Haptic Sequential Monte Carlo Localization for Quadrupedal Locomotion in Vision-Denied Scenarios

Russell Buchanan, Marco Camurri, Maurice Fallon

Abstract— Continuous robot operation in extreme scenarios such as underground mines or sewers is difficult because exteroceptive sensors may fail due to fog, darkness, dirt or malfunction. So as to enable autonomous navigation in these kinds of situations, we have developed a type of proprioceptive localization which exploits the foot contacts made by a quadruped robot to localize against a prior map of an environment, without the help of any camera or LIDAR sensor. The proposed method enables the robot to accurately re-localize itself after making a sequence of contact events over a terrain feature. The method is based on Sequential Monte Carlo and can support both 2.5D and 3D prior map representations. We have tested the approach online and onboard the ANYmal quadruped robot in two different scenarios: the traversal of a custom built wooden terrain course and a wall probing and following task. In both scenarios, the robot is able to effectively achieve a localization match and to execute a desired pre-planned path. The method keeps the localization error down to 10 cm on feature rich terrain by only using its feet, kinematic and inertial sensing.

Fig. 1. ANYmal using its foot to measure the vibrations of a piece of mining equipment. Tasks such as this might require localization even when vision cannot be used. By exploiting prior knowledge of the 3D environment, the robot could use contact events to localize itself. Source: ETH Zurich.

I. INTRODUCTION

Perception is an essential prerequisite for legged robot navigation on challenging terrain. Several works have demonstrated the potential of such machines when they are endowed with effective state estimation [1], [2] and visual perception capabilities [3], [4], [5]. However, vision-denied areas such as dusty underground mines or foggy sewers are places where legged robots are likely to be deployed [6] (see Fig. 1). Additionally, unrecoverable sensor failures are to be expected when a robot is deployed for long-term autonomous missions. Performance should degrade gracefully in such conditions by taking advantage of all proprioceptive sensor measurements and prior information available to the robot.

In this work, we explore haptics as the main source of robot localization information on non-degenerate (i.e., non flat) ground. We present an algorithm for 6 Degrees of Freedom (DoF) localization of a quadrupedal robot based on Sequential Monte Carlo methods.

Given the pose (estimated by a proprioceptive state estimator [1]), the joint kinematics and the foot contact states, the robot is able to identify its most likely location within a prior map and its past trajectory. This is done using only proprioceptive sensors, such as Inertial Measurement Unit (IMU), encoders and torque sensors. To this end, the robot uses its legs to probe the terrain and uses the contact events to estimate its trajectory relative to the terrain map. Unlike previous works, our method can localize in 6-DoF which

The authors are with the Oxford Robotics Institute at the University of Oxford, United Kingdom. {russell, mcamurri, mfallon}@robots.ox.ac.uk

we demonstrate with an experiment using a foot to probe in 3-dimensions. This problem is analogous to a blindfolded person in a room they know well, using their hands to localize within the room through touch.

We tested our algorithm using an ANYbotics ANYmal quadruped robot in two different scenarios. First, we constructed a diverse terrain course with features such as ramps, chevrons and steps. The robot traversed several loops of this course, comparing contact points to the prior map (represented as a 2.5D digital elevation map). In the second scenario, the robot probed against two perpendicular walls to improve its estimated pose, with the map represented as a 3D point cloud.

The remainder of this document is structured as follows: Section II summarizes the recent developments in the field of legged haptic localization and the related problem of in-hand tactile localization; Section III defines the mathematical background of the legged haptic localization problem; Section IV describes our proposed haptic localization algorithm; Section VI presents the experimental results collected using the ANYmal robot; Section VII provides an interpretation of the results and discusses the limitations of the approach; finally, Section VIII concludes the paper with final remarks.

II. RELATED WORKS

Pioneering work on how to use robot's legs not just for locomotion but also to infer terrain information such as friction, stiffness and geometry has been presented by Krotkov [7]. This is a useful capability when legged robots have to operate in vision-denied areas. The idea was recently

revisited by Bednarek et al. [8] to classify different terrain types in order to improve locomotion parameters and by Kolvenbach et al. [6] to inspect concrete in sewage systems by vibration analysis of the sewer floor.

While these works were more focused on using haptic information to improve locomotion or to monitor the terrain, little research has been conducted on how to use haptic information to localize the robot in the environment. However, a similar problem has been explored in grasping: tactile localization.

Tactile localization involves the estimation of the 6-DoF pose of an object (of known shape) in the robot's base frame by means of kinematics of the robot's fingers and its tactile sensors. The pose of the object (in robot base coordinates) is inferred by matching the set of all positions of the sensed contacts (also in base coordinates) with the 3D model of the object being grasped.

Intuitively, this problem is analogous to legged robot haptic localization, if we imagine the entire robot as a hand trying to "grasp" the ground. The objective of legged haptic localization is estimating the pose of the robot relative to a fixed object (in this case the terrain) by means of its "fingers" (in this case the robot's feet), instead of the pose of the object being grasped relative to the robot.

A. Sequential Monte Carlo Methods

Tactile localization has typically been addressed using Sequential Monte Carlo (SMC) methods, a subfamily of which is called *particle filters* [9]. Since the object can have any shape, the probability distribution of its pose given tactile measurements can be multimodal. This excludes all Gaussian estimators such as Kalman filters. In contrast, SMC methods can maintain a discrete approximation of an arbitrary distribution by generating many hypotheses (also called particles) from a known *proposal distribution*. If the number of particles is large enough and the proposal distribution covers the state space well (i.e., it captures the areas of high density of the true distribution), the target distribution is well approximated by a list of state particles and their associated importance weights.

