Global optimal control of quantized systems

Lars Grüne and Florian Müller

***Abstract*— We propose a set oriented approach to the global infinite horizon optimal control of nonlinear systems with quantized state measurement and quantized control values. The algorithm relies on a dynamic programming principle in which the quantization error is modelled as an opponent in a min-max dynamic game formulation. For the solution of the problem we propose a set oriented approach followed by a graph theoretic optimization algorithm. We also discuss a dynamic feedback extension and illustrate the performance of the proposed approach by experimental results.**

I. Introduction

The controller design for quantized systems can be carried out in different ways. One way is to first design a controller ignoring the quantization in state measurements and input variables and then look for a quantizer which ensures good performance (i.e., preservation of stability) under quantization. An example for this approach is the quantizer design proposed in [7], a paper which also gives a good survey on other approaches in the field. Another approach is to consider the quantization of state and input as given and try to design a controller taking the quantization into account.

In this paper, we consider the latter approach for an optimal feedback control problem for nonlinear discrete time systems, i.e., we assume that quantizations of both the state space and the input space are given a priori. Here the discrete time system typically forms a discrete time model of a continuous time sampled-data system. The control task then consists in steering the system to some desired target with minimal costs using a feedback control which is only allowed to use the quantized state measurements and the quantized control values. For this problem, a solution was presented by the authors in [4], based on earlier results from [2], [3], [6]. The procedure relies on modelling the uncertainty induced by the state quantization as a perturbation and extending the original optimal control problem to a zero sum differential game. This leads to a min-max problem in which minimization is performed over the control values and maximization over the perturbation induced by the quantization. Instead of solving the resulting dynamic programming equation directly, the resulting non-determinstic set oriented control system is then represented by a hypergraph on which the optimal control problem can be efficiently solved by a min-max version of Dijkstra's shortest path algorithm [3], [9].

This work was supported by the DFG priority program 1305.

L. Grüne is with Mathematical Institut, University of Bayreuth, 95440 Bayreuth, Germany lars.gruene@uni-bayreuth.de

F. Müller is with Mathematical Institut, University of Bayreuth, 95440 Bayreuth, Germany florian.mueller@uni-bayreuth.de

In this paper, we extend and improve the approach from [4] by constructing the hypergraph in a different way. Essentially, this amounts to a re-interpretation of the results for event-based control from [1] and [5] in a quantized setting. In [4], each edge in the hypergraph corresponds to one evaluation of the discrete time dynamics for a pair of state and control values, i.e., to the state transition in one sampling period. In contrast to this, here each edge corresponds to the state transition until the next quantization region is reached.

Adding this idea, which is developed in Section II, to the algorithm from [4] already significantly improves the results. However, the feedback law resulting from this algorithm only takes into account the quantization region containing the current state. Using a dynamic feedback approach, in which the control value is allowed to depend not only on the current quantization region but also on previous regions, one can narrow down the uncertainty induced by the quantization and thus reduce the conservativity of the approach. In Section III we present this extension, which was motivated by conceptually similar methods in the discrete event system literature, see, e.g., [8] and the references therein, and in Section IV we theoretically compare it with the basic algorithm from Section II. Finally, in order to demonstrate the efficiency of the method, in Section V we illustrate our approach with experimental results obtained at the test plant "VERA" at the Ruhr-Universität Bochum, which extend the results documented in [1]. Here we again compare the approaches from Section II and Section IV.

II. Problem formulation

We consider the discrete-time nonlinear control system

$$\boldsymbol{x}(k+1) = \boldsymbol{f}(\boldsymbol{x}(k), \boldsymbol{u}(k)), \quad k = 0, 1, \dots \ , \tag{1}$$

where $\boldsymbol{f} : \mathcal{X} \times \mathcal{U} \longrightarrow \mathcal{X}$ is continuous, $\boldsymbol{x}(k) \in \mathcal{X}$ is the state of the system, $\boldsymbol{u}(k) \in \mathcal{U}$ is the control input, chosen from compact sets $\mathcal{X} \subset \mathbb{R}^n$ and $\mathcal{U} \subset \mathbb{R}^m$. The set of all control sequences $\underline{\boldsymbol{u}} = \boldsymbol{u}(k)_{k\in\mathbb{N}}$ is denoted by $\mathcal{U}^{\mathbb{N}}$ and for each initial value $\boldsymbol{x}_0$ and control sequence $\underline{\boldsymbol{u}}$ we denote the corresponding trajectory by $x(k, \boldsymbol{x}_0, \underline{\boldsymbol{u}})$. Throughout the paper we interpret (1) as a discrete time model for a continuous time sampled-data system.

The control problem we consider is as follows: Given a target set $\mathcal{X}^* \subset \mathcal{X}$, steer the system into $\mathcal{X}^*$ while minimizing the functional

$$J(\boldsymbol{x}_0, \underline{\boldsymbol{u}}) = \sum_{k=0}^{N(\boldsymbol{x}_0, \underline{\boldsymbol{u}})} c(x(k, \boldsymbol{x}_0, \underline{\boldsymbol{u}}), \boldsymbol{u}(k)) \tag{2}$$

over $\underline{\boldsymbol{u}}$, where $N(\boldsymbol{x}_0, \underline{\boldsymbol{u}})$ denotes the minimal $k \geq 0$ such that $x(k, \boldsymbol{x}_0, \underline{\boldsymbol{u}}) \in \mathcal{X}^*$ holds. Here $c : \mathcal{X} \times \mathcal{U} \to \mathbb{R}$ is a

continuous running cost satisfying $\min_{\boldsymbol{u}\in\mathcal{U}} c(\boldsymbol{x},\boldsymbol{u}) > 0$ for all $\boldsymbol{x} \notin \mathcal{X}^*$.

Our goal now is to find a feedback law which approximately solves this problem, assuming, however, that the system's state is not exactly determinable. In order to formalize this uncertainty, we use a partition P of the state space $\mathcal{X}$ consisting of finitely many connected and disjoint subsets $\mathcal{P}_i \subset \mathcal{X}$ with the properties

$$\begin{aligned} &\textstyle\bigcup_{\mathcal{P}_i\in P} \mathcal{P}_i = \mathcal{X} \quad \text{and} \\ &\mathcal{P}_i \cap \mathcal{P}_j = \emptyset \text{ for all } \mathcal{P}_i, \mathcal{P}_j \in P \text{ with } i \neq j. \end{aligned} \tag{3}$$

In contrast to, e.g., [2], [3] we do not interpret the sets $\mathcal{P} \in P$ as a discretization which we are able to change according to our demands. Rather, the subsets $\mathcal{P}_i$ of this partition model the quantization regions of the state measurements. Here we assume the partition P as given and do not address the question about how to choose good partitions. We assume that our target set $\mathcal{X}^*$ is a union of such regions, i.e., $\mathcal{X}^* = \bigcup_{\mathcal{P}\in P^*} \mathcal{P}$ for some set $\mathcal{P}^* \subset P$.

