

# Robust Model Predictive Control with Least Favorable Measurements

Daniel Lyons, Achim Hekler, Markus Kuderer, and Uwe D. Hanebeck

**Abstract**—Closed-loop model predictive control of nonlinear systems, whose internal states are not completely accessible, incorporates the impact of possible future measurements into the planning process. When planning ahead in time, those measurements are not known, so the closed-loop controller accounts for the expected impact of all potential measurements. We propose a novel conservative closed-loop control approach that does not calculate the expected impact of all measurements, but solely considers the single future measurement that has the worst impact on the control objective. In doing so, the model predictive controller guarantees robustness even in the face of high disturbances acting upon the system. Moreover, by considering only a single dedicated measurement, the complexity of closed-loop control is reduced significantly. The capabilities of our approach are evaluated by means of a path planning problem for a mobile robot.

## I. INTRODUCTION

In contrast to regular control, model predictive control (MPC) does not only try to find optimal control inputs for the current state of the system under control, but also for predicted states of the system [1]. By this prediction of system states, it is possible to react earlier to the anticipated development of the system state and achieve a higher quality of control. Depending on how much model knowledge is incorporated into the planning phase, there is a distinction between *open-loop feedback control* and *closed-loop feedback control*.

In the setting of open-loop feedback control, a stochastic model predictive controller predicts the development of the system under control by solely employing knowledge about the system model and the stochastic model of the noise acting upon it. Exclusively in dependence of these predicted states, the optimal sequence of control inputs is determined [1]–[3]. However, if there are sensors that acquire new information about the system state through measurements, more information about the system under control is available. In particular, this knowledge consists of a model of the measurement process describing the relations between the system state and the received measurements as well as a stochastic characterization of the measurement noise. Open-loop feedback control does not utilize this knowledge about the measurement process and, consequently the stochastic uncertainty about the actual system state may increase drastically.

D. Lyons, A. Hekler, M. Kuderer, and U. D. Hanebeck are with the Intelligent Sensor-Actuator-Systems Laboratory (ISAS), Institute for Anthropomatics, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany. [daniel.lyons@ieee.org](mailto:daniel.lyons@ieee.org), [achim.hekler@ieee.org](mailto:achim.hekler@ieee.org), [markus.kuderer@student.kit.edu](mailto:markus.kuderer@student.kit.edu), [uwe.hanebeck@ieee.org](mailto:uwe.hanebeck@ieee.org).

In contrast, closed-loop feedback control approaches get rid of this problem by actively incorporating future measurements into the planning process [4], [5]. In addition to the prediction of system states, the impact of possible future measurements on the state estimate is considered. Since these measurements are not known in advance, the controller has to consider all potential measurements in the planning phase [4]. As a result, the decision tree has to be branched for every possible measurement at any time instant in the planning horizon.

In the case of a continuous measurement space, there is an infinite number of possible future measurements. So, the insertion of knowledge about the measurement process is computationally infeasible for general systems and, hence, the measurement space has to be discretized. But even finitely many possible measurements result in an increased computational complexity of the predictive planning, as the impact for every measurement has to be evaluated in advance for every time step.

To ensure operability under real-time constraints, approximations of closed-loop feedback control are necessary. One possible approximation technique is to not make the planning dependent on all future measurements, but to choose one specific measurement from all possible ones. The consequence of such an approach is that knowledge about the measurement process is still incorporated into the planning phase, but the complexity is decreased significantly in comparison to closed-loop feedback control. There are different ways to pick a single measurement from all attainable ones at each time step. In [6], the authors propose to always take the nominal measurement for planning, i.e., the measurement that confirms the current state estimate the most.

In this paper, we propose a novel approach to selecting a single future measurement for planning at each time instant in the planning horizon. The key idea of our control algorithm is to always select the single measurement from a discrete set of measurements for further planning that has the worst impact on the control objective. Thus, we do not take the expected reward over all measurements, but always anticipate the least favorable measurement (LFM) in terms of the control objective. Our approach results in solutions to the optimal control problem that are by far more robust than open-loop feedback control and other approximations to closed-loop feedback control. By considering least favorable measurements in the planning process, the controller acts more carefully and avoids the violation of hard constraints like the avoidance of crashes of a robot into a wall.

However, the complexity of calculating the least favorable measurement equals the complexity of the closed-loop control algorithm. We propose an approximation performed by rating the impact of measurements on the systems state through a greedy approach. This impact is expressed in terms of a one-step objective function that evaluates how beneficial or detrimental certain states of the system are. Also, this approximation by a greedy selection of the least favorable measurements leads to a considerable reduction in complexity of the control problem.

