grasping.

We tested our approach in simulation. Since physics-based simulators are hard to calibrate, we use i) a Nvidia Physx-based simulator, [11], and ii) OpenRave-based, [16], Justin Simulator developed at DLR for control and visualisation of the entire robot. Both simulators interface with a Bullet-based simulator, [1], developed at DLR, which is accurately calibrated to simulate hand-object physical interactions. We use the Bullet simulator to simulate contacts and compute grasp quality measure. For proof of concept, we also tested our approach on the Rollin Justin robot. However, in order to guarantee no movement after a contact, the object to be grasped needed to be glued on the desk. Therefore even though our approach converged to a force closure grasp, Rollin Justin could not lift up the object to prove the real stability of the grasp. We now briefly survey other work on information gathering while grasping.

II. Related work

Other authors have attempted to pose the simultaneous grasping and localisation problem (SLAG) in terms of a partially observable Markov decision process (POMDP). Hsiao et al., [7], present a simple blocks world problem, wherein a single finger can execute a small selection of actions (left, right, up, down) in response to a small set of possible sensory detections (contact or no contact below, right or left sides of finger). This simple problem is tractable for solving as a POMDP, producing optimal policies for guiding the finger towards a desired location on a 2D blocks world object. The same authors later addressed real-world grasping with an arm and three-finger Barrett hand equipped with tactile fingertip sensors, [5], [6]. Because the space of possible actions for such a robot is enormous (especially compared to the single finger and 2D blocks-world problem of [7]), in order to solve the problem as a POMDP, the authors restrict the robot's choice of actions to executing a small number of pre-programmed reach-to-grasp motions, described relative to the pose of the target object. Thus the POMDP method can tractably be used to select between this small number of actions, in response to tactile contacts detected by the three fingertip sensors during the previous action. Thus various actions are sequentially selected, with successive refinements of the object pose collected from sensing during each action, until eventually one of the actions achieves a successful grasp.

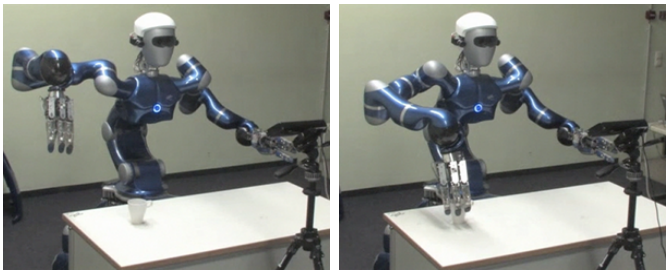

Fig. 2. Rollin Justin at DLR. Justin in a starting configuration, the mug to be grasped (glued on the desk) and the depth camera (left). Justin executing a successful reach-to-grasp trajectory which leads to grasp the mug (right). Since the mug is glued on the desk, Justin cannot lift it up to prove the real stability of the grasp.

In contrast to [5], our method does not select between pre-defined high level actions , but instead we describe how to optimally plan a novel dexterous reach-to-grasp trajectory, in response to recent sensory observations, which maximises the expected information gain from further tactile observations, if the next reach-to-grasp attempt should discover that the target object is not at its expected location.

Petrovskaya [12] also investigated the use of tactile information, derived from collisions of an end effector with an object, to refine localisation estimates for that object. Similar to our method, Petrovskaya represents the location of an object to be grasped, as a collection of particles forming a distribution, and this distribution is refined by successive manipulator contacts with the object. However, the work is limited in that Petrovskaya does not address the problem of planning manipulator trajectories to achieve these contacts. Instead, the author simply begins with an assumption that a selection of collisions will occur, and then shows how to use such collisions to update an object location distribution.

Prentice et al. [15] have shown that motion planning for a mobile robot in belief space can be done efficiently for linear Gaussian systems by using a factored form of the

covariance matrix. The authors incorporate an estimation of localization uncertainty in the cost function of a belief-space variant of the PRM. Similarly to [15] our approach is capable of balancing shorter paths against those which reduce belief uncertainty through sensory information gain. Instead our work plans trajectories that are able to distinguish hypotheses by using the expected sequence of observations, which are computed according to the current belief state. Furthermore our approach represents belief space as non-Gaussian and allows arbitrary implementations of Bayes filters to track belief states.