For the purpose of feedback control, we assume that at each time instant k the region $\mathcal{P}_i$ containing the current value $\boldsymbol{x}(k)$ is known to the controller. This allows to consider feedback maps of the form $\mu : P \to \mathcal{U}$ such that the feedback value for a state $\boldsymbol{x}$ as $\mu(\rho(\boldsymbol{x}))$, using the **correlation function** $\rho : \mathcal{X} \to P$ which to each state assigns the quantization region containing this state, i.e.,

$$\rho(\boldsymbol{x}) := \mathcal{P} \quad \text{ if } \boldsymbol{x} \in \mathcal{P}.$$

In order to extract the essential dynamical information with respect to the quantization, we do not consider the individual sampling times k but only the times at which the state passes from one quantization region to another. This is accomplished by defining the iterates $\boldsymbol{f}^r(\boldsymbol{x},\boldsymbol{u})$ for $r \in \mathbb{N}_0$, $\boldsymbol{x} \in \mathcal{X}$ and $\boldsymbol{u} \in \mathcal{U}$ as

$$\boldsymbol{f}^0(\boldsymbol{x},\boldsymbol{u}) := \boldsymbol{x}, \quad \boldsymbol{f}^{r+1}(\boldsymbol{x},\boldsymbol{u}) := \boldsymbol{f}(\boldsymbol{f}^r(\boldsymbol{x},\boldsymbol{u}),\boldsymbol{u})$$

and define the following value.

Definition 2.1: For each $\boldsymbol{x} \in \mathcal{X}$ with $\boldsymbol{x} \in \mathcal{P}_i$ and each $\boldsymbol{u} \in \mathcal{U}$ we define the value $r(\boldsymbol{x},\boldsymbol{u})$ to be the smallest value $r \in \mathbb{N}$ for which there exists $j \neq i$ with

$$\boldsymbol{f}^{r-1}(\boldsymbol{x},\boldsymbol{u}) \in \mathcal{P}_i, \quad \boldsymbol{f}^r(\boldsymbol{x},\boldsymbol{u}) \in \mathcal{P}_j.$$

In other words, $r(\boldsymbol{x},\boldsymbol{u})$ is the time when the state passes from the quantization region $\mathcal{P}_i$ to the quantization region $\mathcal{P}_j \neq \mathcal{P}_i$.

Formally, we could set $r(\boldsymbol{x},\boldsymbol{u}) = \infty$ if $\boldsymbol{f}^r(\boldsymbol{x},\boldsymbol{u}) \in \mathcal{P}_i$ for all $r \in \mathbb{N}_0$. For the practical implementation, we impose an upper bound $R \in \mathbb{N}_0$ for $r(\boldsymbol{x},\boldsymbol{u})$ and set $r(\boldsymbol{x},\boldsymbol{u}) := R$ whenever $r \geq R$ holds in Definition 2.1.

In order to specify the set valued system for our algorithm, we define the set $2^{\mathcal{X}}$ of all subsets of $\mathcal{X}$ and the set of sequences $(2^{\mathcal{X}})^{\mathbb{N}} := \{\underline{X} = (\mathcal{X}(0),\mathcal{X}(1),\ldots) \,|\, \mathcal{X}(i) \subset \mathcal{X} \text{ for all } i \in \mathbb{N}\}$ and use the following concept of choice functions.

Definition 2.2: A **choice function** $\boldsymbol{\gamma} : (2^{\mathcal{X}})^{\mathbb{N}} \times \mathcal{U}^{\mathbb{N}} \longrightarrow \mathcal{X}^{\mathbb{N}}$ is a function of the form

$$\boldsymbol{\gamma}(\underline{X},\underline{\boldsymbol{u}}) = (\widehat{\gamma}_0(\mathcal{X}(0),\boldsymbol{u}(0)),\widehat{\gamma}_1(\mathcal{X}(1),\boldsymbol{u}(1),\ldots),$$

with component functions $\widehat{\gamma}_i : 2^{\mathcal{X}} \times \mathcal{U} \longrightarrow \mathcal{X}$ satisfying $\widehat{\gamma}_i(\mathcal{X}(j),\boldsymbol{u}(j)) \in \mathcal{X}(j)$ for all $\mathcal{X}(j) \subseteq \mathcal{X}$, $\boldsymbol{u}(j) \in \mathcal{U}$. The set of all choice functions $\boldsymbol{\gamma}$ is denoted by $\mathcal{C}$ and the set of all component functions $\widehat{\gamma}$ by $\widehat{\mathcal{C}}$.
With the components $\widehat{\gamma}$ of the choice function $\boldsymbol{\gamma}$ we model the uncertainty of the state $\boldsymbol{x}$ induced by the quantization by choosing the perturbed state $\widehat{\gamma}(\mathcal{X},\boldsymbol{u}) \in \mathcal{X}$ depending on the control $\boldsymbol{u}$ in the region $\mathcal{X}$ containing $\boldsymbol{x}$. The choice functions $\boldsymbol{\gamma}$ then extend this concept to a sequence of regions and controls.

Using the concept of partitions and choice functions we now define a set valued control system by

$$\mathcal{X}(k+1) = F(\mathcal{X}(k),\boldsymbol{u}(k),\widehat{\boldsymbol{\gamma}}_k(\mathcal{X}(k),\boldsymbol{u}(k))), \tag{4}$$

$k = 0,1,\ldots$, with $F : 2^{\mathcal{X}} \times \mathcal{U} \times \widehat{\mathcal{C}} \to P$ given by

$$\begin{aligned} &F(\mathcal{X}(k),\boldsymbol{u}(k),\widehat{\boldsymbol{\gamma}}_k(\mathcal{X}(k),\boldsymbol{u}(k))) := \\ &\rho(\boldsymbol{f}^{r(\widehat{\boldsymbol{\gamma}}_k(\mathcal{X}(k),\boldsymbol{u}(k)),\boldsymbol{u}(k))}(\widehat{\boldsymbol{\gamma}}_k(\mathcal{X}(k),\boldsymbol{u}(k)),\boldsymbol{u}(k))). \end{aligned}$$

