Preprints of the 18th IFAC World Congress
Milano (Italy) August 28 - September 2, 2011

Comparing Approximations of the Value Function for Entry Time Optimal Control

Matthias Rungger * **Olaf Stursberg** *

* *Institute of Control and System Theory, Dept. of Electrical Engineering and Computer Science, University of Kassel, Germany.*
{matthias.rungger,stursberg}@uni-kassel.de

Abstract: A fundamental quantity of the solution of an optimal control problem is the value function, i.e., the optimal cost-to-go function. In the general case, the function is not known exactly, but need to be approximated numerically. Most approaches to numerical approximation of the value function follow a procedure of three steps: first, the original continuous problem formulation is fully discretized; second, the discretized finite optimal control problem is solved by a shortest-path algorithm, and as third step, the solution is projected back from the finite space onto the continuous state space by using an interpolating function. This paper investigates the differences of discretization schemes and interpolating functions in this context. The performance of the computed approximations is evaluated in terms of an a-posteriori error, which is obtained from the approximating value functions. The convergence of the approximations to the true value function is proved for all considered schemes for the case that the discretization parameters are decreased to zero.

Keywords: Optimal Control, Nonlinear Systems, Dynamic Programming, Value Function.

1. INTRODUCTION

Dynamic programming is a powerful approach to the computation of static feedback laws for general nonlinear systems, which render the closed loop trajectories (approximatively) optimal with respect to a given cost functional. Except for special cases, like e.g. linear systems with quadratic costs in an infinite horizon setting, the value function of the optimal control problem cannot be computed analytically but has to be approximated numerically. While different approximation schemes have been proposed, the question of which scheme is preferable with respect to the approximation error is still largely open – the objective of this paper is to contribute to clarification of this question.

One method being relevant for the work presented here, is the one in Falcone (1987) and in González and Tidball (1991b) for the approximation of the value function for a discounted infinite horizon optimal control problem. The method uses a one-step Euler forward scheme for time discretization and a regular grid for state space discretization, where the grid defines a mesh of simplices covering the state space. For value function approximation, a function is defined that piecewise linearly interpolates between the values in the grid nodes constituting a simplex. The discretized system is interpreted as a controlled Markov chain for which the approximation is obtained by value or policy iteration. As an alternative approach, an undiscounted entry (exit) time optimal control problem was solved by a set oriented sampled-data framework in Grüne and Junge (2008). The sampled-data formulation allows the usage of higher order time discretization schemes like for example Runge-Kutta solvers. The interpolating function was chosen to be piecewise constant, and the fully discretized problem is equivalent to a deterministic shortest path problem on a graph, which can be solved by a Dijkstra-like algorithm. Such an algorithm permits a fast calculation of the discrete value function compared to value or policy iteration. The advantage in computation time is payed for by a generally lower approximation quality of piecewise constant functions compared to piecewise linear approximators.

In this work, the approximation quality and the computation time of different methods are investigated in comparison. The quality of the approximation is measured by an a-posteriori error, which specifies the difference between the sampled-data value function and the computed value function. The paper compares (and illustrates for an example) a one-step Euler forward scheme to Runge-Kutta methods of order three and five for time discretization. These alternatives are combined with piecewise linear and piecewise constant interpolation functions. The convergence of the different schemes is proved in a unified framework – previous results do not cover all the considered combinations. The presented approach follows ideas used in Kreisselmeier and Birkhölzer (1994); Cardaliaguet et al. (1998) and Grüne and Junge (2008) which respect to the fact that the state space discretization is modeled as bounded perturbation of the dynamics.

The paper is organized as follows: After some preliminaries, the considered finite time optimal control problem is introduced, and the discrete-time as well as the perturbed variants are formulated. Subsequently, the convergence of the value function for both variants to the one of the original system are shown for decreasing approximation

parameters. In Sec. 5, the fully discrete problems are formulated for a sampled state space, and the relation to the discrete time perturbed system is derived. The a-posteriori error is introduced in Sec. 6 and then evaluated (together with the computation time) for the different methods using a numerical example. Section 8 concludes the paper.

2. PRELIMINARIES

Throughout the paper, the following notation is used:

- $||x|| := (x^\intercal x)^{1/2}$ denotes the Euclidean norm of $x \in \mathbb{R}^n$.
- $\circ S$, ∂S, $\overline{S}$, S^c, and co$S$ denote the interior, the boundary, the closure, the complement, and the convex hull of the set S.
- $B(\bar{x}, r) := \{x \in \mathbb{R}^n \,|\, ||x - \bar{x}|| < r\}$ denotes the open ball centered at $\bar{x}$ with radius r. Let S be a set, then $B(S, r) := \cup_{x \in S} B(x, r)$.
- $d(x, S) := \inf_{\bar{x} \in S} ||x - \bar{x}||$ is the minimum distance from x to S, while $d(S^*, S) := \inf_{x \in S^*} d(x, S)$ denotes the minimum distance between two points in the sets S and S^*.
- Given $h \in \mathbb{R}^+$, the truncation operator $\lfloor\rfloor : \mathbb{R} \to \mathbb{R}$ is defined by $\lfloor t \rfloor := \max\{hi \,|\, hi < t,\ i \in \mathbb{N}\}$.
- The identity function in $\mathbb{R}^n$ is denoted by id_n.

Definition 1. Given a nonempty set $S \subset \mathbb{R}^n$ with boundary ∂S, let a unit vector p be an *exterior normal* to $\overline{S}$ at $x \in \partial S$ if there exists a ball outside $\overline{S}$ centered at $x + tp$ for some $t > 0$, where the ball touches $\overline{S}$ in x:

$$\bar{B}(x + tp, t) \cap \overline{S} = \{x\}.$$

The set of all unit vectors in x is denoted by $P(x)$. □

3. THE OPTIMAL CONTROL PROBLEM

The considered control system is given by the function $f : \mathbb{R}^n \times U \to \mathbb{R}^n$ with the input space $U \subset \mathbb{R}^m$. Let $\mathcal{U}$ denote the space of measurable functions over $\mathbb{R}_0^+$ with values in U. Given an initial state $x \in \mathbb{R}^n$ and an input function $\nu \in \mathcal{U}$, the *trajectory* of the system $x(t)$ is the *solution* of the initial value problem

$$\dot{x}(t) = f(x(t), \nu(t)),\ x(0) = x$$

at time t and is denoted by $\varphi(t, x, \nu) := x(t)$. φ is also referred to as the *flow* of the continuous dynamics f.