III. PLANNING TRAJECTORIES

A. Problem formulation

This paper is concerned with the problem of planning control actions to reach a goal state in the presence of incomplete or noisy observations. Let us consider a discrete-time system with continuous non-linear deterministic dynamics,

$$x_{t+1} = f(x_t, u_t)$$

where $x_t \in \mathbb{R}^n$ is a configuration state of the robot and $u_t \in \mathbb{R}^l$ is a action vector, both parametrised over time $t \in \{1, 2, ...\}$. Let $p \in SE(3)$ describe the object pose, given an initial prior belief state b_1 we define a set of k hypotheses as $\{p^i\}_{i=1}^k$, where $p^1 = \arg\max b_1$ and $p^i \sim b_1, i \in [2, k]$. We search for a sequence of actions, $u_{1:T-1} = \{u_1, \ldots, u_{T-1}\}$, that distinguish between observations that would occur if the object were in p^1 from any other p^i pose, with $i \in [2, k]$. At each time step, t, the system makes an observation, $y \in \mathbb{R}^m$, that is a non-linear stochastic function of states and hypotheses. Without losing generality, we define y_t to be a column vector of binary values. Each of these values represents whether or not a contact is observed between a given link of the robot and the hypothesis p^i. However, binary values have been shown to be not very informative during the planning phase. Therefore we define,

$$h(x, p^i) = p(y = 1|x, p^i)$$

as a column vector of scores identifying the likehood of observing a contact, $y = 1$, as function of states and hypotheses. More generally, let $F_t(x, u_{1:t-1})$ be the robot configuration at time t if the system begins at state x and takes action $u_{1:t-1}$. Therefore the expected sequence of observations over a trajectory, $u_{1:t-1}$, is:

$$h_t(x, u_{1:t-1}, p^i) = (h(F_2(x, u_1), p^i)^T, h(F_3(x, u_{1:2}), p^i)^T, \ldots, h(F_t(x, u_{1:t-1}), p^i)^T)^T$$

a column vector which describes the likelihood of observing a contact at any time step of the trajectory $u_{1:t-1}$. We then search for a sequence of actions which maximise the difference between observations that are expected to happen in the sampled states, $p^{2:k}$, when the system is actually in the most likely hypothesis, p^1. In other words, we want to find a sequence of action, $u_{1:T-1}$, that minimises

$$J(x, u_{1:T-1}, p^{1:k}) = \sum_{i=2}^{k} N(h(x, u_{1:T-1}, p^i)|h(x, u_{1:T-1}, p^1), \mathbb{Q}) \quad (1)$$

where $N(\cdot|\mu, \Sigma)$ denotes the Gaussian distribution with mean μ and covariance Σ and $\mathbb{Q}$ is the block diagonal of measurement noise covariance matrices of appropriate size. Rather than optimising (1) we follow the suggested simplifications described in [13] by dropping the normalisation factor in the Gaussian and optimising the exponential factor only. Let us define for any $i \in [2, k]$

$$\Phi(x, u_{1:T-1}, p^i) = ||h_t(x, u_{1:T-1}, p^i) - h_t(x, u_{1:T-1}, p^1)||^2_{\mathbb{Q}},$$

then the modified cost function is

$$J(x, u_{1:T-1}, p^{1:k}) = \frac{1}{k} \sum_{i=2}^{k} e^{-\Phi(x, u_{1:T-1}, p^i)} \quad (2)$$

it is worth to notice that when there is a significant difference between the sequence of expected observations, $h_t(x, u_{1:T-1}, p^i)$ and $h_t(x, u_{1:T-1}, p^1)$, the function $\Phi(\cdot)$ is large and therefore $J(\cdot)$ is small. On the other hand if the sequence of expected observations are very similar to each other, their distance measurement tends to 0 and $J(\cdot)$ tends to 1. Equation (2) can be minimised using different planning techniques such as Randomly-exploring Random Trees (RRTs) [10], Probabilistic Roadmap (PRM) [9], Differential Dynamic Programming (DDP) [8] or Sequential Dynamic Programming (SDP) [2]. In Section IV-B, we define a new set of heuristics that can be encoded in a PRM planner for generating more reliable trajectories that explicitly reason over the pose uncertainty, and demonstrate the method with experiments on a virtual testbed.

B. Belief Update

After a trajectory is planned our algorithm executes it. If an unexpected observation occurs at execution-time the algorithm refines the current belief state using an accurate, high-dimensional filter provided by the user.

