Active Multi-Contact Continuous Tactile Exploration with Gaussian Process Differential Entropy

Danny Driess Daniel Hennes Marc Toussaint

***Abstract*— In the present work, we propose an active tactile exploration framework to obtain a surface model of an unknown object utilizing multiple contacts simultaneously. To incorporate these multiple contacts, the exploration strategy is based on the differential entropy of the underlying Gaussian process implicit surface model, which formalizes the exploration with multiple contacts within an information theoretic context and additionally allows for nonmyopic multi-step planning. In contrast to many previous approaches, the robot continuously slides along the surface with its end-effectors to gather the tactile stimuli, instead of touching it at discrete locations. This is realized by closely integrating the surface model into the compliant controller framework. Furthermore, we extend our recently proposed sliding based tactile exploration approach to handle non-convex objects. In the experiments, it is shown that multiple contacts simultaneously leads to a more efficient exploration of complex, non-convex objects, not only in terms of time, but also with respect to the total moved distance of all end-effectors. Finally, we demonstrate our methodology with a real PR2 robot that explores an object with both of its arms.**

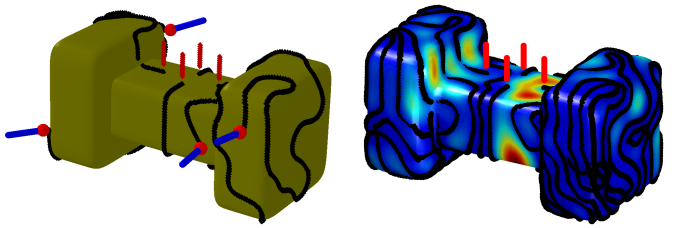
I. Introduction

Tactile perception plays a central role in generating robustness and dexterity when interacting with the environment [1]. By the sense of touch, manifold information like the texture, compliance [2], weight, geometric shape of objects etc. can be perceived, which allows humans and robots to recognize objects rapidly [3] and increases manipulation capabilities [4], [5]. Realizing tactile perception in robotics is challenging both from a sensory and control point of view [6]. In addition, since tactile perception only provides local information at contact locations, an active exploration strategy is required to gather the sensor stimuli.

In particular, knowledge about the shape of an object is important for manipulations such as grasping. The problem of active tactile exploration to obtain a shape model of an unknown object has been considered extensively in previous work, e.g., [7], [4], [8]. These active tactile exploration approaches, however, are mainly based on iteratively selecting discrete locations where the robot should touch the object to improve an underlying model of the surface, which leads to inefficient touch and retract motions, where the robot moves away from the object after each touch.

To overcome this problem, we recently proposed an active learning framework that queries exploration *paths* [9]. By sliding over the unknown surface along these paths, the exploration not only turned out to be more efficient, maintaining the contact during the exploration additionally reduces the uncertainty about the unknown object.

Machine Learning and Robotics Lab, University of Stuttgart, Germany.



(a) True object, end-effectors and exploration paths after 9 m total travel.

(b) Learned surface model. Total travel path length 37 m.

Fig. 1. Exploration of a dumbbell shaped object with four end-effectors.

However, in [9], the experimental setup predominantly considered convex objects. Furthermore, the exploration criteria was developed for a single end-effector only. In the present work we address both non-convexity and exploration with multiple end-effectors at the same time.

When using Gaussian processes to model implicit surfaces to reconstruct the object shape from a tactile point cloud, a natural choice is to use the variance of the GP to base the exploration strategy on, as done, for example, in [9], [4], [8].

However, as we show in the experiments in Sec. VII-B, if multiple end-effectors are employed simultaneously, using the variance could lead to unfavorable exploration behavior. The reason for this is that for a single end-effector, the exploration can be seen as a sequential decision process, in which one observation is added sequentially. Therefore, under some assumptions, due to the submodularity of the information gain, a greedy learning strategy, which leads to maximizing the variance of the GP, is close to optimal [10]. In contrast, for an exploration with multiple contacts, multiple observations have to be chosen simultaneously in realtime, which means that the greedy argument relying on the sequential nature of the problem does not hold anymore.

Therefore, we approach this problem in a more fundamental way by directly considering the entropy of the Gaussian process surface model as the exploration criteria. This not only formally deals with multiple contacts simultaneously in an information theoretic context, but additionally also enables nonmyopic multi-step planning in the context of sliding based tactile exploration.

Sliding along an unknown surface inherently requires a compliant control framework. To this end, we extend the framework we proposed in [9] to multiple end-effectors.
To summarize our main contributions

- Continuous tactile exploration with multiple contacts within an information theoretic context.
- Generalization of [9] to non-convex objects.
- (Nonmyopic) multi-step exploration planning.

The main contribution of our work, the formalization of an

exploration strategy for multiple contacts, is presented in Sec. IV, the extension to non-convex objects in Sec. VI. The background about Gaussian implicit surface models is given in Sec. III and the controller framework is described in Sec. V. Finally in Sec. VII, we demonstrate our approach in simulation and with the real PR2 robot.

II. Related Work

A. Active Tactile Exploration

As mentioned in the introduction, many approaches to tactile shape exploration in the context of active learning have been proposed that touch the unknown object at uncertain locations. Dragiev et al. proposed in [7] to use Gaussian processes (GPs) to reconstruct the shape of the object as an implicit surface and introduced in [4] the idea of basing the exploration on the variance of this GP model to guide the robot to most uncertain locations of the model.

Some authors [11], [12] also consider how visual data can serve as a prior model of the surface which is then actively improved by touching the object where the prior model has the least certainty.

Ottenhaus et al. [13] address one of the issues we identify with touch probing, namely inefficient movements, by taking the path cost into account when selecting the next, still discrete, touch location.

Instead of representing the shape as an implicit surface, Yi et al. [14] and Jamali et al. [8] model the height of an object over a rectangular space. A disadvantage is that many real-world 3D object cannot be represented as height maps and that the exploration area must be specified beforehand.