$\mathbf{A}_1$: It is assumed that f is uniformly continuous, bounded, and Lipschitz-continuous in its second argument, i.e., $L_f, M_f \in \mathbb{R}^+$ exist, such that for all $x, \bar{x} \in \mathbb{R}^n$ and $u \in U$

$$||f(x, u) - f(\bar{x}, u)|| \le L_f ||x - \bar{x}||,\ ||f(x, u)|| \le M_f \qquad \square$$

The *terminal* set $\mathcal{O} \subset \mathbb{R}^n$ specifies the set, where the system is driven to. It defines implicitly the time horizon of the optimal control problem.

$\mathbf{A}_2$: It is assumed that $\mathcal{O} = \overline{\circ\mathcal{O}}$, the boundary $\partial\mathcal{O}$ is compact, and the set of exterior normals is nonempty: $P(x) \neq \emptyset$, $\forall x \in \partial\mathcal{O}$.

Furthermore, some assumptions on the controllability of the system at the terminal set boundary are needed. **$\mathbf{A}_3$:** It is assumed that $\mu \in \mathbb{R}^+$ exist such that

$$\forall x \in \partial\mathcal{O},\ \forall p \in P(x): \qquad \inf_{u \in U} f(x, u) \cdot p < -\mu. \qquad \square$$

The *entry time* in $\mathcal{O}$ (or exit time from $\mathcal{O}^c$) of (1) from $x \in \mathbb{R}^n$ under the control $\nu \in \mathcal{U}$ is defined by

$$\tau(x, \nu) := \begin{cases} \infty & \text{if } \{t \,|\, \varphi(t, x, \nu) \in \partial\mathcal{O}\} = \emptyset \\ \min\{t \,|\, \varphi(t, x, \nu) \in \partial\mathcal{O}\} & \text{otherwise.} \end{cases}$$

The costs along the system trajectory comprises the *running* costs $l : \mathbb{R}^n \times U \to \mathbb{R}^+$ and *terminal* costs $g : \mathbb{R}^n \to \mathbb{R}^+$. The total cost functional is defined as

$$J(x, \nu) := \begin{cases} \infty & \text{if } \tau = \infty \\ \displaystyle\int_0^{\tau(x,\nu)} l(\varphi(t, x, \nu), \nu(t))\mathrm{d}t + g(\varphi(\tau, x, \nu))) & \text{if } \tau < \infty. \end{cases}$$

Regarding the costs functions the following is assumed.

$\mathbf{A}_4$: Let $L_l, L_g, m_l, M_l \in \mathbb{R}^+$ with $L_g \le m_l / M_f$ exist, such that for all $x, \bar{x} \in \mathcal{O}$

$$||g(x) - g(\bar{x})|| \le L_g ||x - \bar{x}||, \quad 0 \le g(x)$$

and for all $x, \bar{x} \in \mathbb{R}^n$ and $u \in U$

$$||l(x, u) - l(\bar{x}, u)|| \le L_l ||x - \bar{x}||,\ m_l \le l(x, u) \le M_l$$

is true. □

The *value function* $v : \mathbb{R}^n \to \mathbb{R}_0^+ \cup \{\infty\}$ for the entry time optimal control problem is obtained by

$$v(x) := \inf_{\nu \in \mathcal{U}} J(x, \nu).$$

A fundamental property, used throughout the paper, is the continuity of the value function, which is ensured by the theorem below.

Theorem 2. Assume (A_1)-(A_4). The value function v is continuous on $\{x \in \mathbb{R}^+ \,|\, v(x) < \infty\}$.

Proof. The proof follows from Propositions 3.3 ii), 3.4, and 3.7 in Bardi and Capuzzo-Dolcetta (1997), Chapter IV. ■

4. TIME DISCRETIZATION AND PERTURBATION

The numerical computation of the value function requires a discretization of the continuous optimal control problem. This section analyzes the time discretization in terms of a sampled-data system and its perturbation.

4.1 Sampled-Data Approximation

In a first step, the optimal input function is restricted to lie in the set of piecewise constant functions over $\mathbb{R}_0^+$ with values in U. Given $h \in \mathbb{R}^+$ as the step size, the set of control functions is defined by:

$$\mathcal{U}_h := \{\nu \in \mathcal{U} \,|\, \nu(t) = \nu(hi),\ \forall t \in [hi,\ h(i+1)),\ i \in \mathbb{N}_0\}.$$

Thus, any input function $\nu_h \in \mathcal{U}_h$ is completely characterized by a sequence $(u_i)_{i \in \mathbb{N}_0}$ with $\nu_h(hi) := u_i$.

The decisions of a controller are now restricted to times $t_i := hi$ with $i \in \mathbb{N}_0$. The system's trajectory at t_i is determined by a finite difference equation with the right hand side obtained by

$$f_h(x, u) := \varphi(h, x, \nu), \quad \nu(t) = u,\ \forall t \in [0, h).$$

The discrete time control system associated to f_h represents a *sampled-data system*. The trajectory $(x_i)_{i \in \mathbb{N}_0}$ initialized at $x \in \mathbb{R}^n$ by the *initialization function* $o_h := \mathrm{id}_n$

under the control sequence $(u_i) \in U^{\mathbb{N}_0}$ is given by the iteration

$$x_{i+1} = f_h(x_i, u_i), \quad x_0 = o_h(x).$$

The sampled-data flow function φ_h is defined such that $x_i = \varphi_h(i, x, (u_i))$ holds for all $i \in \mathbb{N}_0^+$. The purpose of the initialization function o_h in the flow definition is clarified in the next section, when perturbed sampled-data systems are considered.