In order to define an unexpected observation, it may be convenient for the reader to think of a reach-to-grasp trajectory as composed of two parts: i) approaching trajectory which leads to a pre-grasp configuration of the robot in which the fingers generally cage the object to be grasped without generating any contact, and ii) a grasping trajectory which moves the robot in contact and eventually generates a force closure grasp. In this way any contact which occurs during the approaching trajectory is considered as an unexpected observation. Similarly a non sufficient number of contacts for a force closure at the end of a grasping trajectory is considered as unexpected.

In our implementation we update the belief using the Bayes rule assuming deterministic dynamics. In this case we can write the belief update as

$$b_{t+1} = \frac{P(y_{t+1}|x_t, u_t) b_t}{P(y_{t+1})} \quad (3)$$

C. Replanning

Our algorithm plans trajectories assuming only the maximum likelihood observations given the current belief state. Therefore we need to rely on sensory feedback during the execution of the planned trajectory in order to detect whether or not unexpected observations occur. This triggers a belief update, using the observation gathered at execution-time, and consequently a replanning phase, in which the manipulator is moved back to a safe configuration (e.g. outside the uncertain region) and a new reach-to-grasp trajectory is planned.

In our experiments, the algorithm uses torque sensors at each joint of robot's end-effector to detect wheter or not a link of the hand is in contact with the environment. We assume that the sensing abilities of the robot are fine enough to perceive the object without moving it. However even in simulation is difficult to mantain such an assumption. Though our results show that small changes in the configuration of the object do not affect the algorithm which is still able to converge to a force closure grasp.

D. Terminal conditions

Our algorithm terminates its execution when i) no unexpected observations are detected and, ii) the robot achieves a force closure grasp which satisfies an user-defined (minimum) quality. In simulation, knowing the model of the object, the contact points and the executed forces, it is possible to make a force closure analysis using an associated quality measure [17]. In this work, the magnitude of the minimum perturbation wrench that breaks the force closure is used as the grasp quality measure [3].

A possible limitation of our current implementation is that the unexpected observations are treated as binary input (contact, no contact). In other words, a contact in the approaching trajectory will always trigger replanning, even if the contact occured at the very last step of the trajectory and the robot's end-effector is in a configuration where it is still possible to achieve the target grasp. We aim to investigate such cases in future work.

IV. IMPLEMENTATION

Our approach consists in planning informative tactile observations for a dexterous hand while simultaneously reaching a given target grasp. Our innovations are i) implementing a hierarchical PRM algorithm which allows us to plan dexterous reach-to-grasp trajectories, ii) encoding a new set of heuristics for a randomised motion planner and iii) formalising an observational model for contact sensing by a multi-finger hand. We tested our approach for robotic manipulators up to 21 DOFs under non-Gaussian object pose uncertainty in 6 dimensions.

A. Observational model

We assume that a robotic manipulator is composed of parts. These parts are considered collections (or chains) of joints linked somehow together. Without losing generality, we also assume a single part called 'arm' and a set of m parts called 'fingers'. In addition, we describe the observational model as limited only to a given subset of those parts, i.e. only fingers or a subset of them (e.g. finger tips). Let M be the ordered set of parts which compose the manipulator, then $x(j)$ is the configuration in joint space of the j^{th} part, with $j \in M$. In other words, j is the index of a specific chain. We also define $\bar{M}$ to be the set of indices such that the respective part is used in our observational model. In addition, we use the operator $W(x(j))$ to refer to the workspace coordinates in $SE(3)$ of the j^{th} joint w.r.t. a given reference frame.

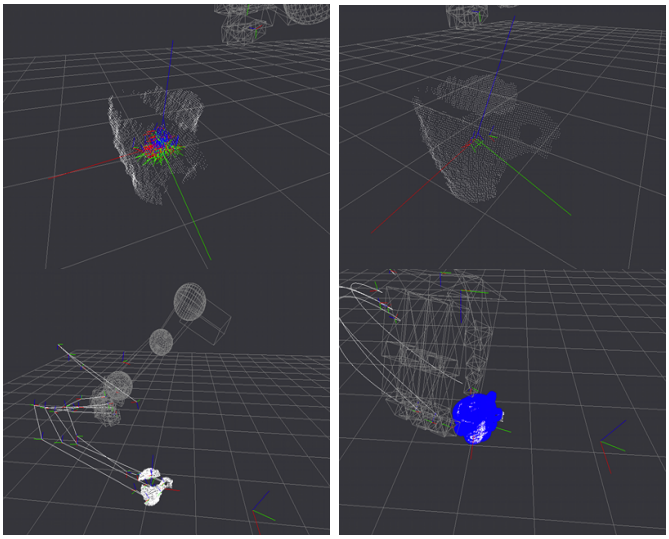

Fig. 3. Top row: The high dimensional belief state used to track object pose (left); the low dimensional filter used for planning (right). Bottom row: The unoptimised PRM plan for fingers and wrist (left); the optimised and smoothed trajectory (right).

