2026 IEEE International Conference on Robotics and Automation (ICRA 2026)
June 1-5, 2026. Vienna, Austria

Touch-Based Object Localisation with Spatially-Aware Belief Entropy Estimation

Lara Brudermüller[1], Julius Jankowski[2], Marc Toussaint[3,4], and Nick Hawes[1]

***Abstract*— Robust robotic manipulation in the real world requires coping with incomplete or unreliable sensory input. While vision provides rich information, it often fails in the presence of occlusions, clutter, or poor lighting. In such cases, touch offers a robust alternative, enabling object localisation through contact alone. We present a touch-only global localisation method that operates in continuous state space with a particle belief. Sparse contact/no-contact signals are turned into informative likelihoods via a proximity-aware measurement model, and contact-aware resampling mitigates particle starvation. An information-gathering controller selects actions that maximise expected information gain using a non-parametric entropy estimator sensitive to both observation updates and dynamics. On real hardware, the system reliably localises and then grasps from broad, multi-modal initial beliefs with mode separations up to 0.4 m, far beyond the narrow uncertainty ranges assumed in related work. Information-aware localisation-actions speed up belief convergence and boost grasp success; and ablations in simulation confirm the benefits of the measurement and resampling components.**

I. Introduction

Humans can manipulate objects using only touch, even in the absence of vision. For robots to approach similar capabilities in unstructured or visually-challenging environments, such as those involving occlusions, clutter, or poor lighting, localising objects through touch alone represents a promising alternative. This has driven the development of algorithms that refine an object's pose estimate through deliberate physical interaction [1]–[3]. As an illustrative example, consider a robot having to retrieve a keyring from inside a bag using a multi-fingered hand. The object pose is initially unknown, contacts are intermittent and ambiguous, and the object may move such that exploratory interactions can either resolve ambiguity or further increase uncertainty in the object state. Such examples highlight the challenges of high-dimensional contact-rich estimation problems, where the robot often starts with little information. Estimating the posterior over object pose in these settings is computationally demanding: the complexity grows rapidly with both the number of degrees of freedom (DOFs) and the size of the initial uncertainty region [4]. As a result, most prior approaches restrict either the problem dimensionality or the scale of initial uncertainty. Yet, contact-rich manipulation is not only characterized by high-dimensional state and action spaces, but also the inherent highly non-linear contact dynamics and the multi-modality of the system state distributions [5]. Discretising the problem space addresses the multi-modality while enabling standard filtering and planning over finite hypotheses [3], [6], but the curse of dimensionality limits the respective resolution and thus fails to capture the rich contact dynamics present in such manipulation tasks. Instead, these tasks require a framework that plans *in continuous state space*, anticipates *what* will be sensed and *how* actions reshape uncertainty through interaction, and operates from *uninformed, non-parametric* priors with broad support.

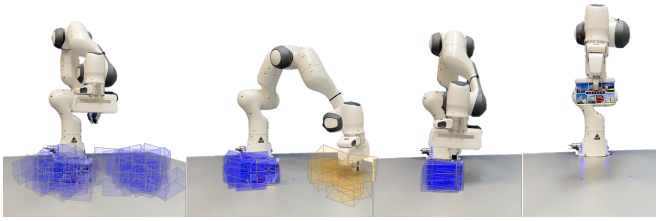

Fig. 1. Experimental setup of blind grasping. *Left to right:* initial particle belief with uniform weights (■); information-gathering trajectory rejecting particle hypotheses (■); converged belief after contact; successful grasp.

Towards this end, we address touch-based localisation with a system that operates in continuous state spaces and actively refines the belief distribution through contact. We propose a particle filter with a proximity-aware measurement model to turn sparse binary proprioceptive contact signals into informative likelihoods combined with contact-aware resampling. An information-gathering controller predicts how actions reshape the belief and selects trajectories that maximise expected information gain using a non-parametric entropy estimator that not only considers the probabilities of different object poses but also their spatial density.

Contributions: We address continuous-space object localisation through contact from uninformed, non-parametric beliefs, making the following contributions:

1) A *touch-only global localisation system* that plans and estimates directly in continuous state space with beliefs from uninformed, non-parametric priors, suitable for operation when vision is unreliable or absent.
2) A *proximity-aware measurement model* for contact that converts sparse binary signals into informative likelihoods, and a *contact-aware resampling* strategy that mitigates particle starvation under discontinuous observations.
3) A *sampling-based information-gathering controller* that

LB was supported by an Amazon Web Services Lighthouse scholarship. NH received EPSRC funding via the "From Sensing to Collaboration" programme grant [EP/V000748/1]. MT was supported by the German Federal Ministry of Research, Technology and Space (BMFTR) under the Robotics Institute Germany (RIG).

[1]Oxford Robotics Institute, University of Oxford, UK; {larab, nickh}@robots.ox.ac.uk
[2]Idiap Research Institute & Ecole Polytechnique Fédérale de Lausanne (EPFL), CH; jankowskijulius@gmail.com
[3]TU Berlin, Germany; toussaint@tu-berlin.de
[4] Robotics Institute Germany

selects candidate probing actions based on a non-parametric *differential entropy estimator* that captures both observation-driven changes (weights) and dynamics-driven changes (spatial density) in the belief.

On real hardware (cf. setup in Fig. 1) and in simulation, the approach localises and grasps reliably under broad, *multi-modal* initial non-parametric beliefs with separations up to $0.4\,\mathrm{m}$, far beyond the narrow uncertainty ranges (typically up to $0.04\,\mathrm{m}$) assumed in related work, highlighting the significant benefit of continuous touch-based object localisation when vision is unreliable or unavailable.