In what follows we will omit the arguments of $\widehat{\boldsymbol{\gamma}}_k$ in order to simplify the notation. The map F describes all possible transitions of a subset $\mathcal{X}_i \subset \mathcal{X}$ of the state space to regions $\mathcal{P} \in P$, parametrized by $\widehat{\boldsymbol{\gamma}}_k$. In other words, for each $\boldsymbol{u} \in \mathcal{U}$ we have the identity

$$\bigcup_{\widehat{\boldsymbol{\gamma}}\in\widehat{\boldsymbol{\mathcal{C}}}} F(\mathcal{X},\boldsymbol{u},\widehat{\boldsymbol{\gamma}}) = \{\mathcal{P} \in P \,|\, \boldsymbol{f}(\boldsymbol{x},\boldsymbol{u}) \in \mathcal{P} \text{ for some } \boldsymbol{x} \in \mathcal{X}\}.$$

A trajectory $\mathcal{X}(k,\mathcal{P}^0,\underline{\boldsymbol{u}},\boldsymbol{\gamma})$, $k \in \mathbb{N}$ of (4) is now a sequence of regions defined by

$$\mathcal{X}(0) = \mathcal{P}^0, \quad \mathcal{X}(k+1) = F(\mathcal{X}(k),\boldsymbol{u}(k),\widehat{\boldsymbol{\gamma}}_k)$$

and depends on the initial set $\mathcal{P}^0 \in P$, the control sequence $\underline{\boldsymbol{u}} \in \mathcal{U}^{\mathbb{N}}$ and the choice function $\boldsymbol{\gamma} \in \mathcal{C}$.

The next object defines the set of regions from which the system (4) can be steered to the target set $\mathcal{X}^*$ regardless of the choice of $\boldsymbol{\gamma}$.

Definition 2.3: The **domain of controllability** of $\mathcal{X}^*$ is defined as

$$\begin{aligned} S = \{\mathcal{P} \in P| \text{ for each } \boldsymbol{\gamma} \in \mathcal{C} \text{ there exists } \underline{\boldsymbol{u}} \in \mathcal{U}^{\mathbb{N}} \text{ and} \\ k \in \mathbb{N} \text{ with } \mathcal{X}(k,\mathcal{P},\underline{\boldsymbol{u}},\boldsymbol{\gamma}) \subset \mathcal{X}^*\}. \end{aligned}$$

and the **first hitting time** is defined as $N(\mathcal{P},\underline{\boldsymbol{u}},\boldsymbol{\gamma}) = \inf\{k \in \mathbb{N} | \mathcal{X}(k,\mathcal{P},\underline{\boldsymbol{u}},\boldsymbol{\gamma}) \in \mathcal{X}^*\}$.
Note that for fixed $\mathcal{P}$ we can interpret $\boldsymbol{\gamma}$ as a map from $\mathcal{U}^{\mathbb{N}}$ to $\mathcal{X}^{\mathbb{N}}$. In the language of dynamic game theory this map defines a nonanticipating strategy, cf. [2], [3].

Using the running cost c we now define a cost function for the set valued control system (4)

$$c_1 : P \times \mathcal{U} \longrightarrow \mathbb{R}_{+,0}, \quad c_1(\mathcal{P},\boldsymbol{u}) := \sup_{\boldsymbol{x}\in\mathcal{P}} c^{r(\boldsymbol{x},\boldsymbol{u})}(\boldsymbol{x},\boldsymbol{u}),$$

with

$$c^{r(\boldsymbol{x},\boldsymbol{u})}(\boldsymbol{x},\boldsymbol{u}) := \sum_{r=0}^{r(\boldsymbol{x},\boldsymbol{u})-1} c(\boldsymbol{f}^r(\boldsymbol{x},\boldsymbol{u}),\boldsymbol{u})$$

and $r(\boldsymbol{x},\boldsymbol{u})$ from Definition 2.1. By this definition we assume the worst case, i.e., the highest cost, over all the uncertain states $\boldsymbol{x} \in \mathcal{P}$. Using c_1 we now define the functional

$$J_1(\mathcal{P},\underline{\boldsymbol{u}},\boldsymbol{\gamma}) := \sum_{k=0}^{N(\mathcal{P},\underline{\boldsymbol{u}},\boldsymbol{\gamma})} c_1(\mathcal{X}(k,\mathcal{P},\underline{\boldsymbol{u}},\boldsymbol{\gamma}),\boldsymbol{u}_k)$$

with values in $\mathbb{R}_{+,0} \cup \{+\infty\}$ and the optimal value function

$$V_1(\mathcal{P}) = \sup_{\boldsymbol{\gamma}\in\mathcal{C}} \inf_{\underline{\boldsymbol{u}}\in\mathcal{U}^{\mathbb{N}}} J_1(\mathcal{P},\underline{\boldsymbol{u}},\boldsymbol{\gamma}).$$

By standard arguments one sees that V_1 fulfills the optimality principle

$$V_1(\mathcal{P}) = \inf_{\boldsymbol{u}\in\boldsymbol{\mathcal{U}}} \left\{ c_1(\mathcal{P},\boldsymbol{u}) + \sup_{\widehat{\boldsymbol{\gamma}}\in\widehat{\boldsymbol{\mathcal{C}}}} V_1(F(\mathcal{P},\boldsymbol{u},\widehat{\boldsymbol{\gamma}})) \right\} \quad (5)$$

for all $\mathcal{P} \not\subset \mathcal{X}^*$ and $V(\mathcal{P}) = 0$ for all $\mathcal{P} \subset \mathcal{X}^*$.

Since P consists of finitely many sets, from this it is easy to see by induction that

$$S = \{\mathcal{P} \in P \,|\, V_1(\mathcal{P}) < \infty\}.$$

In particular, the domain of controllability is easily obtained once V_1 is computed.

We will now investigate the behavior of V_1 along an optimal trajectory for the original system (1). To this end, observe that the optimal feedback law $\mu : P \to U$ is the control value realizing the minimum in (5), i.e.,

$$\mu(\mathcal{P}) = \operatorname*{argmin}_{\boldsymbol{u}\in\mathcal{U}} \left\{ c_1(\mathcal{P},\boldsymbol{u}) + \sup_{\widehat{\boldsymbol{\gamma}}\in\widehat{\mathcal{C}}} V_1(F(\mathcal{P},\boldsymbol{u},\widehat{\boldsymbol{\gamma}})) \right\}$$

Using this μ we get the following theorem.