Mathematically we formalise the likelihood of observing a contact for each finger of the robot as an exponential distribution over the Euclidian distance, d_{ji}, between the finger tip's pose, $W(x(j))$, and the closest surface of the object assumed to be in pose p^i. Note that, for the aim of this work, we limited our observational model to contacts which may occur on the internal surface of fingers. This affects directly our planner which rewards trajectories that would generate contacts on the finger tips rather than on the back side of the fingers. Therefore for any $j \in \bar{M}$ we write

$$p(y(j) = 1 | x(j), p^i) = \begin{cases} \eta \exp(-\lambda d_{ji}) & \text{if } d_{ji} \leq d_{max} \\ & \text{and } \langle n_{xj}, \hat{n_{p^i}} \rangle < 0 \\ 0 & \text{otherwise} \end{cases}$$

where $\langle n_{xj}, \hat{n_{p^i}} \rangle$ is the inner product of, respectively, j^{th} finger tip's normal and the estimated object surface's normal, and d_{max} describes a maximum range in which the likelihood of reading a contact is not zero, η is a normaliser. That allows us to rewrite the likelihood of reading a contact on the force/torque sensors of the robot, $h(x, p^i), i \in [1, k]$ with $j_1, \ldots, j_m \in \bar{M}$ as follows,

$$h(x, p^i) = [p(y(j_1) = 1 | x(j_1), p^i), \\ \ldots, p(y(j_n) = 1 | x(j_m), p^i)]^T$$

B. Planning a trajectory to maximise information gain

The implementation of our planner uses a modified version of Probabilistic Roadmap (PRM), [9], to plan trajectories and detect collisions. Generally the PRM method is composed by two phases: i) Learning phase, in which a connected graph G of obstacle-free configurations of the robot is generated and, ii) Query phase, in which a path is searched for a given pair of configuration x_{root}, x_{goal}. However the computational cost for the learning phase grows very fast with respect to the dimensionality of the problem. We therefore incrementaly build connections between neighbouring nodes during the query phase. Given a pair $\langle x_{root}, goal \rangle$ which describes the root state in configuration space, $\mathbb{R}^n$, and goal state in workspace, $SE(3)$, of the trajectory, we use A* algorithm to find a minimum cost trajectory in obstacle-free joint space according to:

$$c(x) = c_1(x, x_{root}) + c_2(x, x', \hat{x}_{goal}) \tag{4}$$

where $x, x' \in \mathbb{R}^n$ and $x' \in Neighbour(x)$, $\hat{x}_{goal}$ is a reachable goal configuration for the robot computed by inverse kinematic, $c_1(x, x_{root})$ is the cost-to-reach x from x_{root} and $c_2(x, x', \hat{x}_{goal})$ is a linear combination of the cost-to-go from x to a neighbour node x' and the expected cost-to-go from x' to the target. We implemented $c_1(\cdot)$ as a commulative discounted and rewarded travelled distances. More in details, we define

$$\begin{aligned} c_2(x, x', \hat{x}_{goal}) = \alpha d_{bound}(x, x') + \beta d(x', \hat{x}_{goal}) \\ + \gamma d_{cfg}(x) \end{aligned} \tag{5}$$

where $\alpha, \beta, \gamma \in \mathbb{R}$, $d(\cdot)$ is a distance function in $SE(3)$ which linearly combine rotational and transitional distances in workspace[1]. For $d_{bound}(\cdot)$, let $\mathcal{B}_n(r) = \{x \in \mathbb{R}^n | x^T x \leq r^2\}$ and $\mathcal{B}(r_l, r_a) = \{A = \left[\begin{smallmatrix} R & p \\ 0 & 1 \end{smallmatrix}\right] \in SE(3) | p^T p \leq r_l^2$ and $1 - \langle Q(R), Q(R) \rangle \leq r_a^2\}$[2] denote repectively the r-ball in $\mathbb{R}^n$ and in $SE(3)$, then $b_{bound}(x, x')$ is defined as