II. Related Work

a) Object Localisation through Contact: Early work in robot manipulation primarily relied on visual feedback, but as we move towards more robust and dexterous manipulation in unstructured environments, tactile sensing becomes increasingly important. In this work, we focus on object localisation through contact, where the robot must infer object state from sparse and noisy contact signals during physical interaction. This problem is particularly challenging due to continuous, high-dimensional state and action spaces, the high computational cost of physics simulation (required to accurately model contact dynamics), and the inherently discontinuous nature of contact sensor observations [6]. A common approach is to plan a sequence of "move-until-touch" actions for exploration [3], [4], [7], followed by a goal-directed phase, such as grasping. Extensions of this idea in belief-space control compute exploratory actions by optimising information-theoretic costs [5], [8], or switch between exploration and exploitation based on the belief uncertainty [9]. More principled approaches frame the task as a Partially-Observable Markov decision process (POMDP) [6], [10], [11], though tractability often requires coarse state discretisations. However, most of these methods are subject to two major limitations: *(i)* belief representations and planning typically rely on discretisation, which limits resolution and scalability; and *(ii)* object dynamics are often ignored or severely simplified, assuming static objects or restricted interaction effects. The latter is particularly problematic in contact-rich settings, where robot actions can increase uncertainty, e.g., by pushing or perturbing the object. Our work addresses both limitations by planning directly in continuous state and action spaces, and by capturing the effect of interaction dynamics on belief evolution through a non-parametric particle filter.

b) Informative Action Selection: A key challenge in planning under uncertainty is to evaluate which actions are most informative—typically by estimating their expected information gain (IG), defined as the reduction in entropy between belief states before and after an observation [3], [7], [12]. In manipulation settings, the inherent multi-modality typically requires non-parametric belief representations, such as weighted particle sets, complicating the computation of entropy. Prior work has largely relied on two approximations. One approach discretises the state space and computes discrete entropy, using occupancy grids or histograms where the entropy is defined over categorical probabilities [3]. Another approach assumes a continuous belief but approximates differential entropy using only particle weights, e.g., [9], [12], [13]. While weight-based measures reflect changes in belief confidence due to observations, they neglect the spatial distribution of particles. As a result, a belief with widely scattered particles may be assigned the same entropy as one with tightly clustered particles, despite reflecting fundamentally different levels of uncertainty. In the context of touch, this limitation becomes especially relevant: contact not only yields an observation (e.g., contact/no-contact) but can also physically alter the object's pose. Thus, touch simultaneously refines the belief and shifts the underlying state, making it essential to account for both observation-driven and dynamics-induced changes in the belief's spatial structure. In contrast, more expressive estimators of differential entropy explicitly account for the spatial distribution of the belief. For instance, kernel-based methods approximate the underlying density using a smooth distribution over particle locations [14], allowing them to capture how spread out or concentrated the belief is in state space. However, these methods are often computationally demanding and require careful tuning of kernel hyperparameters, such as bandwidth [14], [15]. [16] also derive an entropy estimate specifically for particle filters via Bayes' theorem, but this requires known transition and observation models. An alternative class of estimators based on k-nearest neighbours (kNN) density estimates provides a principled and efficient way to estimate differential entropy from samples, including for non-parametric distributions [17], [18]. Extensions to weighted particle sets make these estimators well-suited for non-parametric belief representations used in robotics [19]. These kNN density estimators can be regarded as a kernel estimator with a bandwidth that adapts to local data density [20]. Thus, kNN-based differential entropy estimators directly incorporate spatial density and have been widely adopted in reinforcement learning as intrinsic reward signals [21]–[23]. To the best of our knowledge, our work is the first to propose the use of such estimators for action selection in active tactile object localisation, where belief evolution is driven by both observations and complex interaction dynamics. This setting presents unique challenges, as actions can increase uncertainty, making spatial sensitivity essential for accurate entropy estimation. For a broader overview of non-parametric entropy estimation methods, we refer the reader to [24].

III. Problem Formulation

This paper addresses the problem of localising a rigid object o through sparse proprioceptive contact measurements using probing actions that aim to maximise the estimated expected information gain under the current belief state. We define successful localisation as reaching a state from which a grasp will succeed. To achieve this, the robot must perform information-gathering actions that reduce uncertainty about the object's location. Let $\boldsymbol{x}_t = (\boldsymbol{q}^r, \boldsymbol{q}^o) \in \mathbb{R}^{(n^r_{\mathrm{dof}}+n^o_{\mathrm{dof}})}$ describe the state of the underactuated robotic system, which includes the pose of the object $\boldsymbol{q}^o_t$, and the robot configuration

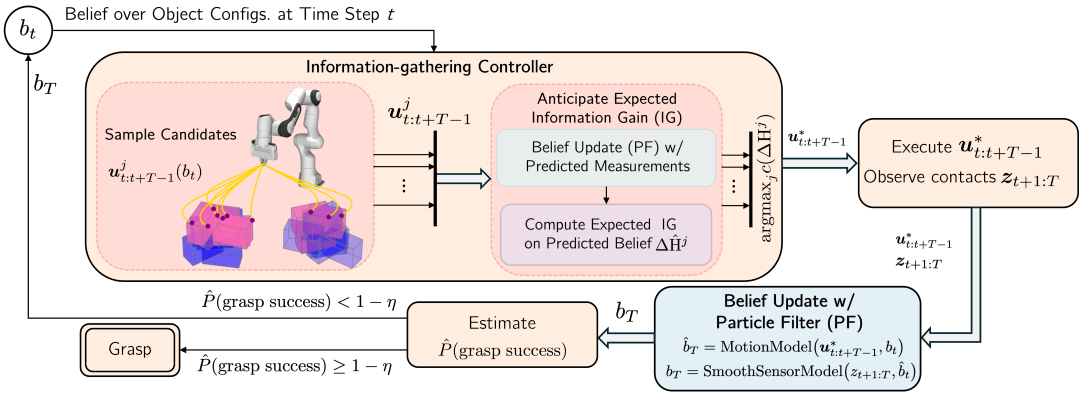

Fig. 2. Proposed pipeline for contact-rich localisation and manipulation.

$\boldsymbol{q}^r_t$ at time step t. In this work, $\boldsymbol{q}^o \in SE(2)$ denotes the object's planar pose, parameterised by (x, y, θ). The robot applies control inputs $\boldsymbol{u}_t \in \mathbb{R}^{n^r_{\text{dof}}}$, which influence the object state only indirectly through contact and interaction. The system evolves according to stochastic dynamics $\boldsymbol{x}_{t+1} \sim p(\cdot|\boldsymbol{x}_t, \boldsymbol{u}_t)$ and generates stochastic measurements $\boldsymbol{z}_t \sim p(\cdot|\boldsymbol{x}_t)$, where $\boldsymbol{z}_t$ represents sparse, binary contact observations. Due to the partial observability of the system and it's complex dynamics, we maintain a *continuous* belief over the system state, represented as a probability distribution $b_t{=}p(\boldsymbol{x}_t|\boldsymbol{u}_{0:t-1}, \boldsymbol{z}_{1:t}, b_0)$, which depends on the full history of actions, measurements, and the initial belief b_0. Note that we explicitly include the robot configuration in the state because it directly governs the contact interactions that drive the object's stochastic dynamics, as well as the measurements.