Likewise, the one-step running costs l_h are computed by

$$l_h(x,u) := \int_0^h l(\varphi(t,x,u)), \nu(t))\mathrm{d}t, \quad \nu(t) = u,\ \forall t \in [0,h).$$

The Lipschitz-constants for f_h and l_h are bounded by $L_{f_h} \leq e^{L_f h}$ and $L_{l_h} \leq L_l e^{L_f h}$.

The entry time of the sampled-data system, from the initial state $x \in \mathbb{R}^n$ under the control sequence $u \in U^{\mathbb{N}_0}$, is given by

$$\tau_h(x,\nu_h) := \begin{cases} \infty & \text{if } \{i \,|\, \varphi_h(i,x,\nu_h) \in \mathcal{O}\} = \emptyset \\ \min\{i \,|\, \varphi_h(i,x,\nu_h) \in \mathcal{O}\} & \text{otherwise.} \end{cases}$$

The sampled-data version of the cost functional results in

$$J_h(x,\nu_h) = \sum_{i=0}^{\tau_h(x,\nu_h)-1} l_h(\varphi_h(i,x,\nu_h),\nu_h(i)) + g(\varphi_h(\tau_h,x,\nu_h))$$

and the corresponding value function in

$$v_h(x) := \min_{\nu \in \mathcal{U}_h} J_h(x,\nu).$$

In the next theorem the difference between the original and the sampled-data value functions v and v_h is estimated on a sub-level set $\mathcal{K}_c$, obtained by

$$\mathcal{K}_c := \{x \in \mathbb{R}^n \,|\, v(x) \leq c\}.$$

Theorem 3. Let f, l, g, $\mathcal{O}$ be givne according to (A_1)-(A_4) and $c \in \mathbb{R}^+$ such that $\mathcal{K}_c \neq \emptyset$. For any $\epsilon \in \mathbb{R}^+$, there exists $h \in \mathbb{R}^+$ such that the following inequality holds for all $x \in \mathcal{K}_c$:

$$||v(x) - v_h(x)|| \leq \epsilon.$$

Before proving the theorem, the reader is reminded of a lemma by González and Tidball (1991a).

Lemma 4. Given f, l, according to (A_1)-(A_4), let $\tilde{\varphi}$ denote the flow for the extended continuous system $f_l(x,u) := (f(x,u) \times l(x,u))$ and $\tilde{\varphi}_h$ the corresponding flow of the sampled-data system. There exists $K \in \mathbb{R}^+$ such that for any $x \in \mathbb{R}^n$, $\nu \in \mathcal{U}$, and $h \in \mathbb{R}^+$, a control sequence $\nu_h \in \mathcal{U}_h$ exists that satisfies the inequality

$$||\tilde{\varphi}(t,x,\nu) - \tilde{\varphi}(t,x,\nu_h)|| \leq K e^{L_f t} h^{1/2}. \qquad \diamondsuit$$

Proof. See Lemma 4.3 in González and Tidball (1991a). ■

Proof of Theorem 3. i) If $x \in \mathcal{O}$ $v_h(x) = g(x) = v(x)$. Thus, $x \in \mathcal{K}_c \backslash \mathcal{O}$ is considered, and an input function $\nu \in \mathcal{U}$ is chosen, such that holds: $J(x,\nu) \leq v(x) + \epsilon_1$.

ii) An extended ϵ-optimal control function

$$\nu_\epsilon(t) := \begin{cases} \nu(t) & 0 \leq t \leq \tau \\ u & \tau < t \leq \tau + \epsilon_2 \end{cases}$$

with $\tau := \tau(x,\nu)$ is constructed. The input $u \in U$ and $\epsilon_2 \in \mathbb{R}^+$ are chosen such that $x(\tau + \epsilon_2) \in \circ\mathcal{O}$. Due to (A_3) such an input with ϵ_2 (arbitrary small) always exists.

The abbreviation $\tau_\epsilon := \tau(x,\nu) + \epsilon_2$ is used in the following.

iii) A neighborhood $\delta_x \in \mathbb{R}^+$, along with a time step $h \in \mathbb{R}^+$, and $\epsilon_3 \in \mathbb{R}^+$ are chosen such that for all $\bar{x} \in B(x,\delta_x)$

$$B(\varphi(\lfloor \tau_\epsilon \rfloor, \bar{x}, \nu_\epsilon), \epsilon_3) \subseteq \mathcal{O}$$

with the bound

$$||\varphi(\lfloor \tau_\epsilon \rfloor, \bar{x}, \nu_\epsilon) - \varphi_h(\lfloor \tau_\epsilon \rfloor, \bar{x}, \nu_h^\epsilon)|| \leq K e^{L_f \lfloor \tau_\epsilon \rfloor} h^{1/2} \leq \epsilon_3$$

holds. The input function ν_h^ϵ is chosen according to Lemma 4.

iv) Let $\delta(x) := \min\{\delta_x, \bar{\delta}_x\}$, where $\bar{\delta}_x$ is chosen, such that for all $\bar{x} \in B(x,\bar{\delta}_x) \Rightarrow ||v(\bar{x}) - v(x)|| \leq \epsilon$ holds. Then, the value functions satisfy for all $\bar{x} \in B(x,\delta_x)$ the following estimate:

$$\begin{aligned} v_h(\bar{x}) - v(x) &\leq v_h(\bar{x}) - v(\bar{x}) + ||v(\bar{x}) - v(x)|| \\ &\leq \int_0^{\lfloor \tau_\epsilon \rfloor} l_h(\bar{x}_h(t), \nu_h^\epsilon(t)) - l(\bar{x}(t), \nu(t))\mathrm{d}t \\ &\quad + g(\bar{x}_h(\lfloor \tau_\epsilon \rfloor)) - g(\bar{x}(\tau)) + M_l \epsilon_2 + \epsilon_1 + \epsilon \\ &\leq K e^{L_f \lfloor \tau_\epsilon \rfloor} h^{1/2} + L_g ||\bar{x}_h(\lfloor \tau_\epsilon \rfloor) - \bar{x}(\tau)|| \\ &\quad + M_l \epsilon_2 + \epsilon_1 + \epsilon \\ &\leq L_g(||\bar{x}_h(\lfloor \tau_\epsilon \rfloor) - \bar{x}(\lfloor \tau_\epsilon \rfloor)|| + ||\bar{x}(\lfloor \tau_\epsilon \rfloor) - \bar{x}(\tau)||) \\ &\quad + K e^{L_f \lfloor \tau_\epsilon \rfloor} h^{1/2} + M_l \epsilon_2 + \epsilon_1 + \epsilon \\ &\leq L_g(K e^{L_f \lfloor \tau_\epsilon \rfloor} h^{1/2} + M_f \epsilon_2) \\ &\quad + K e^{L_f \lfloor \tau_\epsilon \rfloor} h^{1/2} + M_l \epsilon_2 + \epsilon_1 + \epsilon \\ &\leq (1 + L_g) K e^{L_f \lfloor \tau_\epsilon \rfloor} h^{1/2} + (L_g M_f + M_l)\epsilon_2 + \epsilon_1 + \epsilon. \end{aligned}$$

The estimate from Lemma 4 is used to bound the integral expression in the first inequality and to estimate $||\bar{x}_h(\lfloor \tau_\epsilon \rfloor) - x(\lfloor \tau_\epsilon \rfloor)||$ in the third inequality.

Then, ϵ_1, ϵ_2, and h, can be chosen, such that for all $\bar{x} \in B(x,\delta(x))$, the inequality $v_h(\bar{x}) - v(\bar{x}) \leq 2\epsilon$ holds.

v) The upper bound for the reverse inequality is uniformly given by $v(x) - v_h(x) \leq (M_l + L_g M_f)h$. It is the maximal cost of one step.

vi) The compactness of $\mathcal{K}_c$ follows from the continuity of v. Thus, there exists a finite cover

$$\mathcal{K}_c \subset \bigcup_{i=0}^{m} B(x_i, \delta(x_i)).$$

Let h_x denote the time step associated to every $x \in \mathcal{K}_c$ (as in the steps **i-iv)**), Now, with $h := \min_{i \in \{0,\ldots,m\}}\{h_{x_i}\}$, for any $x \in \mathcal{K}_c$, there exists x_i with $x \in B(x_i, \delta(x_i))$ and

$$||v_h(x) - v(x)|| \leq ||v_h(x) - v(x_i)|| + ||v(x_i) - v(x)|| \leq 3\epsilon.$$

■

The following functional results from the dynamic programming principle:

$$v_h(x) = \begin{cases} \min\limits_{u \in U}\{l_h(o_h(x),u) + v_h(f_h(o_h(x),u))\} & \text{if } x \in \mathcal{K}_{\epsilon_h} \backslash \mathcal{O} \\ g(o_h(x)) & \text{if } x \in \mathcal{O}. \end{cases} \tag{1}$$

This equation will be used again in Section 6 to determine the error of the different approximation schemes with respect to v_h.

4.2 Perturbation

The right hand side of the sampled-data system and the one-step costs are often obtained by numerical simulation

when the exact solutions of the integrals are not known. Thus, an inaccurate computation of f_h and l_h is considered by introducing disturbance terms. Furthermore, the perturbations model the finite representation of the state space e.g. obtained from a grid-like partitioning.

The computational errors are modelled by the *perturbation* functions $p_f : \mathbb{R}^n \times U \to \mathbb{R}^n$, $p_l : \mathbb{R}^n \times U \to \mathbb{R}$, $p_g : \mathbb{R}^n \to \mathbb{R}$, and $p_o : \mathbb{R}^n \to \mathbb{R}^n$. The data of the optimal control problem results in

$$o_h^p := \mathrm{id}_h + p_o, \quad f_h^p := (\mathrm{id}_n + p_o) \circ (f_h + p_f),$$
$$l_h^p := l_h + p_l, \qquad g^p := g + p_g.$$

The trajectory $(x_i^p)_{i\in\mathbb{N}}$, the entry time τ_h^p, the cost functional J_h^p, and the value function v_h^p of the *perturbed sampled-data system* are obtained identical to the corresponding quantities of the sampled-data system, only that o_h^p, f_h^p, l_h^p, g^p instead of o_h, f_h, l_h, g are used in the definitions.

Note that the usual approach for a perturbed optimal control problem is the definition of the value function in a *min-max* formulation Huang et al. (2005); Grüne and Junge (2008), or Bardi and Capuzzo-Dolcetta (1997) Chapter VIII. As this contribution focuses on the approximation quality of the value function, this is not considered here. In consequence, in order to obtain a stabilizing feedback from the value function, a rather fine state space discretization is necessary.

Let the bounds of the perturbation functions be denoted by

$$\sup_{x\in\mathbb{R}^n} ||p_o|| \le b_1, \qquad \sup_{x\in\mathbb{R}^n, u\in U} ||p_l|| \le b_2,$$
$$\sup_{x\in\mathbb{R}^n, u\in U} ||p_l|| \le b_3. \qquad \sup_{x\in\mathbb{R}^n} ||p_g|| \le b_3.$$

Theorem 5. Let f, l, g, $\mathcal{O}$ according to (A_1)-(A_4), and $\mathcal{K}_c \neq \emptyset$. For every $\epsilon \in \mathbb{R}^+$, there exist h and b_i, $i \in \{1,2,3\}$, such that for all $x \in \mathcal{K}_c$:

$$||v(x) - v_h^p(x)|| \le \epsilon.$$

Proof. The proof follows by similar arguments as the proof of Thm. 3 with the following distance estimates of the trajectories $(x_i)_{i\in\mathbb{N}}$ and $(x_i^p)_{i\in\mathbb{N}}$, $i \in \mathbb{N}$: $||x_i - x_i^p|| \le i(b_1 + b_2) + b_1$. ■

When a numerical integration scheme with consistency of order m is used to obtain f_h^p, the perturbation function p_f is bounded by $b_2 = Kh_i^m$ for some $K \in \mathbb{R}^+$. h_i are the inter-sample integration steps with $\sum h_i = h$.