The paper is structured as follows: In the following section, we describe the considered optimal control problem in detail. Our novel approach of conservative planning with least favorable measurements is presented in Section III. The capabilities and the efficiency of our approach are evaluated by means of a path planning problem in Section IV. Conclusions and an outlook to future work conclude the paper.

## II. PROBLEM FORMULATION

A model predictive controller predicts the system state over a planning horizon in dependence on a sequence of control inputs. The predicted states within the planning horizon are rated by an objective function that models how advantageous or disadvantageous specific states and control inputs are. The sequence of control inputs is determined that maximizes this objective and the first element of the optimal sequence is applied to the system. In the next time step, this procedure is repeated.

### A. System Properties

The random variable  $\underline{x}_k$  characterizes the controller's estimate about the current state of the system. Throughout the paper, we assume that the dynamic behaviour of the system under control can be described by the discrete-time nonlinear dynamic system

$$\underline{x}_{k+1} = \underline{u}_k(\underline{x}_k, \underline{u}_k, \underline{w}_k), \quad (1)$$

in dependence of applied control inputs  $\underline{u}_k$ . The control input  $\underline{u}_k$  that can be applied to the system is taken from a finite discrete set

$$\mathcal{U}_k := \{\underline{u}_k^{(1)}, \underline{u}_k^{(2)}, \dots, \underline{u}_k^{(L)}\}.$$

The nonlinear measurement model

$$\underline{z}_k = \underline{h}_k(\underline{x}_k, \underline{v}_k), \quad (2)$$

describes how the outcome of the measurement process is linked to the system state.

Furthermore, the random variables  $\underline{w}_k$  and  $\underline{v}_k$  subsume the stochastic noises acting upon the system and the measurement process. Both the system state  $\underline{x}_k \in \mathcal{X}$  and the measurements  $\underline{z}_k \in \mathcal{Z}$  are continuous random variables characterized by their probability density functions  $f_k^x$  and  $f_k^z$ , respectively.

For the specific class of systems, whose internal states are not completely accessible, these states can be estimated on the basis of noisy measurements and the previously

applied control inputs. The current state estimate  $\underline{x}_k$  can be calculated by a recursive Bayesian estimator, which will be explained in the following subsection.

### B. Bayesian Estimation

A Bayesian estimator consists of two alternating processing steps called *prediction* and *filtering step*.

In the prediction step, the system state  $\underline{x}_k$  is updated to the system state  $\underline{x}_{k+1}^p$  at the next time step by means of the Chapman-Kolmogorov equation

$$f_{k+1}^p(\underline{x}_{k+1}) = \int f^T(\underline{x}_{k+1}|\underline{x}_k, \underline{u}_k) f_k(\underline{x}_k) d\underline{x}_k, \quad (3)$$

where the transition density  $f_k^T(\underline{x}_{k+1}|\underline{x}_k, \underline{u}_k)$  is a probabilistic representation of (1).

The filtering step incorporates a new measurement  $\hat{z}_k$  into a prior state estimate  $f_k^p$ , which is the result of the previously executed prediction step, by employing Bayes' law

$$f_k^e(\underline{x}_k) = \frac{1}{\hat{c}} \cdot f_k^L(\hat{z}_k|\underline{x}_k) \cdot f_k^p(\underline{x}_k). \quad (4)$$

Here,  $\hat{c}_k$  is a normalization constant and the likelihood  $f_k^L(\hat{z}_k|\underline{x}_k)$  can be derived from Eq. (2).

### C. Considered Optimal Control Problem

In a setting with imperfect knowledge about the system state, the model predictive controller has to deal with random variables  $\underline{x}_k$  characterizing the current state estimate. The desired behavior of the system is modeled via a reward function  $r(\underline{x}_k, \underline{u}_k)$ , which acts on probability densities over the system state and assigns a scalar reward to the considered state estimate  $\underline{x}_k$  and the control input  $\underline{u}_k$ . The controller plans predictively into the future and aims at finding control inputs that are optimal over a horizon of length  $N$ . This optimality is defined in terms of a cumulative reward function, the sum of all one step rewards given at each time instant in the planning horizon. At every time step  $k$ , the cumulative reward function

$$\begin{aligned} V_k(\underline{x}_k, \underline{u}_k, \underline{\mu}_{k,1:N-1}) \\ = \mathbb{E} \left\{ r(\underline{x}_k, \underline{u}_k) + \sum_{i=1}^{N-1} r(\underline{x}_{k+i}, \underline{\mu}_{k,i}(\underline{x}_{k+i})) \right\} \end{aligned} \quad (5)$$

is maximized.