Theorem 2.4: For all $\boldsymbol{x} \in \mathcal{X}$ with $\rho(\boldsymbol{x}) \in S$ the inequality

$$c^r(\boldsymbol{x},\mu(\rho(\boldsymbol{x}))) + V_1(\rho(\boldsymbol{f}^r(\boldsymbol{x},\mu(\rho(\boldsymbol{x}))))) \le V_1(\rho(\boldsymbol{x})) \quad (6)$$

holds for $r = r(\boldsymbol{x},\mu(\rho(\boldsymbol{x})))$.

Proof: Using the optimality principle (5) and the definition of μ, $\widehat{\boldsymbol{\gamma}}$ and c_1 we get

$$\begin{aligned} &V_1(\rho(\boldsymbol{x})) = \\ &\quad \inf_{\boldsymbol{u}\in\mathcal{U}} \left\{ c_1(\rho(\boldsymbol{x}),\boldsymbol{u}) + \sup_{\widehat{\boldsymbol{\gamma}}\in\widehat{\mathcal{C}}} V_1(F(\rho(\boldsymbol{x}),\boldsymbol{u},\widehat{\boldsymbol{\gamma}})) \right\} \\ &\quad = c_1(\rho(\boldsymbol{x}),\mu(\rho(\boldsymbol{x}))) + \sup_{\widehat{\boldsymbol{\gamma}}\in\widehat{\mathcal{C}}} V_1(F(\rho(\boldsymbol{x}),\mu(\rho(\boldsymbol{x})),\widehat{\boldsymbol{\gamma}})) \\ &\quad \ge c^r(\boldsymbol{x},\mu(\rho(\boldsymbol{x}))) + V_1(\rho(\boldsymbol{f}^r(\boldsymbol{x},\mu(\rho(\boldsymbol{x})))))) \end{aligned}$$

which shows the assertion. ∎

The result has an immediate consequence for the trajectory $x(k,\boldsymbol{x}_0,\mu)$ of (1) with feedback control μ defined by

$$\boldsymbol{x}(k+1) = \boldsymbol{f}(\boldsymbol{x}(k),\mu(\rho(\boldsymbol{x}(k))))).$$

The value $V_1(\rho(x(k,\boldsymbol{x}_0,\mu)))$ is decreasing in k until $\mathcal{X}^*$ is reached and strictly decreasing for each k in which the state passes from one quantization region to another. This implies that $x(k,\boldsymbol{x}_0,\mu)$ eventually reaches $\mathcal{X}^*$ provided $\rho(\boldsymbol{x}_0) \in S$ (or equivalently $V_1(\rho(\boldsymbol{x}(0)) < \infty)$ holds.

Remark 2.5: The advantage of defining the set oriented dynamics via the times when the system passes from one quantization region to another compared to the sampled data approach in [2]–[4] can be explained as follows: In these references the set valued map F is constructed directly from (1). Hence, if there exists $\mathcal{P}_i \in P$ and $\boldsymbol{x} \in \mathcal{P}_i$ with $\boldsymbol{f}(\boldsymbol{x},\boldsymbol{u}) \in \mathcal{P}_i$ for all $\boldsymbol{u} \in \mathcal{U}$, then $F(\mathcal{P}_i,u,\widehat{\gamma}(\mathcal{P}_i,u)) = P_i$ holds for $\widehat{\gamma}(\mathcal{P}_i,u) = \boldsymbol{x}$. In this case, the optimality principle (5) immediately implies $V(\mathcal{P}) = \infty$. Using $\boldsymbol{f}^{r(\boldsymbol{x},\boldsymbol{u})}(\boldsymbol{x},\boldsymbol{u})$ instead of $\boldsymbol{f}(\boldsymbol{x},\boldsymbol{u})$ for constructing F resolves this problem, because — unless $\boldsymbol{f}^r(\boldsymbol{x},\boldsymbol{u}) \in \mathcal{P}_i$ for all $r \ge 0$ or, in our practical implementation, for $r \in \{0,\dots,R\}$ — the set valued map F will always satisfy $F(\mathcal{P}_i,u,\widehat{\gamma}(\mathcal{P}_i,u)) \neq P_i$.

III. Including past information

The approach described in the previous section is conservative because by maximizing over $\boldsymbol{\gamma}$ we implicitly assume the worst case in each step along the trajectory, i.e., that for each k among all the possible states in $\mathcal{X}(k)$ the actual state $\boldsymbol{x}(k)$ is the one which produces the largest cost. Of course, this is not necessarily the case. The approach we propose in order to reduce the conservatism relies on the idea that at time k we consider the last m measurements in order to compute the feedback μ. This way we can collect more information, thus reduce the uncertainty of the system and consequently obtain a less conservative result. In other words, we are now looking at an approximately optimal feedback map of the form $\mu_{m+1}(\mathcal{X}(k-m),\dots,\mathcal{X}(k))$. Note that this construction resembles the dynamic feedback concept well known in observer design.

In order to keep the exposition simple, we restrict ourselves to $m = 1$. All arguments can, however, be extended to the more general setting $m \ge 1$. Our goal in this case is to find a feedback law $\mu_2(\mathcal{X}(k-1),\mathcal{X}(k))$.

To this end, we define $P_2 := (P \cup \{\delta\}) \times P$ and introduce a new set valued state $\mathcal{Z}(k) = (\mathcal{Z}_1(k),\mathcal{Z}_2(k))^T \in P_2$ which represents $(\mathcal{X}(k-1),\mathcal{X}(k))^T$.

For $\mathcal{Z}$ we define the set valued control system as

$$\begin{aligned} \mathcal{Z}(k+1) &= F_2(\mathcal{Z}(k),\boldsymbol{u}(k),\widehat{\boldsymbol{\gamma}}_k) \\ &:= \begin{pmatrix} \mathcal{Z}_2(k) \\ F(X(\mathcal{Z}(k)),\boldsymbol{u}(k),\widehat{\boldsymbol{\gamma}}_k) \end{pmatrix} \end{aligned} \quad (7)$$

with F from (4) and

$$X(\mathcal{Z}) := \begin{cases} \mathcal{Z}_2, & \text{if } \mathcal{Z}_1 = \delta \\ \bigcup\limits_{\boldsymbol{x}\in\mathcal{Z}_1 \boldsymbol{u}\in\mathcal{U}} f^{r(\boldsymbol{x},\boldsymbol{u})}(\boldsymbol{x},\boldsymbol{u}) \cap \mathcal{Z}_2, & \text{else} \end{cases} \quad (8)$$

with $r(x,u)$ from Definition (2.1). Here the symbol δ represents the "undefined" region, which appears when the system is started at time $k = 0$ with initial region $\mathcal{P}^0 \in P$ but undefined previous region $\mathcal{P}^{-1}$. Therefore, at time $k = 0$ a trajectory starts with the vector $\mathcal{Z}(0) = (\delta,\mathcal{P}^0)^T$.