For example for using the forward Euler scheme with one-step per sample time h, it holds that: $b_2 = L_f M_f h^2/2$.

The input set U is assumed to be finite in the following. The error which results from this simplification is also representable by p_f and p_l. Due to the uniform continuity of f and l, the error can be made arbitrary small by increasing the number of samples in U.

5. COMPUTATION OF THE VALUE FUNCTION

The domain of calculation $\mathcal{R} \subset \mathbb{R}^n$ is spanned by the set of N sample states or *nodes* denoted by Ξ. The set of nodes $\Xi_\mathcal{O} := \Xi \cap \mathcal{O}$ and $\Xi_\mathcal{R} := \Xi \cap \overline{\mathcal{R}^c}$ represent the terminal set $\mathcal{O}$ and the boundary set $\partial\mathcal{R}$ respectively.

The discretization parameter k associated with the set Ξ is determined by

$$k := \sup_{x\in\mathcal{R}} \min_{\xi\in\Xi} ||x - \xi||.$$

The value function for the fully discretized system is computed based on the state space discretization. Depending on the chosen interpolation scheme, the algorithm to calculate the value function differs. In the following, a piecewise constant interpolation and a linear interpolation are analyzed.

5.1 Piecewise Constant Interpolation

The piecewise constant interpolation is based on a Voronoi partition Δ obtained by the node set Ξ. Each node $\xi \in \Xi$ is associated to a partition element δ_ξ of Δ.

A graph $G = (\Xi, E)$ is constructed from the Voronoi partition. The set of edges is given by

$$E := \{(\xi, u, \bar{\xi}) \in \Xi \times U \times \Xi \,|\, f_h(\xi, u) + p_f(\xi, u) \in \delta_{\bar{\xi}}\}.$$

and the costs along the edges are $l_h^p(\xi, u)$. The costs of the terminal nodes $\xi \in \Xi_\mathcal{O} \cup \Xi_\mathcal{R}$ are set to $g(\xi)$ for $\xi \in \Xi_\mathcal{O}$ and to ∞ for $\xi \in \Xi_\mathcal{R}$.

The value function V based on the graph G is computed by solving a *deterministic* shortest path problem for all $\xi \in \Xi$.

The consideration is restricted to the set $\Xi_\mathcal{K} := \{\xi \in \Xi \,|\, V(\xi) < \infty\}$. It is known, see for example Bertsekas (1995), that the value function $V : \Xi_\mathcal{K} \to \mathbb{R}^+$ satisfies the functional equation

$$V(\xi) = \begin{cases} \min\limits_{(\xi,u,\bar{\xi})\in E} \{l_h^p(\xi, u) + V(\bar{\xi})\} & \text{if } \xi \in \Xi_\mathcal{K} \backslash \Xi_\mathcal{O} \\ g(\xi) & \text{if } \xi \in \Xi_\mathcal{O}. \end{cases}$$

The computational complexity of computing V is of order $O(|\Xi|^2)$ using the Dijkstra algorithm and can be reduced to $O(|\Xi| \log(|\Xi| + |E|))$ using a binary heap.

The projection of $V(\xi)$ back onto the continuous state space is defined for the set

$$\mathcal{K}_\Xi^0 := \{x \in \delta_\xi \,|\, \delta_\xi \in \Delta,\ V(\xi) < \infty\}$$

by the function $v_\Xi^0 : \mathcal{K}_\Xi^0 \to \mathbb{R}^+$. Given $x \in \delta_\xi$, the value simply follows to:

$$v_\Xi^0(x) := V(\xi).$$

5.2 Linear Interpolation

The linear interpolation is based on a triangulation of the nodes into a set of simplices Γ. Each state x, covered by the triangulation, may be written by a linear combination of the nodes $\xi_0, \ldots, \xi_n$ spanning the simplex $\gamma_{\xi_0,\ldots,\xi_n} \in \Gamma$ containing x. The coefficients of this linear combination are referred to as *barycentric coordinates*.

Let $[\lambda_i(\xi, u)]_{i=1,\ldots,N}$ denote an N dimensional vector containing the barycentric coordinates of $f_h^p(\xi, u)$. Since all $\lambda_i(\xi, u) \ge 0$ and $\sum_i \lambda_i(\xi, u) = 1$, it is possible to interpret the barycentric coordinates as transition probabilities of a controlled Markov chain. The states of the chain are given by the set Ξ and the input set by U. See the monograph Kushner and Dupuis (1992) Chapter 4 and 5 for a detailed treatment of the Markov chain approach with piecewise linear function approximation.

The value function V of the Markov chain is computed by solving a *stochastic* shortest path problem. The standard algorithms solving a stochastic shortest path problem are *value* - and *policy iteration*, see Bertsekas (1995), Vol II, Chapter 2.2.

On the set $\Xi_\mathcal{K} := \{\xi \in \Xi \mid V(\xi) < \infty\}$, the value function $V : \Xi_K \to \mathbb{R}^+$ satisfies the functional equation

$$V(\xi) = \begin{cases} \min\limits_{u\in U}\{l_h^p(\xi,u) + \sum\limits_{i=0}^{N} \lambda_i(\xi,u)\, V(\xi_i)\} & \text{if } \xi \in \Xi_\mathcal{K}\backslash\Xi_\mathcal{O} \\ g(\xi) & \text{if } \xi \in \Xi_\mathcal{O}. \end{cases}$$

Under the assumption that no self transitions exist, the complexity of value iteration is of order $O(|\Xi|^2|U|)$ and for policy iteration $O(|\Xi|^2 + X|\Xi||U|)$, where X is the number of policy improvement iterations. In special cases, when there exist only transitions to states with lower cost-to-go, Dijkstra like methods can be used to calculate the value function (see Tsitsiklis (1995)).