Since at every time instant, the stochastic estimator gives feedback about the state of the system in terms of an updated state estimate, the reward function at the planning time step  $i$  has to consider a control policy  $\underline{\mu}_i(\underline{x}_i)$ , which maps the state estimate  $\underline{x}_i$  to a control input  $\underline{u}_i$ . In other words, the optimal control input at time instant  $i > k$  depends on the previously applied control inputs and the estimate of the state  $\underline{x}_i$ .

The optimal control input

$$\underline{u}_k^*(\underline{x}_k) = \arg \max_{\underline{u}_k} \left\{ \max_{\underline{\mu}_{k,1:N-1}} \{ V_k(\underline{x}_k, \underline{u}_k, \underline{\mu}_{k,1:N-1}) \} \right\} \quad (6)$$

is applied to the system and the planning algorithm is repeated at the next time step.

Since the planning should reflect the real estimation procedure as described in Section II-B as accurately as possible,

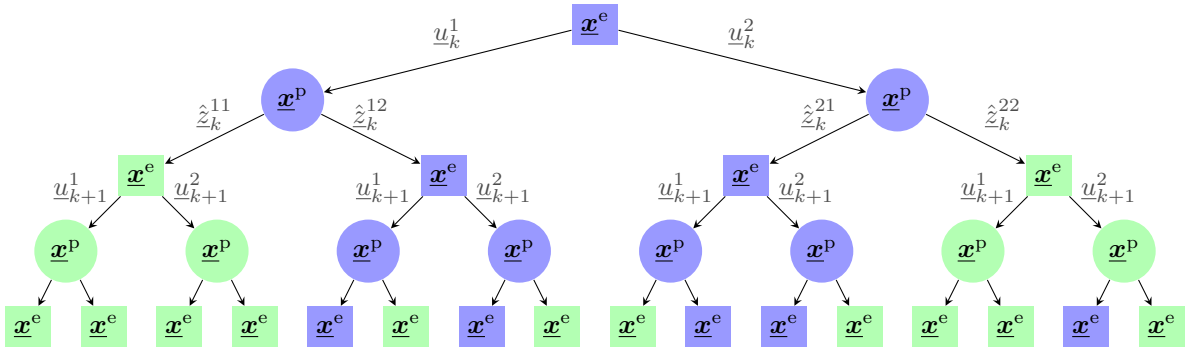


Fig. 1. Exemplary search tree of a closed-loop planning for  $L = 2$  different control inputs,  $M = 2$  discrete measurements and  $N = 2$  planning time steps. Approaches considering all virtual measurements have to build the whole tree. In contrast, the proposed LFM approach reduces the complexity significantly by selecting a specific single measurement. The pruning of subtrees is illustrated by light green nodes.

the controller not only has to incorporate the nonlinear system dynamics, but also future measurements and properties of the stochastic estimator.

### III. LEAST FAVORABLE MEASUREMENTS

In this section, we will introduce our novel concept of incorporating only one measurement into the planning process. This will be a measurement that leads to the least desirable posterior state estimate  $\underline{x}_k^e$  in terms of the reward function. From a purely technical standpoint, we want the planning to be as robust as possible, i.e., we want to account for future possible measurements that may be undesirable in terms of an application specific reward function.

In open-loop feedback control, it is assumed that the controller does not receive observations about the system under control. For the expected reward to be maximized in (5) this means that future system states  $\underline{x}_{k+i}$  are generated by prediction using (3). The cumulative reward function then is

$$V_k(\underline{x}_k, \underline{u}_k, \underline{u}_{k,1:N-1}) = r(\underline{x}_k, \underline{u}_k) + \sum_{i=1}^{N-1} r(\underline{x}_{k+i}, \underline{u}_{k,i})$$