Given a sequence of control actions $\boldsymbol{u}_{0:t-1}$, we can estimate the expected information gain $\text{IG}(b_0, \boldsymbol{u}_{0:t-1})$ using dynamics and measurement models, as well as a model predicting future observations. We formulate the information gathering problem as trajectory optimisation in belief space:

$$\max_{\boldsymbol{u}_{0:T-1}} \quad \text{IG}(b_0, \boldsymbol{u}_{0:T-1}), \tag{1}$$

where T is the planning horizon.

The following sections outline our approach to the information-gathering problem, with a summary illustrated in Fig. 2. In Sec. IV, we introduce a sampling-based controller that selects the next best action based on a spatially-aware estimator for the expected information gain. In Sec. V, we present our particle-based state estimator for object localisation through touch that is able to handle the discontinuous and sparse nature of the contact measurement signal. The overall approach is evaluated in Sec. VI in robot experiments and ablation studies.

Notation. In the remainder of the paper, we use subscripts for the time index and superscripts for the particle indices.

IV. Information-Gathering Controller

We propose a sampling-based controller that generates candidate probing trajectories and selects the one with the highest expected information gain (IG). The idea of sampling trajectories and ranking them by IG follows prior work in active perception [3], but here we extend it to continuous belief space and contact-rich dynamics.

A. Differential Entropy of a Weighted Particle Set

The differential entropy of a random, continuous variable $\boldsymbol{x}$ following a probability distribution $p(\boldsymbol{x})$ is defined as

$$\text{H}[\boldsymbol{x}] = -\int_{\mathcal{X}} p(\boldsymbol{x}) \log p(\boldsymbol{x}) d\boldsymbol{x}. \tag{2}$$

However, the definition of differential entropy does not directly extend to finite particle sets [24]. This is due to the fact that there is no unique way to define a probability density function that is parameterised by N_p weighted samples $\{\boldsymbol{x}^i, w^i\}_{i=1}^{N_p}$. Without considering dynamics, the particle-based belief may be treated as a discrete probability distribution with the weights representing the probability of each particle. The corresponding *weight-based entropy approximation* is then given by $\hat{\text{H}}_w[\boldsymbol{x}] = -\sum_i w^i \log(w^i)$ [9]. While this approximation captures information gained through observations, it does not capture the spatial density of particles, i.e., the local distance between the states of the particles. In contrast, *kernel-based* approaches approximate the underlying belief distribution b with a kernel density estimate (KDE) computed from the weighted particle set, i.e.

$$\hat{b}(\boldsymbol{x}^i) = p\left(\boldsymbol{x} \middle| \{\boldsymbol{x}^i, w^i\}_{i=1}^{N_p}\right) = \sum_i w^i \, k(\boldsymbol{x}, \boldsymbol{x}^i), \tag{3}$$

where k represents the corresponding kernel function. Fischer et al. [14] show that for particle-based belief representations the integral in the differential entropy, as defined in Eq. (2) is commonly approximated as

$$\hat{\text{H}}[\hat{b}] = -\sum_{i=1}^{m} w^i \log \hat{b}(\boldsymbol{x}^i). \tag{4}$$

While this approximation accounts for the spatial density information, it is typically computationally more demanding and requires careful choice of kernel hyperparameters. In this work, we choose a uniform kernel with shared support Ω estimated from local ρ-nearest-neighbour bounding boxes for all particles (more details on how we compute Ω are provided in A). As a result, the integral in Eq. (4) simplifies

to a sum over a weight-based term and a spatially-aware density term. This yields the entropy estimate

$$\hat{\mathrm{H}}[\boldsymbol{x}] \;=\; -\sum_i w^i \log w^i \;+\; \log V(\Omega), \tag{5}$$

where the first term reflects observation-driven weight changes and the second captures *spatial density* via the hypervolume $V(\Omega)$. While conceptually similar to knn-based entropy estimators [17], this formulation computes a *shared* support for all particles instead of individual supports, which is more robust to outliers and splits the entropy into two separate terms, which renders the computation of the expected information gain more efficient (cf. Sec. IV-B). The only hyperparameter is ρ for locality, and no bandwidth tuning is required. Further details on the derivation of the estimator, its properties, and a discussion of its computational complexity are provided in A.

B. Expected Information Gain

To evaluate candidate actions for information gathering, we estimate their *expected information gain (IG)*, i.e. the expected reduction in belief entropy after executing a trajectory and incorporating new observations. This process is summarised in Alg. 1. The trajectory is then truncated at the time step with maximal expected information gain.

Formally, for an initial belief b_0 and a control sequence $\boldsymbol{u}_{0:T-1}$, we define the IG as

$$\mathrm{IG}(b_0, \boldsymbol{u}_{0:T-1}) = \mathrm{H}[b_0] - \mathbb{E}_{\hat{\boldsymbol{z}}_{1:T}}\big[\mathrm{H}[b_T]\big] = \Delta \mathrm{H}_{0,T}, \tag{6}$$

where $\hat{\boldsymbol{z}}_{1:T}$ are hypothetical measurements generated along the trajectory. A common approach to estimating this expectation is to use the maximum likelihood (ML) state from the belief to simulate observations [5], [25]. However, when particle weights are uniform, such as immediately after resampling, this biases the IG estimate incorrectly towards a single hypothesis. Instead, we marginalise over all particles in the belief to generate hypothetical measurements and compute a weighted expectation of the resulting entropy. At each time step, each particle is treated as a potential ground truth state of the system, producing a hypothetical observation from which all particle weights are updated. The expected entropy is then computed as the weighted average over all such hypotheses. Since the observations generated along the candidate trajectory are only a guess of what could occur upon execution, we estimate the information gain at time step t as if no observations have been made in previous time steps. Therefore, the update of a particle weight w^i_{t+1} is done assuming that the previous weight w^i_t is equal to w^i_0. In other words, we predict the belief states along a candidate trajectory as $b_t{=}p(\boldsymbol{x}_t|\boldsymbol{u}_{0:t-1}, \boldsymbol{z}_t, b_0)$ instead of $b_t{=}p(\boldsymbol{x}_t|\boldsymbol{u}_{0:t-1}, \boldsymbol{z}_{1:t}, b_0)$.