By including the extra information in the definition of F_2 the uncertainty of the system is reduced. Instead of using

$F(\mathcal{X}(k), \boldsymbol{u}(k), \widehat{\boldsymbol{\gamma}}_{\boldsymbol{k}})$ as in the previous section we use now $F(\mathcal{X}(\mathcal{Z}(k)), \boldsymbol{u}(k), \widehat{\boldsymbol{\gamma}}_{\boldsymbol{k}})$, where $X(\mathcal{Z}(k))$ is a subset of the current region $\mathcal{X}(k)$. The set $X(\mathcal{Z}(k))$ contains only those states which can be reached from the past region $\mathcal{Z}_1(k) = \mathcal{X}(k-1)$, i.e., we exclude those states from $\mathcal{Z}_2$ which the system cannot reach.

Clearly, not all the pairs $\mathcal{Z} = (\mathcal{P}_i, \mathcal{P}_j)^T \in P_2$ are actually attained by the systems dynamics. In fact, only those pairs with $X(\mathcal{Z}) \neq \emptyset$ can appear on the left hand side of (7) which is why we define the active state regions $P_2^a := \{\mathcal{Z} \in P_2 \,|\, X(\mathcal{Z}) \neq \emptyset\}$. We denote the trajectories of (7) by $Z(k, Z_0, \mathbf{u}, \beta)$ and adapt the definitions from the previous section to our new setting.

The target set now becomes $\mathcal{Z}^* = \{\mathcal{Z} \in P_2 \,|\, \mathcal{Z}_2 \subseteq \mathcal{X}^*\}$ and the definition of the domain of controllability S and the first hitting time N changes accordingly. For the cost function

$$c_2 : P_2 \times \mathcal{U} \to \mathbb{R}_{+,0}, \quad c_2(\mathcal{Z}, \boldsymbol{u}) = \sup_{\boldsymbol{x} \in X(\mathcal{Z})} c^{r(\boldsymbol{x}, \boldsymbol{u})}(\boldsymbol{x}, \boldsymbol{u})$$

we define the functional

$$J_2(\mathcal{Z}, \underline{\boldsymbol{u}}, \boldsymbol{\gamma}) = \sum_{k=0}^{N(Z, \underline{\boldsymbol{u}}, \boldsymbol{\gamma})} c_2(\mathcal{Z}(k, \mathcal{Z}, \underline{\boldsymbol{u}}, \boldsymbol{\gamma}), \boldsymbol{u}_k) \ \in \mathbb{R}_{+,0} \cup \{+\infty\}$$

and the optimal value function

$$V_2(\mathcal{Z}) = \sup_{\boldsymbol{\gamma} \in \mathcal{C}} \inf_{\underline{\boldsymbol{u}} \in \mathcal{U}^{\mathbb{N}}} J_2(\mathcal{Z}, \underline{\boldsymbol{u}}, \boldsymbol{\gamma}).$$

V_2 again fulfills the optimality principle

$$V_2(\mathcal{Z}) = \inf_{\boldsymbol{u} \in \mathcal{U}} \left\{ c_2(\mathcal{Z}, \boldsymbol{u}) + \sup_{\widehat{\boldsymbol{\gamma}} \in \widehat{\boldsymbol{\mathcal{C}}}} V_2(F_2(\mathcal{Z}, \boldsymbol{u}, \widehat{\boldsymbol{\gamma}})) \right\} \tag{9}$$

The optimal feedback $\mu_2(\mathcal{Z})$ is given by the argmin of this expression. The following theorem is the counterpart of Theorem 2.4.

Theorem 3.1: For all $\boldsymbol{x} \in \mathcal{X}$ and all $\mathcal{Z} \subset S$ with $\boldsymbol{x} \in \mathcal{X}(\mathcal{Z})$ the inequality

$$c^r(\boldsymbol{x}, \mu_2(\mathcal{Z})) + V_2((\rho(\boldsymbol{x}), \rho(\boldsymbol{f}^r(\boldsymbol{x}, \mu_2(\mathcal{Z})))^T) \leq V_2(\mathcal{Z})$$

holds for $r = r(\boldsymbol{x}, \mu_2(\mathcal{Z}))$. In particular, the inequality holds for $\mathcal{Z} = (\delta, \rho(\boldsymbol{x}))^T$.

Proof: Completely analogous to Theorem 2.4. ∎

IV. Comparison of the two approaches

In the preceding sections we have introduced the optimal value functions V_1 and V_2 and the corresponding feedback laws μ and μ_2. In this section we now show that V_1 is an upper bound for V_2. In [4] a similar theorem for the sampled data approach is proven.

Theorem 4.1: The optimal value functions V_1 and V_2 satisfy

$$V_2(\mathcal{Z}) \leq V_1(\mathcal{P}) \quad \text{for all } \mathcal{Z} \in P_2^a,\ \mathcal{P} \in P \text{ with } Z_2 = \mathcal{P}.$$

Proof: We prove the theorem by induction over the elements $\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_l \in P$ which we number according to their values in the optimal value function V_1, i.e., $V_1(\mathcal{P}_i) \leq V_1(\mathcal{P}_j)$ for all $1 \leq i < j \leq l$. We will frequently use the obvious inclusion $X(\mathcal{Z}) \subseteq \mathcal{Z}_2$ for $X(\mathcal{Z})$ from (8) and all $\mathcal{Z} = (\mathcal{Z}_1, \mathcal{Z}_2)^T \in P_2$.

Induction start $n = 1$:
Since $V_1(\mathcal{P}) = 0$ holds if and only if $\mathcal{P} \subseteq \mathcal{X}^*$ we obtain $\mathcal{P}_1 \subseteq \mathcal{X}^*$. Since $\mathcal{Z} \subseteq \mathcal{Z}^*$ for all $\mathcal{Z} \in P_2$ with $\mathcal{Z}_2 = \mathcal{P}_1 \subseteq \mathcal{X}^*$ we obtain $V_2(\mathcal{Z}) = 0 = V_1(\mathcal{P}_1)$ and thus the assertion for $\mathcal{P}_1$.