The closed-loop feedback controller anticipates future measurements and future beliefs are calculated by (3) and (4). The cumulative reward function to be maximized becomes

$$\begin{aligned} V_k(\underline{x}_k, \underline{u}_k, \underline{\mu}_{k,1:N-1}) \\ = \mathbb{E}_{\underline{z}_k, \dots, \underline{z}_{k+N}} \left\{ r(\underline{x}_k, \underline{u}_k) + \sum_{i=1}^{N-1} r(\underline{x}_{k+i}, \underline{\mu}_{k,i}(\underline{x}_{k+i})) \right\} \end{aligned}$$

where  $\underline{\mu}(\underline{x}_{k+i})$  is defined as in Section II. Instead of considering the weighted average over future measurements, we consider

$$\begin{aligned} V_k(\underline{x}_k, \underline{u}_k, \underline{\mu}_{k,1:N-1}) \\ = \inf_{\underline{z}_k, \dots, \underline{z}_{k+N}} \left\{ r(\underline{x}_k, \underline{u}_k) + \sum_{i=1}^{N-1} r(\underline{x}_{k+i}, \underline{\mu}_{k,i}(\underline{x}_{k+i})) \right\} \end{aligned} \quad (7)$$

Since the infimum over future measurements in Eq. (7) may become arbitrarily small, we assume in the following that

either the measurement space  $\mathcal{Z}$  is a compact or a discrete set. Ideally, one would evaluate the impact of possible measurements on the complete planning process in Eq. (7). However, this would not lead to a reduction in complexity compared to closed-loop control because in order to evaluate the impact of a measurement on future decisions, the whole following decision tree would have to be expanded over all successive control inputs and future measurements.

We propose a greedy approximation to judge on the impact of a measurement at any instance in the planning process by only evaluating the one step reward function of the posterior state estimate. Thus, not the whole decision tree has to be expanded for all future measurements.

We will first describe our novel approach for myopic planning with time horizon of length one, the generalization to arbitrary time horizons will be conducted afterwards.

#### A. Myopic Control with Least Favorable Measurements

For predictive control the current state estimate  $\underline{x}_k$  is first predicted using the system dynamics depending on the applied control input by equation (3). Thus, we obtain a predicted state estimate  $\underline{x}_{k+1}^p$  linked to a control input  $\underline{u}_k$ . In the following, we assume that we have discretized the space of possible measurements  $\mathcal{Z}_{k+1}$  into a set of discrete measurements

$$\mathcal{Z}_{d,k+1} := \{\hat{z}_{k+1}^1, \dots, \hat{z}_{k+1}^M\}.$$

Methods for obtaining this discretized measurement space are described in Subsection III-C. For each prospective measurement  $\hat{z}_{k+1}^i \in \mathcal{Z}_{d,k+1}$ , we can calculate the posterior state estimate  $\underline{x}_{k+1}^e$  linked to the control input  $\underline{u}_k$  and the measurement  $\hat{z}_{k+1}^i$ .

The quality of this posterior state estimate can then be evaluated in the one step reward function defined in Section II-C by calculating  $r(\underline{x}_{k+1}^e, \underline{u}_k | \hat{z}_{k+1}^i)$ , which denotes the reward given for the estimate of the state  $\underline{x}_{k+1}^e$  linked to the control input  $\underline{u}_k$  and the measurement  $\hat{z}_{k+1}^i$ . For each measurement in  $\mathcal{Z}_{d,k+1}$ , the corresponding one step reward  $r(\underline{x}_{k+1}^e, \underline{u}_k | \hat{z}_{k+1}^i)$  can be calculated to obtain a set of one step rewards

$$\{r(\underline{x}_{k+1}^e, \underline{u}_k | \hat{z}_{k+1}^1), \dots, r(\underline{x}_{k+1}^e, \underline{u}_k | \hat{z}_{k+1}^M)\}.$$

The reward we then attach to the control input  $\underline{u}_k$  is

$$\begin{aligned} r(\underline{\mathbf{x}}_{k+1}^e, \underline{u}_k) \\ := \min_{\hat{z}_{k+1}^i} \{r(\underline{\mathbf{x}}_{k+1}^e, \underline{u}_k | \hat{z}_{k+1}^1), \dots, r(\underline{\mathbf{x}}_{k+1}^e, \underline{u}_k | \hat{z}_{k+1}^M)\} , \end{aligned} \quad (8)$$

i.e., the minimal achievable greedy reward considering the possible measurements. The objective of the myopic controller is then to maximize this minimal reward over all possible control inputs.

The minimum in (8) is always assumed, as we only consider a discrete and finite subset of all measurements.

