

Approximation of Stochastic Nonlinear Closed-Loop Feedback Control with Application to Miniature Walking Robots

Christof Chlebek and Uwe D. Hanebeck

Abstract— We consider stochastic nonlinear time-variant systems with imperfect state information in the context of model predictive control. The optimal control performance can only be achieved by closed-loop feedback (CLF) policies, which in fact anticipate future behavior. However, the computation of these policies is in general not tractable due to the presence of the dual effect, i.e., the control actions not only influence the state but also the uncertainty of its estimate. Thus, we propose an approximation to closed-loop control. We use a forward calculation approach, which is derived from an open-loop feedback (OLF) control setup, but implements the fundamental property of closed-loop control that future measurement feedback is considered in the optimization. By using a finite set of representative measurements, the feedback behavior is anticipated only on basis of current information. The proposed optimization scheme is based on a continuation method, which implements an effective calculation to obtain a sequence of control inputs. The effectiveness of the presented approach is demonstrated by means of a miniature walking robot.

I. INTRODUCTION

In Model Predictive Control (MPC), the main objective is to compute and apply control inputs such that the behavior of a system quantified by means of a cost function is optimized over a certain control horizon. Modelling errors or disturbances affecting the system can be considered in a stochastic fashion leading thereby to stochastic MPC (SMPC). The optimal solution, or strictly speaking the closed-loop optimal solution, is given by the Bellman equation, which unfortunately is only computable in few special cases, such as the linear-quadratic-Gaussian (LQG) control problem or the control of systems with a finite number of states and control inputs. This is mostly due to the curse of dimensionality and the fact that separation of estimation and control does not hold in the case of stochastic nonlinear systems [1], [2].

Literature Review: Assuming state feedback, scenario-based approaches solve the control problem by sampling the underlying uncertainty space. This has been implemented for linear systems [3], [4], [5] and nonlinear systems [6], [7], where for convex optimization problems a bound has been established on the number of required random samples for a guaranteed quality of control [8]. An alternative approach is the usage of vector quantization, where the scenarios are represented as part of a code book. This approach is implemented for finite input sets and nonlinear systems [9]. Other methods include an approach based on a minimum

C. Chlebek, and U. D. Hanebeck are with the Intelligent Sensor-Actuator-Systems Laboratory (ISAS), Institute for Anthropomatics and Robotics, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, christof.chlebek@kit.edu, uwe.hanebeck@ieee.org.

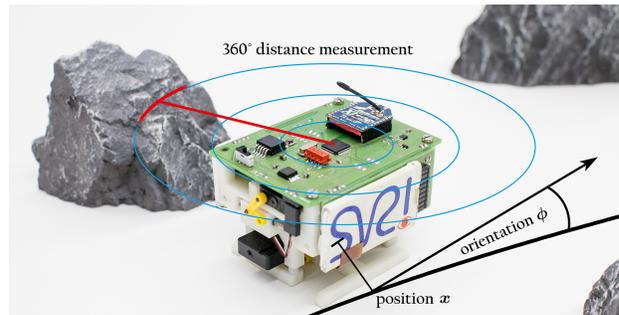


Fig. 1. Example system miniature walking robot.

principle, where the control problem is converted into a two-point boundary-value problem and solved by homotopy continuation [10], and a two-step approach based on open-loop feedback (OLF) pre-calculation and closed-loop feedback (CLF) post-processing on a discretized restricted state space [11]. Perturbed measurement feedback has been considered for finite sets of control inputs and finite sets of representative measurements, by generating a search tree by means of all combinations of inputs and measurements over the control horizon [12]. Partially observable Markov decision processes (POMDPs) traditionally consider discrete state spaces. Here the feedback information is represented in form of estimated states in the so-called belief space. Relevant work assumes an infinite control horizon with discounted cost using reinforcement learning methods to explore the belief space and compute an explicit form of the cost function. This leads to the calculation of stationary policies, which implies time-invariant systems [13]. In recent years continuous-state spaces have been considered, by discretization of the belief space and thereby diffusing the separating line between POMDPs and SMPC. Using Gaussian representations of the belief space, continuous state POMDPs with finite sets of inputs [14], [15] and continuous input spaces [16] are solved. Alternatively the belief space can be represented by a finite set of discrete probability densities. This has been considered using a Monte-Carlo approach and Kullback-Leibler divergence as similarity measure [17] and deterministic state propagation together with cost interpolation based on the Wasserstein distance [18].

Contribution: In our approach, we present an MPC approach with CLF control properties for stochastic time-variant systems with disturbed measurement feedback. More specifically, a continuous-valued state space and continuous-valued control inputs are considered. Two key ideas render a feasible solution. First, we approximate the CLF control

problem by a forward calculation derived from an OLF control approach. This is made possible by the assumptions that future feedback can be anticipated by a finite set of representative measurements and measurement feedback is not handled directly by the controller, but given in form of a state estimate. Second, an effective optimization method is employed, which exploits the solution of a simplified problem. Using homotopy continuation, this simplification is progressively transferred back into the original problem. The optimum is tracked by a gradient descent method.