Induction step $n \to n+1$:
We use the induction hypothesis $V_2(\mathcal{Z}) \leq V_1(\mathcal{P}_j)$ for all $j = 0, \ldots, n$ and all $\mathcal{Z} \in P_2^a$ with $\mathcal{Z}_2 = \mathcal{P}_j$ in order to show $V_2(\mathcal{Z}) \leq V_1(\mathcal{P}_{n+1})$ for all $\mathcal{Z} \in P_2^a$ with $\mathcal{Z}_2 = \mathcal{P}_{n+1}$.
The optimality principle for V_1 yields

$$\begin{aligned} V_1(\mathcal{P}_{n+1}) &= \\ &\inf_{\boldsymbol{u} \in \mathcal{U}} \left\{ c_1(\mathcal{P}_{n+1}, \boldsymbol{u}) + \sup_{\widehat{\boldsymbol{\gamma}} \in \widehat{\mathcal{C}}} V_1(F(\mathcal{P}_{n+1}, \boldsymbol{u}, \widehat{\boldsymbol{\gamma}})) \right\} \\ &= c_1(\mathcal{P}_{n+1}, \mu(\mathcal{P}_{n+1})) \\ &\quad + \sup_{\widehat{\boldsymbol{\gamma}} \in \widehat{\mathcal{C}}} V_1(F(\mathcal{P}_{n+1}, \mu(\mathcal{P}_{n+1}), \widehat{\boldsymbol{\gamma}})). \end{aligned}$$

By positivity of c_1 this implies

$$V_1(F(\mathcal{P}_{n+1}, \mu(\mathcal{P}_{n+1}), \widehat{\boldsymbol{\gamma}})) < V_1(\mathcal{P}_{n+1})$$

for all $\widehat{\boldsymbol{\gamma}}$ and thus the numbering of the $\mathcal{P}_j$ yields

$$F(\mathcal{P}_{n+1}, \mu(\mathcal{P}_{n+1}), \widehat{\boldsymbol{\gamma}}) \in \{\mathcal{P}_1, \ldots, \mathcal{P}_n\}. \tag{10}$$

Now the optimality principle for V_2 yields

$$\begin{aligned} V_2(\mathcal{Z}) &= \inf_{\boldsymbol{u} \in \mathcal{U}} \left\{ c_2(\mathcal{Z}, \boldsymbol{u}) + \sup_{\widehat{\boldsymbol{\gamma}} \in \widehat{\mathcal{C}}} V_2(F_2(\mathcal{Z}, \boldsymbol{u}, \widehat{\boldsymbol{\gamma}})) \right\} \\ &\leq c_2(\mathcal{Z}, \mu(\mathcal{P}_{n+1})) + \sup_{\widehat{\boldsymbol{\gamma}} \in \widehat{\mathcal{C}}} V_2(F_2(\mathcal{Z}, \mu(\mathcal{P}_{n+1}), \widehat{\boldsymbol{\gamma}})) \\ &= c_2(\mathcal{Z}, \mu(\mathcal{P}_{n+1})) + V_2(\mathcal{Z}_{\max}), \end{aligned} \tag{11}$$

where $\mathcal{Z}_{\max} = (\mathcal{P}_{n+1}, \mathcal{P}_i)^T$ denotes the element from $\{F_2(\mathcal{Z}, \mu(\mathcal{P}_{n+1}), \widehat{\boldsymbol{\gamma}}) \,|\, \widehat{\boldsymbol{\gamma}} \in \widehat{\mathcal{C}}\}$ realizing the supremum, which exists because F_2 can only assume finitely many values.

Now $X(\mathcal{Z}) \subseteq \mathcal{P}_{n+1}$ implies $\mathcal{P}_i = F(\mathcal{P}_{n+1}, \mu(\mathcal{P}_{n+1}), \widehat{\boldsymbol{\gamma}})$ for some suitable $\widehat{\boldsymbol{\gamma}}$ and thus from (10) we can conclude $i \leq n$. Furthermore, from the optimality principle for V_1 we obtain

$$\begin{aligned} V_1(\mathcal{P}_{n+1}) &= c_1(\mathcal{P}_{n+1}, \mu(\mathcal{P}_{n+1})) \\ &\quad + \sup_{\widehat{\boldsymbol{\gamma}} \in \widehat{\mathcal{C}}} V_1(F(\mathcal{P}_{n+1}, \mu(\mathcal{P}_{n+1}), \widehat{\boldsymbol{\gamma}})) \\ &\geq c_1(\mathcal{P}_{n+1}, \mu(\mathcal{P}_{n+1})) + V_1(\mathcal{P}_i). \end{aligned}$$

Using the induction assumption $V_1(\mathcal{P}_i) \geq V_2(\mathcal{Z}_{\max})$ (which is applicable since $i \leq n$) and

$$\begin{aligned} c_2(\mathcal{Z}, \mu(\mathcal{P}_{n+1})) &= \sup_{\boldsymbol{x} \in X(\mathcal{Z})} c^{r(\boldsymbol{x}, \mu(\mathcal{P}_{n+1}))}(\boldsymbol{x}, \mu(\mathcal{P}_{n+1})) \\ &\leq \sup_{\boldsymbol{x} \in \mathcal{P}_{n+1}} c^{r(\boldsymbol{x}, \mu(\mathcal{P}_{n+1}))}(\boldsymbol{x}, \mu(\mathcal{P}_{n+1})) \\ &= c_1(\mathcal{P}_{n+1}, \mu(\mathcal{P}_{n+1})) \end{aligned}$$

we can continue to estimate

$$\begin{aligned} V_1(\mathcal{P}_{n+1}) &\geq c_1(\mathcal{P}_{n+1}, \mu(\mathcal{P}_{n+1})) + V_1(\mathcal{P}_i) \\ &\geq c_2(\mathcal{Z}, \mu(\mathcal{P}_{n+1})) + V_2(\mathcal{Z}_{\max}) \end{aligned}$$

which together with (11) yields the assertion. ∎

In practice, we expect V_2 to be considerably smaller than V_1 and the corresponding controller to be much more efficient, as the experimental example in the following section confirms. Theorem 4.1, however, only yields $V_2 \leq V_1$ because system (7) may not contain any useful additional information compared to (4), which is theoretically possible but appears to be an exceptional case.

V. Example

We illustrate our approach with experimental results for a temperature and fill level control of a tank model which extend the results documented in [1]. The experiment is part of the experimental plant "VERA" at the Ruhr-Universität Bochum. Figure 1 shows a schematic image.

