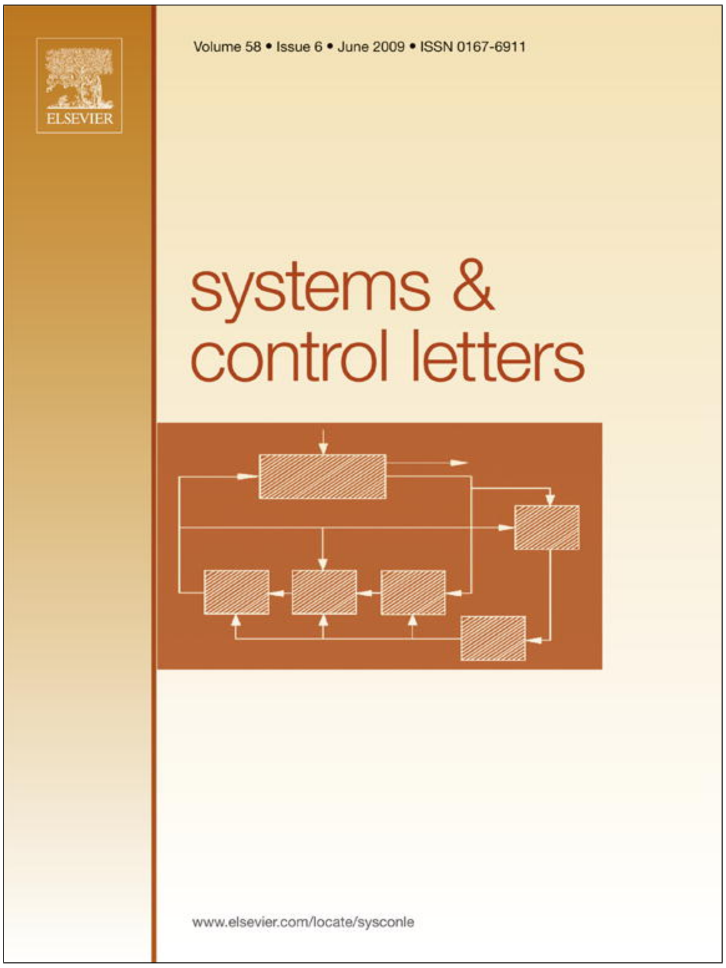
Volume 58 • Issue 6 • June 2009 • ISSN 0167-6911
ELSEVIER
systems & control letters
www.elsevier.com/locate/sysconle

Systems & Control Letters 58 (2009) 395–399

Contents lists available at ScienceDirect

Systems & Control Letters

journal homepage: www.elsevier.com/locate/sysconle

On the discretization of switched linear systems[☆]

Alain Pietrus[a], Vladimir M. Veliov[b,c,*]

[a] Laboratoire AOC, Dépt. de Mathématiques, Université des Antilles et de la Guyane, Campus de Fouillole, BP 250, F-97157 Pointe-à-Pitre, Guadeloupe
[b] Institute Mathematical Methods in Economics, Vienna University of Technology, Argentinierstrasse 8/119, A-1040 Vienna, Austria
[c] Institute of Mathematics and Informatics, BAS, 1113 Sofia, Bulgaria

ARTICLE INFO

Article history:
Received 4 March 2008
Received in revised form
20 January 2009
Accepted 20 January 2009
Available online 20 February 2009

Keywords:
Switched control systems
Bilinear control systems
Discrete approximations
Relaxation

ABSTRACT

A bilinear single-control system which can be viewed as a control formulation of a linear switched system is considered. The control is restricted to take values either (i) in {0, 1} (switched system), or (ii) in [0, 1] (relaxed system). In more practical considerations the control is often allowed to change only at the points of a given time-net with a step length h. The paper investigates what is the approximation error in terms of the reachable set in the two cases (i) and (ii). The error estimates that follow directly from known results are of order $\sqrt{h}$ and h, respectively. In the present paper estimations of order h and $h^{1.5}$ are proved in a constructive way. The second one makes use of the effect of non-accumulation of errors established earlier by the second author.