Apart from our previous work [9], Rosales et al. [15] are the only ones we know of who consider sliding based exploration within an active learning framework. An interesting aspect of their work is that the object is hold by one hand, such that the other end-effector can touch it from other sides. However, they do not realize the sliding variant on the real robot, due to the lack of a suitable control framework. Therefore, most of their experiments are limited to touch probing as well. Such a controller that is closely coupled with the surface estimation is a crucial aspect of our work to enable the robot to slide over the unknown object.

The work of Lepora et al. [16] differs from the ones mentioned so far in the way that they explore by actively following an edge of the object. A key component is the sensor they developed, which is able to detect edges. However, their approach is limited to only two dimensional, planar objects.

B. Tactile Exploration with Multiple Contacts

The methodologies discussed in the last paragraph have been developed for a single contact point/end-effector only. Sommer et al. [17] and Bierbaum et al. [18] consider tactile exploration with multiple fingers of a robotic hand and therefore multiple contacts. These approaches, however, do not derive their exploration strategies from information theoretic criteria. For example, in [17], the exploration targets are specified beforehand manually. While suitable for multiple contacts on a hand, these methods are also only developed for exploration with a single robotic arm.

C. Active Decision Planning / Active Learning

Apart from tactile exploration, active learning methodologies play an important role in other areas like field robotics or machine learning [19], [10], [20]. In field robotics, information efficient path planning is important since a robot often has to travel long distances [21], [22]. However, in this area, robots usually do not have to deal with unknown constraints like finding or maintaining the contact with an object.

Research in so-called Bayesian optimization, which is suitable for robotic applications [23] due to its sample efficiency, also focuses on active decision making to reason about the minimum of a black-box function. But as we mentioned in the introduction, the methods of Bayesian optimization cannot directly be transferred to our problem, because exploration with multiple contacts simultaneously is not a sequential decision process.

In summary, existing approaches either perform touch probing, consider single contacts/end-effectors only or do not derive the exploration from active learning principles. To our knowledge, the present work is the first to formalize sliding based active tactile exploration with multiple contacts at the same time and multi-step planning from an information theoretic point of view as well as realizing it on the real robot through a compliant controller framework.

III. Gaussian Process Implicit Surfaces

In this work, we assume that the robot has $p \in \mathbb{N}$ end-effectors, which are equipped with tactile sensors that measure the locations $\mathbf{x}_i \in \mathbb{R}^3$ where the end-effectors are in contact with an object. If there is no contact on an end-effector, the center location of this end-effector is used as $\mathbf{x}_i$. During the whole exploration process, this data is collected with a certain sampling rate in form of the tactile point cloud $\mathscr{D} = \{(\mathbf{x}_i, c_i)\}_{i=1}^{w}$, where $c_i = 0$ if there is contact at $\mathbf{x}_i$, otherwise $c_i = 1$. In case of multiple contacts at one end-effector, all observations are added. Based on this dataset, an implicit surface model is constructed with Gaussian processes (GPs), as proposed by [7] and [24]. For a positive definite kernel $k : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ and a constant prior mean $m \in \mathbb{R}$, this GP models the probability of the implicit surface function F conditioned on the tactile data as $P(F(\mathbf{x})|\mathscr{D}) = \mathcal{N}(F(\mathbf{x})|\mu_F(\mathbf{x}), \mathbb{V}_F(\mathbf{x}))$ with mean function $\mu_F(\mathbf{x}) = m + \boldsymbol{\kappa}(\mathbf{x})^T \mathbf{G}^{-1}(\mathbf{Y} - m)$ and variance $\mathbb{V}_F(\mathbf{x}) = k(\mathbf{x}, \mathbf{x}) - \boldsymbol{\kappa}(\mathbf{x})^T \mathbf{G}^{-1} \boldsymbol{\kappa}(\mathbf{x})$ with $\boldsymbol{\kappa}(\mathbf{x}) = (k(\mathbf{x}, \mathbf{x}_i))_{i=1}^{w}$, $\mathbf{G} = \mathbf{K}_{ww} + \sigma^2 \mathbf{I}_w$, $\mathbf{K}_{ww} = (k(\mathbf{x}_i, \mathbf{x}_j))_{i=1,j=1}^{w,w}$, $\mathbf{Y} = (c_i)_{i=1}^{w}$. The surface is then approximated as the zero level set of the GP mean function, i.e. $\mathcal{S} \approx \left\{\mathbf{x} \in \mathbb{R}^3 | \mu_F(\mathbf{x}) = 0\right\}$. The choice $m = 1$ incorporates the prior knowledge that most of the $\mathbb{R}^3$ space does not contain an object [9], [7]. Instead of the often used squared exponential kernel, we choose the inverse-multiquadric kernel with length scale $l \in \mathbb{R}$, which we observed produces more favorable implicit surfaces even for small length scales without producing holes in them [9].

IV. Active Multi-Contact Exploration Framework

As discussed in the introduction and as we show in the experiments in Sec. VII-B, naively extending exploration

strategies based on the variance to multiple end-effectors simultaneously could lead to unfavorable exploration behavior. Therefore, we approach this problem in a more fundamental way. We first describe an information theoretic exploration objective suitable for multiple contacts, which is then turned into a path based exploration strategy.

As common in probabilistic modeling, the (differential) entropy of a random variable can be used to derive a notion of the information gained when adding new observations. The differential entropy of the GP implicit surface model with dataset $\mathscr{D} = \{(\mathbf{x}_i, c_i)\}_{i=1}^{w}$ is given by

$$h(\mathscr{D}) = \frac{1}{2} \log\left((2\pi e)^n \det\left(\mathbf{G}(\mathscr{D})\right)\right). \tag{1}$$