### B. General Predictive Control with Least Favorable Measurements

For predictive control with longer planning horizons, the objective is to find a sequence of control inputs that maximizes the cumulative reward (5). For non-myopic planning, the myopic approach explained above is repeated at each time step  $j > k$  in the planning horizon. At each node in the decision tree the controller considers control inputs in  $\mathcal{U}$  and for each input  $\underline{u}_j$  predicts its effect on the system via (3) and obtains the predicted state estimate  $\underline{\mathbf{x}}_{j+1}^p$  linked to the control input  $\underline{u}_j$ . Then, just like above, the controller determines the set of greedy rewards

$$\{r(\underline{\mathbf{x}}_{j+1}^e, \underline{u}_j | \hat{z}_j^1), \dots, r(\underline{\mathbf{x}}_{j+1}^e, \underline{u}_j | \hat{z}_j^M)\}$$

and the minimal reward

$$\begin{aligned} r(\underline{\mathbf{x}}_{j+1}^e, \underline{u}_j) \\ := \min_{\hat{z}_j^i} \{r(\underline{\mathbf{x}}_{j+1}^e, \underline{u}_j | \hat{z}_j^1), \dots, r(\underline{\mathbf{x}}_{j+1}^e, \underline{u}_j | \hat{z}_j^M)\} . \end{aligned} \quad (9)$$

The planning for further time instants in the planning horizon is then repeated with the reward (9). Just as in the open-loop case, the decision tree now has to be branched only over the possible control inputs (see Fig. 1). At each time instant in the planning process, the one step reward attached to a control input is defined via (9).

### C. Obtaining Representative Measurements

There are several different techniques to determine a suitable discretization of the continuous measurement space at each time step. In this subsection, we will briefly describe two methods to obtain such a discretization depending on the probability of obtaining a measurement.

The most intuitive approach would be to employ random sampling from the probability density  $f(z_j)$  of the measurement process. If we want to receive  $M$  possible measurements, we can randomly draw  $M$  samples  $\hat{z}_j^i \sim f(z_j)$  from this probability density. For time instants that lie in the future, i.e.,  $j > k$ , the probability density  $f(z_j)$  can be constructed from the likelihood and the density describing the state estimate  $f(\underline{\mathbf{x}}_j)$  at time step  $j$  via

$$f(z_j) = \int_{\mathcal{X}_j} f^L(z_j | \underline{\mathbf{x}}_j) \cdot f(\underline{\mathbf{x}}_j) d\underline{\mathbf{x}}_j .$$

Another way to obtain  $M$  possible measurements would be through deterministic approximation of the measurement

density  $f(z_j)$  with the methods proposed in [4] and [7]. The advantage of deterministic sampling is that through the systematic approach no outliers will occur and a certain quality of representation can be guaranteed.

### D. Complexity Analysis

At each time instant where a decision can be made, the controller has to branch over all possible decisions in form of possible control inputs. Let  $|\mathcal{U}|$  denote the number of (discrete) control inputs that can be applied to the system under control and denote the maximal number of possible measurements by  $|\mathcal{Z}_d|$ .

The computational complexity of expanding the whole decision tree for the open-loop algorithm is  $\mathcal{O}(|\mathcal{U}|^N)$ , where  $N$  is the length of the planning horizon. For closed-loop planning, the complexity becomes  $\mathcal{O}(|\mathcal{Z}_d| \cdot |\mathcal{U}|^N)$ , as the controller not only branches over all possible control inputs but also over all possible future measurements. For the least favorable measurement algorithm, the controller has to calculate the minimum of the rewards for every possible future observation  $\hat{z}_j^i \in \mathcal{Z}_d$  at each time step  $j$  which is  $\mathcal{O}(|\mathcal{Z}_d|)$ . Since the branching is now only conducted over all control inputs, the complexity of the approach suggested above is reduced to  $\mathcal{O}(|\mathcal{Z}_d| \cdot |\mathcal{U}|^N)$ .

## IV. SIMULATIONS

To demonstrate the capabilities of the proposed conservative planning approach, we applied the LFM algorithm to a robot control problem. In the considered simulation setting, a robot [8] is supposed to reach a target area without colliding with obstacles. The pose of the robot can be estimated by means of the previous applied control inputs and distance measurements to two landmarks. At every time step, the controller determines the optimal control input with regard to less favorable measurements. This specific input is applied to the robot and the whole procedure is repeated in the next time step.