Notation: Throughout this work, random variables \mathbf{x} are denoted in bold face letters, while deterministic quantities x are in normal letters. Vector-valued quantities \mathbf{x} are made distinguishable from scalar quantities x by underlining the corresponding variables. The notation $\underline{\mathbf{x}}_k \sim f_k^{\mathbf{x}}(\underline{\mathbf{x}}_k)$ denotes that $\underline{\mathbf{x}}_k$ is characterized by the probability distribution $f_k^{\mathbf{x}}(\underline{\mathbf{x}}_k)$, where the index k assigns a variable or a function to a specific time-step. Finally, a matrix \mathbf{A} is written in bold face capital letters.

Outline: In the following section, the formal problem is stated. We introduce the considered system and present the general closed-loop control problem. In Sec. III, we approximate the original problem using a finite set of representative measurements and derive a forward calculation approach. This forward calculation approach is used to minimize the cost function with respect to the control inputs by means of an optimization based on homotopy continuation method. Finally, in Sec. IV, we evaluate the presented work based on an experiment using a miniature walking robot and Sec. V concludes this work.

II. PROBLEM FORMULATION

In this work, we consider time-variant stochastic nonlinear systems of the form

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

where the system state $\underline{\mathbf{x}}_k \in \mathbb{R}^x$ is a random variable characterized by the probability density $f_k^{\mathbf{x}}(\underline{\mathbf{x}}_k)$. The initial state $\underline{\mathbf{x}}_0 \sim f_0^{\mathbf{x}}(\underline{\mathbf{x}}_0)$ is known and the control input vector $\underline{\mathbf{u}}_k \in \mathbb{R}^u$ is chosen from the bounded set $\mathcal{U} = [\underline{u}_k^{\min}, \underline{u}_k^{\max}]$. State propagation is performed by the system function $a: \mathbb{R}^x \times \mathcal{U} \rightarrow \mathbb{R}^x$, which is affected by independent and identically distributed (i.i.d.) noise $\underline{\mathbf{w}}_k \in \mathbb{R}^x$, where $\underline{\mathbf{w}}_k$ is characterized by $f_k^{\mathbf{w}}(\underline{\mathbf{w}}_k)$. We assume the state $\underline{\mathbf{x}}_k$ not to be directly accessible. It can only be observed according to the following equation

$$\underline{\mathbf{y}}_k = h(\underline{\mathbf{x}}_k) + \underline{\mathbf{v}}_k, \quad (2)$$

where $\underline{\mathbf{y}}_k \in \mathbb{R}^y$ is the measurement output characterized by $f_k^{\mathbf{y}}(\underline{\mathbf{y}}_k)$. The measurement function $h: \mathbb{R}^x \rightarrow \mathbb{R}^y$ is disturbed by the i.i.d. noise $\underline{\mathbf{v}}_k \in \mathbb{R}^y$ with $\underline{\mathbf{v}}_k \sim f_k^{\mathbf{v}}(\underline{\mathbf{v}}_k)$.

Example System: Miniature Walking Robot

Let us visualize this class of systems by a simple example. We assume an autonomous walking robot moving with constant speed following a given path. We are only interested in the

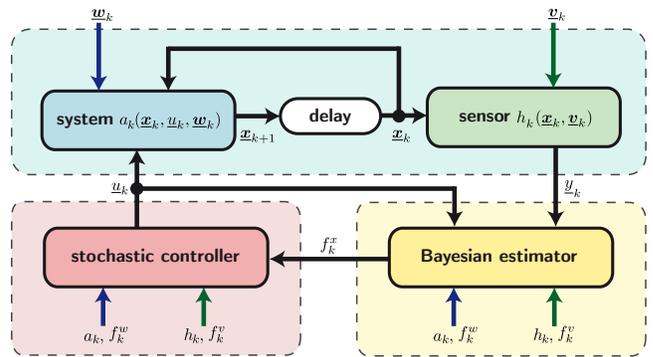


Fig. 2. The system structure considered in this work.

distance deviation to the path and the relative orientation of the robot, i.e, the system state $\underline{\mathbf{x}}_k = [\mathbf{x}_k, \phi_k]^T$. The robot can be controlled, by changing its orientation by the control input u_k , which leads to the two-dimensional process model

$$\underline{\mathbf{x}}_{k+1} = \begin{bmatrix} \mathbf{x}_{k+1} \\ \phi_{k+1} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_k + \sin(\phi_k + u_k) \\ \phi_k + u_k \end{bmatrix} + \begin{bmatrix} \mathbf{w}_k^x \\ \mathbf{w}_k^\phi \end{bmatrix},$$

where $\underline{\mathbf{w}}_k = [\mathbf{w}_k^x, \mathbf{w}_k^\phi]^T$ is a disturbance affecting the position and orientation of the state, respectively. The state itself cannot be observed directly. Instead, the robot is equipped with a 360° distance sensor, which measures the closest distance to objects in the robots environment. Therefore a one-dimensional sensor model is given by

$$y_k = d(x_{obs}, \mathbf{x}_k) + v_k,$$

where $d()$ is a distance function, x_{obs} is the positions of closest obstacle, and v_k is an additive noise term. We have visualized the given example by depicting the robot in his environment in Fig. 1.