When adding new contact observations at locations $\mathbf{x}_{1:p}$, we are interested in the new entropy $h(\mathscr{D} \cup \{(\mathbf{x}, c)_{1:p}\})$, which involves the kernel matrix $\mathbf{G}_p \in \mathbb{R}^{(w+p)\times(w+p)}$ after adding the contact observations $\mathbf{x}_{1:p}$

$$\mathbf{G}_p(\mathbf{x}_{1:p}) = \begin{pmatrix} \mathbf{G} & \mathbf{K}_{wp}(\mathbf{x}_{1:p}) \\ \mathbf{K}_{pw}(\mathbf{x}_{1:p}) & \mathbf{K}_{pp}(\mathbf{x}_{1:p}) + \sigma^2 \mathbf{I}_p \end{pmatrix} \tag{2}$$

with $\mathbf{K}_{wp} = (k(\mathbf{x}_i, \mathbf{x}_j))_{i=1,j=1}^{w,p} \in \mathbb{R}^{w\times p}$, where $\mathbf{x}_i$ are the old and $\mathbf{x}_j$ the new observations ($\mathbf{K}_{pw}, \mathbf{K}_{pp}$ analog). Ignoring irrelevant constants and factors, the quantity we are interested in is therefore

$$\bar{H}(\mathbf{x}_{1:p}) = \log\left(\det\left(\mathbf{G}_p(\mathbf{x}_{1:p})\right)\right). \tag{3}$$

To calculate the determinant term efficiently, we utilize block Cholesky decompositions. Given the Cholesky decompositions $\mathbf{L} = \mathrm{chol}(\mathbf{G}) \in \mathbb{R}^{w\times w}$ and $\mathbf{N}(\mathbf{x}_{1:p}) = \mathrm{chol}(\mathbf{M}(\mathbf{x}_{1:p}))$, the determinant of the new kernel matrix can be obtained as

$$\det\left(\mathbf{G}_p(\mathbf{x}_{1:p})\right) = \prod_{i=1}^{w} \mathbf{L}_{ii}^2 \prod_{j=1}^{p} \mathbf{N}(\mathbf{x}_{1:p})_{jj}^2 \tag{4}$$

where with $\mathbf{A}(\mathbf{x}_{1:p}) = \mathbf{L}^{-1}\mathbf{K}_{wp}(\mathbf{x}_{1:p}) \in \mathbb{R}^{w\times p}$

$$\mathbf{M}(\mathbf{x}_{1:p}) = \mathbf{K}_{pp}(\mathbf{x}_{1:p}) + \sigma^2 \mathbf{I}_p - \mathbf{A}(\mathbf{x}_{1:p})^T \mathbf{A}(\mathbf{x}_{1:p}). \tag{5}$$

Inserting this into (3) and again dropping irrelevant constant factors, leads to the **exploration objective**

$$H(\mathbf{x}_{1:p}) = \sum_{i=1}^{p} \log\left(\mathrm{chol}\left(\mathbf{M}(\mathbf{x}_{1:p})\right)_{ii}\right), \tag{6}$$

which can be interpreted as the information gained about the shape of the unknown object when observing its surface at $\mathbf{x}_{1:p}$ *simultaneously*.

Based on this objective, the exploration strategy can be derived. For each of the p current contact points the idea is to find $T \in \mathbb{N}$ subsequent on-surface points that maximize (6). To obtain sliding paths, those points are constrained to be $\alpha \in \mathbb{R}$ apart from their predecessors. This is formalized in the central **Active Exploration Optimization Problem**

$$\max_{\mathbf{x}_{1:p,1:T}} \; H(\mathbf{x}_{1:p,1:T}) \tag{7a}$$

$$\text{s.t.} \quad \mu_F(\mathbf{x}_{i,t}) = 0 \qquad \forall_{i=1,\dots,p}\ \forall_{t=1,\dots,T} \tag{7b}$$

$$\|\mathbf{x}_{i,t} - \mathbf{x}_{i,t-1}\|_2^2 = \alpha^2 \quad \forall_{i=1,\dots,p}\ \forall_{t=1,\dots,T} \tag{7c}$$

$$\mathbf{x}_{i,0} = \mathbf{x}_i \qquad \forall_{i=1,\dots,p}. \tag{7d}$$

In the following, we discuss a one step ($T = 1$) and multi step ($T > 1$) solution. For both cases, the exploration criteria (6) has to be differentiated. The derivative with respect to the j-th contact point, $j = 1, \dots, p$, is

$$\frac{\partial}{\partial \mathbf{x}_j} H(\mathbf{x}_{1:p}) = \sum_{i=1}^{p} \sum_{k\le i} \mathbf{N}(\mathbf{x}_{1:p})_{ij}^{-1} \mathbf{N}(\mathbf{x}_{1:p})_{ik}^{-1} \left[\partial_1 k(\mathbf{x}_j, \mathbf{x}_k) - \left(\mathbf{G}^{-1}\boldsymbol{\kappa}(\mathbf{x}_k)\right)^T \frac{\partial}{\partial \mathbf{x}_j}\boldsymbol{\kappa}(\mathbf{x}_j)\right] \tag{8}$$

The derivation of this compact looking result is given in the appendix. The derivative of the involved Cholesky decomposition can be obtained with the results presented in [25].

1) One Step Solution: If $T = 1$, i.e. only one step is planned ahead for each end-effector simultaneously, then the solution to (7) can be approximated efficiently. Linearizing both the objective (7a) and the on-surface constraint (7b) at the current contact positions $\mathbf{x}_{1:p}$, the new position reference as the solution of (7) for the i-th end-effector is given by

$$\mathbf{x}_i^{\mathrm{ref}} = \mathbf{x}_i + \alpha \frac{\mathbf{P}_{\mathbf{n}_F}(\mathbf{x}_i) \frac{\partial^T}{\partial \mathbf{x}_i} H(\mathbf{x}_{1:p})}{\left\| \mathbf{P}_{\mathbf{n}_F}(\mathbf{x}_i) \frac{\partial^T}{\partial \mathbf{x}_i} H(\mathbf{x}_{1:p}) \right\|_2} \tag{9}$$

with the tangent space projector

$$\mathbf{P}_{\mathbf{n}_F}(\mathbf{x}) = \mathbf{I}_3 - \mathbf{n}_F(\mathbf{x})\mathbf{n}_F(\mathbf{x})^T. \tag{10}$$

The term $\mathbf{n}_F(\mathbf{x})^T = \frac{\partial}{\partial \mathbf{x}}\mu_F(\mathbf{x}) / \left\|\frac{\partial}{\partial \mathbf{x}}\mu_F(\mathbf{x})\right\|$ is the surface normal at $\mathbf{x}$, estimated from the GP model. In the case of a single contact point ($p = 1$), (9) is equivalent to Eq. (28) of [9]. One can see that the coupling between the multiple end-effectors enters in $\frac{\partial}{\partial \mathbf{x}_i} H(\mathbf{x}_{1:p})$. The computational complexity of this one step solution is $\mathcal{O}(w^2p + p^3 + wp^2)$, which basically is $\mathcal{O}(w^2 p)$, since the number of contact points p is usually much smaller than the observed tactile observations w. Comparing this to the complexity of calculating the gradient of the variance of the GP as in [9], it turns out that one step planning with multiple contacts simultaneously has the same asymptotic effort than a single contact only.