To compute the differential entropy for a belief state b_t, we use the estimator introduced above in Eq. (5), which decomposes the entropy into two terms: a weight-based term $\hat{\mathrm{H}}_w$ and a spatial density term $\hat{\mathrm{H}}_\Omega$. Only the weight term depends on the simulated measurements; the density term depends solely on the particle locations after propagating

Algorithm 1: Predict Expected Information Gain

Input: Initial particles & weights $\{(\boldsymbol{x}^i_0, w^i_0)\}_{i=1}^{N_p}$, controls $\boldsymbol{u}_{0:T-1}$
Output: Max. expected IG $\Delta\hat{\mathrm{H}}^\star_{0,t}$ at time step $t^\star$

// Precompute constants at $t{=}0$
$\hat{\mathrm{H}}_{w,0} \leftarrow -\sum_{i=1}^{N_p} w^i_0 \log w^i_0$
$\Omega_0 \leftarrow \textsc{SharedSupport}\big(\{\boldsymbol{x}^i_0\}_{i=1}^{N_p}\big)$; $\hat{\mathrm{H}}_{\Omega,0} \leftarrow \log V(\Omega_0)$
$\Delta\hat{\mathrm{H}}^\star_{0,t} \leftarrow -\infty$; $t^\star \leftarrow 0$
for $t \leftarrow 0$ **to** $T{-}1$ **do**
  // Simulate particle dynamics under $\boldsymbol{u}_t$
  $\boldsymbol{x}^i_{t+1} \sim p(\cdot|\boldsymbol{x}^i_t, \boldsymbol{u}_t)$
  // Expected weight entropy (via Eq. (7))
  $\mathbb{E}\big[\hat{\mathrm{H}}_{w,t+1}\big] \leftarrow$
    $\textsc{ExpectedWeightEntropy}\big(\{\boldsymbol{x}^i_t, w^i_t\}_{i=1}^{N_p}, \boldsymbol{u}_t\big)$
  // Spatial entropy term
  $\Omega_{t+1} \leftarrow \textsc{SharedSupport}\big(\{\boldsymbol{x}^i_{t+1}\}_{i=1}^{N_p}\big)$
  $\hat{\mathrm{H}}_{\Omega,t+1} \leftarrow \log V(\Omega_{t+1})$
  // Predicted entropy reduction relative to $t{=}0$
  $\Delta\hat{\mathrm{H}}_{0,t+1} \leftarrow \big(\hat{\mathrm{H}}_{w,0} - \mathbb{E}[\hat{\mathrm{H}}_{w,t+1}]\big) + \big(\hat{\mathrm{H}}_{\Omega,0} - \hat{\mathrm{H}}_{\Omega,t+1}\big)$
  // Keep track of max. IG and respective time step
  $(\Delta\hat{\mathrm{H}}^\star_{0,t},\, t^\star) \leftarrow \max\big\{(\Delta\hat{\mathrm{H}}^\star_{0,t},\, t^\star),\; (\Delta\hat{\mathrm{H}}_{0,t+1},\, t{+}1)\big\}$
end
return $\Delta\hat{\mathrm{H}}^\star_{0,t}$, $t^\star$

dynamics. Therefore we only need to marginalise the weight term over hypothetical measurements, while the density term can be computed directly from the predicted particles. The expected weight entropy at time step t is given by

$$\mathbb{E}\left[\hat{\mathrm{H}}_{w,t}\right] = \sum_j \left(-\sum_i \hat{w}^{i,j}_t \log \hat{w}^{i,j}_t\right) w^j_0, \tag{7}$$

where $\hat{w}^{i,j}_t$ is the weight of particle i at time step t when particle j is assumed to be the true state, i.e.,

$$\hat{w}^{i,j}_t = \frac{w^i_0\; P(z_t|\boldsymbol{x}^j_t)}{\sum_l w^l_0\; P(z_t|\boldsymbol{x}^l_t)}. \tag{8}$$

Consequently, we define the predicted entropy reduction from time step t_1 to t_2 as:

$$\Delta\hat{\mathrm{H}}_{t_1,t_2} = \left(\hat{\mathrm{H}}_{w,t_1} - \mathbb{E}[\hat{\mathrm{H}}_{w,t_2}]\right) + \left(\hat{\mathrm{H}}_{\Omega,t_1} - \hat{\mathrm{H}}_{\Omega,t_2}\right). \tag{9}$$

In our case, we always evaluate changes relative to the initial belief at $t{=}0$, i.e., $\Delta\hat{\mathrm{H}}_{0,t}$.

C. Belief-Dependent Candidate Sampling

To efficiently solve the information-gathering problem in Eq. (1), it is important to generate candidate trajectories ($\boldsymbol{u}^l_{0:T-1}$) that are likely to result in contact with the object. We adopt the spline-based trajectory representation from [26] (cf. Appendix C This representation ensures that the trajectory is smooth and kinodynamically feasible. Generating trajectories with this representation reduces to sampling the control points of the spline, which we refer to as *via-points*. There are multiple ways to design the sampling strategy for the via-points based on the current belief state, but our information-gathering controller is more effective if the sampling strategy results in a diverse set of candidate trajectories that are likely to make contact with the object.

In our experiments, we show an example of how this can be achieved.

V. Particle Filter for Continuous Object Pose Estimation through Contact

The effectiveness of the information-gathering controller depends on having a reliable state estimator, particularly given the discontinuous and sparse nature of contact measurements. To estimate the belief state in the control problem of Eq.(1), we use a particle filter with N_p particles $\{\boldsymbol{x}^i, w^i\}_{i=1}^{N_p}$, where each $\boldsymbol{x}^i$ is a state sample and $w^i \in [0,1]$ is its weight, with $\sum_i w^i = 1$. The belief is updated by propagating particles through a dynamics model and re-weighting them based on measurement likelihoods, following standard particle filtering [27]. Assuming known object properties, we simulate contact dynamics using a physics engine such as MuJoCo [28]. These dynamics are deterministic given fixed initial conditions. Unlike prior work [9], [29], we do not assume the object remains static during contact.