We consider closed-loop feedback model predictive control, which takes into account that future control decisions will be based on more available feedback information. In general, this leads to the optimization of functions $\mu_k()$ mapping the initial condition, available measurements, and already applied control inputs to new control inputs, i.e., $\mu_k(f_0^{\mathbf{x}}, \underline{y}_{1:k}, \underline{u}_{0:k-1}) = \underline{u}_k$. In the following, we assume that the state conditional probability density $f_k^{\mathbf{x}}(\underline{\mathbf{x}}_k | f_0^{\mathbf{x}}, \underline{y}_{1:k}, \underline{u}_{0:k-1})$ characterizing the state estimate $\underline{\mathbf{x}}_k$ is a sufficient statistic. We assume, the controller has no direct access to the measurement feedback $\underline{\mathbf{y}}_k$, but can only make decisions based on $\underline{\mathbf{x}}_k$, which is calculated recursively via Bayes' law. We have illustrated the system structure in Fig. 2.

Let us denote the admissible control policy consisting of a sequence of functions by $\pi = \{\mu_0(), \mu_1(), \dots\}$. The objective is to find an optimal policy π^* , minimizing the expected cumulative cost function

$$J^\pi = \mathbf{E} \left\{ g_N(\underline{\mathbf{x}}_N) + \sum_{k=0}^{N-1} g_k(\underline{\mathbf{x}}_k, \mu_k(\underline{\mathbf{x}}_k)) \right\} \quad (3)$$

over a control horizon N , where $g_k: \mathbb{R}^x \times \mathcal{U} \rightarrow \mathbb{R}^+$ is a one-step cost function mapping every state-input pair to a real number and $g_N: \mathbb{R}^x \rightarrow \mathbb{R}^+$ denotes the terminal cost.

The expectation is subject to the noise $\underline{w}_{0:N-1}$ and $\underline{v}_{0:N-1}$, and the initial condition \underline{x}_0 . Thus, we are looking for π^* such that $J^{\pi^*} = \min_{\pi \in \Pi} J^\pi$, where Π is the set of all admissible policies.

Using Bellmans principle of optimality, the minimal cost J^{π^*} and consequently the closed-loop optimal input sequence $U^* = \underline{u}_{0:N-1}^*$, can be obtained by nested optimization problems minimizing the expected cost conditioned on available information at each time step. Solving this optimization problem has to be worked from inside out, i.e., by starting the calculation from the innermost expectation, which leads to the well-known backward-recursive formulation based on the concept of dynamic programming. Unfortunately, this is computationally not feasible in this general formulation [1].

III. APPROXIMATION OF CLOSED-LOOP FEEDBACK CONTROL

The CLF MPC problem implies an optimization of functions, resulting from the principle of optimality. Thus, CLF optimal policy can be computed by

$$J^{CLF} = \min_{u_0} \mathbf{E} \left\{ g(\underline{x}_0, \underline{u}_0) + \min_{u_1} \mathbf{E} \left\{ g(\underline{x}_1, \underline{u}_1) + \min_{u_2} \dots \dots + \min_{u_{N-1}} \mathbf{E} \left\{ g(\underline{x}_{N-1}, \underline{u}_{N-1}) + g(\underline{x}_N) | \underline{x}_{k-1} \right\} \dots \dots | \underline{x}_2 \right\} | \underline{x}_1 \right\} | \underline{x}_0 \right\},$$

which are nested optimization problems minimizing expected cost conditioned on the available information at each time step. Due to this nested formulation, this computation is not solvable for nonlinear systems [2].

A. Derivation from an Open-Loop Feedback Policy

We propose to minimize the cost over the control horizon conditioned on the available information at the given time step. This is described by the policy

$$\begin{aligned} J^{OLF} &= \min_{u_0:u_{N-1}} \mathbf{E} \left\{ g_N(\underline{x}_N) + \sum_{k=0}^{N-1} g_k(\underline{x}_k, \underline{u}_k) \mid \underline{x}_0, u_{0:N-1} \right\} \\ &= \min_{u_0:u_{N-1}} \mathbf{E} \left\{ g_{0:N}(\underline{x}_{0:N}, \underline{u}_{0:N-1}) \mid \underline{x}_0, u_{0:N-1} \right\}, \end{aligned} \quad (4)$$