2) Multi Step Planning: If $T > 1$, then we use sequential quadratic programming to solve (7). To address the non-convexity of (7), the solver is restarted four times, initialized with the one step solution and three other, randomly selected directions in the tangent plane of each contact point.

V. Controller Concept

Sliding over a priori unknown surfaces requires a compliant controller framework that is closely coupled with the current surface estimation. To achieve this, we develop a task space controller, which tasks are parameterized directly by the GP surface estimation. The following controller framework is a sightly modified combination of [23] and [9]. The desired behavior of the robot with n joints is described in terms of task maps $\phi : \mathcal{D} \subset \mathbb{R}^n \to \mathbb{R}^d$, $\mathbf{y} = \phi(\mathbf{q})$ that map the robot configuration $\mathbf{q} \in \mathcal{D}$ to a d-dim. space like the position, which task map will be called ϕ_{pos}, or orientation of an end-effector etc. Given desired references $\mathbf{y}^{\mathrm{ref}}$, $\dot{\mathbf{y}}^{\mathrm{ref}} \in \mathbb{R}^d$, we define in that task space a linear control law $\dot{\mathbf{y}}^* = \mathbf{K}(\mathbf{y}^{\mathrm{ref}} - \phi(\mathbf{q})) + \dot{\mathbf{y}}^{\mathrm{ref}}$ with gain matrix

$\mathbf{K} \in \mathbb{R}^{d\times d}$ as well as a matrix $\mathbf{C} \in \mathbb{R}^{d\times d}$ which specifies how important this task space is for the controller. For each of the p end-effectors, mainly two special task spaces $\phi^i_{\mathbf{n}_F}$, ϕ^i_m, $i = 1, \dots, p$ are necessary:

1) Maintaining Contact Task: A velocity reference in the 1D task space with map $\phi^i_{\mathbf{n}_F}(\mathbf{q}) = -{\mathbf{n}^i_F}^T \phi^i_{\text{pos}}(\mathbf{q})$ moves each end-effector towards the surface to maintain or re-establish the contact during the exploration. Here, $\mathbf{n}^i_F = \mathbf{n}^i_F(\mathbf{q}, \mathbf{x}_i, c_i) \in \mathbb{R}^3$ is the (normalized) surface normal estimation. In this work, the robots have ball-shaped end-effectors, which allows the estimation of the surface normal via the contact location $\mathbf{x}_i$. If there is no contact, then the normalized gradient of the GP mean function is used to guide the end-effector back to the object to re-establish the contact, i.e.

$$\mathbf{n}^i_F = \begin{cases} \frac{\phi^i_{\text{pos}}(\mathbf{q}) - \mathbf{x}_i}{\left\|\phi^i_{\text{pos}}(\mathbf{q}) - \mathbf{x}_i\right\|} & \text{if contact, i.e. } c_i = 0 \\ \frac{\frac{\partial^T}{\partial \mathbf{x}} \mu_F(\phi^i_{\text{pos}}(\mathbf{q}))}{\left\|\frac{\partial^T}{\partial \mathbf{x}} \mu_F(\phi^i_{\text{pos}}(\mathbf{q}))\right\|} & \text{otherwise} \end{cases} \tag{11}$$

This normal estimation is also used in the tangent space projector (10) for the one step exploration strategy (9).

2) Moving Task: In order to slide to the references $\mathbf{x}_{i,t}$, a special position task ϕ^i_m for the i-th end-effector is utilized with importance matrix

$$\mathbf{C}^i_m(\mathbf{q}, \mathbf{x}_i, c_i) = \mathbf{V}\left(\mathbf{n}^i_F\right) \mathbf{\Lambda}_i \mathbf{V}\left(\mathbf{n}^i_F\right)^T \in \mathbb{R}^{3\times 3} \tag{12}$$

consisting of the orthogonal matrix of eigenvectors

$$\mathbf{V}\left(\mathbf{n}^i_F\right) = \begin{pmatrix} \mathbf{n}^i_F & \mathbf{t}_1\left(\mathbf{n}^i_F\right) & \mathbf{t}_2\left(\mathbf{n}^i_F\right) \end{pmatrix} \in \mathbb{R}^{3\times 3} \tag{13}$$

and eigenvalues $\mathbf{\Lambda}_i = \text{diag}(0,\ \beta_i,\ \beta_i) \in \mathbb{R}^{3\times 3}$, $\beta_i \in \mathbb{R}$, such that $(\mathbf{t}_1 \perp \mathbf{t}_2) \perp \mathbf{n}^i_F$. This ensures that the end-effector can move in the tangent space at its current contact position, without interfering with the contact maintenance task.

Based on those (and additional) M task maps and their references/parameters, the corresponding reference in the state space of the robot is then found via

$$\min_{\dot{\mathbf{q}}} \sum_{i=1}^{M} \left\| \dot{\phi}_i(\mathbf{q}) - \dot{\mathbf{y}}^*_i \right\|^2_{\mathbf{C}_i} \tag{14}$$

with solution

$$\dot{\mathbf{q}}_{\text{ref}} = \mathbf{A}^{-1} \sum_{i=1}^{M} \mathbf{J}^T_{\phi_i} \mathbf{C}_i \left(\mathbf{K}_i \left(\mathbf{y}^{\text{ref}}_i - \phi_i(\mathbf{q}) \right) + \dot{\mathbf{y}}^{\text{ref}}_i \right) \tag{15}$$

where $\mathbf{A} = \sum_{i=1}^{M} \mathbf{J}^T_{\phi_i} \mathbf{C}_i \mathbf{J}_{\phi_i}$ and $\mathbf{J}_{\phi_i} \in \mathbb{R}^{d_i \times n}$ is the Jacobian of ϕ_i at the current robot configuration $\mathbf{q}$. Singularity robustness is achieved via a regularizing posture task $\phi_{\text{reg}}(\mathbf{q}) = \mathbf{q}$ with $\mathbf{y}^{\text{ref}}_{\text{reg}} = \mathbf{q}$, $\dot{\mathbf{y}}^{\text{ref}}_{\text{reg}} = 0$.