1. Introduction

In this paper we address the following two well-known approximation issues for control systems on a finite time interval with the usual set of admissible controls consisting of all measurable functions with values in a given constraining set. What is the approximation error (in a reasonable sense) if the controls are additionally restricted: (i) to be piece-wise constant functions with jumps only on a uniform time-net; (ii) to be as in (i) and to take only values that are extreme points of the control constraining set (see e.g. [1–4]). These issue are of key importance for time-discretization and for time-and-control-discretization of control/uncertain systems, respectively. The answer of each of these questions is not simple, and the questions are open, in general.

In the present paper the above two questions are investigated for a bilinear system resulting from a *switched linear system* (for applications of switched systems see e.g. [5,6]). Namely, the dynamics of the switched system is defined by two $n \times n$-matrices A_1 and A_2: a trajectory x satisfies for a.e. $t \in [0, 1]$ either the equation $\dot{x}(t) = A_1 x(t)$, or the equation $\dot{x}(t) = A_2 x(t)$. An equivalent formulation involves the bilinear control system

$$\dot{x}(t) = u(t)A_1 x(t) + (1 - u(t))A_2 x(t), \quad x(0) = x^0, \tag{1}$$

where the measurable function u takes values 0 or 1. Although simple, this system exhibits a rather complex behavior, as it will be pointed out below.

Denote $\mathcal{U} = \{u \in L_\infty(0, 1) : u(t) \in \{0, 1\} \text{ for a.e. } t\}$ – the set of *admissible controls* on [0, 1]. Let $x[u]$ denote the solution of (1) corresponding to some $u \in \mathcal{U}$. The reachable set of (1) on [0, 1] in the class $\mathcal{U}$ of admissible controls is defined as $R = \{x[u](1) : u \in \mathcal{U}\}$.

In practice it is often convenient to allow a jump of the control function only on a uniform mesh of time moments with a step length $h > 0$. This also enables application of a variety of single step discretization methods for simulation or for solving control problems for (1). Therefore, for a natural number N and $h = 1/N$ we define the set $\mathcal{U}^h$ of admissible controls, consisting of all functions with values 0 or 1 that are constant on each subinterval $[ih, (i + 1)h)$. Respectively, we denote by R^h the reachable set of (1) in this set of admissible controls. In Section 2 we address the problem of estimating the Hausdorff distance $H(\mathrm{cl}\, R, R^h)$ between the closure of R and $R^h \subset R$ (R needs not be closed, while R^h is always closed). This problem is not trivial for the following reasons: (i) the control set $\mathcal{U}^h$, as well as $\mathrm{cl}\, R$ and R^h are not convex; (ii) for any number m, the admissible controls with a variation not exceeding m may not be enough to generate the set R ([7, Proposition 2]). On the other hand, from the result in [3,4] one can obtain the estimation $H(\mathrm{cl}\, R, R^h) \le c\sqrt{h}$. We prove an

☆ This research was supported by the Austrian Science Foundation (FWF) under grant P18161. The work was completed during the visit of the second author at Université des Antilles et de la Guyane.

* Corresponding author at: Institute Mathematical Methods in Economics, Vienna University of Technology, Argentinierstrasse 8/119, A-1040 Vienna, Austria.
E-mail addresses: apietrus@univ-ag.fr (A. Pietrus), veliov@tuwien.ac.at (V.M. Veliov).

0167-6911/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.sysconle.2009.01.005

estimation of first order with respect to h, which is apparently the exact order. Namely, we constructively define a mapping $\mathcal{D}^h : \mathcal{U} \mapsto \mathcal{U}^h$ such that

$$\left|x[\mathcal{D}^h(u)](t) - x[u](t)\right| \le Ch, \quad t \in [0, 1],$$

where C is independent of u and h. Moreover, the mapping $\mathcal{D}^h$ is non-anticipative: the restriction $\mathcal{D}^h(u)_{|[0,ih]}$ is independent of $u_{|(ih,1]}$, $i = 1, \dots, N-1$. The proof utilizes the Volterra series expansion presented in [8].

In Section 3 we consider the so-called *relaxed* formulation of the switched system considered above, where the control function in (1) is allowed to take all values in [0, 1]. We denote by $\hat{\mathcal{U}}$ this set of admissible controls, and by $\hat{R}$ – the corresponding reachable set.