The projection of V back onto the continuous state space is defined for the set

$$\mathcal{K}_\Xi^1 := \{x \in \gamma_{\xi_0,\dots,\xi_n} \,|\, \gamma_{\xi_0,\dots,\xi_n} \in \Gamma,\, \forall \xi_i : V(\xi_i) < \infty\}$$

by: $v_\Xi^1 : \mathcal{K}_\Xi^1 \to \mathbb{R}^+$. Given $x \in \gamma_{\xi_0,\dots,\xi_n}$ and the barycentric coordinates $\lambda_0,\dots,\lambda_n$ with $x = \sum_{i=0}^n \lambda_i \xi_i$, the value $v_\Xi^1(x)$ is determined by

$$v_\Xi^1(x) := \sum_{i=0}^{n} \lambda_i\, V(\xi_i).$$

5.3 Perturbation and State Space Discretization

The projection of the continuous system onto a finite state space can be understood as a perturbation which is bounded by the discretization parameter k.

Definition 6. Given the sets $\mathcal{K}_\Xi^0$ and $\mathcal{K}_\Xi^1$ along with the value functions v_Ξ^0 and v_Ξ^1 (according to the derivation in Sec. 5.1 and 5.2 for a grid Ξ), candidate perturbation functions

$$p_0 : \mathcal{K}_\Xi^0 \to \mathbb{R}^n \quad \text{and} \quad p_i : \mathcal{K}_\Xi^1 \to \mathbb{R}^n, \quad i \in \{1,2\}.$$

are defined in the following.

For $x \in \mathcal{K}_\Xi^0$, let $\delta_\xi \subset \mathcal{K}_\Xi^0$ denote the partition element containing $x \in \delta_\xi$, such that one obtains:

$$p_0(x) := \xi - x.$$

For $x \in \mathcal{K}_\Xi^1$, let $\gamma_{\xi_0,\dots,\xi_n} \subset \mathcal{K}_\Xi^1$ denote the partition element containing $x \in \gamma_{\xi_0,\dots,\xi_n}$. Let $\xi_{\min}, \xi_{\max} \in \{\xi_0,\dots,\xi_n\}$ denote the nodes for which $v_\Xi^1(\xi_{\min}) \le v_\Xi^1(\xi_j)$ and $v_\Xi^1(\xi_{\max}) \ge v_\Xi^1(\xi_j)$ holds for all $\xi_j \in \{\xi_0,\dots,\xi_n\}$. The perturbation functions are defined by

$$p_1(x) := \xi_{\min} - x \quad \text{and} \quad p_2(x) := \xi_{\max} - x. \qquad \square$$

The next theorem relates the value functions of the perturbed sampled-data system to the value functions of the fully discretized system.

Theorem 7. Given f_h^p, l_h^p, g^p, $\mathcal{O}$, and $\mathcal{R}$ under the conditions of Theorem 5. The sets $\mathcal{K}_\Xi^0$, $\mathcal{K}_\Xi^1$ and functions v_Ξ^0, v_Ξ^1, p_0, p_1, and p_3 from Definition 6 are considered.

Let $v_h^{p_i}$, $i \in \{0,1,2\}$ denote the value functions for the perturbed sampled-data systems, defined by $o_h^{p_i} := \mathrm{id}_n + p_i$, $f_h^{p_i} := (\mathrm{id}_n + p_i) \circ (f_h + p_f)$. The equations

$$\forall x \in \mathcal{K}_\Xi^0 : \qquad v_h^{p_0}(x) = v_\Xi^0(x) \tag{2}$$

$$\forall x \in \mathcal{K}_\Xi^1 : \qquad v_h^{p_1}(x) \le v_\Xi^1(x) \le v_h^{p_2}(x). \tag{3}$$

are satisfied, with the functions $p_i \le k$, $p_l \le L_l k$, $p_g \le L_g k$.

6. ERROR MEASUREMENT

The error estimates used in the previous convergence proofs are conservative upper bounds and thus are not immediately useful in the performance evaluation of the value function approximation.

Motivated by the formulation for v_h in the Equation (1), the error due to the discretization of the problem is reinterpreted in terms of a new error function $e : \mathbb{R}^n \to \mathbb{R}$.

Given the set $\mathcal{K}_\Xi$ and the function v_Ξ, referring to $\mathcal{K}_\Xi^0$ or $\mathcal{K}_\Xi^1$ and v_Ξ^0 or v_Ξ^1 respectively, the following functional equation is satisfied:

$$v_\Xi(x) = \begin{cases} \min\limits_{u\in U}\{l_h(x,u) + e(x) + v_\Xi(f_h(x,u))\} & \text{if } x \in \mathcal{K}_\Xi\backslash\mathcal{O} \\ g(x) + e(x) & \text{if } x \in \mathcal{O}. \end{cases} \tag{4}$$

The second quantity which is examined in the numerical tests is related to the descent property of the value function with respect to the optimal trajectories. The control law $u_h^a : \mathcal{K}_\Xi \to U$ is defined by

$$u_h^a(x) := \operatorname*{argmin}_{u\in U}\{l_h^a(x,u) + v_\Xi(f_h^a(x,u))\}. \tag{5}$$

The dynamics f_h^a and one-step costs l_h^a are obtained by the numerical simulation which is used in the computation of the value function (the true quantities f_h and l_h are unknown for the controller). In case the inequality

$$l_h(x, u_h^a(x)) + e(x) > 0 \tag{6}$$

holds on $\mathcal{K}_\Xi$, the value function is a Lyapunov function of the closed loop system since

$$v_\Xi(x) > v_\Xi(f_h(x, u_h^a(x)))$$

follows. Thus, the stability of the control system under the control law (5) can be deduced in this case.

In general, the error function e is unknown. Thus, in the numerical tests, the error function e is evaluated at some sample states $\eta \in \mathcal{K}_\Xi$. The sampled errors $e(\eta)$ are used as *indicators* of the performance of the closed loop system with respect to optimality and stability.