which, on the other hand, in an open-loop feedback (OLF) policy. In this case the measurement feedback is not considered in the optimization process. As a matter of fact, consideration of all future measurements using this forward formulation is equivalent to not considering any future feedback. The future possible measurements are removed by taking the expectation, i.e., in short form, the expectation

in eq. (4) can be calculated by

$$\begin{aligned} &\int_x g_{0:N}(\underline{x}_{0:N}, \underline{u}_{0:N-1}) f^P(\underline{x}_{1:N} | \underline{x}_0, \underline{u}_{0:N-1}) d\underline{x}_{0:N} \\ &= \int_x \int_y f^y(\underline{y}_{1:N} | \underline{x}_0, \underline{u}_{0:N-1}) g_{0:N}(\underline{x}_{1:N}, \underline{u}_{0:N-1}) \\ &\quad \cdot f^e(\underline{x}_{0:N} | \underline{x}_0, \underline{y}_{1:N}, \underline{u}_{0:N-1}) d\underline{x}_{0:N} d\underline{y}_{1:N}, \end{aligned} \quad (5)$$

where we emphasized the prior conditional state densities by f^P as in *predicted state*, the posterior densities comprising all measurements by f^e as in *estimated state*, and the density of possible measurements by f^y . The proof is given in the Appendix.

In order to solve the problem of removing measurement considerations from the presented forward formulation, we propose to approximate all measurement probability densities $f^y(\underline{y}_k | \underline{x}_0, \underline{u}_{0:k})$ by a set of specific measurements, i.e.,

$$f^y(\underline{y}_k | \underline{x}_0, \underline{u}_{0:k}) \approx \sum_{j=1}^L \frac{1}{L} \cdot \delta(\underline{y}_k - \underline{y}_k^{(j)}),$$

where $\delta(\cdot)$ is a Dirac delta function and $\underline{y}_k^{(j)}$ denotes the position of the j -th Dirac component. This approximation can be used in various ways. First, it is possible to consider only the expected measurement $\hat{\underline{y}}_k = h_k(\mathbf{E}\{\underline{x}_k | \underline{u}_{0:k}\}, \underline{0})$. Second, in the case of bounded measurement noise, the worst measurement realization can be chosen, i.e., the one maximizing the cost function. A similar approach has been implemented for systems with finite sets of inputs and finite sets considered measurements in [19]. Finally, the density $f^y(\underline{y}_k | \underline{x}_0, \underline{u}_{0:k})$ can be approximated by means of a discrete probability density (e.g. by [20], [21], [22]). Note that the expected measurement may lead to a loss of robustness, while computation of the worst measurement realization results in an expensive min-max optimization problem. Finally, utilization of several specific measurements results in a tree like optimization structure, where the computational cost strongly depends on the number of measurement samples.

Using this formulation, the calculation can be performed in a forward prediction-filtering approach minimizing the cost function by optimizing over the control inputs, instead of functions. This can be interpreted as an OLF optimization with the property of considering future information feedback, which is an approximation of the closed-loop optimal policy.

B. Progressive Optimization

The main challenge of the now stated forward calculation is to find an optimal sequence of control inputs by minimizing the cost over the control horizon. This means that we have to solve a high-dimensional optimization problem. The evaluation of the expected cost implies Bayesian estimation over the control horizon and integration of the state estimates over the cost function. Considering the already high computational cost for prediction and estimation of nonlinear systems, we use expected measurements $\hat{\underline{y}}_k$ in the optimization method.

The key idea of the proposed optimization method is to use an easy computable control problem as an initial solution, which is done by system simplification. This simplified system is progressively transferred by means of a continuation method into the original system, while the solution is tracked by a gradient descent method.

First, we use a successive linearization approach for the system simplification, where the system is linearized around a nominal trajectory. Each step consists of system linearization around the nominal state $\hat{\underline{x}}_k$ and input \underline{u}_k by

$$\mathbf{A}_k = \frac{\partial a(\hat{\underline{x}}_k, \underline{u}_k)}{\partial \underline{x}_k}, \quad \mathbf{B}_k = \frac{\partial a(\hat{\underline{x}}_k, \underline{u}_k)}{\partial \underline{u}_k}, \quad \mathbf{H}_k = \frac{\partial h(\hat{\underline{x}}_k)}{\partial \underline{x}_k}.$$