Similarly as before we consider the "practical" formulation in which the set of admissible controls is

$$\hat{\mathcal{U}}^h = \{u \in \hat{\mathcal{U}} : u \text{ is constant on each } [ih, (i+1)h), \quad i = 0, \dots, N-1\}.$$

Denote by $\hat{R}^h$ the corresponding reachable set of (1). Apparently $R^h \subset \hat{R}^h \subset \hat{R} = \operatorname{cl} R$, where the last equality follows from the Filippov–Ważewki relaxation theorem (see e.g. [9]). From the result in [10] one can easily obtain a first order estimate for $H(\hat{R}, \hat{R}^h)$. On the other hand, it was shown in [7, Example 1] that it may happen that some points from $\hat{R}$ are reachable only by controls $u \in \hat{U}$ with unbounded variations. This suggests that the first order estimation may be sharp. Nevertheless, we prove in Section 3 the estimation $H(\hat{R}, \hat{R}^h) \le ch^{\frac{3}{2}}$. The proof implicitly uses the effect of *non-accumulation of errors* established in [11,12], the essence of which is that in the discretization of multi-valued (or control) dynamic systems the error in the end of the time horizon may happen to be of higher order than the sum of the errors made at each step. In the proof we define a mapping $\hat{\mathcal{D}}^h : \hat{\mathcal{U}} \mapsto \hat{\mathcal{U}}^h$ such that

$$\left|x[\hat{\mathcal{D}}^h(u)](1) - x[u](1)\right| \le Ch^{1.5}.$$

We stress that, in contrast to $\mathcal{D}^h$, the mapping $\hat{\mathcal{D}}^h$ is anticipative. We expect that a non-anticipative mapping $\hat{\mathcal{D}}^h$ with the above property does not exist, in general, although a proof is not available.

Our analysis is constructive (see Lemmas 1 and 2 for the switched and for the relaxed system, respectively) and suggests approximation procedures that can be used either for simulation of frequently switching systems, or for solving optimal control problems for switched systems. This subject, however, is not discussed in the present paper.

2. Error analysis for the switched system

In this section we analyze what approximation of an arbitrary trajectory of the switched system (1) in the set $\mathcal{U}$ of admissible controls can one obtain by using controls from the set $\mathcal{U}^h$ only. As a consequence we obtain an estimation for the Hausdorff distance between the reachable sets R and R^h. To point out the principle difficulty, we recall two results from [13,7].

As usual, for two $n \times n$ matrices D and E we denote by $[D, E] = DE - ED$ the commutator (Lie bracket) of D and E. By definition, the pair of matrices (A_1, A_2) is nilpotent of order k if all Lie brackets containing A_1 and A_2 more than k times vanish. In this case the minimal number k with this property is called the *order of nilpotency* of (A_1, A_2) and is denoted by $\mathcal{N}(A_1, A_2)$. In [13] it was proved that in the case $\mathcal{N}(A_1, A_2) = 2$ every point of the reachable set R of (1) is reachable by a piece-wise constant control having not more than $m = 4$ switches.

In the cases where a number m with the above property exists, for every point of the reachable set one needs only to modify the corresponding (according to the theorem) $u \in \mathcal{U}$ on m intervals $[ih, (i+1)h)$ in order to obtain a $u^h \in \mathcal{U}^h$. Obviously the so-obtained u^h provides a trajectory deviating by no more than Cmh from the one for u. Thus we have the estimation $H(R, R^h) \le Cmh$. On the other hand, an example with $\mathcal{N}(A_1, A_2) = 3$ is given in ([7, Proposition 2]), where for any number m, the admissible controls with not more than m jumps are not enough to generate a dense subset of $\operatorname{cl} R$ (which formally corresponds to $m = \infty$). Therefore the first order estimate obtained below in the general case (which is obviously sharp) is not straightforward.

Theorem 1. *Let $L = \max\{\|A_2\|, \|A_2 - A_1\|\}$, where $\|A\|$ is the operator norm of the matrix A. Then for every trajectory x of (1) in the set $\mathcal{U}$ of admissible controls and for every $h = 1/N$ (with $N \ge 1$) there is a trajectory x^h of (1) generated by some $u^h \in \mathcal{U}^h$ such that*

$$\|x^h - x\|_{C[0,1]} \le 6Le^{3L}|x^0|h.$$

The following simple lemma is of key importance.