To translate this reference to motor commands on the real robot, we extend the idea we proposed in [9] to multiple end-effectors. Starting from a joint space PD control law

$$\mathbf{u} = \widehat{\mathbf{K}}_p \left(\mathbf{q}_{\text{ref}} - \mathbf{q} \right) + \widehat{\mathbf{K}}_d \left(\dot{\mathbf{q}}_{\text{ref}} - \dot{\mathbf{q}} \right) + \mathbf{u}_f \tag{16}$$

with joint space stiffness $\widehat{\mathbf{K}}_p \in \mathbb{R}^{n\times n}$ and damping matrix $\widehat{\mathbf{K}}_d \in \mathbb{R}^{n\times n}$, the goal is that the end-effectors are compliant in the task spaces $\phi^i_{\mathbf{n}_F}$, i.e. in the direction of the surface, in order to safe- and robustly slide over the object. This is realized by first using a hand-tuned diagonal joint stiffness matrix $\mathbf{K}^{\text{base}}_p$, which is modified $\widehat{\mathbf{K}}_p = \mathbf{S}\mathbf{K}^{\text{base}}_p\mathbf{S}$ with

$$\mathbf{S} = \mathbf{I}_n - \mathbf{T}\left(\mathbf{T}^T\mathbf{T}\right)^{-1}\mathbf{T}^T \tag{17}$$

to generate compliance in the space spanned by the columns of $\mathbf{T} = \begin{pmatrix} \mathbf{p}_1 & \cdots & \mathbf{p}_p \end{pmatrix} \in \mathbb{R}^{n\times p}$. These are given by

$$\mathbf{p}_i = \mathbf{J}_{\phi^i_{\text{pos}}}(\mathbf{q})^T \mathbf{n}^i_F, \tag{18}$$

such that compliance in the direction of the surface normal estimation $\mathbf{n}^i_F$ at each end-effector is accomplished. The last important quantity $\mathbf{u}_f$ in (16) comes from an additional limit force controller, which we proposed in [23], that limits the interaction force to reduce friction and to increase safety.

VI. Handling Non-Convex Objects

So far the derivation of the exploration strategy and the control framework assumed that the number of contacts is equal to the number p of end-effectors. When exploring complex objects, it could, however, happen that there are multiple contact points on a single end-effector. Let $\mathcal{C}_i \subset \mathbb{R}^3$ be the set of all contact positions on the i-th end-effector. Then the contact location $\mathbf{x}_i$ for the exploration strategy and the task maps for the controller is selected by

$$\mathbf{x}_i = \underset{\mathbf{x}\in\mathcal{C}_i}{\text{argmax}}\, \mathbb{V}_F(\mathbf{x}). \tag{19}$$

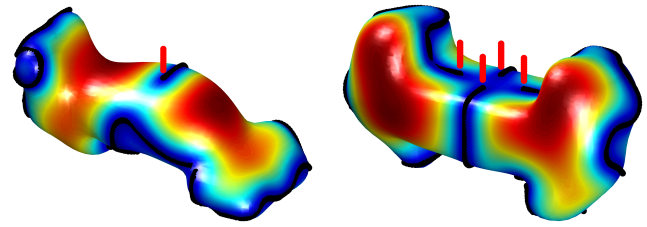
This way, if for example an end-effector moves into a corner leading to multiple contacts, the point with highest uncertainty is chosen, which resolves the non-convexity.

VII. Experiments

To focus on the exploration strategy, for the experiments performed in Sec. VII-A, -B, -C the robot consisted of a different number p of independent ball-shaped end-effectors that can translate and rotate arbitrarily, i.e. $n = 6 \cdot p$. In Sec. VII-D, we apply the proposed method to both a real and simulated PR2 robot, which has two arms. The hyperparameters for all experiments were $\alpha = 0.01$ m, $l = 0.15$, $\sigma = 0.3$. In all surface plots, the color on the surface indicates uncertainty, red high, blue low. Black points are on-surface observations, red off-surface ones.

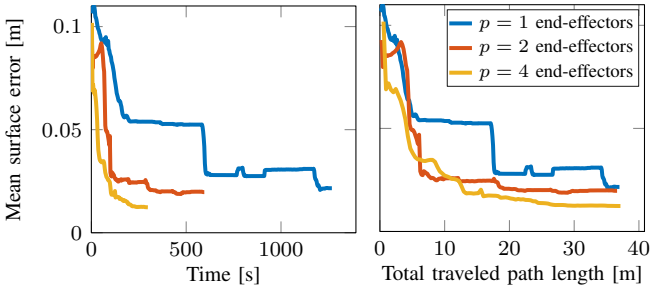
A. Influence of the Number of Contacts/End-Effectors

The experimental setup is shown in Fig. 1a. The unknown object has a dumbbell like shape with outer dimensions of 40 x 40 x 112 cm. In Fig. 2c, the evolution of the mean surface error over time for a single, two and four end-effectors is shown. Whereas it may not be surprising that more contact points at the same time lead to a faster exploration, looking at Fig. 2d, the exploration with multiple end-effectors is even more efficient when compared by the total summed traveled path of all end-effectors (meaning also the same number of contact observations). This can be explained by the fact that using multiple end-effectors, the exploration is spread more evenly on the object, as visualized in Fig. 2a and 2b, where the surface reconstruction is shown after 5 m total



(a) One end-effector. Surface model after 5 m exploration path in total.