7. NUMERICAL TESTS

The various discretization and interpolation methods are applied to compute the value function of an inverted pendulum with quadratic cost functions. The dynamics is given by

$$\dot{x}_1 = x_2,\ \dot{x}_2 = -0.01x_2 + 9.81\sin(x_1) + u,$$

where x_1 and x_2 are the angle and angular velocity. The cost functions are set to

$$g(x) = 0.1x_1^2 + 0.1x_2^2$$

$$l(x,u) = 0.1x_1^2 + 0.1x_2^2 + 0.01u^2.$$

The domain of computation is chosen to $\mathcal{R} := [-\pi, \pi] \times [-6, 6]$ and the terminal set to $\mathcal{O} := [-0.2, 0.2]^2$. The state

space of the first coordinate is unified at $\pm\pi$ and the input space is given by $U = [-3\ 3]$. The box constraints on the input prevent a direct swing-up of the pendulum.

7.1 Computational Setup and Results

Three time discretization schemes are considered: the one-step Euler forward method (E) and two Runge-Kutta methods of order three (RK3) and five (RK5). The set of different examined time steps is given by

$$h \in \{0.5,\ 0.2,\ 0.07,\ 0.03\}.$$

For the Runge-Kutta methods, the number of intersample time steps is chosen to 5.

Each time discretization method is combined with the piecewise constant (PC) and piecewise linear (PL) interpolation function approximations. For both methods a uniform gird is used with grid node distances

$$k \in \{0.1,\ 0.05,\ 0.02,\ 0.01\}.$$

A standard value iteration method is applied to solve the stochastic shortest path problem, and the Dijkstra algorithm with a binary heap is used to solve the deterministic shortest path problem. The input space is approximated by 21 equidistantly distributed samples for all conducted computations.

The computation times of the value function, applying the various combinations of h and k, are illustrated in Figure 1. The plots show the overall time, i.e., the time needed for computing the discretized system as well as the time spend for solving the shortest path problem. The legend of the plots, listing the method assigned to each symbol, is given in the following table.

△ E/PC	□ RK3/PC	◊ RK5/PC
▲ E/PL	■ RK3/PL	♦ RK5/PL

The results are plotted only if the area $\mathcal{K}_\Xi$, from which the terminal set is reachable for the discretized system, is greater then 50 % of $\mathcal{R}$.

Results in terms of the error, cf. (4), are illustrated in Figure 2. The plots display the mean of the absolute error $e(\eta)$ evaluated in the center $\eta = \xi + k/2$ of each partition element of the uniform grid. Matlab's variable step size solver `ode45` with error control parameter `AbsTol` and `RelTol` set to 10^{-10} is used to compute f_h, l_h in (4).

The descent property of the value function v_Ξ, cf. (6), is also evaluated at the center η of each partition element. Figure 3 shows the percentage of partition elements for which (6) holds. The circle around the symbols in the plots indicate that the control law (5) stabilizes the upper position of the pendulum initialized at $x_0 = [\pi,\ 0]^\mathrm{T}$. The closed loop system is simulated again by using Matlab's solver `ode45` with the same settings as above.

All the computations are performed on an AMD Phenom II X4 920 processor with 4GB memory.

7.2 Interpretation and Discussion

To conclude this section, the presented results and the experimental observations are discussed.

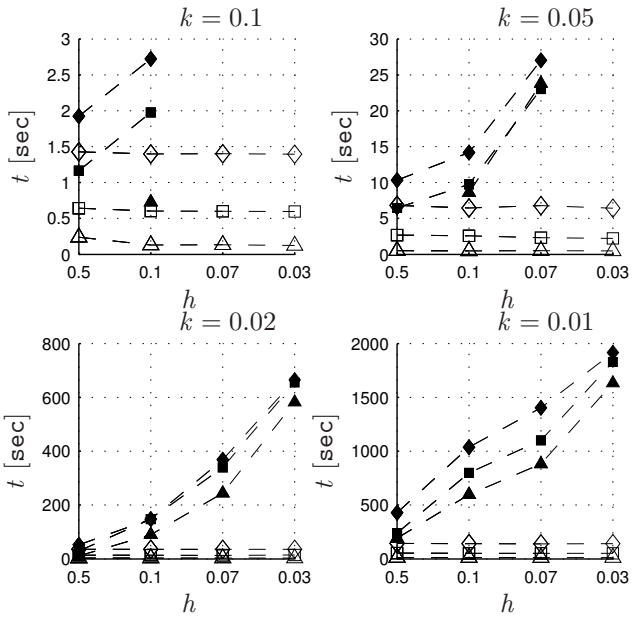

Fig. 1. Computation times in seconds.

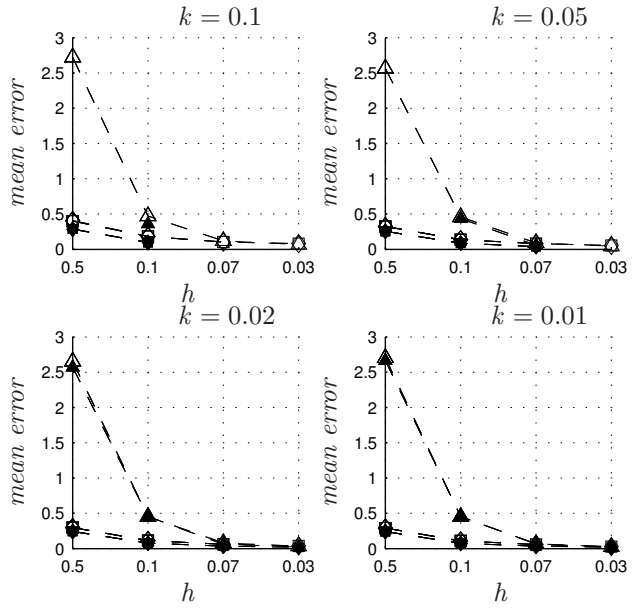

Fig. 2. The plot shows the mean of the absolute error $e(\eta)$ sampled at the center η of each partition element of the uniform grid.