Lemma 1. *For every $u \in \mathcal{U}$ there exists $u^h \in \mathcal{U}^h$ such that for every $t \in [0, 1]$*

$$\left|\int_0^t (u(s) - u^h(s))\, ds\right| \le h.$$

Proof. The function u^h can be defined inductively by setting

$$u^h(t) = \begin{cases} 1 & \text{if } \int_0^{ih} (u(s) - u^h(s))\, ds + \int_{ih}^{(i+1)h} u(s)\, ds > h/2 \\ 0 & \text{else} \end{cases}$$

on $[ih, (i+1)h)$, $i = 0, \dots, N-1$. One can easily verify that

$$\left|\int_0^{ih} (u(s) - u^h(s))\, ds\right| \le h/2,$$

which implies the claim of the lemma since $\int_0^t (u(s) - u^h(s))\, ds$ is Lipschitz with a constant equal to 1. □

Proof of Theorem 1. Let $u \in \mathcal{U}$ be arbitrarily chosen and let $x = x[u]$. Let u^h be defined as in Lemma 1 and let $x^h = x[u^h]$.

We expand the function x in a Volterra series (one may use the representation in [8], which simplifies in the linear case):

$$x(t) = x^0 + \left[\sum_{k=1}^{\infty} \int_0^t \int_0^{s_1} \cdots \int_0^{s_{k-1}} F(s_1) \cdots F(s_k)\, ds_k \cdots ds_1\right] x^0,$$

where $F(t) = u(t)A_1 + (1 - u(t))A_2$. Let us consider the k-th member of the sum, S_k. Denote $A = A_2$, $D = A_1 - A_2$, so that $F = A + uD$. Also we denote

$$J_k = \{j = (j_1, \dots, j_k) : j_i \in \{0, 1\}\},$$
$$v_{j_i}(s) = \begin{cases} u(s) & \text{if } j_i = 1, \\ 1 & \text{if } j_i = 0, \end{cases} \quad B_{j_i}(s) = \begin{cases} A & \text{if } j_i = 1, \\ D & \text{if } j_i = 0. \end{cases}$$

With these notations

$$S_k := \sum_{j \in J_k} B_{j_1} \cdots B_{j_k} \times \int_0^t \int_0^{s_1} \cdots \int_0^{s_{k-1}} v_{j_1}(s_1) \cdots v_{j_k}(s_k)\, ds_k \cdots ds_1. \tag{2}$$

Using an analogous expression for the solution x^h, with $v_{j_i}^h$ and S_k^h defined as above for the control u^h we obtain

$$S_k - S_k^h = \sum_{j \in J_k \setminus (0,\dots,0)} B_{j_1} \cdots B_{j_k} \times \int_0^t \int_0^{s_1} \cdots \int_0^{s_{k-1}} [v_{j_1}(s_1) \cdots v_{j_k}(s_k) - v_{j_1}^h(s_1) \cdots v_{j_k}^h(s_k)]\, ds_k \cdots ds_1.$$

We represent

$$v_{j_1}(s_1)\cdots v_{j_k}(s_k) - v^h_{j_1}(s_1)\cdots v^h_{j_k}(s_k) = (v_{j_1}(s_1) - v^h_{j_1}(s_1)) \times v_{j_2}(s_2)\cdots v_{j_k}(s_k) + v^h_{j_1}(s_1)(v_{j_2}(s_2) - v^h_{j_2}(s_2))v_{j_3}(s_3)\cdots v_{j_k}(s_k) + \cdots + v^h_{j_1}(s_1)\cdots v^h_{j_{k-1}}(s_{k-1})(v_k(s_k) - v^h_{j_k}(s_k)).$$

Each of the summands contains a factor $(v_{j_p}(s_p) - v^h_{j_p}(s_p))$, which is zero if $j_p = 0$, and equals $(u(s_p) - u^h(s_p))$ if $j_p = 1$. The multiple integral $\int_0^t \int_0^{s_1} \cdots \int_0^{s_{k-1}}$ of each of these terms is either zero, or can be estimated by

$$\frac{2t^{k-1}}{(k-1)!}\left|\int_0^t (u(s) - u^h(s))\,\mathrm{d}s\right|.$$

To prove this, one has to integrate by parts replacing $(u(s_p) - u^h(s_p))$ by $\mathrm{d}_{s_p}\int_0^{s_p}(u(s) - u^h(s))\,\mathrm{d}s$ and to use $v_{j_i}(s) \in [0,1]$. We skip the details.