SMC is sensitive to the dimension of the state space, which should be low enough to avoid combinatorial explosions or particle depletion. State-of-the-art methods aim to reduce this dimensionality and also to sample the state space in an efficient manner. In the SLAM community, this has been typically achieved by careful design of the proposal distribution and adaptive importance resampling to avoid particle depletion [10].

B. Tactile Localization in Manipulation

Vezzani et al. [11] proposed an algorithm for tactile localization using the Unscented Particle Filter (UPF) [12] and tested it on the iCub robot (equipped with contact sensors at the fingertips) to localize four different objects in the robot's reference frame. The algorithm is recursive and can process data in real-time as new measurements become available. The object and the robot's base are assumed to be static, therefore the pose to be estimated is a fixed parameter. This assumption allows the use of a window of past measurements to better address the sparsity of the measurements, which consist of a series of single finger touches. For legged haptic localization, the assumption of both a static robot and terrain does not hold in general and more general methods are required.

Chalon et al. [13] proposed a particle filtering method for online in-hand object localization that tracks the pose of an object while it is being manipulated by a fixed base arm. The pose is subsequently used to improve the performance of pick and place tasks (e.g., by placing a bottle in a upright position). The particles representing the object's pose are initialized by a Gaussian distribution around the true initial pose of the object, acquired by a vision system (used only at start). The particle weights are taken from the measurements according to a function which penalizes both finger/object co-penetration and the distance between the object and the fingertip that detected the contact. The best estimate is the one with highest weight.

This work, as with many others in the field of grasping, assumes the hand can envelop the object, since they have comparable sizes. This is not the case with legged robots touching the ground, which makes the problem generally harder and involves more uncertainty. For example, a hand touching a very large object has fewer clues as to the object's pose, because the contact points would not be equally spread across the object but concentrated in a small area of it.

C. Haptic Localization of Walking Robots

To the best our knowledge, Chitta et al. [14] is the only significant prior example of legged haptic localization to date. In their work, they presented a proprioceptive localization algorithm based on a particle filter for the LittleDog, a small electric quadruped. The robot was commanded to perform a statically stable gait over a known irregular terrain course, using a motion capture system to feed the controller. While walking, the algorithm approximated the probability distribution of the base state with a set of particles. The state was reduced from six to three dimensions: linear position on the xy-plane and yaw. Each particle was sampled from the uncertainty of the odometry, while the weight of a particle was determined by the L2 norm of the height error between the map and the contact location of the feet. The algorithm was run offline on eight logged trials of $50\,\mathrm{s}$ each.

In our proposed work, we improve upon [14] both theoretically and experimentally. First, we estimate the full past trajectory at every step instead of the most recent pose. This allows us to estimate poses which are globally consistent over the entire trajectory of the map. Second, we perform localization in the full 6-DoF of the robot, instead of limiting the estimation to the x, y and yaw dimensions. Third, we experimentally demonstrate the effectiveness of the algorithm by running it online on our quadruped robot. This is the first time such an algorithm has been successfully executed online on real hardware and used in a closed loop navigation system to successfully execute a planned path.

Finally, we demonstrate the algorithm in three dimensions in online experiments.

III. Problem Statement

Let $\boldsymbol{x}_k = [\mathbf{p}_k, \boldsymbol{\theta}_k]$ be a robot's pose at time k, composed of the base position in the world frame $\mathbf{p} \in \mathbb{R}^3$ and orientation $\boldsymbol{\theta} \in SO(3)$. With a slight abuse of notation, we will use the same symbol for its homogeneous matrix form $\boldsymbol{x}_k \in SE(3)$. We assume that for each timestep k, an estimate of the robot pose $\tilde{\boldsymbol{x}}_k$ and its covariance $\Sigma_k \in \mathbb{R}^{6\times6}$ is available from an odometric estimator[1]. We also assume that the location of the robot's end effectors in the base frame $(\mathbf{d}_{\mathrm{LF}}, \mathbf{d}_{\mathrm{RF}}, \mathbf{d}_{\mathrm{LH}}, \mathbf{d}_{\mathrm{RH}}) \in \mathbb{R}^{3\times4}$ are known from forward kinematics and their binary contact states $C \in \mathbb{B}^4$ from inverse dynamics.

For simplicity, we neglect any uncertainties due to inaccuracy in joint encoder readings or limb flexibility. Therefore, the propagation of the uncertainty from the base to the end effectors is straightforward to compute. For brevity, the union of the aforementioned states (pose and contacts) at time k will be referred as the *quadruped state* $\mathcal{Q}_k = \{\tilde{\mathbf{p}}_k, \tilde{\boldsymbol{\theta}}_k, \Sigma_k, \mathbf{d}_{\mathrm{LF}}, \mathbf{d}_{\mathrm{RF}}, \mathbf{d}_{\mathrm{LH}}, \mathbf{d}_{\mathrm{RH}}, C_k\}$.

Our approach can localize against both 2.5D terrain elevation maps and 3D point clouds. These prior maps are denoted with $\mathcal{M}$.