The resulting system together with its nominal state

$$\begin{aligned} \underline{\mathbf{x}}_{k+1} &= \mathbf{A}_k \underline{\mathbf{x}}_k + \mathbf{B}_k \underline{u}_k + \underline{\mathbf{w}}_k, \\ \underline{\mathbf{y}}_k &= \mathbf{H}_k \underline{\mathbf{x}}_k + \underline{\mathbf{v}}_k \end{aligned}$$

are then used to calculate the next control input \underline{u}_{k+1} by means of a linear-quadratic-Gaussian (LQG) controller. The resulting input sequence $\underline{u}_{0:N-1}$, together with the sequence of linear systems $\mathbf{A}_{0:N-1}$, $\mathbf{B}_{0:N-1}$, $\mathbf{C}_{1:N}$, are then used as the starting point for the optimization procedure. Second, we progressively transfer the linear system to the original nonlinear system by means of a linear continuation method, given by

$$\begin{aligned} \underline{\mathbf{x}}_{k+1} &= \tilde{a}_k(\underline{\mathbf{x}}_k, \underline{u}_k, \gamma) + \underline{\mathbf{w}}_k, \\ \underline{\mathbf{y}}_k &= \tilde{h}_k(\underline{\mathbf{x}}_k, \gamma) + \underline{\mathbf{v}}_k, \end{aligned}$$

where $\gamma \in [0, 1]$ is the progression parameter. The linear progression functions for state prediction $\tilde{a}_k(\cdot, \cdot, \cdot)$ and measurement $\tilde{h}_k(\cdot, \cdot)$ are given by

$$\begin{aligned} \tilde{a}_k(\underline{\mathbf{x}}_k, \underline{u}_k, \gamma) &= \gamma \cdot \tilde{a}_k(\underline{\mathbf{x}}_k, \underline{u}_k) + (1 - \gamma) \cdot (\mathbf{A}_k \underline{\mathbf{x}}_k + \mathbf{B}_k \underline{u}_k), \\ \tilde{h}_k(\underline{\mathbf{x}}_k, \gamma) &= \gamma \cdot \tilde{h}_k(\underline{\mathbf{x}}_k) + (1 - \gamma) \cdot \mathbf{H}_k \underline{\mathbf{x}}_k. \end{aligned}$$

Thus, for $\gamma = 0$ the system is equal to the linearized starting point and for $\gamma = 1$ the original nonlinear system is considered. During the optimization γ is stepwise increased. Finally, we update the sequence of control inputs $\underline{u}_{0:N-1}$ for each step of the progression by a gradient descent approach. We generate for each calculated control sequence the corresponding trajectory and assumed measurements. This implies Bayesian filtering for nonlinear systems (e.g., see [23], [24], [25]). The gradient of the cost function is calculated around this trajectory and the control sequence is updated. This procedure is repeated until the optimization for $\gamma = 1$ is finished.

IV. EVALUATION

We have evaluated the presented approximate closed-loop MPC approach by implementing the example system described in Sec. II on a simulation of the miniature walking robot depicted in Fig. 1. The used robot is a 3D printed advanced version of the omnidirectional miniature walking robot first introduced in [26]. The implementation consists of the following setup. The robot can apply control inputs $u_k \in [-0.2, 0.2]$ updating its orientation in radian. The

TABLE I
EVALUATION RESULTS

Method	ϕ cost	# safety stops before crash
Approx. CLF (proposed)	8.47	0
Discrete Input CLF	9.39	0
Linearized LQG	16.75	12

initial position is set to $\underline{x}_0 = [5\text{cm}, 0]^T$. State estimation is conducted by the Unscented Kalman Filter (UKF) [23] and is initialized by $\underline{x}_0 \sim \mathcal{N}(\hat{\underline{x}}_0, \mathbf{C}_0)$, where

$$\hat{\underline{x}}_0 = \begin{bmatrix} 5 \\ 0 \end{bmatrix}, \quad \mathbf{C}_0 = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.03 \end{bmatrix}.$$

In order to construct a challenging control problem, we consider a tunnel scenario, where the path followed by the robot is enclosed by walls to both sides, which is depicted in Fig. 3. The walls are positioned at $x_{w1}, x_{w2} = \pm 17\text{cm}$. Thus considering the robot's center and rotation, the robot must not cross $x_{max} = \pm 12\text{cm}$ to avoid crashes. The system noise and measurement noise are modeled as Gaussian, zero mean, white-noise with $\underline{\mathbf{w}}_k \sim \mathcal{N}(\underline{0}, \mathbf{C}_k^w)$ and $\underline{\mathbf{v}}_k \sim \mathcal{N}(0, \mathbf{C}_k^v)$, respectively, where the covariance matrices are set to

$$\mathbf{C}_k^w = \begin{bmatrix} 0.3 & 0 \\ 0 & 0.001 \end{bmatrix}, \quad \mathbf{C}_k^v = [0.5].$$

We employ the quadratic cost function

$$g(\underline{\mathbf{x}}_k, \underline{u}_k) = \underline{\mathbf{x}}_k^T \begin{bmatrix} 0.1 & 0 \\ 0 & 1 \end{bmatrix} \underline{\mathbf{x}}_k + u_k^T [1] u_k$$

and evaluate the expected cost by stochastic numerical integration by approximation of the state estimate $\underline{\mathbf{x}}_k$ by a discrete density representation utilizing the technique introduced in [21].