### A. System and Measurement Model

We model the pose of the robot as a three-dimensional continuous state  $\underline{\mathbf{x}}_k = [\mathbf{x}_k, \mathbf{y}_k, \phi_k]^T$ , where  $\mathbf{x}_k$  and  $\mathbf{y}_k$  are the coordinates of the robot's position in the plane and  $\phi_k$  is its orientation. The motion of the robot is described by the following system model

$$\begin{bmatrix} \mathbf{x}_{k+1} \\ \mathbf{y}_{k+1} \\ \phi_{k+1} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_k \\ \mathbf{y}_k \\ \phi_k \end{bmatrix} + \begin{bmatrix} (u_k^s + \mathbf{w}_k^s) \cdot \cos(\phi_k) \\ (u_k^s + \mathbf{w}_k^s) \cdot \sin(\phi_k) \\ u_k^\phi + \mathbf{w}_k^\phi \end{bmatrix} , \quad (10)$$

motivated by the robots introduced in [8]. The two-dimensional control input  $\underline{u}_k = [u_k^s, u_k^\phi]^T$  consists of the step size  $u_k^s$  in forward direction and a turning angle  $u_k^\phi$ . We assume that the robot can either execute a forward step with a fixed length of 10 cm or does not move forward. Additionally, the robot can superimpose turns of  $\frac{\pi}{8}$  rad in both directions onto the forward motion or turn without

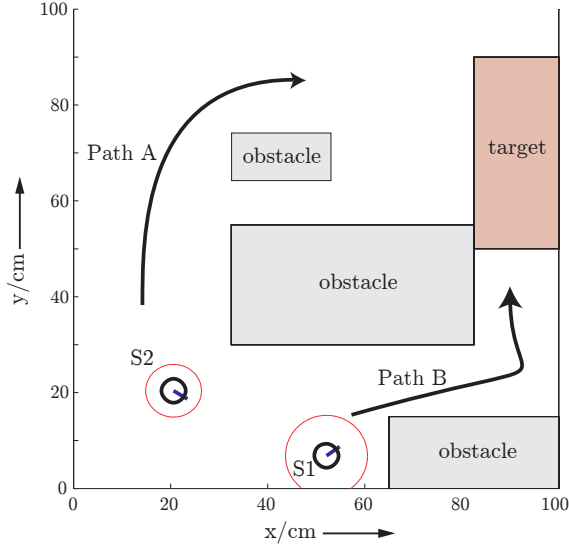


Fig. 2. Environment setting. The grey shaded areas represent obstacles, the red rectangle is the target area.  $S1$  and  $S2$  are the initial robot configurations for both scenarios where the line within the circle denotes the orientation. The two landmarks used for measurements are located at  $[50 \text{ cm}, -50 \text{ cm}]^T$  and  $[-50 \text{ cm}, 50 \text{ cm}]^T$

taking a forward step. Hence, the entire action space consists of the finite discrete set

$$\mathcal{U} = \left\{ \begin{bmatrix} 0 \text{ cm} \\ \pm \frac{\pi}{8} \text{ rad} \end{bmatrix}, \begin{bmatrix} 0 \text{ cm} \\ 0 \text{ rad} \end{bmatrix}, \begin{bmatrix} 10 \text{ cm} \\ \pm \frac{\pi}{8} \text{ rad} \end{bmatrix}, \begin{bmatrix} 10 \text{ cm} \\ 0 \text{ rad} \end{bmatrix} \right\}.$$

The random variables  $w_k^s$  and  $w_k^\phi$  in Eq. (10) subsume the noise acting upon the forward step and the rotational movement.

For state estimation, the robot can take noisy distance measurements to two landmarks. The  $i$ -th entry of the measurement vector  $\underline{z}_k$  is given by

$$z_k^i = \sqrt{(\mathbf{x}_k - \tilde{x}^i)^2 + (\mathbf{y}_k - \tilde{y}^i)^2} + v_k, \quad (11)$$

where  $[\tilde{x}^i, \tilde{y}^i]^T$  is the position of landmark  $i = 1, 2$  and  $v_k$  denotes zero-mean white Gaussian measurement noise.

### B. Reward Function

The robot's task is to reach the target area without colliding with obstacles. Therefore, we implemented a reward function that assigns negative values to all state-action combinations that cause the robot to crash. Additionally, each step that does not result in a target state, results in a small penalty value in order to prefer shorter paths over longer ones. Finally, steps that let the robot reach the target area safely have zero penalty. In detail, the values are set as following:

$$r(\mathbf{x}_k, \underline{u}_k) = \begin{cases} -40, & \text{if action leads to collision} \\ 0, & \text{if target is reached} \\ -1, & \text{otherwise} \end{cases} \quad (12)$$