$$d_{bound}(x, x') = \begin{cases} \psi(x, x') & \text{if } W(x) - W(x') \in \mathcal{B}(r_l, r_a) \\ & \text{and } x - x' \in \mathcal{B}_n(r) \\ +\infty & \text{otherwise} \end{cases}$$

where $Q(\cdot)$ is the Quaternion operator for $R \in SO(3)$, $\langle q_1, q_2 \rangle$ is the inner product of two quaternions, $r_l, r \in \mathbb{R}, r_a \in (0, 1)$, and $\psi(x, x') = \zeta d(x, x') + (1 - \zeta)||x - x'||_2$ with $\zeta \in (0, 1)$. Finally, $d_{cfg}(\cdot)$ is a function which penalises dangerous configurations of the robot (i.e. close to joint limits).

We redefine the heuristic $c_2(\cdot)$ in order to reward informative tactile explorations while attempting to reach the goal state (described as a target configuration of the manipulator).

$$\begin{aligned} \bar{c}_2(x, x', \hat{x}_{goal}, A, p^{1:k}) = & \alpha J(x, x', p^{1:k}) d_{bound}(x, x') \\ & + \beta d_A(x', \hat{x}_{goal}) + \gamma d_{cfg}(x') \end{aligned} \tag{6}$$

[1]For the sake of simplicity, we abuse of the mathematical notation by writing $d(x, x')$ instead of d(W(x),W(x')).

[2]We simplified the notation $\mathcal{B}_{SE(3)}(\cdot)$ in $\mathcal{B}(\cdot)$ for pratical reasons.

where A is the diagonal covariance matrix of our sampled states, for any column vector $a, \mu \in \mathbb{R}^n$, $d_A(a, \mu) = \sqrt{(a - \mu)^T A^{-1} (a - \mu)}$ is the Mahalanobis distance centered in μ and $J(x, x', p^{1:k}) \in (0, 1]$ is a factor which rewards trajectories with a large difference between expected observations if the object is at the expected location, p^1, versus observations that would be expected if the object is at other poses, $p^{2:k}$, sampled from the distribution of poses associated with the object's positional uncertainty:

$$\bar{J}(x, x', p^{1:k}) = \frac{1}{k - 1} \sum_{i=2}^{k} e^{-\Phi(x, x', p^i)} \tag{7}$$

where:

$$\Phi(x, x', p^i) = ||h_t(x, x', p^i) - h_t(x, x', p^1)||_2$$

for each $i \in [2, k]$ and $h_t(x, x', p^i)$ is sequence of probability of reading a contact travelling from state x to x'. In our implementation $h_t(x, x', p^i) = h(x', p^i)$. In other words, we evaluate the likelihood of making a contact while moving from state x to x' as the likelihood of making a contact only in the next state x'. Note that our current observational model is designed to conserve (6) as in (5) when the likelihood of observing a tactile contact is zero. In fact, for robot configurations in which the distance to the sampled poses is larger than a threshold, d_{max}, the cost function $J(\cdot)$ is equal to 1. However we also encode uncertainty in the second factor of our heuristics, $d_{\mathcal{A}}(\cdot)$, which evaluates the expected distance to the goal configuration. In this way the planner also copes with pose uncertainty at the early stages of the trajectory, when the robot is still too far away from the object to observe any contacts.

C. Planning for Dexterous manipulator

In order to compute a dexterous trajectory which allows us to plan movement for both arm and fingers we need to break down the 'curse of dimensionality' or, equivalently, increase the number of sampled configuration to properly cover the configuration space.

Our proposed solution is to build a hierarchical planner. First we generate a PRM only in the arm configuration space in order to find a global path between the $x_{root}, \hat{x}_{goal}$. It is worth noticing that in this phase the rest of the joints of the manipulator are interpolated in order to have a smooth passage from x_{root} to $\hat{x}_{goal}$. We then refine the planned trajectory generating a new PRM in the entire configuration space of the manipulator (e.g. arm + hand joint space) along the global path. In other words, we limited the new PRM to explore only the subspace nearby the configurations which compose the global path. Subsequently an optimisation proceedure is executed along the trajectory to generate a smoother transition from one configuration to the next.