Smooth Measurement Model from Binary Contact: The measurement model updates the belief state after applying control $\boldsymbol{u}_t$ and observing measurement $\boldsymbol{z}_t$. We infer binary contact (contact or no-contact) from proprioceptive torque signals on the robot joints D). However, modeling contact as a strict binary process, as in [3], is limiting as it does not account for the fact that the robot might be close to the object without actually making contact. In addition, unlike prior work [10], [29], we do not assume that contact measurements are perfectly discriminative. To smooth this discontinuous signal, we introduce a proximity-based measurement model where the probability of contact decays exponentially with the robot-object distance. Specifically, the likelihood of contact given particle pose $\boldsymbol{x}_k^i$ is:

$$P(z_t = 1|\boldsymbol{x}_t^i) = \alpha_{\text{fp}} + (\alpha_{\text{tp}} - \alpha_{\text{fp}}) \exp(-\gamma\, d(\boldsymbol{q}_t^r, \boldsymbol{q}_{i,t}^o)), \quad (10)$$

where α_{tp} and α_{fp} are the true/false positive contact rates, γ is a decay rate, and $d(\cdot)$ is the minimum Euclidean distance between robot and object configurations (not limited to the end-effector, unlike [30]). The no-contact likelihood is $1 - P(z_t = 1|\boldsymbol{x}_t^i)$. This smoothing improves belief tracking and reduces particle starvation, but adds computational overhead due to distance evaluations. As this reduces but not eliminates the likelihood of particle starvation, we also introduce a proximity-aware resampling strategy that replaces low-weight particles with samples better matching the latest observation, while using standard resampling for no-contact cases (cf. Appendix B).

VI. Experiments

In order to evaluate all core components of our proposed framework, we conduct two sets of experiments. First, we demonstrate the proposed framework in a robot setup for object localisation and subsequent grasping through sparse contact measurements in Sec. VI-A. Second, we conduct an ablation study to evaluate the impact of the proposed smooth measurement model and resampling strategy on the belief tracking performance in Sec. VI-B.

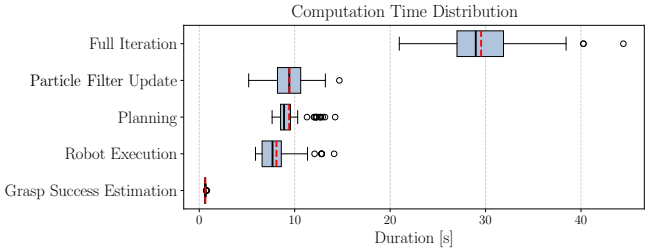

Fig. 3. Computation times for one iteration and different sub-steps of the algorithmic framework across experiments. Boxplots for sub-steps are sorted from longest average duration to shortest from top to bottom. Dashed red and solid black lines correspond to mean and median, respectively.

A. Robot Experiments

a) Setup: We use a Franka Emika robot manipulator to localise and grasp a box in its workspace (see Fig. 1). The binary contact measurements are computed from the measured external torques acting on the robot joints (see D). We show an overview of the different steps in the proposed system consisting of the information-gathering controller and the state estimation in Fig. 2. Given only an initial set of particles sampled from a bi-modal Gaussian mixture distribution, we repeatedly sample candidate trajectories, estimate their information gain, execute the highest information gain trajectory, update the particle filter belief and finally estimate the probability of successfully grasping the box given the updated belief (cf. D). If the estimated probability of success is above a given threshold, the robot attempts an open-loop grasp of the object. In order to ensure that we generate candidate trajectories that are likely to result in contact events, we sample the via-points such that the corresponding trajectories poke into the object belief. More precisely, we generate the final via-points by importance sampling particles from the current belief. Then, for each particle, we uniformly sample a point within the object's volume and compute the inverse kinematics solution such that the robot's end-effector reaches the sampled point. We constrain our motions to poking due to the unreliability of the external torque measurements, which makes it difficult to distinguish between contact and no-contact when the robot pushes an object over longer distances. For each experiment, we report the grasp success rate (with 95% Wilson confidence intervals) and the mean number of iterations required to achieve a successful grasp. The number of iterations is limited to 15. We run each experiment 30 times with random initial box poses. More implementation details for the robot experiments can be found in Appendix D.

b) Baselines: We compare our information-gathering controller against two baselines: *i) Uninformed baseline:* the framework described above but without an information-gain metric to choose trajectories; and *ii) Maximum contact baseline:* the framework described above, but that chooses trajectories that maximise the number of particles contacted in the belief state.

c) Results: The results, summarised in Table I, show that the proposed method significantly outperforms both baselines in terms of success rate and efficiency. A compilation of all experiments, showcasing our approach alongside

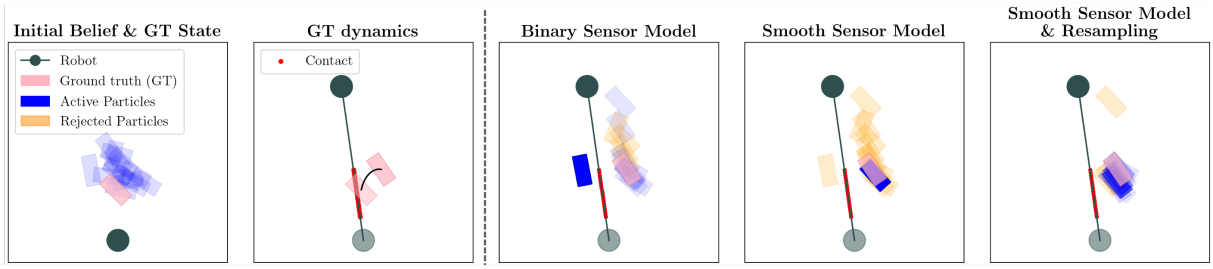

Fig. 4. *Ablation Study of Sensor Models and Resampling Strategy.* We use the following setup to compare three particle-filter variants. A circular robot (dark green) pushes a rectangular object (pink) along a fixed straight-line trajectory (light green) in a planar *quasi-static* setting. The left-most panel shows the initial particle distribution and ground-truth (GT) state; the next panel shows the GT rollout with contact events in red. To the right of the dashed line, we compare the binary sensor model [3], our smooth sensor model, and our smooth model in combination with contact-aware resampling. The smooth model with resampling best preserves particle diversity and while still accurately tracking the belief.