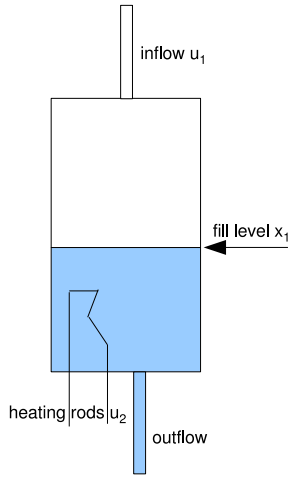

Fig. 1. Model of the tank

We have a two-dimensional state, x_1 is the fill level and x_2 is the temperature of the water in the tank. The fill level is constrained to the interval $[0.26\,\mathrm{m}; 0.45\,\mathrm{m}]$ and the temperature to $[293.15\,\mathrm{K}; 323.15\,\mathrm{K}]$ which defines our state space $\mathcal{X} = [0.26, 0.45] \times [293.15, 323.15]$. The first component of the two-dimensional control input regulates the inflow with a continuous adjustable valve with values between 0 and 1, quantized with 7 equidistant control values in our algorithm. The second component decides how many (0 to 6) heating rods are turned on. The system dynamics of the model are

$$\dot{x}_1 = \frac{1}{70 \cdot 10^{-3}} \left(q(u_1) \underbrace{-1.5876 \cdot 10^{-5} \sqrt{2 g x_1}}_{(*)} \right),$$

$$\dot{x}_2 = \frac{1}{0.07 x_1 - 1.9 \cdot 10^{-3}} \left(q(u_1)(\vartheta_{ext} - x_2) + \frac{P_{el} k_h u_2}{\varrho c_p} \right)$$

where

$$q(u_1) = \begin{cases} 0.07 \cdot 10^{-4}(11.1 u_1^2 + 13.1 u_1 + 0.2), & u_1 > 0.2 \\ 0, & \text{else} \end{cases}$$

and $(*)$ is the outflow of the tank. The constants of the systems dynamics are

$P_e l$	$3000\,\mathrm{W}$	Electric power of the heating rods
k_h	0.7	Heating coefficient
ϱ	$998\,\frac{\mathrm{kg}}{\mathrm{m}^3}$	Density of water
c_p	$4180\,\frac{\mathrm{J}}{\mathrm{kgK}}$	Specific heat capacity of water
g	$9.81\,\frac{\mathrm{m}}{\mathrm{s}^2}$	Gravitational constant
ϑ_{ext}	$293.15\,\mathrm{K}$	Temperature of inflowing water.

The discrete time system (1) has been obtained by sampling the continuous time system with sampling period $T = 1.0$ and as cost function we use

$$c(\boldsymbol{x}, \boldsymbol{u}) = \frac{1}{0.19^2}(x_1 - 0.349)^2 + \frac{1}{30^2}(x_2 - 310.56)^2 \quad (12)$$

which penalizes the distance of the state to the target. Simulation results for this example can be found in [4]. Here we show experimental results obtained by applying the feedback laws μ_1 and μ_2 obtained from our algorithms at the plant.

In numerical simulations we were able to stabilize the system with 8^2 rectangular quantization regions of equal size regular boxes for μ_2 and with 16^2 such regions for μ_1. However, due to unmodeled system reaction times, for the control of the real process we need 16^2 quantization regions for μ_2 and 32^2 regions for μ_1. With fewer regions the constraint set $\mathcal{X}$ turned out not to be positively invariant for the experimental trajectories. In the subsequent experiments we have used these partitions.

Figures 2 and 3 show the trajectories for the initial state $x(0) = (0.29, 298)$ for μ_2 and μ_1, respectively. With the dynamic feedback μ_2, which makes use of the past information, the target region (indicated by the blue lines) is reached in about half the time. The main difference between μ_1 and μ_2 lies in the control values for the inflow. Without using past information the fill level is kept constant until the specified temperature is reached and then the inflow valve is opened, while the feedback μ_2 using past information increases the fill level and the temperature simultaneously.

In the bottom diagram in each figure the changes of the quantized control inputs are shown. Here for μ_1 we observe 9 changes while for μ_2 16 changes occurred.

For the feedback law μ_1 without past information not only the time to reach the target is significantly longer, but also the cost — measured by summing up the cost function (12) along the experimental trajectories — is considerably higher, as Figure 4 shows.

In Figures 5–7 we repeat the experiment with intial state $x(0) = (0.4, 320)$.

Comparing the trajectories in Figure 5 for μ_1 and Figure 6 for μ_2, the main difference is again the inflow in the tank. While for μ_1 we have constant inflow on the whole time interval, for μ_2 the control input changes between three levels: the maximal inflow, no inflow and an intermediate value.

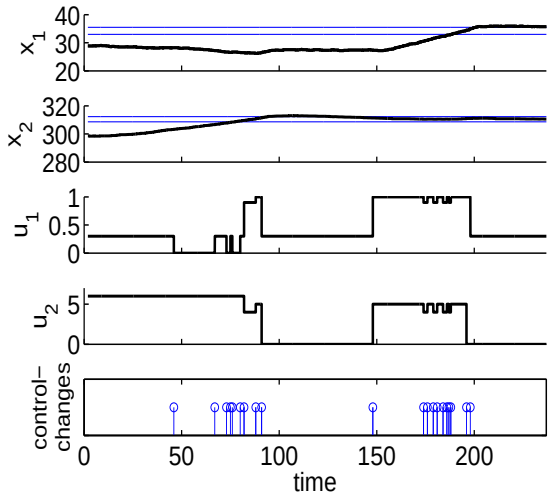

Fig. 2. Experimental trajectory and controls for initial state $x(0) = (0.29, 298)$ using μ_1 with 32×32 quantization regions

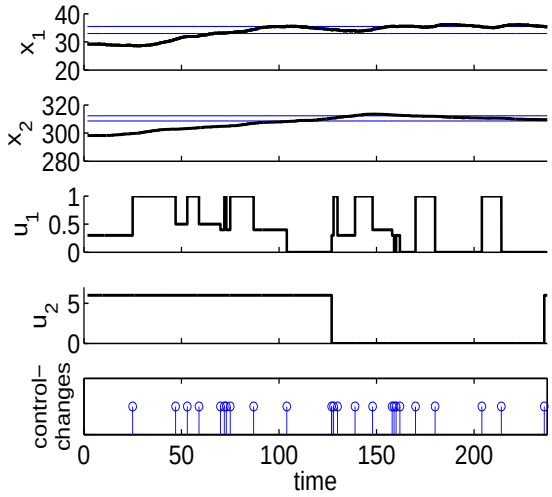

Fig. 3. Experimental trajectory and controls for initial state $x(0) = (0.29, 298)$ using μ_2 with 16×16 quantization regions

This way the controller uses the inflowing water in order to reduce the temperature of the water and consequently reaches the target faster. This refined control is also the main reason that there are 13 changes of the quantized control value for μ_2 compared to only 3 for μ_1.