Our goal is to use a sequence of quadruped states, and their corresponding uncertainties Σ_k, to estimate the most likely sequence of robot states up to time k as:

$$\mathcal{X}^*_{0:k} = [\boldsymbol{x}^*_0, \boldsymbol{x}^*_1, \ldots, \boldsymbol{x}^*_k] \tag{1}$$

such that the likelihood of the contact points to be on the map is maximized.

IV. Proposed Method

To perform localization, we sample a predefined number of particles at regular intervals from the pose distribution provided by the odometry, as described in Section IV-B, and we compute the likelihoods of the measurements by comparing each particle to the prior map so as to update the weight of the particle filter (see Section IV-C). For convenience, we give a brief summary of particle filter theory in Section IV-A.

A. Particle Filter Theory

In a particle filtering framework, the objective is to approximate the posterior distribution of the state $\boldsymbol{x}_k$ given a history of measurements $\boldsymbol{z}_0, \ldots, \boldsymbol{z}_k = \boldsymbol{z}_{0:k}$ as follows:

$$p(\boldsymbol{x}_k|\boldsymbol{z}_{0:k}) = \sum_i w^i_{k-1} p(\boldsymbol{z}_k|\boldsymbol{x}_k) p(\boldsymbol{x}_k|\boldsymbol{x}^i_{k-1}) \tag{2}$$

where w^i is the *importance weight* of the i-th particle; $p(\boldsymbol{z}_k|\boldsymbol{x}_k)$ is the measurement likelihood function and $p(\boldsymbol{x}_k|\boldsymbol{x}^i_{k-1})$ is the dynamic model for the i-th particle state.

Since $p(\boldsymbol{x}_k|\boldsymbol{z}_{0:k})$ is normally unknown, the state $\boldsymbol{x}_k$ is typically sampled from $p(\boldsymbol{x}_k|\boldsymbol{x}^i_{k-1})$, yielding:

$$p(\boldsymbol{x}_k|\boldsymbol{z}_{0:k}) = \sum_i w^i_{k-1} p(\boldsymbol{z}_k|\boldsymbol{x}^i_{k-1}) \delta(\boldsymbol{x}_k - \boldsymbol{x}^i_k) \tag{3}$$

where $\delta(\cdot)$ is the Dirac delta function.

B. Locomotion Control and Sampling Strategy

In the considered situation, the robot would be locomoting blindly. As a result, it is likely to do so with a statically stable gait while using the terrain to achieve a position fix.

Statically stable gaits alternate between three and four points of support by moving a single foot. When a new four-point support phase occurs at time k, the estimated robot state $\mathcal{Q}_k$ from the onboard state estimator is collected. Similarly to [14], particles are sampled from the following proposal distribution:

$$p(\boldsymbol{x}_k|\boldsymbol{x}^i_{k-1}) = \mathcal{N}(\boldsymbol{x}_k, \Delta\tilde{\boldsymbol{x}}_k \boldsymbol{x}^i_{k-1}, \Sigma_k) \tag{4}$$

where $\Delta\tilde{\boldsymbol{x}}_k = \tilde{\boldsymbol{x}}^{-1}_{k-1} \tilde{\boldsymbol{x}}_k$ is the relative pose between the estimated states at times $k-1$ and k.

The new samples $\boldsymbol{x}^i_k$ are drawn from a Gaussian centered at the pose of the previous iteration composed of the relative pose from the odometry, with the covariance of the odometry at time k. Since the estimator [1] fuses both IMU and leg odometry, the covariance Σ_k has very low uncertainty on roll and pitch because these two quantities are observable due to gravity. This would result in all particles having virtually the same roll and pitch which permits reduction of the sample space to the four other dimensions. This allows us to localize in the full 6-DoF space without incurring problems due to high dimensionality.

C. 2.5-Dimensional Measurement Likelihood Model

To account for uncertainties in the prior map, we model the measurement likelihood as a univariate Gaussian centered at the local elevation of each cell, with a manually selected variance σ_z (1 cm in our case). In contrast with [14], we evaluate each contact point individually. Given a particle state $\boldsymbol{x}^i_k$, the estimated position of a contact in world coordinates for foot f is defined as the concatenation of the estimated robot base pose and the location of the end effector in base coordinates (Fig. 2 left):

$$\mathbf{d}^i_f = (d^i_{xf}, d^i_{yf}, d^i_{zf},) = \boldsymbol{x}^i_k \mathbf{d}_f \tag{5}$$

Thus, the measurements and their relative likelihood functions for the i-th particle and a specific foot f are:

$$z_k = d^i_{zf} - \mathcal{M}(d^i_{xf}, d^i_{yf}) \tag{6}$$

$$p(z_k|\boldsymbol{x}^i_k) = \mathcal{N}(z_k, 0, \sigma_z) \tag{7}$$

where: d^i_{zf} is the vertical component of the estimated contact point location in world coordinates of foot f, according to the i-th particle; $\mathcal{M}(d^i_{x,f}, d^i_{y,f})$ is the corresponding map elevation at the xy coordinates of $\mathbf{d}^i_f$.