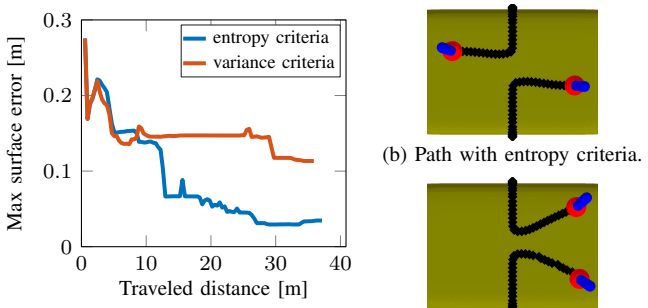
(b) Four end-effectors. Surface model after 5 m exploration path in total.



(c) Mean surface error over time.

(d) Mean surface error over total path.

Fig. 2. Influence of the number of end-effectors on the surface error.



(a) Max surface error over total path.

(b) Path with entropy criteria.

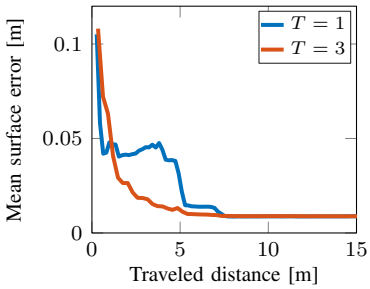
(c) Path with variance criteria.

Fig. 3. Comparison between the proposed exploration strategy that is based on entropy and a variance based approach. Four end-effectors, same dumbbell object as shown in Fig. 1.

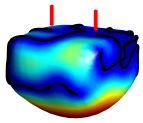
exploration path with one and four end-effectors. The final learned surface explored with four end-effectors after a total of 37 m travel path is visualized in Fig. 1b.

B. Entropy vs. Variance

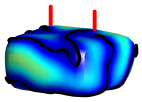
One could argue, especially if the contact points are far away from each other, that basing the exploration on the gradient of the variance of the GP as in [9] should be sufficient. In this experiment, we show that this is not the case. The setup is the same as in Sec. VII-A, i.e. same dumbbell shaped object, four end-effectors (see Fig. 1a). As can be seen in Fig. 3a, the maximum surface error when using the variance is much larger than with the proposed entropy criteria. The reason for this is that with the variance, two of the four end-effectors moved in the same direction, as shown in Fig. 3c, therefore, they both explored the same part of the object, while with the entropy strategy (Fig. 3b), the end-effectors moved in opposite directions. For simpler objects like a box, however, we still observed a difference, but less significant as for this dumbbell shape like object.



(a) Max surface error over total path.



(b) Surface after 4.7 m travel with one step planning.



(c) Surface after 4.7 m travel with multi step planning.

Fig. 4. Comparison between one step exploration planning ($T = 1$) and multi step ($T = 3$) with two end-effectors. Box object (40 x 40 x 20 cm).

C. Multi-Step Predictions

Fig. 4a shows that even for a simpler object like a box as used in this experiment, multi step planning (in this case three steps, $T = 3$) is more efficient than the one step solution ($T = 1$). Comparing the surface model after 4.7 m total travel length, one can see that the three step planning strategy already looks like a box (Fig. 4c), whereas with the one step solution, the bottom is still round (Fig. 4b). However, it has to be clearly stated that solving (7) for $p = 2$, $T = 3$ could not be done in realtime, compared to the one step multi contact solution (Sec. IV-.1), whose computation time for the considered task is neglectable.

D. PR2 Robot

In Fig. 5, the exploration of a box shaped object with a simulated PR2 robot with two end-effectors (one on each arm) is shown. The dimensions of the object are 28 x 54 x 12 cm. Fig. 6 presents the proposed formalism with a real PR2 robot exploring a salad bowl (30 cm diameter) with both arms simultaneously. The experimental setup is similar to the one of [9], in which the same object is explored, but with only one end-effector. Since the end-effectors are not equipped with contact location sensors, whether the end-effector is in contact or not is inferred based on force-torque measurements in the wrists of the PR2. The surface normal is therefore always estimated based on the GP surface model.

In contrast to the experiments of the last paragraphs, the end-effectors of the PR2 cannot move arbitrarily in space. In order to account for those kinematic limits and to prevent that the two arms could interlock each other, the optimization problem (7) is modified by including a function $P : \mathbb{R}^{3 \cdot p \cdot T} \to \mathbb{R}$, $P \in C^1$ that acts like a potential

$$\max_{\mathbf{x}_{1:p,1:T}} \; H(\mathbf{x}_{1:p,1:T}) - P(\mathbf{x}_{1:p,1:T}) \tag{20a}$$

$$\text{s.t.} \quad \text{(7b)}, \text{(7c)}, \text{(7d)}, \tag{20b}$$

where P is a quadratically increasing function if $\mathbf{x}_{i,t}$ is below a certain height or if an arm leaves its part of the body. Otherwise, P is zero. This way, the two arms stay mostly on their side of the body, which prevents the interlocking. Furthermore, a joint limit task is added to the controller, such that the robot does not try to explore outside its reachable limits. In Fig. 7, it is shown that both in simulation and with the real robot, the exploration with two arms simultaneously is faster in terms of mean surface error reduction than with

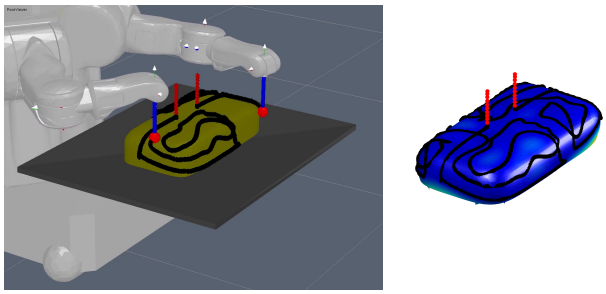


Fig. 5. Dual arm object exploration with simulated PR2 robot. Right: learned surface model after a total of 7 m traveled path.