TABLE I

ROBOT EXPERIMENTS: PROPOSED APPROACH COMPARED TO BASELINES

Models	Success rate ↑	Avg. num. iterations (± std.) ↓
Proposed	**0.8** [0.63, 0.90]	9.37 ±3.47
Uninformed	0.23 [0.12, 0.41]	12.53 ±3.68
Max. Contact	0.33 [0.19, 0.51]	13.77 ±2.76

TABLE II

ABLATION RESULTS OF PARTICLE FILTER

Method	MSE ↓		Active Particles ↑ (%)
	Position	Orientation	
Binary	0.023 ± 0.01	0.208 ± 0.09	0.038 ± 0.02
Smooth	0.019 ± 0.01	0.185 ± 0.10	0.077 ± 0.03
Smooth & Resampled	**0.018** ± 0.01	**0.164** ± 0.17	**0.511** ± 0.14

the baselines, is provided in the accompanying s[1]. Notably, for the given object, the margin of error in the object localisation, allowing for a successful grasp, is very small, as the end-effector width is slightly larger than the object width. We also report the computation times spent on different parts in the algorithmic pipeline across iterations of the algorithm and experiment runs for our approach in Fig. 3.

B. Ablation Studies

We compare three particle-filter variants: (i) the binary sensor model of [3], (ii) our smooth sensor model, and (iii) the smooth model with resampling. We evaluate their performance in a planar, quasi-static simulation with a circular robot and a rectangular object with each run using 100 particles and a straight-line robot trajectory (see the qualitative example in Fig. 4). We report the *Mean Squared Error (MSE)* ± std over 1,000 runs in position and orientation computed over the top-10 weighted particles, and the final *share of active particles (AP)*, defined as the fraction of particles with weights $> 10^{-4}$ (Table II). Relative to the binary baseline, the smooth model yields better position estimates, and adding resampling further improves position accuracy. While resampling slightly worsens orientation error, it substantially increases particle diversity, consistent with the qualitative example in Fig. 4. The share of active particles is highest with resampling. Note that we do not resample at the final step to obtain a realistic estimate of particle diversity; immediately after resampling, the share of active particles would trivially be one.

[1]Note that the AR-tag in the videos was only used to add the initial and final ground truth object position to the Mujoco renderings in the videos. Video available at: https://youtu.be/RJhPJi790BM

VII. LIMITATIONS

We acknowledge several limitations and outline promising future directions. First, the applicability of our method to replanning is limited by the computational cost of rolling out candidate control trajectories in a physics engine. Replacing this with a learned contact dynamics model from real-world data could reduce the current computational complexity, while also eliminating the need for known object properties (e.g., mass, friction), and better capturing the inherent stochasticity of contact interactions. Second, our current setup assumes an uncluttered scene and uses sparse binary contact signals. Richer tactile sensing would enable detection of lighter or more diverse objects. Future work will consider more complex, cluttered environments and varied object geometries. While our approach is not tied to specific geometries, sim-to-real discrepancies may grow with non-rigid or more complex objects. Incorporating residual dynamics learning could enhance robustness to these challenges. Lastly, a compelling extension would be to broaden the information-theoretic framework beyond sensing-driven actions to include actions that actively funnel uncertainty into smaller regions of the state space, as explored in prior work [31]–[33].

VIII. CONCLUSION

We presented a continuous-state, touch-only localisation framework capable of operating under high initial uncertainty. We combine a proximity-aware contact model with contact-aware resampling to maintain informative, non-parametric particle beliefs during physical interaction. A central component of our approach is the use of a k-nearest-neighbour-based entropy estimator that accounts for both observation-driven weight changes and interaction-driven shifts in particle distribution. This enables effective evaluation of candidate actions based on how they reshape the belief through both sensing and dynamics. To the best of our knowledge, this is the first use of knn-based estimators for action selection in active tactile object localisation, where

the belief evolution is fundamentally shaped by contact and physical interaction. We validate our method in simulation and real world, demonstrating reliable localisation and grasping under large, multi-modal uncertainty regions, far beyond those addressed in prior work. Compared to baselines relying on heuristic contact maximisation or weight-only entropy, our approach yields higher success rates and more efficient exploration. These results highlight the importance of explicitly modelling both the sensory and physical structure of contact for information-driven manipulation.

REFERENCES

[1] N. Fazeli, M. Oller, J. Wu, Z. Wu, J. B. Tenenbaum, and A. Rodriguez, "See, feel, act: Hierarchical learning for complex manipulation skills with multisensory fusion," *Science Robotics*, vol. 4, no. 26, 2019.

[2] A. Rodriguez, "The unstable queen: Uncertainty, mechanics, and tactile feedback," *Science Robotics*, vol. 6, no. 54, 2021.

[3] P. Hebert, T. Howard, N. Hudson, J. Ma, and J. W. Burdick, "The next best touch for model-based localization," in *2013 IEEE International Conference on Robotics and Automation*. IEEE, 2013, pp. 99–106.

[4] A. Petrovskaya and O. Khatib, "Global localization of objects via touch," *IEEE Transactions on Robotics*, vol. 27, no. 3, pp. 569–585, 2011.

[5] R. Platt, L. Kaelbling, T. Lozano-Perez, and R. Tedrake, "Efficient planning in non-gaussian belief spaces and its application to robot grasping," in *Robotics Research*. Springer, 2017, pp. 253–269.

[6] M. C. Koval, N. S. Pollard, and S. S. Srinivasa, "Pre-and post-contact policy decomposition for planar contact manipulation under uncertainty," *The International Journal of Robotics Research*, vol. 35, no. 1-3, pp. 244–264, 2016.

[7] S. Javdani, M. Klingensmith, J. A. Bagnell, N. S. Pollard, and S. S. Srinivasa, "Efficient touch based localization through submodularity," in *2013 IEEE International Conference on Robotics and Automation*. IEEE, 2013, pp. 1828–1835.

[8] R. Platt, L. Kaelbling, T. Lozano-Perez, and R. Tedrake, "Simultaneous localization and grasping as a belief space control problem," in *International Symposium on Robotics Research*, vol. 2, 2011.

[9] E. Nikandrova, J. Laaksonen, and V. Kyrki, "Towards informative sensor-based grasp planning," *Robotics and Autonomous Systems*, vol. 62, no. 3, pp. 340–354, 2014.