---

[1]uncertainties for the rotation manifold are maintained in the Lie tangent space, as in [15]

$$d_{z,f} - \mathcal{M}(d_{x,f}, d_{y,f}) \qquad \|\mathbf{d}_f - \mathrm{kd}(\mathcal{M}, \mathbf{d}_f)\|$$
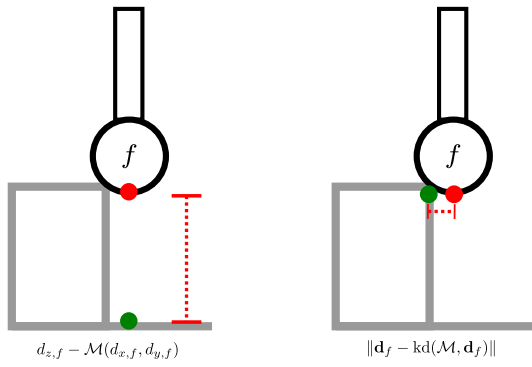
Fig. 2. Comparison between contact measurements for 2.5D (left) and 3D (right) map representations. The red dots indicate the contact point as sensed by the robot, while the green dots indicate the corresponding location returned by the map. The red line shows the magnitude of the measurement. Note how the same contact interaction generates a much lower likelihood in the 2.5D map than the 3D map, which is more representative. This is a practical limitation of using a 2.5D map.

D. 3-Dimensional Measurement Likelihood Model

Our method can incorporate contact events from 3D probing. This is useful for areas where the floor does not provide enough information to localize. Instead, the robot can probe walls and 3D objects. In this case, we instead represent the prior map $\mathcal{M} \in \mathbb{R}^{3 \times N}$ by a 3D point cloud with N points. The likelihood of a particular contact point is computed using the Euclidean distance between the foot and the nearest point in the map. This likelihood is again modeled as a zero-mean Gaussian evaluated at the Euclidean distance between the estimated contact point $\mathbf{d}_f^i$ and its nearest neighbor on the map, with variance σ_z:

$$z_k = \|\mathbf{d}_f^i - \mathrm{kd}(\mathcal{M}, \mathbf{d}_f^i)\| \tag{8}$$

$$p(z_k|\boldsymbol{x}_k^i) = \mathcal{N}(z_k, 0, \sigma_z) \tag{9}$$

where $\mathrm{kd}(\mathcal{M}, \mathbf{d}_f^i)$ is the function that returns the nearest neighbor of $\mathbf{d}_f^i$ on the map $\mathcal{M}$, computed from its k-d tree.

V. IMPLEMENTATION

Pseudocode for our method is listed in Algorithm 1. We update the filter when the robot enters a new four-support configuration. At this time k, the estimated robot pose $\tilde{\boldsymbol{x}}_k$ is collected and used to compute the relative motion $\Delta\tilde{\boldsymbol{x}}_{k-1:k}$. This is used to propagate forward the state of each particle $\boldsymbol{x}_k^i$ and draw a sample from the distribution centered in $\Delta\tilde{\boldsymbol{x}}_k\boldsymbol{x}_{k-1}^i$ with covariance $\Sigma_k = (\sigma_{x,k}, \sigma_{y,k}, \sigma_{z,k})$.

The weight of a particle w^i is then updated by multiplying it by the likelihood that each foot is in contact with the map. In our implementation, we modify the likelihood functions from Equations 7, 9 as:

$$p(z_k|\boldsymbol{x}_k^i) = \min(\rho, \mathcal{N}(z_k, 0, \sigma_z)) \tag{10}$$

where ρ is a minimum weight threshold, so that outlier contact measurements do not immediately lead to degeneracy. Resampling is triggered when the variance of the weights rises above a certain threshold. This was manually tuned once and used for all experiments. Resampling is

$\boldsymbol{x}_0^i \sim \mathcal{N}(\boldsymbol{x}_0, \tilde{\boldsymbol{x}}_0, \Sigma_0) \quad \forall i \in N$
foreach *four-support phase k* **do**
  $\Delta\tilde{\boldsymbol{x}}_k \leftarrow \tilde{\boldsymbol{x}}_{k-1}^{-1}\tilde{\boldsymbol{x}}_k$
  foreach *particle* $i \in N$ **do**
    $\boldsymbol{x}_k^i \sim \mathcal{N}(\boldsymbol{x}_k, \Delta\tilde{\boldsymbol{x}}_k\boldsymbol{x}_{k-1}^i, \Sigma_k)$
    $w_k^i \leftarrow w_{k-1}^i$
    foreach *foot f* **do**
      if *2.5D* **then**
        $z_k \leftarrow d_{zf}^i$
      else
        $z_k \leftarrow \|\mathrm{kd}(\mathcal{M}, \mathbf{d}_f^i) - \mathbf{d}_f^i\|$
      end
      $w_k^i \leftarrow w_k^i p(z_k|\boldsymbol{x}_k^i)$
    end
    $x_k^* \leftarrow \mathrm{WeightedMean}(x_k^0 \ldots x_k^N, w_k^0, \ldots w_k^N)$
    $\mathcal{X}_k^* \leftarrow [\boldsymbol{x}_0^j, \ldots, \boldsymbol{x}_k^j]$
    if $variance(w_k^i) > threshold$ **then**
      $resample(x_k^i)$
    end
  end
end

Algorithm 1: Haptic Sequential Monte Carlo Localization

necessary otherwise the particle set would disperse across the state space and many particle would have low weight. By triggering this process when the variance of the weights increases, the particles can first track the dominant modes of of the underlying distribution.