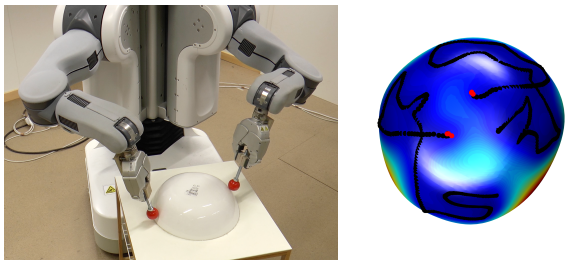
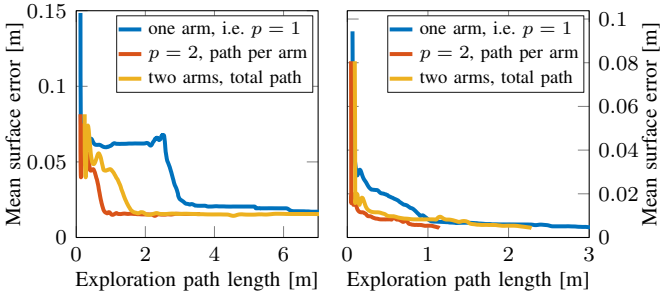


Fig. 6. Dual arm object exploration with real PR2 robot. Right: learned surface model after a total of 2.3 m exploration path.



(a) Simulated PR2 robot. (b) Real PR2 robot.

Fig. 7. Comparison between exploration with one and two arms with respect to the mean surface error. Simulation and with real PR2 robot.

a single end-effector only. Note that for the real salad bowl object we do not have the ground truth surface model and therefore, in order to allow for a fair comparison, the surface error in Fig. 7b is calculated against the mean of the two surface estimations (with one and two end-effectors) at the end of their exploration. Another advantage of using multiple arms is that the possible exploration range is larger. For example, the object in the simulated experiment (Fig. 5) has a small part that is unreachable if only one arm is used, but it had no significant influence on the surface error.

VIII. Conclusion

We presented a novel formalism for the active tactile exploration of non-convex objects with multiple contacts simultaneously. It turned out that using more than one end-effector at the same time leads to a more efficient exploration, even when considering the total traveled path length of all end-effectors. Furthermore, naively extending concepts based on the variance of the GP to multiple contacts showed to perform inferior than the proposed method that is based on the entropy of the GP. While in principle the introduced framework enables multi-step planning, which also leads to an even more efficient exploration, the computational effort in our current implementation is not suitable for realtime. We belief that this work lays the formal basis for advanced active tactile shape exploration solutions.

Appendix

Here we derive the derivatives (8) of the exploration objective (6). For better readability, we abbreviate $\mathbf{M}_{kl} = \mathbf{M}(\mathbf{x}_{1:p})_{kl}$, $\mathbf{N}_{kl} = \mathbf{N}(\mathbf{x}_{1:p})_{kl} = \mathrm{chol}(\mathbf{M}(\mathbf{x}_{1:p}))_{kl}$. $\mathbf{N}_{kl}$ is lower triangular. The term $\frac{\partial \mathbf{N}_{ij}}{\partial \mathbf{M}_{kl}}$ denotes the derivative of the ij-th entry of the Cholesky decomposition of $\mathbf{M}$ with respect to the kl-th entry in $\mathbf{M}$. $\mathbf{N}^{-1}_{kl}$ means the kl-th entry of $\mathbf{N}^{-1}$. Applying the chain rule to (6) gives

$$\frac{\partial}{\partial \mathbf{x}_j} H(\mathbf{x}_{1:p}) = \sum_{i=1}^{p} \frac{1}{\mathbf{N}_{ii}} \sum_{k} \sum_{l \leq k} \frac{\partial \mathbf{N}_{ii}}{\partial \mathbf{M}_{kl}} \frac{\partial \mathbf{M}_{kl}}{\partial \mathbf{x}_j}. \tag{21}$$

Now with (11) of [25] the partial derivatives of the Cholesky decomposition $\frac{\partial \mathbf{N}_{ii}}{\partial \mathbf{M}_{kl}}$ can be obtained, which yields

$$= \sum_{i=1}^{p} \sum_{k} \sum_{l \leq k} \mathbf{N}^{-1}_{ik} \mathbf{N}^{-1}_{il} \left(1 - \tfrac{1}{2}\delta_{kl}\right) \frac{\partial \mathbf{M}_{kl}}{\partial \mathbf{x}_j}. \tag{22}$$

The expression δ_{kl} is the Kronecker delta. Then after inserting the derivative of (5), we arrive at

$$\begin{aligned} = \sum_{i=1}^{p} \sum_{k} \sum_{l \leq k} & \mathbf{N}^{-1}_{ik} \mathbf{N}^{-1}_{il} \left(1 - \tfrac{1}{2}\delta_{kl}\right) \\ & \Big[\partial_1 k(\mathbf{x}_j, \mathbf{x}_l)\delta_{jk} + \partial_2 k(\mathbf{x}_k, \mathbf{x}_j)\delta_{jl} \\ & \left(\boldsymbol{\kappa}(\mathbf{x}_k)\delta_{jl} + \boldsymbol{\kappa}(\mathbf{x}_l)\delta_{jk}\right)^T \mathbf{G}^{-1} \frac{\partial}{\partial \mathbf{x}_j} \boldsymbol{\kappa}(\mathbf{x}_j)\Big]. \end{aligned} \tag{23}$$

Here, ∂_1 denotes the derivative of the kernel with respect to its first argument, ∂_2 to its second. Resolving some of the sums via the Kronecker deltas yields

$$\begin{aligned} = & \sum_{i=1}^{p} \sum_{l \leq j} \mathbf{N}^{-1}_{ij} \mathbf{N}^{-1}_{il} \left(1 - \tfrac{1}{2}\delta_{jl}\right) \\ & \qquad \left[\partial_1 k(\mathbf{x}_j, \mathbf{x}_l) - \boldsymbol{\kappa}(\mathbf{x}_l)^T \mathbf{G}^{-1} \frac{\partial}{\partial \mathbf{x}_j} \boldsymbol{\kappa}(\mathbf{x}_j)\right] \\ & + \sum_{i=1}^{p} \sum_{k \geq j} \mathbf{N}^{-1}_{ij} \mathbf{N}^{-1}_{ik} \left(1 - \tfrac{1}{2}\delta_{jk}\right) \\ & \qquad \left[\partial_2 k(\mathbf{x}_k, \mathbf{x}_j) - \boldsymbol{\kappa}(\mathbf{x}_k)^T \mathbf{G}^{-1} \frac{\partial}{\partial \mathbf{x}_j} \boldsymbol{\kappa}(\mathbf{x}_j)\right]. \end{aligned} \tag{24}$$