We compare the proposed work to a CLF controller for systems with finite sets of control inputs proposed in [12]. Here, we use the finite set $\mathcal{U}_d = \{-0.2, -0.1, 0, 0.1, 0.2\}$ and use the control optimization horizon $N = 3$. Furthermore, we compare these approaches to a linear controller, where the system is linearized locally at the point of the nominal state estimate $\mathbf{E}\{\underline{\mathbf{x}}_k^e\}$ and is controlled by a linear quadratic Gaussian (LQG) regulator. Using this method on the other hand, separation of estimation and control for the linearized system holds and the CLF-optimal policy is equal to the

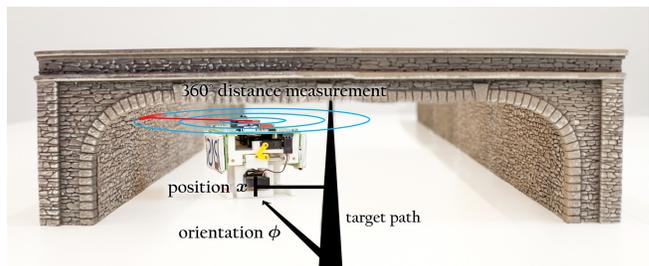


Fig. 3. The evaluated tunnel scenario.

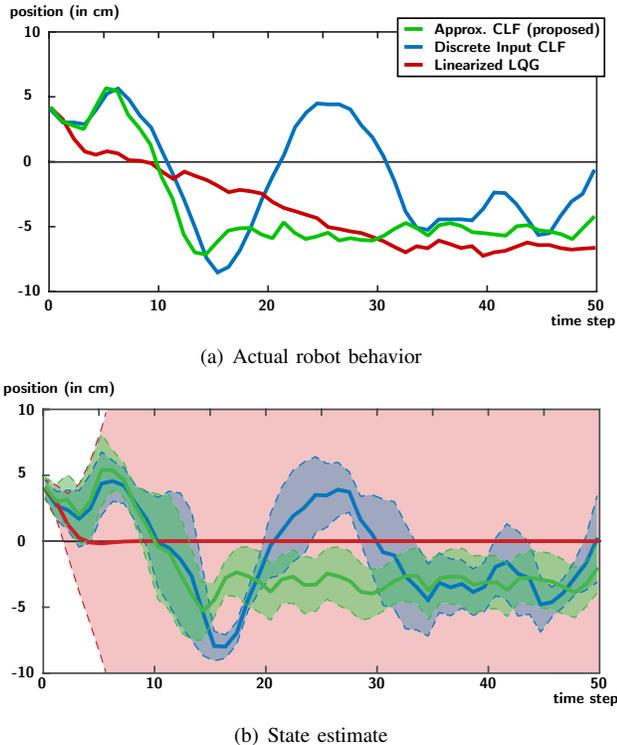


Fig. 4. An exemplary evaluation of the example system for the three compared approaches. (a) depicts the behavior of the real system, whereas (b) shows the corresponding state estimate. The solid lines represent the mean and the dashed lines together with the shaded colors indicate the uncertainty showing the $\sigma = 3$ bound.

deterministic policy. This unfortunately does not hold for the original nonlinear system.

We have summarized the evaluation of 50 performed runs over 50 time-steps in Tab. I. As can be seen, the proposed approximation of closed loop feedback control, as well as the CLF control with discretized control input is able to control the robot without leading it into a safety hold. The slightly worse quality of the second is due to the application of discretized control inputs and thus, has less precise control options. On the other hand, the assumed CE in the LQG approach leads to poor control performance and safety holds caused by imminent crashes.

For further analysis of the control behavior, we have plotted the trajectory of the robot of one exemplary outcome for each of the three compared approaches in Fig. 4. Prominent in the estimation plot is that the LQG control approach, holds the control target perfectly, but the uncertainty rises beyond the scale of the plot. This is given rise to by the fact that the state cannot be reconstructed correctly in the tunnel center. With both walls equally far away, in this position it is undecidable which obstacle wall has been measured. This interpretation is also supported by the actual robot trajectory, which can be seen as random walk. In blue, the discrete input CLF control uses some oversteering, resulting in frequent changes of the approached side. This results in the increased control cost. On the other hand it implicitly implements a

good trade-off between information gain and control cost, leading to a good overall control quality. A slightly better result can be seen with the proposed method in green. After an initial phase, the robot finds a good line on one side and tries to optimize the distance of the desired path such that the information acquired by the feedback sensor is sufficient for the control loop. This keeps the uncertainty of the state estimate low and the resulting control cost low.

Summarizing the evaluation results, we can say that the approximation of CLF control solves the problem of simultaneously optimizing information, which is crucial in systems, where observability is dependent on the control strategy. Thus, both approaches considering the dual effect show a good coverage of estimated state and the real system behavior. This is also emphasized by Fig. 5, where we plotted the average determinant of the state estimate over 50 Monte-Carlo simulation runs. In contrast, the control policy of the CE approach is to stay at a point, where no detectability is given and thus, the uncertainty rises after few steps beyond the range of the plot.