A. Particle Statistics

The optimal value for the k-th iteration, $\boldsymbol{x}_k^*$, is computed from the weighted average of all the particles, as in [14]. However, this method does not take into account the multimodality of the particle distribution. There can be several hypotheses resulting in multiple clusters of particles whose weighted average is not a reliable estimate of the robot's location. To account for this, we check the variance of the particle positions in the x and y axes. If they are low (i.e., $\ll \sigma_{x,k}, \sigma_{y,k}$), we assume a good estimate and update the robot's full pose. However, if they are high, we update only the z component of the robot's location (the main drifting dimension). The other dimensions drift at the same rate as the state estimator, but the upward drift is corrected since the robot is always in contact with the ground.

VI. EXPERIMENTAL RESULTS

Our localization algorithm has been tested in two different scenarios: a walking task crossing a custom built terrain course and a task involving probing and following of a wall. The prior maps were captured with a Leica BLK-360 laser scanner, which provides registered point clouds with sub-centimeter accuracy.

Before testing on real hardware, the 3D model of the terrain course was imported as a simulated environment (in Gazebo) allowing us to validate the localization algorithm
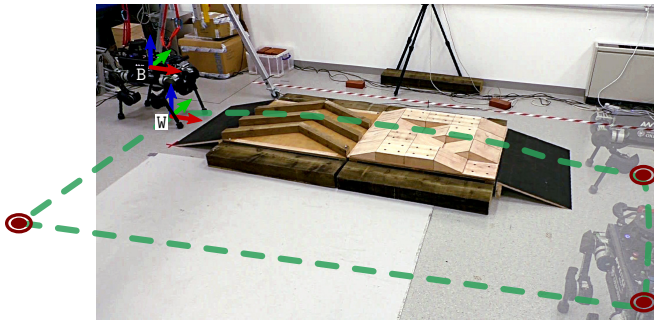
Fig. 3. ANYmal haptic localization experiments. The robot traverses the terrain, turns 90 deg right and comes back to the initial position passing trough the flat area. The goals given to the planner are marked by the dark red disks, while the planned route is a dashed green line (one goal is out of the camera field of view). The world frame W is fixed to the ground, while the base frame B is rigidly attached to the robot's chassis. The mutual pose between the robot and the terrain course is bootstrapped with the Vicon.
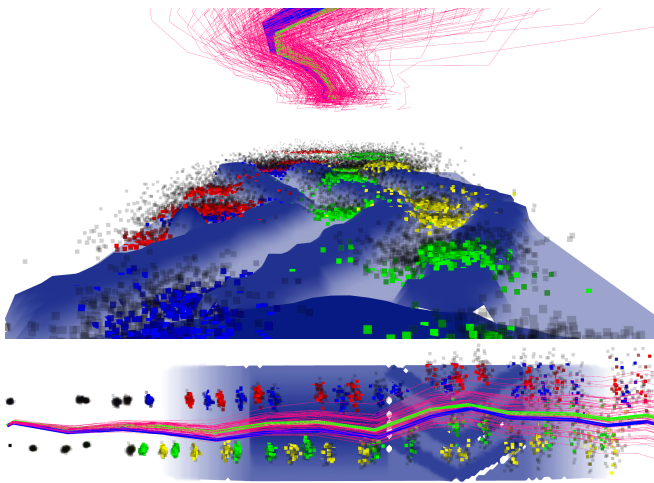

Fig. 4. *Top:* Perspective view of a simulated traversal of similar terrain to Fig. 3. At the top is the 1000 sampled trajectories shown by thin magenta lines, while the ground truth and estimated trajectory are shown with a green and blue line, respectively. *Bottom:* Top down view of the same scene. The top perspective is from the center the terrain facing right. Each squared dot represents one of 4000 contacts, colored by leg type (LF: red, RF: green, LH: blue, RH: yellow). The tone and alpha of each square is associated to the likelihood: darker and more transparent relates to a lower likelihood. The grid map is colored by height, from white to dark blue.

with different values of noise and particle set dimensions. Fig. 4 shows an example with 1000 particles depicted as candidate trajectories (in magenta) with the associated contacts colored by their likelihood (brighter/more solid color = higher likelihood). The most likely trajectory is highlighted in blue and is the closest to the ground truth (in green).

In real world experiments, the ground truth robot poses were collected at $100\,\mathrm{Hz}$ with millimetric accuracy using a Vicon motion capture system. At start of the experiment, the relative position of the robot within the map was measured using Vicon and used for initialization only. Thereafter, the pose of the robot was estimated using the particle filter and the Vicon was used only to compute error statistics. To account for initial errors, particles at the start were sampled from a Gaussian centered at the initial robot pose with a covariance of $20\,\mathrm{cm}$.

Exp.	Dist. [m]	Time [s]	Loops	ATE TSIF	ATE HL
1	66.42	525	2	0.63	**0.13**
2	145.31	1097	4	2.57	**0.40**
3	55.67	557	2	0.52	**0.19**
4	68.71	604	2	0.65	**0.32**
5	172.65	1606	4	2.00	**0.61**

TABLE I

ESTIMATION PERFORMANCE ON THE TERRAIN COURSE DATASETS. ATE = ABSOLUTE TRANSLATION ERROR; TSIF = TWO-STATE IMPLICIT FILTER [1] ; HL = HAPTIC LOCALIZATION.