This approach enable us to plan dexterous reach-to-grasp trajectories up to 21 DoFs with only 1,000 sampled configurations. Note that this is the same order of magnitude that we use in practise for planning trajectories of 6 DoFs robot manipulators.

D. Belief update

Once a trajectory is executed and a real (unexpected) observation y is detected, we update our belief state according to the Bayes' rule. We represent our belief state as a set of N particles $b_t = \{b_t^z\}_{z=1}^N$. In a particle filter fashion we update the weight of each particle b_t^z, $z \in \{1, \ldots, N\}$, as follows

$$p(y|x, b_t^z) = \prod_{j \in \hat{M}} p(y_t(j)|x_t(j), b_t^z)$$

and then we execute a resampling which generates a posterior distribution b_{t+1} as new set of particles $\{b_{t+1}^z\}_{z=1}^N$.

In simulation we assume that there are no false dectections. However it is possible to distinguish whether or not a contact occurs between the object to be grasped and the robot's end-effector. For example, in case a contact with the environment is detected we skip the belief update step and we move the robot back to a safe configuration before triggering the re-planning.

V. EXPERIMENTS

In this work we aim to show that sequential re-planning is capable of achieving higher successful grasp rates than single grasp attempts in presence of non-Gaussian object-pose uncertainty in 6 dimensions. We also show that planning trajectories that maximise information gain requires fewer re-planning iterations to achieve a grasp with the same order of magnitude of grasp quality.

We run 12 trials in a virtual environment. Each trial has a different initial probability density over the object pose and we tested the ability of different strategies to achieve a grasp configuration in which it is possible to obtain force closure grasp despite the pose uncertainty. In these experiments, after an unexpected observation occurs, the robot is always moved back to the initial configuration shown in Fig. 2 (left image).

Table I summarises the data collected in our experiments. First we computed the error in the initial estimation of the object pose with respect to the ground truth, which is available in the simulation environment. We decomposed this error into translational and rotational components. The translational component measures displacement as Euclidian distance in a 3 dimensional space between the estimated location and the ground truth. The rotational or angular displacement is evaluated using the quaternion representation.

We then performed three different algorithms: i) an open-loop trajectory towards the expect object pose without re-planning, ii) a sequential re-planning algorithm without information gathering (ReGrasp) and iii) a sequential re-planning algorithm with information gathering (ReGrasp+IG). Column 3 in table I shows the rate of successful attempts for the open-loop trajectory. Columns 4 and 5 show that successive re-planning enables us to achieve a grasp in cases when the open-loop strategy fails. We present both the number of (re-)planning iterations and the grasp quality value once a force closure grasp is achieved. Finally, columns 6 and 7 show that planning trajectories which maximise information gain reduce the number of re-planning iterations required to achieve a successful grasp while producing the same order of magnitude of grasp quality. For all the trials, the nominal quality for the goal grasp configuration with the object in the nominal pose is 0.006920.