The most obvious observation is the immense reduction of the computation time due to the use of the Dijkstra algorithm (see Fig. 1). Although, there exist improved algorithms to solve stochastic shortest path problems, a speed up comparable to the Dijkstra algorithm's performance is not expected. In particular, for the piecewise constant approach, most time of the solution is spend in building the graph, while for the piecewise linear approach, most time is consumed for the value iteration. Since the number of iterations is proportional to the maximum length of a path in the Markov chain, the computation time increases with decreasing time step h.

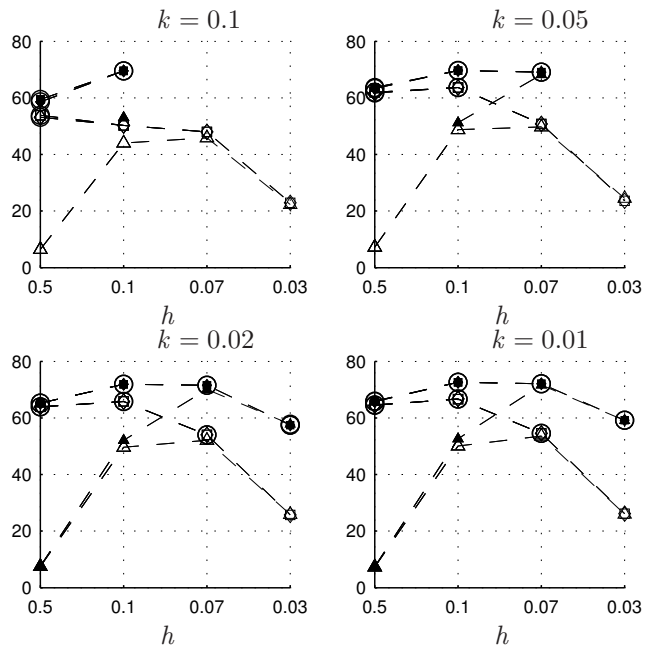

Fig. 3. Percentage of the partition elements on which (6) holds.

The reduction of the computation time has already been observed in Grüne and Junge (2005). However, the more interesting detail is that the accuracy of approximation of the two interpolations schemes to not differ considerable.

The second pertinent observation is the poor performance of the Euler solver compared to the Runge-Kutta schemes. While it is slightly better in terms of the computation time, the test trajectory initialized at x_0 converged only twice to the origin when the piecewise linear function approximation has been used. For the piecewise constant function approximation no combination of k and h resulted in a pendulum swing up under the control law (5). This supports the theoretical results with respect to optimization based synthesis of stabilizing feedbacks reported in Grüne and Nešić (2003).

The third observation, is the loss of the stabilizing property of the control law when the time step h is chosen too small for the Runge-Kutta solvers (see Figure 3). Although a slightly better behavior is observed in this respect for the piecewise linear function approximations, the methods fail to swing up the pendulum for the cases $k = 0.1$, $h \in \{0.07, 0.03\}$ and $k = 0.05$, $h = 0.03$.

8. SUMMARY AND CONCLUSION

This work described the numerical approximation of the value function for entry time optimal control problems. The convergence of the approximation to the true value function was shown for the case that the discretization parameters are decreased to zero. The convergence proof considered a perturbed sampled data time discretization and two function approximation methods – a piecewise constant and piecewise linear approximation.

The performance of three different solvers of the sampled data one-step dynamics was compared for an example. Each solver was combined with one of the two function approximation methods. The best performance with respect to a balance between computation time and error measurements was obtained for a Runge-Kutta solver of order three in combination with the piecewise constant function approximation.

REFERENCES

Bardi, M. and Capuzzo-Dolcetta, I. (1997). *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*. Birkhäuser.

Bertsekas, D.P. (1995). *Dynamic Programming and Optimal Control*. Athena Scientific.

Cardaliaguet, P., Quincampoix, M., and Saint-Pierre, P. (1998). Set-valued numerical analysis for optimal control and differential games. In M. Bardi, T.E.S. Raghavan, and T. Parthasarathy (eds.), *Stochastic and differential games: theory and numerical methods*. Birkhäuser.

Falcone, M. (1987). A numerical approach to the infinite horizon problem of deterministic control theory. *Applied Mathematics and Optimization*, 15, 1–13.

González, R.L.V. and Tidball, M.M. (1991a). On a discrete time approximation of the hamilton-jacobi equation of dynamic programming. Technical Report 1375, Institut National de Recherche en Informatique et en Automatique (INRIA).

González, R.L.V. and Tidball, M.M. (1991b). On the rates of convergence of fully discrete solutions of hamilton-jacobi equations. Technical Report 1376, Institut National de Recherche en Informatique et en Automatique (INRIA).

Grüne, L. and Junge, O. (2005). A set oriented approach to optimal feedback stabilization. *Systems and Control Letters*, 54.

Grüne, L. and Junge, O. (2008). Global optimal control of perturbed systems. *Journal of Optimization Theory and Applications*, 136, 411–429.

Grüne, L. and Nešić, D. (2003). Optimization-based stabilization of sampled-data nonlinear systems via their approximate discrete-time models. *SIAM Journal on Control and Optimization*, 42, 98–122.

Huang, S., James, M.R., Nešić, D., and Dower, P. (2005). A unified approach to controller design for achieving iss and related properties. *IEEE Trans. Automat. Cont.*, 50, 1681 – 1697.

Kreisselmeier, G. and Birkhölzer, T. (1994). Numerical nonlinear regulator design. *IEEE Trans. Automat. Cont.*, 39, 33–46.

Kushner, H.J. and Dupuis, P. (1992). *Numerical Methods for Stochastic Control Problems in Continuous Time*. Springer-Verlag.

Tsitsiklis, J.N. (1995). Efficient algorithms for globally optimal trajectories. *IEEE Trans. Automat. Cont.*, 42, 1528–1538.