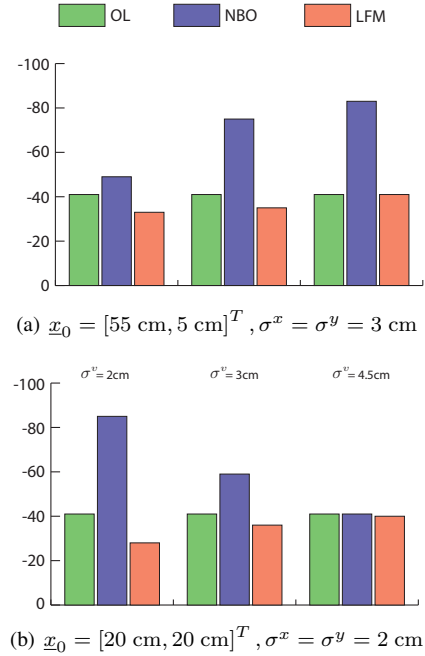


Fig. 3. Average rewards from 50 Monte-Carlo runs for different start configurations and noise settings. Diagram (a) shows the achieved reward from initial configuration  $S1$  in Fig. 2, diagram (b) denotes results of  $S2$ . For each configuration, LFM was compared to open-loop control (OL) and Nominal Belief State Optimization (NBO) with different measurement noise levels.

The planning horizon in our simulation is set to 40 time steps as well, so any collision-free trajectory is preferred, even if the target is not reached. Thus, the best a robot can achieve is a cumulative reward of zero. Each additional step and especially collisions decrease the attained reward. So far, the function assigns values to combinations of states and actions. The extension of the deterministic reward function (12) to the probabilistic state description is conducted via

$$r(\underline{\mathbf{x}}_k, \underline{u}_k) = \mathbb{E}\{r(\mathbf{x}_k, \underline{u}_k)\}, \quad (13)$$

the expected reward given an initial state estimate  $\underline{\mathbf{x}}_k$  and a fixed action  $\underline{u}_k$ .

### C. Calculation of Optimal Policies

For efficiency reasons, the optimal solution to the planning problem in Section II is calculated via the principle of dynamic programming [9]. In order to ensure computational feasibility even for long time horizons, we employed a parametrized fitted value iteration algorithm described in [10]. In this approach, the authors propose a parametrization of the belief space, i.e., the space of probability density functions  $f_j$  characterizing the state estimates  $\underline{\mathbf{x}}_j$ . In our simulation, we used an extended Kalman filter for state estimation that implicitly allows the parametric representation of the belief space in form of Gaussian densities.

### D. Test Scenarios

The following section describes test scenarios that differ in the start configuration of the robot and in the measurement

	OL	NBO	LFM
Steps:	-	21	<b>28</b>
Target Reached:	0	54	<b>68</b>
Collisions:	0	9	<b>5</b>

(a)  $\underline{x}_0 = [55 \text{ cm}, 5 \text{ cm}]^T$ ,  $\sigma^x = \sigma^y = 3 \text{ cm}$

	OL	NBO	LFM
Steps:	-	23	<b>30</b>
Target Reached:	2	66	<b>63</b>
Collisions:	0	6	<b>2</b>

(b)  $\underline{x}_0 = [20 \text{ cm}, 20 \text{ cm}]^T$ ,  $\sigma^x = \sigma^y = 2 \text{ cm}$

Fig. 4. Average test data over all tested noise settings. Each table lists average values for one initial configuration, where ‘Target Reached’ shows percentage of successful runs, i.e., target was reached without crash. ‘Collisions’ denotes percentage of runs that included collisions and ‘Steps’ are average number of steps needed to reach target in successful runs. The maximal allowed number of steps was 40.

noise. Fig. 2 shows the target area and obstacles in an arena of  $100 \text{ cm} \times 100 \text{ cm}$ . Two landmarks were positioned at  $[50 \text{ cm}, -50 \text{ cm}]^T$  and  $[-50 \text{ cm}, 50 \text{ cm}]^T$ .

In all scenarios, the standard deviation of the system noise is set to  $\sigma^{w_s} = 3 \text{ cm}$  and  $\sigma^{w_\phi} = 0.2 \text{ rad}$ . For the measurement noise, three different standard deviations  $\sigma^v$  of 2 cm, 3 cm, and 4.5 cm were applied. The  $3\text{-}\sigma$  ellipses of the start configurations  $S1$  and  $S2$  are as shown in Fig. 2. We compared the proposed approach to open-loop planning (OL) [1] and Nominal Belief State Optimization (NBO) [6].