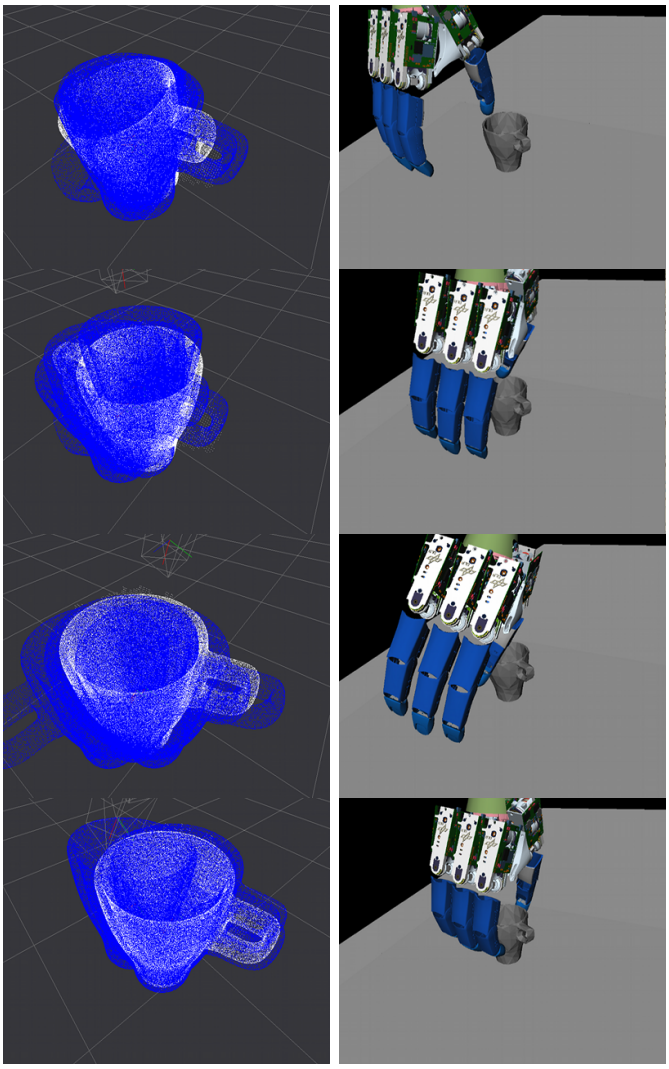

Fig. 4. All belief states are the low dimensional belief states sub-sampled from the corresponding high dimensional belief states. Top row: Initial belief state (left), first contact (right). Second row: updated belief state from first contact (left), second contact (right). Third row: updated belief after second contact (left), third contact (right). Bottom row: updated belief after third contact (left), executed reach to grasp pose (right).

An interesting case is shown in the 8th row of Table I. In this trial a single attempt fails while ReGrasp requires 7 iterations to generate a quite poor force closure. Instead using ReGrasp+IG produces a successful trajectory at the very first attempt, although the grasp quality does not improve in this specific case.

To illustrate the behaviour of our re-planning system we show a typical sequence generated by one simulation trial. We assume a known point cloud model of the object shape, and uncertainty in the object pose. In this trial the robot observes the object as a point cloud and applies a model fitting process described in [4]. The model fitting process is stochastic and so the resulting pose of the object is uncertain.

TABLE I

EXPERIMENTAL RESULTS IN SIMULATION

Initial error		Single attempt	ReGrasp		ReGrasp+IG	
Linear (m)	Angular (quaternion)		Iterations	FCA	Iterations	FCA
0.055909	0.006344	success	1	0.006301	2	0.005876
0.05343	0.012268	success	1	0.006072	2	0.00649
0.057804	0.013443	success	1	0.005996	2	0.005308
0.060809	0.016942	failure	2	0.000452	1	0.000911
0.058412	0.017696	success	1	0.008286	1	0.006266
0.05996	0.019115	failure	4	0.005679	1	0.008111
0.05755	0.019915	success	1	0.005936	1	0.003597
0.05815	0.020103	failure	7	0.003868	1	0.003737
0.059598	0.021463	failure	3	0.007579	1	0.006105
0.061404	0.023431	success	1	0.007397	1	0.006409
0.063758	0.025339	success	1	0.007959	2	0.002738
0.059935	0.054883	failure	2	0.005933	2	0.00752

We sample multiple poses by repeating this process, and obtain a belief state consisting of the resulting set of possible poses of the object. A trajectory is then planned to achieve a given grasp on the object. At each step a trajectory for the wrist and fingers are generated that will move to the desired grasp, while deviating from a minimum length trajectory to maximise information gathered through tactile observations. The belief state is updated and re-planning occurs each time a tactile contact is made. In this example three separate contacts are made during reach-to-grasp trajectories. In each instance the trajectory is re-planned given the new belief state. After three contacts the fourth trajectory achieves a configuration suitable for grasping.

A drawback of ReGrasp+IG is the computational time. In our experiments ReGrasp requires in average ~ 7 seconds to plan a trajectory against ~ 200 seconds of ReGrasp+IG. ReGrasp+IG adds extra computation during the query phase of the PRM algorithm in order to compute the likehood of reading a contact, which requires finding the closest surface to the finger tips for each hypothesis everytime a node in the PRM is in a neighborhood of the uncertain region. We aim to reduce the computational time required by parallelising the query phase of our PRM algorithm.

VI. CONCLUSION AND FUTURE WORK

A. Conclusion

In this paper we have shown how to solve the problem of dexterous grasping of objects with uncertain pose by using information gain re-planning.

We have proposed a method for tactile information gain planning for dexterous, high DoFs manipulators by using i) a hierarchical PRM planner which encodes informational measure in its segments, and ii) a method for sequentially refining pose uncertainty, by using tactile observations gathered during unsuccessful grasp attempts. We have shown that this approach enables planning for robot manipulators with 21 DoFs and non-Gaussian object pose uncertainty in 6 dimensions.