Using the estimation in Lemma 1, the definition of the number L, and the inequality $k2^k \le 3^k$ we obtain that

$$\left|S_k - S_k^h\right| \le 2^k L^k k \frac{2t^{k-1}}{(k-1)!}h \le 3^k L^k \frac{2t^{k-1}}{(k-1)!}h.$$

Then

$$|x(t) - x^h(t)| \le \sum_{k=1}^{\infty} 6L\frac{(3Lt)^{k-1}}{(k-1)!}|x^0|h,$$

which implies the claim of the theorem. □

Corollary 1. *Under the conditions of Theorem 1 it holds that*

$$H(R^h, R) \le 6Le^{3L}|x^0|h.$$

3. Error analysis for the relaxed system

In this section we investigate the distance between the reachable sets $\hat{R}$ and $\hat{R}^h$ of (1), for the convexified sets $\hat{\mathcal{U}}$ and $\hat{\mathcal{U}}^h$ of admissible controls, respectively. From the main result in [10] one can easily obtain the estimation $H(\hat{R}^h, R) \le Ch$. The aim of this section is to prove that the estimation

$$H(\hat{R}^h, R) \le Ch^{1.5}. \tag{3}$$

The proof makes an essential implicit use of the effect of non-accumulation of errors established in [11,12].

Theorem 2. *For any matrices A_1 and A_2 there exists a constant C such that for every $N \ge 1$ the estimation (3) holds true with $h = 1/N$.*

Proof. Let N and $M \in [2, N]$ be natural numbers, and let $h = 1/N$.

Lemma 2. *For every $u \in \hat{\mathcal{U}}$ there exists $u^h \in \hat{\mathcal{U}}^h$ such that*

$$\left|\int_0^t (u(s) - u^h(s))\,\mathrm{d}s\right| \le h \quad \forall\, t \in [0, Mh],$$

$$\left|\int_0^{Mh} (u(s) - u^h(s))\,\mathrm{d}s\right| \le \frac{1}{2}h^2,$$

$$\left|\int_0^{Mh} s(u(s) - u^h(s))\,\mathrm{d}s\right| \le \frac{1}{2}h^2.$$

Proof. Denote $Z = \{z(u) = (z_0(u), z_1(u), hz_2(u), \ldots, hz_M(u)) : u \in \hat{\mathcal{U}}\}$, where

$$z_0(u) = \int_0^{Mh} tu(t)\,\mathrm{d}t, \quad z_i(u) = \int_0^{(M+1-i)h} u(t)\,\mathrm{d}t, \quad i = 1, \ldots, M.$$

Similarly we denote $Z^h = \{z(u) = (z_0(u), z_1(u), hz_2(u), \ldots, hz_M(u)) : u \in \hat{\mathcal{U}}^h\}$. We shall prove that

$$H_\infty(Z^h, Z) \le \frac{1}{2}h^2, \tag{4}$$

where H_∞ is the Hausdorff distance with respect to the norm $|z|_\infty = \max_i |z_i|$ in $\mathbf{R}^{M+1}$. This would imply the lemma, since for every $u \in \mathcal{U}$ and $u^h \in \mathcal{U}^h$ the function $t \longrightarrow \int_0^t (u(s) - u^h(s))\,\mathrm{d}s$ is Lipschitz with constant 1.

Both sets Z and Z^h are convex and compact, and obviously $Z^h \subset Z$. Therefore, by a standard convex analysis, there exists a vector $l = (l_0, l_1, \ldots l_M)$ with $|l|_1 := \sum_{i=0}^M |l_i| = 1$ such that

$$H_\infty(Z^h, Z) = \max_{u\in\hat{\mathcal{U}}}\langle l, z(u)\rangle - \max_{u\in\hat{\mathcal{U}}^h}\langle l, z(u)\rangle. \tag{5}$$

Let $u \in \hat{\mathcal{U}}$ be a maximizer of the first member in the right-hand side. Then one can represent

$$\langle l, z(u)\rangle = l_0\int_0^{Mh} tu(t)\,\mathrm{d}t + l_1\int_0^{Mh} u(t)\,\mathrm{d}t + h\sum_{i=2}^{M} l_i \int_0^{(M+1-i)h} u(t)\,\mathrm{d}t = \int_0^{Mh}\lambda(t)u(t)\,\mathrm{d}t,$$

where

$$\lambda(t) = \begin{cases} tl_0 + l_1 + hl_2 + \cdots + hl_M & \text{for } t \in [0, h),\\ tl_0 + l_1 + hl_2 + \cdots + hl_{M-1} & \text{for } t \in [h, 2h),\\ \ldots & \ldots\\ tl_0 + l_1 & \text{for } t \in [(M-1)h, Mh]. \end{cases}$$