A. Terrain Course Experiments

In the first scenario, the robot was commanded to navigate between four walking goals at the corners of a rectangle. One of the edges required crossing a $4.2\,\mathrm{m}$ terrain course composed of a $12°$ ascending ramp, a $13\,\mathrm{cm}$ high chevron pattern, an asymmetric composition of uneven square blocks and a $12°$ descending ramp (Fig. 3). After crossing the wooden course, the robot returned to the starting position across a portion of flat ground, which tests how the system behaves in feature-deprived conditions.

While impressive, blind reactive locomotion has been developed by a number of groups including [16], [4], unfortunately our haptic controller was not sufficiently reliable to cross the terrain course so we resorted to use of the statically stable gait from [3] which used a depth camera to aid footstep planning. Using our approach, localization was performed without access to any camera information.

To demonstrate repeatability, we have performed five experiments of this type, for a total distance traveled of more than $0.5\,\mathrm{km}$ and $1\,\mathrm{h}13\,\mathrm{min}$ duration. A summary of the experiments is presented in Table I, where the haptic localization shows an overall improvement between $50\,\%$ to $85\,\%$ in the Absolute Translation Error (ATE). With the haptic localization, the ATE is $33\,\mathrm{cm}$ on average, which reduces to $10\,\mathrm{cm}$ when evaluating only the feature-rich portion of the experiments (i.e., the terrain course traversal).

For Experiments 1 and 2, the robot was manually operated to traverse the terrain course, completing two and four loops, respectively. In Experiments 3–5, the robot was commanded to follow the rectangular path autonomously. In these experiments, the haptic localization algorithm was run online and in closed-loop and effectively guided the robot towards the goals (Fig. 5). Using only the prior knowledge of the terrain geometry and the contact events, the robot stayed localized in all the runs and successfully tracked the planned goals. This can be seen in Fig. 5, where the estimated trajectory (in blue) drifts on the xy-plane when the robot is walking on the flat ground, but it becomes re-aligned with the ground truth when the robot crosses the terrain course again.

Fig. 6 shows in detail the estimator performance for each of the three linear dimensions and yaw. Since position and yaw are unobservable, the drift on these states is unbounded.
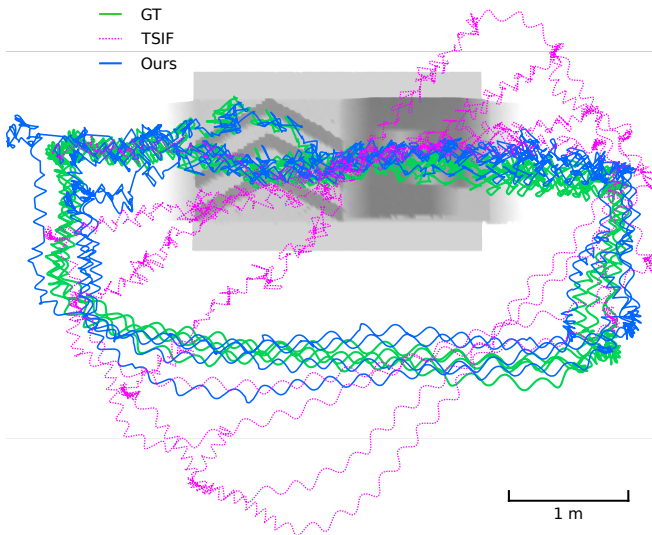
Fig. 5. Top view of the estimated trajectories from TSIF (dashed magenta), haptic localization (blue) and ground truth (green) for Experiment 2.

In particular, the error on the odometry filter (TSIF [1], magenta dashed line) is dominated by upward drift (due to kinematics errors and impact nonlinearities, see third plot) and yaw drift (due to IMU bias, see bottom plot). This drift is estimated and compensated for by the haptic localization (blue solid line), allowing accurate tracking of the ground truth (green solid line) in all dimensions. This can be noted particularly at the four peaks in the z-axis plot, where the estimated trajectory and ground truth overlap. These times coincide with the robot's summit of the terrain course.

B. Online Haptic Exploration on Vertical Surfaces

The second scenario involved a haptic wall following task: the robot started in front of a wall with an uncertain location. The particles were again initialized with $20\,\mathrm{cm}$ position covariance. To test the capability to recover from an initial error, we applied a $10\,\mathrm{cm}$ offset in both x and y from the robot's true position in the map. At start, the robot was commanded to walk $1\,\mathrm{m}$ to the right (negative y direction) and press a button on the wall. To accomplish the task, the robot needed to "feel its way" by alternating probing motions with its right front foot and walking laterally to localize inside the room. The fixed number of probing motions was pre-scripted so with each step to the right, the robot probed both in front and to its right. In this scenario, the prior map is represented as a 3D point cloud with a minimum distance of $1\,\mathrm{cm}$ between its points. The whole experiment was executed blindly with the static controller from [3].