V. CONCLUSION AND FUTURE WORK

We present a novel control approach for stochastic nonlinear time-variant systems with imperfect state information. The system is considered for continuous state and control input spaces. The presented method effectively exploits information about future measurement feedback into an otherwise open-loop feedback optimization. This enables the consideration of the dual effect of control and estimation similar to a closed-loop optimal controller. The optimization is performed by a continuation method, which progressively transfers a linearized starting point into the original nonlinear system, while keeping track of the minimum. The minimum is tracked by a gradient descent method that calculates the trajectory and the assumed measurements for a given input sequence and updates the sequence by linearizing the cost function around this trajectory.

Major benefit of the presented approach is that the dual effect of control and estimation is considered in an effective way and issues involving the observability can be handled implicitly. Though the introduced optimization method al-

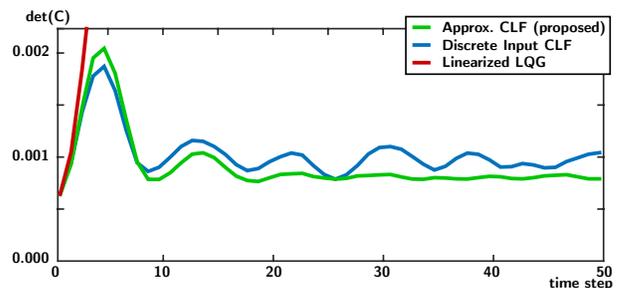


Fig. 5. Visualization of the estimation quality performed during the control. The Bayesian filtering was performed for all methods by the UKF [23]. The plot shows the development of the covariance matrix over time, by means of its determinant averaged over all 50 runs.

lows for an efficient calculation, the heavy computational load from recurring evaluation of nonlinear Bayesian filtering restricts this approach to a relatively short control horizon.

Future work includes the improvement of the continuation method by a prediction-correction approach and the implementation of more effective optimum tracking methods to increase the performance. Furthermore, effective optimization methods for more robust measurement strategies, as introduced in Sec. III-A, have to be developed.

APPENDIX

Proof of eq. (5): Let us consider all possible measurements of an arbitrary time step based on the predicted state probability density f^p . Omitting the time index, we can write this as a probability density

$$f^y(y) = \int_x f(y|x) \cdot f^p(x) dx. \quad (6)$$

Furthermore, we can write the posterior f^e in terms of the prior f^p by the Bayesian filtering step

$$f^e(x|y) = c(y) \cdot f(y|x) \cdot f^p(x), \quad (7)$$

where $c(y) = 1/\int f(y|x) \cdot f^p(x) dx$ is a normalization constant. Using the filter step (7) we can rewrite the right-hand side of eq. (5) by

$$\int_x \int_y c(y_{1:N}) \cdot f^y(\underline{y}_{1:N} | \underline{x}_0, \underline{u}_{0:N-1}) \cdot g_{0:N}(\underline{x}_{1:N}, \underline{u}_{0:N-1}) \cdot f(y_{1:N} | x_0, u_{0:N-1}) \cdot f^p(x_{0:N} | x_0, u_{0:N-1}) dx_{0:N} dy_{1:N}.$$

Using eq. (6) to rewrite the measurement density f^y , we obtain

$$f^y(\underline{y}_{1:N} | \underline{x}_0, \underline{u}_{0:N-1}) = 1/c(y_{1:N}),$$

which then leads to

$$\int_x g_{0:N}(\underline{x}_{1:N}, \underline{u}_{0:N-1}) \cdot f^p(x_{0:N} | x_0, u_{0:N-1}) dx_{0:N} \cdot \int_y f(y_{1:N} | x_0, u_{0:N-1}) dy_{1:N},$$

where $f(y_{1:N} | x_0, u_{0:N-1})$ is a probability density and thus, the second part integrates to one. This yields the desired result.

REFERENCES

- [1] D. Bertsekas, "Dynamic Programming and Optimal Control", Athena Scientific, Vol. 1., 3rd ed., 1995.
- [2] Y. Bar-Shalom and E. Tse, "Dual Effect, Certainty Equivalence, and Separation in Stochastic Control", in *IEEE Transactions on Automatic Control*, vol. 19, no. 5, pp.494–500, Oct. 1974.
- [3] D. Bernardini and A. Bemporad, "Scenario-based Model Predictive Control of Stochastic Constrained Linear Systems", in *Proc. of the 48rd IEEE Conference on Decision and Control (CDC 2009) & the 28th Chinese Control Conference*, Dec. 2009.
- [4] G.C. Calafiore and L. Fagiano, "Robust Model Predictive Control via Scenario Optimization", in *IEEE Transactions on Automatic Control*, vol.58, no.1, pp. 219–224, Jan. 2013.
- [5] G. Schildbach, L. Fagiano, C. Frei, and M. Morari, "The scenario approach for Stochastic Model Predictive Control with bounds on closed-loop constraint violations", in *Automatica*, Vol. 50, No. 12, pp. 3009–3018, Dec. 2014.