Define the averages

$$u_i = \frac{1}{h}\int_{ih}^{(i+1)h} u(t)\,\mathrm{d}t$$

and the function $\bar{u}^h(t) = u_i$ on $[ih, (i+1)h)$. Obviously $\bar{u}^h \in \hat{\mathcal{U}}^h$. Moreover, $z_i(\bar{u}^h) = z_i(u)$, $i = 1, \ldots, M$. If $l_0 = 0$, we would have $H_\infty(Z^h, Z) = 0$ according to (5). Let us consider the alternative case: $l_0 < 0$ (the other possibility, $l_0 > 0$, can be treated in the same way).

Assuming $l_0 < 0$, the function λ is linear and decreasing on every interval $[ih, (i+1)h)$. Hence, the structure of the maximizer u is as follows:

$$u(t) = \begin{cases} 1 & \text{for } t \in [ih, ih + \tau_i),\\ 0 & \text{for } t \in [ih + \tau_i, (i+1)h), \end{cases} \tag{6}$$

where $\tau_i \in [0, h]$. We shall define $u^h \in \hat{\mathcal{U}}^h$ so that

$$\langle l, z(u)\rangle - \langle l, z(u^h)\rangle \le \frac{1}{2}h^2, \tag{7}$$

which will prove (4) according to (5) and the choice of u.

We shall define the function $u^h \in \hat{\mathcal{U}}$ by modifying the stepwise averaged function $\bar{u}^h$ on some intervals. Denote

$$d_i = \int_{ih}^{(i+1)h} t(u(t) - \bar{u}^h(t))\,\mathrm{d}t.$$

According to (6) and the definition of $\bar{u}^h(t) = u_i = \tau_i/h$ we have

$$d_i = \int_{ih}^{ih+\tau_i} t\,\mathrm{d}t - \int_{ih}^{(i+1)h} t\frac{\tau_i}{h}\,\mathrm{d}t = -\frac{1}{2}\tau_i(h - \tau_i).$$

Clearly $|d_i| \le h^2/8$. For every i we have (due to $u_i = \frac{\tau_i}{h}$) that

$$u_i + v_i \in [0,1] \text{ for every } v_i \text{ with } |v_i| \le \Delta_i := \min\{\tau_i/h, 1-\tau_i/h\}. \quad (8)$$

Comparing Δ_i with d_i we easily obtain that $\Delta_i \ge 2|d_i|h^{-2}$. We redefine $\Delta_i = 2|d_i|h^{-2}$, which does not affect the validity of (8).

Denote by $\Gamma_j, j \ge 1$ the set of all i such that

$$|d_i| \in \frac{h^2}{8}\left[\frac{1}{2^j}, \frac{1}{2^{j-1}}\right], \quad j = 1, \dots,$$

and by k_j – the number of elements of Γ_j. We shall modify the function $\bar{u}^h$ in the following way. Denote by J_1 the set of those j for which Γ_j is non-empty, and by J_2 – the set of those j for which Γ_j contains at least 2 elements. For $j \in J_2$ we consider the first interval $(i_1(j)h, (i_1(j)+1)h)$ and the last interval $(i_2(j)h, (i_2(j)+1)h)$ in Γ_j, and modify $\bar{u}^h$ by subtracting $v_j = \alpha \min\{\Delta_{i_1(j)}, \Delta_{i_2(j)}\}$ from $\bar{u}^h$ in the first interval, and adding the same value to $\bar{u}^h$ on the second interval. Here $\alpha \in [0,1]$ is a parameter. Making this modification for all $j \in J_2$ we obtain a function $u^h \in \hat{\mathcal{U}}^h$.