As shown in Fig. 7 the robot was able to correct its localization and complete the task of touching the button. The initial probe to the front reduced uncertainty in the robot's x and z directions, which reduces the particles to an ellipsoidal elongated along y. As the robot moves, uncertainty in the x direction increases slightly. By touching the wall on the right, the robot re-localized in all three dimensions in much the same way as a human following a wall in the dark would. The re-localization allows the robot to press the button,
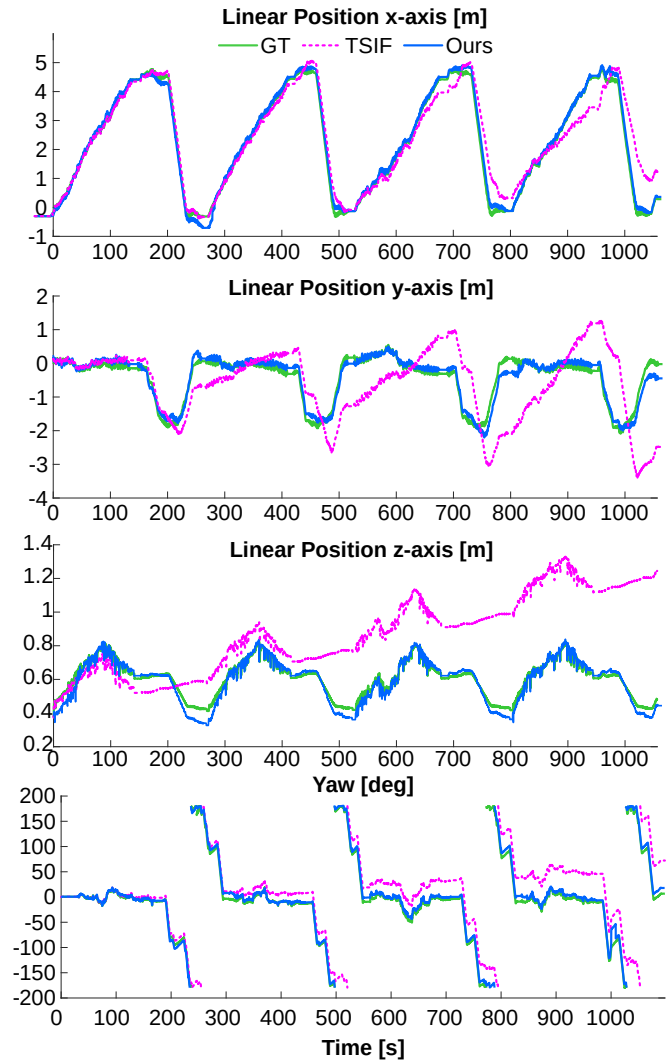

Fig. 6. Comparison between the estimated position from TSIF (dashed magenta) and haptic localization (blue) against Ground Truth (green) for Experiment 2. After $200\,\mathrm{s}$, the estimation error in TSIF has drifted significantly upward and in yaw. In particular, the upward drift is noticeable in the third plot, where the values grow linearly. The drift is eliminated by the re-localization against the prior map.

demonstrating the generalization of our algorithm to 3D. The final position error was: $[7.7, -3.7, -0.2]$ centimeters in the x, y and z directions[2].

VII. DISCUSSION

The results presented in Section VI demonstrate that a terrain with moderate complexity such as the one shown in Fig. 3 already provides enough information to bound the uncertainty of the robot's location. The effectiveness is obviously limited by the actual terrain morphology in a real world situation, which would need to contain enough features such that all the DoF of the robot are constrained once the robot has touched them. Fig. 8 shows the evolution of the particles up to the first half of the terrain course for Experiment 2. As the robot walks through, the particle swarm

[2]A video showing these experiments is available online: https://youtu.be/VfA8DAyr8t0
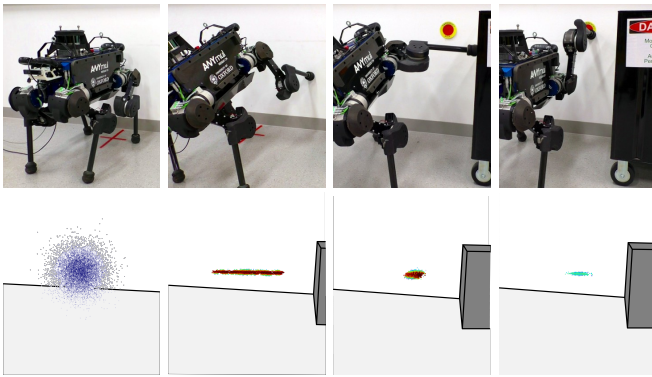
Fig. 7. 3-dimensional haptic probing experiment. The top row shows the robot performing the experiment while the bottom row shows the particle distribution and localization estimate. The particle set is colorized by normalized weight according to the jet colormap (i.e., dark blue = lowest weight, dark red = highest weight). First, on the bottom left, the robot has an initial distribution with poses equally weighed. The robot makes a forward probe and then moves to the right. Now the particles are distributed as an ellipse with high uncertainty to the left and right of the robot. Then, the robot makes a probe to the right and touches an obstacle; the particle cloud collapses into a tight cluster. Since the robot is now localized, it is able to complete the task of pressing the button on the wall.
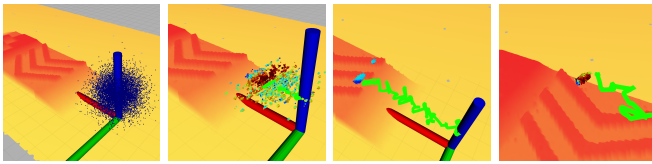

Fig. 8. Evolution of the particle set during a trial on the terrain course. The particle set is colorized by normalized weight according to the jet colormap (i.e., dark blue = lowest weight, dark red = highest weight). The green line indicates the ground truth trajectory. *A)* At start, all the particles have the same weight and are normally distributed at the starting position. *B)* After a few steps on the ramp, the robot pose is well estimated on x and z directions, but there is uncertainty on y. *C)* When the robot approaches the chevron the particle set divides into two clusters, indicating two strong hypotheses as to the robot pose. *D)* After a few more steps on the chevron, the robot is fully localized and the particles are tightly clustered.

gets tighter, indicating good convergence to the most likely robot pose.