Again, the use of past information in μ_2 leads to reduced cost along the experimental trajectory as Figure 7 shows, although the advantage here for initial state $x(0) = (0.4, 320)$ is not as pronounced as in Figure 4 for initial state $x(0) = (0.29, 298)$.

Summarizing, we can say with the use of past information we can stabilize systems on a coarser partition and at the same time obtain trajectories which are cheaper in terms of our optimization criterion and reach the target in considerably smaller time. On the other hand, the use of past data

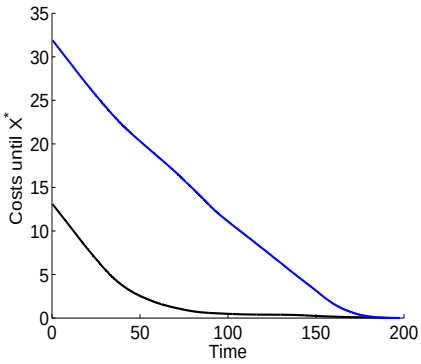

Fig. 4. Comparison of the costs for μ_1 with 32×32 quantization regions (blue) and μ_2 with 16×16 regions (black), $x(0) = (0.29, 298)$

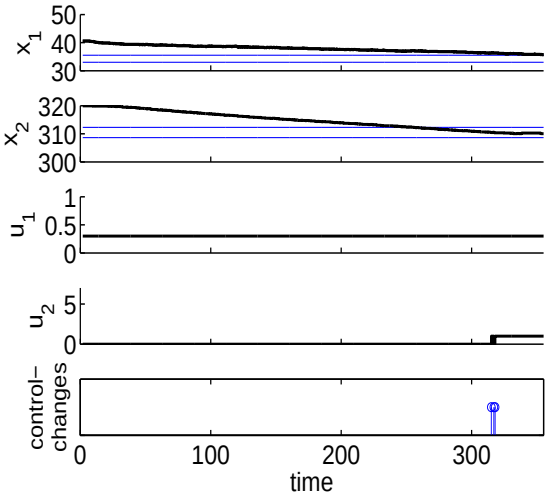

Fig. 5. Experimental trajectory and controls for initial state $x(0) = (0.4, 320)$ using μ_1 with 32×32 quantization regions

requires more computation time in the offline contruction of the hypergraph and results in trajectories with a higher number of control switches.

VI. Conclusion

In this paper we have introduced an algorithm for the optimal feedback control of nonlinear systems with coarse quantization. Compared to similar approaches for sampled data systems, the algorithm is able to obtain stabilizing feedback laws on much coarser quantizations. Using a dynamic feedback concept which takes into account information about past quantized states further significantly improves the results.

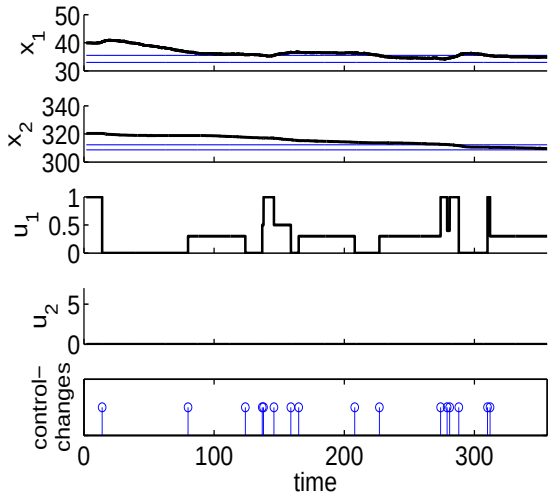

Fig. 6. Experimental trajectory and controls for initial state $x(0) = (0.4, 320)$ using μ_2 with 16×16 quantization regions

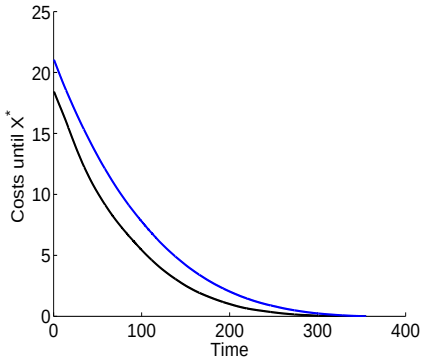

Fig. 7. Comparison of the costs for μ_1 with 32×32 quantization regions (blue) and μ_2 with 16×16 regions (black), $x(0) = (0.4, 320)$

REFERENCES

[1] L. Grüne, S. Jerg, O. Junge, D. Lehmann, J. Lunze, F. Müller, and M. Post. Two complementary approaches to event-based control. *at-Automatisierungstechnik*, 58(4):173–182, 2010. Special Issue on Networked Control Systems.

[2] L. Grüne and O. Junge. Approximately optimal nonlinear stabilization with preservation of the lyapunov function property. In *Proceedings of the 46th IEEE Conference on Decision and Control*, pages 702–707, New Orleans, Louisiana, 2007.

[3] L. Grüne and O. Junge. Global optimal control of perturbed systems. *Journal of Optimization Theory and Applications*, 136:411–429, 2008.

[4] L. Grüne and F. Müller. Set oriented optimal control using past information. In *Proceedings of the 18th International Symposium on Mathematical Theory of Networks and Systems MTNS2008*, Blacksburg, Virginia, 2008.

[5] L. Grüne and F. Müller. An algorithm for event-based optimal feedback control. In *Proceedings of the 48th IEEE Conference on Decision and Control*, pages 5311–5316, Shanghai, China, 2009.

[6] O. Junge and H. M. Osinga. A set oriented approach to global optimal control. *ESAIM Control Optim. Calc. Var.*, 10(2):259–270, 2004.

[7] D. Liberzon. Hybrid feedback stabilization of systems with quantized signals. *Automatica*, 39(9):1543–1554, 2003.

[8] T. Moor, J. Raisch, and S. O'Young. Discrete supervisory control of hybrid systems by l-complete approximations. *Journal of Discrete Event Dynamic Systems*, 12(1):83–107, 2002.

[9] M. von Lossow. A min-max version of dijkstra's algorithm with application to perturbed optimal control problems. In *Proceedings of the 6th International Congress on Industrial and Applied Mathematics (ICIAM 07) and GAMM Annual Meeting, PAMM 7(2007)*, pages 4130027–4130028, Zürich, Switzerland, 2007.