We have also shown how sequential re-planning can achieve better quality grasps than single attempts to directly grasp an object at its expected pose, and that re-planning with trajectories designed to maximise tactile information gain, achieves successful grasps with fewer iterations than sequential attempts to grasp directly towards the object's (sequentially updated) expected pose.

This work extends that of [13], which offers a way to avoid the complexity of planning in a high dimensional belief space. It does this in two ways, i) by approximating the informational value of actions from a low-dimensional subspace of the belief state; and ii) by embedding that informational value into the physical space. This enables standard motion planning techniques to trade off directly between information gain and achievement of the goal pose for the manipulator.

B. Future Work

We aim to extend our observational model in order to cope with non-static objects by using a physics simulator as a forward model which enable us to predict how the object behaves under a hand-object interaction. In addition, we are interested in extending sensing abilities of the robot by using tactile sensors on the robotic hand. We also aim to investigate our re-planning approach with different withdrawing policies, instead of moving the robot back to its starting configuration at each iteration, e.g. adjusting fingers and wrist configurations to gain even more contact information.

REFERENCES

[1] Bullet physics simultator, http://bulletphysics.org/wordpress/.

[2] J. Betts. *Pratical methods for optimal control using nonlinear programming*. Siam, 2001.

[3] C. Ferrari and J. Canny. Planning optimal grasps. In *IEEE Proc. Int. Conf. on Robotics and Automation*, page 22902295, 1992.

[4] U. Hillenbrand and A. Fuchs. An experimental study of four variants of pose clustering from dense range data. In *Computer Vision and Image Understanding*, 2011.

[5] K. Hsiao and L. Kaelbling. Task-driven tactile exploration. In *Proc. of Robotics: Science and Systems (RSS)*, 2010.

[6] K. Hsiao, L. P. Kaelbling, and T. Lozano-Perez. Robust grasping under object pose uncertainty. In *Autonomous Robots*, number 31 in (2-3), pages 253–268, 2011.

[7] K. Hsiao, L.P. Kaelbling, and T. Lozano-Perez. Grasping pomdps. In *IEEE Proc. Int. Conf. on Robotic and Automation (ICRA)*, 2007.

[8] D. Jacobson and D. Mayne. *Differential dynamic programming*. Elsevier, 1970.

[9] L.E. Kavraki and P. Svestka. Probabilistic roadmaps for path planning in high-dimensional configuration spaces. In *IEEE Transactions on Robotics and Automation*, 1996.

[10] S. M. LaValle. Rapidly-exploring random trees: A new tool for path planning. Technical report, CS Dept, Iowa State University, 1998.
[11] NVIDIA PhysX. Physics simulation for developers. http://developer.nvidia.com/object/physx.html, 2009.
[12] A. Petrovskaya and Khati. Global localization of objects via touch. *IEEE Transactions on Robotics*, 27(3):569–585, 2011.
[13] R. Platt, L. Kaelbling, T. Lozano-Perez, and R. Tedrake. Simultaneous localization and grasping as a belief space control problem. Technical report, CSAIL, MIT, 2011.
[14] R. Platt, L. Kaelbling, T. Lozano-Perez, and R. Tedrake. Non-gaussian belief space planning: Correctness and complexity. In *IEEE Proc. Int. Conf. on Robotic and Automation (ICRA)*, 2012.
[15] S. Prentice and N. Roy. The belief roadmap: Efficient planning in linear pomdps by factoring the covariance. In *12th International Symposium of Robotics Research*, 2008.
[16] D. Rosen and J. Kuffner. Openrave: A planning architecture for autonomous robotics. Technical Report CMU-RI-TR-08-34, Robotics Institute, Carnegie Mellon University, 2008.
[17] R. Suarez, M. A. Roa, and J. Cornella. Grasp quality measures. Technical Report IOC-DT-P-2006-10, Technical University of Catalunya (UPC), 2006.
[18] C. Zito, R. Stolkin, M. S. Kopicki, M. Di Luca, and J. L. Wyatt. Exploratory reach-to-grasp trajectories for uncertain object poses. In *Proc. Workshop on Beyond Robot Grasping: Modern Approaches for Dynamic Manipulation. Intelligent Robots and Systems (IROS)*, 2012.