In the third subfigure, it can be noted how the probability distribution over the robot's pose follows a bimodal distribution which is visible as two distinct clusters of particles. This situation justifies the use of particle filters, as they are able to model non-Gaussian distributions which can arise from a particular terrain morphology. In this case, the bimodal distribution is caused by the two identical gaps in between the chevrons. In such situations, a weighted average of the particle swarm would lead to a poor approximation of the true pose distribution. Therefore, the particle analysis described in Section V-A is crucial to reject such an update.

VIII. CONCLUSION

We have presented a haptic localization algorithm for quadruped robots based on Sequential Monte Carlo methods. The algorithm can reduce the uncertainty of the estimated robot's trajectory walking over a non-degenerate terrain course, with an average pose estimation accuracy up to $10\,\mathrm{cm}$. This is fundamental for repetitive autonomous tasks

in vision-denied conditions such as inspections in sewage systems. With routine inspection tasks in mind, we have also implemented a haptic exploration motion to localize against walls. In future work, we are planning to extend the localization algorithm to incorporate more features into the measurement model, such as estimated slope and friction, to further improve the localization performance as well as active planning for direct the probing actions.

IX. ACKNOWLEDGMENTS

REFERENCES

[1] M. Bloesch, M. Burri, H. Sommer, R. Siegwart, and M. Hutter, "The two-state implicit filter recursive estimation for mobile robots," *IEEE Robotics and Automation Letters*, vol. 3, no. 1, pp. 573–580, 2018.

[2] S. Nobili, M. Camurri, V. Barasuol, M. Focchi, D. Caldwell, C. Semini, and M. Fallon, "Heterogeneous sensor fusion for accurate state estimation of dynamic legged robots," in *Robotics: Science and Systems (RSS)*, Cambridge, Massachusetts, 2017.

[3] P. Fankhauser, M. Bjelonic, C. Dario Bellicoso, T. Miki, and M. Hutter, "Robust rough-terrain locomotion with a quadrupedal robot," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2018, pp. 1–8.

[4] M. Focchi, R. Orsolino, M. Camurri, V. Barasuol, C. Mastalli, D. G. Caldwell, and C. Semini, *Heuristic Planning for Rough Terrain Locomotion in Presence of External Disturbances and Variable Perception Quality*. Cham: Springer International Publishing, 2020, pp. 165–209.

[5] M. F. Fallon, P. Marion, R. Deits, T. Whelan, M. Antone, J. McDonald, and R. Tedrake, "Continuous humanoid locomotion over uneven terrain using stereo fusion," in *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)*, 2015, pp. 881–888.

[6] H. Kolvenbach, D. Wisth, R. Buchanan, G. Valsecchi, R. Grandia, M. Fallon, and M. Hutter, "Towards autonomous inspection of concrete deterioration in sewers with legged robots," *J. of Field Robotics*.

[7] E. Krotkov, "Active perception for legged locomotion: every step is an experiment," in *Proc. 5th IEEE International Symposium on Intelligent Control*, 1990, pp. 227–232 vol.1.

[8] J. Bednarek, M. Bednarek, L. Wellhausen, M. Hutter, and K. Walas, "What am I touching? Learning to classify terrain via haptic sensing," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2019, pp. 7187–7193.

[9] D. Fox, S. Thrun, W. Burgard, and F. Dellaert, *Particle Filters for Mobile Robot Localization*. New York, NY: Springer New York, 2001, pp. 401–428.

[10] G. Grisetti, C. Stachniss, and W. Burgard, "Improving grid-based slam with rao-blackwellized particle filters by adaptive proposals and selective resampling," in *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, 2005, pp. 2432–2437.

[11] G. Vezzani, U. Pattacini, G. Battistelli, L. Chisci, and L. Natale, "Memory unscented particle filter for 6-DOF tactile localization," *IEEE Transactions on Robotics*, vol. 33, no. 5, pp. 1139–1155, 2017.

[12] R. Van Der Merwe, A. Doucet, N. De Freitas, and E. Wan, "The unscented particle filter," in *Proc. of the 13th Intl. Conf. on Neural Information Processing Systems (NeurIPS)*, 2000, pp. 563–569.

[13] M. Chalon, J. Reinecke, and M. Pfanne, "Online in-hand object localization," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2013, pp. 2977–2984.

[14] S. Chitta, P. Vernaza, R. Geykhman, and D. Lee, "Proprioceptive localization for a quadrupedal robot on known terrain," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2007, pp. 4582–4587.

[15] C. Forster, L. Carlone, F. Dellaert, and D. Scaramuzza, "On-manifold preintegration for real-time visual–inertial odometry," *IEEE Trans. Robotics*, vol. 33, no. 1, pp. 1–21, 2017.

[16] J. Di Carlo, P. M. Wensing, B. Katz, G. Bledt, and S. Kim, "Dynamic locomotion in the mit cheetah 3 through convex model-predictive control," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2018, pp. 1–9.