Since a kernel is symmetric, we have $\partial_1 k(\mathbf{x}, \mathbf{x}') = \partial_2 k(\mathbf{x}', \mathbf{x})$ and therefore

$$\begin{aligned} = \sum_{i=1}^{p} \sum_{k=1}^{p} & \mathbf{N}^{-1}_{ij} \mathbf{N}^{-1}_{ik} \underbrace{\left(1 - \tfrac{1}{2}\delta_{jk}\right)\left(1 + \delta_{jk}\right)}_{=1} \\ & \left[\partial_1 k(\mathbf{x}_j, \mathbf{x}_k) - \left(\mathbf{G}^{-1} \boldsymbol{\kappa}(\mathbf{x}_k)\right)^T \frac{\partial}{\partial \mathbf{x}_j} \boldsymbol{\kappa}(\mathbf{x}_j)\right]. \end{aligned} \tag{25}$$

Finally, since the inverse of a lower triangular matrix is lower triangular, the result (8) follows.

REFERENCES

[1] S. J. Lederman and R. L. Klatzky, "Haptic perception: A tutorial," *Attention, Perception, & Psychophysics*, 2009.

[2] Z. Su, J. A. Fishel, T. Yamamoto, and G. E. Loeb, "Use of tactile feedback to control exploratory movements to characterize object compliance," *Frontiers in neurorobotics*, vol. 6, p. 7, 2012.

[3] M. Kaboli, K. Yao, D. Feng, and G. Cheng, "Tactile-based active object discrimination and target object search in an unknown workspace," *Autonomous Robots*, 2018.

[4] S. Dragiev, M. Toussaint, and M. Gienger, "Uncertainty aware grasping and tactile exploration," in *Proc. of the Int. Conf. on Robotics and Automation (ICRA)*, 2013.

[5] F. Veiga and A. Bernardino, "Active tactile exploration for grasping," in *ICRA Workshop on Autonomous Learning*, 2013.

[6] Q. Li, C. Schuermann, R. Haschke, and H. Ritter, "A control framework for tactile servoing," in *Robotics: Science and Systems*, 2013.

[7] S. Dragiev, M. Toussaint, and M. Gienger, "Gaussian process implicit surfaces for shape estimation and grasping," in *Proc. of the Int. Conf. on Robotics and Automation (ICRA)*, 2011.

[8] N. Jamali, C. Ciliberto, L. Rosasco, and L. Natale, "Active perception: Building objects' models using tactile exploration," in *Proc. of the Int. Conf. on Humanoid Robots (Humanoids)*, 2016.

[9] D. Driess, P. Englert, and M. Toussaint, "Active learning with query paths for tactile object shape exploration," in *Proc. of the Int. Conf. on Intelligent Robots and Systems (IROS)*, 2017.

[10] A. Krause and C. E. Guestrin, "Near-optimal nonmyopic value of information in graphical models," arXiv:1207.1394, 2012.

[11] M. Björkman, Y. Bekiroglu, V. Högman, and D. Kragic, "Enhancing visual perception of shape through tactile glances," in *IEEE Int. Conf. on Intelligent Robots and Systems (IROS)*, 2013.

[12] S. Wang, J. Wu, X. Sun, W. Yuan, W. T. Freeman, J. B. Tenenbaum, and E. H. Adelson, "3D shape perception from monocular vision, touch, and shape priors," arXiv:1808.03247, 2018.

[13] S. Ottenhaus, L. Kaul, N. Vahrenkamp, and T. Asfour, "Active tactile exploration based on cost-aware information gain maximization," *International Journal of Humanoid Robotics*, vol. 15, no. 01, 2018.

[14] Z. Yi, R. Calandra, F. Veiga, H. van Hoof, T. Hermans, Y. Zhang, and J. Peters, "Active tactile object exploration with gaussian processes," in *Proc. of the IEEE Int. Conf. on Intelligent Robots and Systems (IROS)*, 2016.

[15] C. Rosales Gallegos, F. Spinelli, M. Gabiccini, C. Zito, and J. Wyatt, "GPAtlasRRT: A local tactile exploration planner for recovering the shape of novel objects," *International Journal of Humanoid Robotics*, 2018.

[16] N. F. Lepora, K. Aquilina, and L. Cramphorn, "Exploratory tactile servoing with active touch," *Robotics and Automation Letters*, vol. 2, no. 2, 2017.

[17] N. Sommer and A. Billard, "Multi-contact haptic exploration and grasping with tactile sensors," *Robotics and Autonomous Systems*, 2016.

[18] A. Bierbaum, M. Rambow, T. Asfour, and R. Dillmann, "Grasp affordances from multi-fingered tactile exploration using dynamic potential fields," in *Proc. of the Int. Conf. on Humanoid Robots*, 2009.

[19] J. Schreiter, D. Nguyen-Tuong, M. Eberts, B. Bischoff, H. Markert, and M. Toussaint, "Safe exploration for active learning with gaussian processes," in *Proc. of the Conf. on Machine Learning (ECML)*, 2015.

[20] P. Hennig and C. J. Schuler, "Entropy search for information-efficient global optimization," *Journal of Machine Learning Research*, vol. 13, 2012.

[21] G. A. Hollinger, B. Englot, F. S. Hover, U. Mitra, and G. S. Sukhatme, "Active planning for underwater inspection and the benefit of adaptivity," *The Int. Journal of Robotics Research*, 2012.

[22] R. Marchant and F. Ramos, "Bayesian optimisation for informative continuous path planning," in *Proc. of the Int. Conf. on Robotics and Automation (ICRA)*, 2014.

[23] D. Drieß, P. Englert, and M. Toussaint, "Constrained bayesian optimization of combined interaction force/task space controllers for manipulations," in *Proc. of the Int. Conf. on Robotics and Automation (ICRA)*, 2017.

[24] O. Williams and A. Fitzgibbon, "Gaussian process implicit surfaces," June 2006.

[25] I. Murray, "Differentiation of the cholesky decomposition," arXiv:1602.07527, 2016.