### E. Results

The results listed in Fig. 3 show the achieved rewards after 50 Monte-Carlo runs with a controller employing open-loop control, NBO control, and LFM control from two different start configurations and different measurement noise settings. Control with LFM performed best in our test scenarios. The controller employing NBO tended to take the narrow path B from both start positions. This plan can be too risky under conditions where disturbances act upon the system. This is the reason, why the robot crashed mostly during the passing of the narrow section on path B. Open-loop control resulted in a reward of  $-40$  in all the test runs. The state estimation after some planning steps without filtering gets highly imprecise. After 40 prediction steps, which was the planning horizon in our scenario, the expected probability for reaching the target is relatively low compared to the probability of crashing into obstacles. Thus, staying at the current position was rated best among all available actions. Fig. IV-D shows the percentage of successful runs, i.e., runs where the target was reached without colliding with obstacles, percentage of collisions, and steps needed to reach the target as average over the different noise settings. It can be seen that open-loop control is not suitable for this kind of planning problem, where measurements are mandatory for successful behaviour, as the robot never reached the target at

all. In successful runs, the NBO-Algorithm needed slightly less steps to bring the robot home safely but this advantage comes at a price, as the percentage of collisions is higher than by employing LFM.

## V. CONCLUSIONS

We have proposed a novel control algorithm that is robust against disturbances in the form of noisy and detrimental measurements. Our approach allows for a conservative and robust control even in the face of large measurement noise. Furthermore, this algorithm is an efficient approximation to the general model predictive closed-loop control framework. The approximation consists of considering at each time instant in the planning horizon the single measurement that is the most detrimental to the control objective. We have evaluated our approach against other approximative planning algorithms and have shown that a controller employing control with least favorable measurements is more careful and the probability of a violation of hard constraints is significantly lower. Future work will be concerned with evaluating our approach on existing hardware [8].

## ACKNOWLEDGMENTS

This work was partially supported by the German Research Foundation (DFG) within the Research Training Group GRK 1194 ‘‘Self-organizing Sensor-Actuator-Networks’’.

## REFERENCES

- [1] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, 2nd ed. Athena Scientific, Belmont, Massachusetts, USA, 2007, vol. 1.
- [2] E. F. Camacho and C. Bordons, *Model Predictive Control*, 2nd ed. Springer-Verlag London Ltd., June 2004.
- [3] L. Blackmore and B. C. Williams, ‘‘Optimal, Robust Predictive Control of Nonlinear Systems under Probabilistic Uncertainty Using Particles,’’ in *Proceedings of the American Control Conference (ACC 2007)*, vol. 17, no. 2, New York, USA, Jul. 2007.
- [4] F. Weissel, T. Schreiter, M. F. Huber, and U. D. Hanebeck, ‘‘Stochastic Model Predictive Control of Time-Variant Nonlinear Systems with Imperfect State Information,’’ in *Proceedings of the 2008 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2008)*, Seoul, Republic of Korea, Aug. 2008, pp. 40–46.
- [5] L. Kaelbling, M. Littman, and A. Cassandra, ‘‘Planning and acting in partially observable stochastic domains,’’ *Artificial Intelligence*, vol. 101, no. 1-2, pp. 99–134, 1998.
- [6] S. A. Miller, Z. A. Harris, and E. K. P. Chong, ‘‘Coordinated Guidance of Autonomous UAVs via Nominal Belief-State Optimization,’’ in *American Control Conference, 2009. ACC’09.*, 2009, pp. 2811–2818.
- [7] V. Klumpp and U. D. Hanebeck, ‘‘Dirac Mixture Trees for Fast Sub-optimal Multi-Dimensional Density Approximation,’’ in *Proceedings of the 2008 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2008)*, Seoul, Republic of Korea, Aug. 2008, pp. 593–600.
- [8] F. Weissel, M. F. Huber, and U. D. Hanebeck, ‘‘Test-Environment based on a Team of Miniature Walking Robots for Evaluation of Collaborative Control Methods,’’ in *Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2007)*, San Diego, California, Nov. 2007, pp. 2474–2479.
- [9] R. E. Bellman, *Dynamic Programming*. Princeton University Press, Princeton, New Jersey, 1957.
- [10] A. Brooks, S. Williams, and H. Durrant-Whyte, ‘‘Parametric POMDPs for Planning in Continuous State Spaces,’’ *Robotics and Autonomous Systems*, vol. 54, pp. 887–897, 2006.