[10] K. Hsiao, L. P. Kaelbling, and T. Lozano-Pérez, "Task-driven tactile exploration." in *Robotics: science and systems*, vol. 12, 2010.

[11] M. Horowitz and J. Burdick, "Interactive non-prehensile manipulation for grasping via pomdps," in *2013 IEEE International Conference on Robotics and Automation*. IEEE, 2013, pp. 3257–3264.

[12] B. Saund, S. Chen, and R. Simmons, "Touch based localization of parts for high precision manufacturing," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2017, pp. 378–385.

[13] N. Miyazawa, D. Kato, Y. Kobayashi, K. Hara, and D. Usui, "Optimal action selection to estimate the aperture of bag by force-torque sensor," *Advanced Robotics*, vol. 38, no. 2, pp. 95–111, 2024.

[14] J. Fischer and Ö. S. Tas, "Information particle filter tree: An online algorithm for pomdps with belief-based rewards on continuous domains," in *International Conference on Machine Learning*. PMLR, 2020, pp. 3177–3187.

[15] B. W. Silverman, *Density estimation for statistics and data analysis*. Routledge, 2018.

[16] Y. Boers, H. Driessen, A. Bagchi, and P. Mandal, "Particle filter based entropy," in *2010 13th International Conference on Information Fusion*. IEEE, 2010, pp. 1–8.

[17] H. Singh, N. Misra, V. Hnizdo, A. Fedorowicz, and E. Demchuk, "Nearest neighbor estimates of entropy," *American journal of mathematical and management sciences*, vol. 23, no. 3-4, pp. 301–321, 2003.

[18] J. Jiao, W. Gao, and Y. Han, "The nearest neighbor information estimator is adaptively near minimax rate-optimal," *Advances in neural information processing systems*, vol. 31, 2018.

[19] J. Ajgl and M. Šimandl, "Differential entropy estimation by particles," *IFAC Proceedings Volumes*, vol. 44, no. 1, pp. 11 991–11 996, 2011.

[20] T. B. Berrett, R. J. Samworth, and M. Yuan, "Efficient multivariate entropy estimation via k-nearest neighbour distances," 2019.

[21] H. Liu and P. Abbeel, "Behavior from the void: Unsupervised active pre-training," *Advances in Neural Information Processing Systems*, vol. 34, pp. 18 459–18 473, 2021.

[22] Y. Seo, L. Chen, J. Shin, H. Lee, P. Abbeel, and K. Lee, "State entropy maximization with random encoders for efficient exploration," in *International conference on machine learning*. PMLR, 2021, pp. 9443–9454.

[23] M. Mutti, L. Pratissoli, and M. Restelli, "Task-agnostic exploration via policy gradient of a non-parametric state entropy estimate," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 35, 2021, pp. 9028–9036.

[24] J. Beirlant, E. J. Dudewicz, L. Györfi, E. C. Van der Meulen *et al.*, "Nonparametric entropy estimation: An overview," *International Journal of Mathematical and Statistical Sciences*, vol. 6, no. 1, pp. 17–39, 1997.

[25] R. Platt, R. Tedrake, L. P. Kaelbling, and T. Lozano-Perez, "Belief space planning assuming maximum likelihood observations," in *Robotics: Science and Systems*, 2010.

[26] J. Jankowski, L. Brudermüller, N. Hawes, and S. Calinon, "Vp-sto: Via-point-based stochastic trajectory optimization for reactive robot behavior," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 10 125–10 131.

[27] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*, ser. Intelligent Robotics and Autonomous Agents series. MIT Press, 2005.

[28] E. Todorov, T. Erez, and Y. Tassa, "Mujoco: A physics engine for model-based control," in *2012 IEEE/RSJ international conference on intelligent robots and systems*. IEEE, 2012, pp. 5026–5033.

[29] M. C. Koval, N. S. Pollard, and S. S. Srinivasa, "Pose estimation for planar contact manipulation with manifold particle filters," *The International Journal of Robotics Research*, vol. 34, no. 7, pp. 922–945, 2015.

[30] C. Zito, V. Ortenzi, M. Adjigble, M. Kopicki, R. Stolkin, and J. L. Wyatt, "Hypothesis-based belief planning for dexterous grasping," *arXiv preprint arXiv:1903.05517*, 2019.

[31] M. A. Erdmann and M. T. Mason, "An exploration of sensorless manipulation," *IEEE Journal on Robotics and Automation*, vol. 4, no. 4, pp. 369–379, 1988.

[32] M. R. Dogar, K. Hsiao, M. T. Ciocarlie, and S. S. Srinivasa, "Physics-based grasp planning through clutter." in *Robotics: Science and systems*, vol. 8, 2012, pp. 57–64.

[33] J. Jankowski, L. Brudermüller, N. Hawes, and S. Calinon, "Planning for robust open-loop pushing: Exploiting quasi-static belief dynamics and contact-informed optimization," *arXiv preprint arXiv:2404.02795*, 2024.

[34] L. Zhang and J. C. Trinkle, "The application of particle filtering to grasping acquisition with visual occlusion and tactile sensing," in *2012 IEEE International Conference on Robotics and Automation*. IEEE, 2012, pp. 3805–3812.

[35] G. Garofalo, N. Mansfeld, J. Jankowski, and C. Ott, "Sliding mode momentum observers for estimation of external torques and joint acceleration," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 6117–6123.

APPENDIX

A. Differential Entropy Estimate

To estimate the differential entropy of a particle-based belief, we approximate the belief as a mixture of uniform distributions, each centered at a particle $\boldsymbol{x}^i$ with weight w^i and shared support $\Omega \subseteq \mathcal{X}$, a bounded region in state space with hypervolume $V(\Omega) = \int_\Omega 1 d\boldsymbol{x}$:

$$\hat{b}(\boldsymbol{x}) = \sum_i w^i \, \mathcal{U}(\boldsymbol{x} - \boldsymbol{x}^i, \Omega), \tag{11}$$

where $\mathcal{U}(\boldsymbol{x}, \Omega) = \frac{1}{V(\Omega)}$ if $\boldsymbol{x} \in \Omega$, and 0 otherwise. We determine the support Ω as follows: *i)* For each particle i, fit a tight bounding box ${}^{\rho}\Omega_i$ to its ρ-nearest neighbors, *ii)* compute the average bounding box ${}^{\rho}\bar{\Omega}$ by averaging box dimensions, and *iii)* scale it to obtain $\Omega = \frac{{}^{\rho}\bar{\Omega}}{\left(\sqrt[d]{\rho}-1\right)}$, where d is the number of dimensions. This procedure captures local

particle density: higher densities lead to smaller $V(\Omega)$ and lower entropy. The hyperparameter ρ controls the locality of the approximation: small values capture multimodality; large values yield global smoothing. The local uniform mixture distribution does not admit a closed-form evaluation due to overlaps of the uniform components. Instead, we derive an upper bound based on the inequality $\log(p(\boldsymbol{x}))\geq\log(p_i(\boldsymbol{x}))$, where $p_i(\boldsymbol{x})$ is the probability density contribution of the i-th component:

$$\hat{\mathrm{H}}[\boldsymbol{x}] \leq -\sum_i w^i \log w^i + \log V. \tag{12}$$