- [6] X. Zhang, S. Grammatico, K. Margellos, P.J. Goulart, and J. Lygeros, "Randomized nonlinear MPC for uncertain control-affine systems with bounded closed-loop constraint violations", in *Proc. of the 19th IFAC World Congress (IFAC 2014)*, Aug. 2014.
- [7] A. Mesbah, S. Streif, R. Findeisen, and R. D. Braatz, "Stochastic nonlinear model predictive control with probabilistic constraints," in *Proc. of the 2014 American Control Conference (ACC)*, Jun. 2014.
- [8] R. Tempo, G. Calafiore, and F. Dabbene, "Randomized Algorithms for Analysis and Control of Uncertain Systems", Springer, 2005.
- [9] G. C. Goodwin, J. Østergaard, D. E. Quevedo, and A. Feuer, "A Vector Quantization Approach to Scenario Generation for Stochastic NMPC", in *Nonlinear Model Predictive Control*, Lecture Notes in Control and Information Sciences, vol. 384, pp 235–248, Springer, 2009.
- [10] M. P. Deisenroth, T. Ohtsuka, F. Weissel, D. Brunn, and U. D. Hanebeck, "Finite-Horizon Optimal State-Feedback Control of Nonlinear Stochastic Systems Based on a Minimum Principle", in *Proc. of the 2006 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2006)*, Sep. 2006.
- [11] M. P. Deisenroth, F. Weissel, T. Ohtsuka, and U. D. Hanebeck, "Online-Computation Approach to Optimal Control of Noise-Affected Nonlinear Systems with Continuous State and Control Spaces", in *Proc. of the 2007 European Control Conference (ECC 2007)*, Jul. 2007.
- [12] 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 *Proc. of the 2008 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2008)*, Aug. 2008.
- [13] L. P. Kaelbling, M. L. Littman, and A. R. Cassandra, "Planning and acting in partially observable stochastic domains", in *Artificial Intelligence Journal*, vol. 101, no. 1-2, pp 99–134, May 1998.
- [14] A. Brooks, A. Makarenko, S. Williams, and H. Durrant-Whyte, "Parametric POMDPs for Planning in Continuous State Spaces", in *Robotics and Autonomous Systems*, vol 54, no. 11, pp. 887–897, 2006.
- [15] E. Zhou, M. C. Fu, and S. I. Marcus., "Solving Continuous-State POMDPs Via Density Projection", *IEEE Transactions on Automatic Control*, vol. 55, no. 5, pp. 1101-1116, 2010.
- [16] J. van den Berg, S. Patil, and R. Alterovitz, "Efficient Approximate Value Iteration for Continuous Gaussian POMDPs", in *Proc. of the AAAI Conference on Artificial Intelligence*, 2012.
- [17] S. Thrun, "Monte Carlo POMDPs", *Neural Information Processing Systems (NIPS) Foundation*. Vol. 12., pp. 1064–1070, 1999.
- [18] C. Chlebek and U. D. Hanebeck, "Stochastic Nonlinear Model Predictive Control Based on Deterministic Scenario Generation", in *Proc. of the 18th International Conference on Information Fusion (Fusion 2015)*, Jul. 2015.
- [19] D. Lyons, A. Hekler, M. Kuderer, and U. D. Hanebeck, "Robust Model Predictive Control with Least Favorable Measurements", in *Proc. of the 2010 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2010)*, Sep. 2010.
- [20] U. D. Hanebeck and V. Klumpp, "Localized Cumulative Distributions and a Multivariate Generalization of the Cramér-von Mises Distance", in *Proc. of the 2008 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2008)*, Aug. 2008.
- [21] U. D. Hanebeck, M. F. Huber, and V. Klumpp, "Dirac Mixture Approximation of Multivariate Gaussian Densities", in *Proc. of the 2009 IEEE Conference on Decision and Control (CDC 2009)*, Dec. 2009.
- [22] I. Gilitschenski and U. D. Hanebeck, "Efficient Deterministic Dirac Mixture Approximation", in *Proc. of the 2013 American Control Conference (ACC 2013)*, Jun. 2013.
- [23] S. J. Julier and J. K. Uhlmann, "Unscented Filtering and Nonlinear Estimation", in *Proc. of the IEEE*, Vol. 92, No. 3, pp. 401–422, Mar. 2004.
- [24] J. Steinbring and U. D. Hanebeck, "S2KF: The Smart Sampling Kalman Filter", in *Proc. of the 16th International Conference on Information Fusion (Fusion 2013)*, Jul. 2013.
- [25] U. D. Hanebeck, "PGF 42: Progressive Gaussian Filtering with a Twist" in *Proc. of the 16th International Conference on Information Fusion (Fusion 2013)*, Jul. 2013.
- [26] 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 *Proc. of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2007)*, Nov. 2007.