This bound assumes disjoint uniform components, i.e., $\Omega_i \cap \Omega_j = \emptyset$ for $i \neq j$. Though the bound may be loose when components overlap, it is deterministic, captures trends in the ground-truth entropy, and is sensitive to small changes in particle positions. For completeness, we provide an extended derivation along with a proof that the estimator serves as an upper bound on the differential entropy of the induced uniform mixture distribution in the supplementary material[2].

a) Computational Complexity: The main cost is computing ρ-nearest neighbors for N_p particles, with complexity $\mathcal{O}(N_p D \log(N_p + \rho))$ using a ball tree. This is more efficient than kernel density estimation with complexity $\mathcal{O}(N_p^2 D^3)$ [14].

b) Limitation: The differential entropy may diverge to $-\infty$ if the distribution collapses in any dimension, which corresponds to $V(\Omega) \to 0$. To mitigate this, we compute entropy only over relevant dimensions (e.g., omitting z-dimension if particles lie on a table).

B. Particle Filter: Resampling Strategy

Despite smoothing sparse binary contact signals with proximity estimates, particle starvation remains challenging in contact-rich manipulation [29], [34]. Starvation occurs when no particles lie near the true state, often due to a loss of diversity, which can cause filter divergence. Standard resampling methods (e.g., importance sampling [27]) help but are less effective under contact, since sampling near the contact manifold (zero-distance configurations) is unlikely [29]. Although contact manifolds can be explicitly constructed, doing so is difficult in high-dimensional spaces. We address this with a proximity-based resampling strategy. When contact is observed, we sample M particles near the current robot pose, evaluate them using the smooth measurement model, and replace low-weight particles with better-matching ones. For no-contact measurements, standard importance sampling is used. Resampling is triggered when particle weights drop below a threshold.

C. Spline-based Trajectory Representation

We adopt the spline-based trajectory representation of [26], which maps a phase variable $s\in[0,1]$ and parameters $\boldsymbol{\theta}$ to a reference position $\boldsymbol{u}(s) = \boldsymbol{\Phi}_{\boldsymbol{\theta}}(s)\boldsymbol{\theta} + \boldsymbol{\Phi}_{\boldsymbol{\lambda}}(s)\boldsymbol{\lambda}$, $\boldsymbol{\theta}=\boldsymbol{\theta}_{1:N_{\text{via}}}$ denotes N_{via} via-points such that $\boldsymbol{u}(s={}^{n}/_{N_{\text{via}}})=\boldsymbol{\theta}_n$, and $\boldsymbol{\lambda}=\left[\boldsymbol{u}_0^\top, \dot{\boldsymbol{u}}_0^\top, \dot{\boldsymbol{u}}_{T-1}^\top\right]^\top$ encode the initial set-point $\boldsymbol{u}(0)=\boldsymbol{u}_0$ along with initial and final velocities. In our experiments, velocities at the first and last time steps are set to zero. The matrices $\boldsymbol{\Phi}_{\boldsymbol{\theta}}(s)$ and $\boldsymbol{\Phi}_{\boldsymbol{\lambda}}(s)$ contain the spline basis functions evaluated at s.

D. Implementation Details

a) Contact Measurement Signal: We infer contact from the robot's torque sensors. Yet, the Franka robot's torque sensors are too noisy to use directly, so we convert the signals into a binary contact indicator. Following [35], we filter out gravity, friction, and actuation effects, compute the norm of the first five joint torques (excluding the last two due to noise), and threshold the result to obtain a binary contact signal $z_t \in \{0,1\}$ for the particle filter.

b) Low-level Control: We ensure moderate contact forces throughout the robot operation via an impedance controller on the low-level. While the control gains are higher during the localisation phase, we reduce the gains in the moment of grasping to allow for a more robust grasp. In parallel, we keep track of the contact forces acting on the robot's end effector by projecting the torques measured in the robot's joints onto the end effector frame. If the contact forces exceed a given threshold, the robot stops the current action early and transitions to the particle filter update phase.

c) Initial Belief: The initial belief about the object pose is represented as a set of particles, where each particle is a 6D pose of the object. The particles are sampled from a Gaussian mixture model with two components. For each component, we use a standard deviation of σ_{pos}=0.06 [m] for the box position and a standard deviation of σ_{ori}=0.5 [rad] for the yaw orientation. All weights are set to ${}^{1}/_{N_p}$, where N_p=100 is the number of particles.

d) Estimation of the Probability of Grasp Success: The grasp success probability is estimated using the maximum likelihood object pose after localization. A grasp primitive is planned for this pose and evaluated over the full weighted particle belief. Success probability is the weighted fraction of particles in which the grasp succeeds (i.e., the object's z-position is above the ground).

e) Hyperparameters: In all experiments, the threshold on the probability of grasp success is set to 0.8. The number of candidates sampled in each iteration of the predictive sampling-based planner is N_{samples}=20. For our hardware experiments, this corresponded to the maximum number of threads that could be run in parallel. The support Ω for the entropy approximation in Eq. (5) was computed using $\rho = 5$ neighbours. We empirically found this value offers a favourable balance between local and global density. The parameters for the proximity-based measurement model are $\alpha_{tp} = 0.5$, $\alpha_{fp} = 0.1$ and $\gamma = 1000$. Moreover, the maximum number of localising actions is 15. Afterwards, the robot stops the localisation phase and transitions to the grasping phase, regardless of the probability of grasp success.

[2] https://github.com/brudermueller/icra2026_supplementary_material