Obviously $z_1(u^h) = z_1(\bar{u}^h) = z_1(u)$. Moreover, for $k > 1$

$$|z_k(u^h) - z_k(u)| \le \sum_{j\in J_2} h\alpha\Delta_{i_1(j)} \le h\sum_{j\ge 1} 2|d_{i_1(j)}|h^{-2}$$
$$\le 2h^{-1}\sum_{j\ge 1}\frac{h^2}{8}\frac{1}{2^{j-1}} \le \frac{h}{2}. \quad (9)$$

For $k = 0$ we have

$$z_0(u^h) - z_0(u) = z_0(\bar{u}^h) - z_0(u) + z_0(u^h) - z_0(\bar{u}^h)$$
$$= \sum_{i=0}^{M-1} d_i + \sum_{j\in J_2}\left[-\int_{i_1(j)h}^{(i_1(j)+1)h} t v_j + \int_{i_2(j)h}^{(i_2(j)+1)h} t v_j\right]$$
$$= -\sum_{i=0}^{M-1}|d_i| + \sum_{j\in J_2} h^2 v_j (i_2(j) - i_1(j)) = -A + B.$$

Here

$$A = \sum_{i=0}^{M-1}|d_i| = \sum_{j\in J_1}\sum_{i\in\Gamma_j}|d_i| \le \frac{h^2}{8}\sum_{j\in J_1}\frac{1}{2^{j-1}}k_j =: \frac{h^2}{8}\gamma.$$

On the other hand

$$B \ge \alpha\frac{h^2}{4}\sum_{j\in J_2}\frac{1}{2^j}(k_j - 1) = \alpha\frac{h^2}{8}\sum_{j\in J_1}\frac{1}{2^{j-1}}(k_j-1) = \alpha\frac{h^2}{8}(\gamma - 2).$$

For $\alpha = 1$ we have $A \le B + h^2/4$, hence we may find $\alpha \in [0,1]$ so that $z_0(u^h) - z_0(u) = B - A = -h^2/4$. Using this and (9) we obtain

$$\langle l, z(u) - z(u^h)\rangle$$
$$= l_0(z_0(u) - z_0(u^h)) + l_1(z_1(u) - z_1(u^h))$$
$$+ h\sum_{k=2}^{M} l_k(z_k(u) - z_k(u^h))$$
$$\le |l|_1\, |(z_0(u-u^h), z_1(u-u^h), hz_2(u-u^h), \dots,$$
$$hz_M(u-u^h))|_\infty \le \frac{h^2}{2},$$

which proves (7), hence (4) and the lemma. □

We continue with the proof of the theorem. Let $u \in \hat{\mathcal{U}}$ be arbitrarily chosen and let x be the corresponding solution of (1) on $[0,1]$. We shall define $u^h \in \hat{\mathcal{U}}^h$ such that for $x^h := x[u^h]$ it holds that $|x^h(1) - x(1)| \le Ch^{1.5}$, with C independent of $u \in \hat{\mathcal{U}}$ and h, which will prove the theorem.

Let K be the largest natural number such that $K^2 \le N$. Define $\tau_i = iKh$, $i = 0, \dots, K$, $\tau_{K+1} = 1$. In each $[\tau_i, \tau_{i+1})$ we define u^h as in Lemma 2 applied for $M = K$, excepting the last interval $[\tau_K, \tau_{K+1}]$, where M may be smaller.

On each subinterval $[\tau_i, \tau_{i+1})$ we expand x and x^h in a Volterra series truncating the terms of order higher than three. Since (1) is stationary, we make the local analysis below only on the interval $[0, Mh]$, where M is defined above. We use expressions (2) for S_k, $k = 1, 2, 3$. Regrouping the terms and changing the order of integration where appropriate we represent

$$x(t) = \Big[f_0(t) + f_1(t)\int_0^t u(s)\,ds + f_2(t)\int_0^t su(s)\,ds$$
$$+ f_3(t)\int_0^t s^2 u(s)\,ds$$
$$+ f_4(t)\int_0^t u(s)\int_0^s u(\tau)\,d\tau\,ds + f_5(t)\int_0^t\int_0^s u(\tau)$$
$$\times\int_0^\tau u(\theta)\,d\theta\,d\tau\,ds$$
$$+ f_6(t)\int_0^t u(s)\int_0^s \tau u(\tau)\,d\tau\,ds + f_7(t)\int_0^t u(s)\int_0^s u(\tau)$$
$$\int_0^\tau u(\theta)\,d\theta\,d\tau\,ds + t^4 g(t)\Big]x^0,$$

where $f_0, \dots, f_7$ are bounded functions depending on A_1 and A_2 and $g(\cdot)$ is bounded, uniformly in u. A similar expansion we have also for x^h, with the same functions f_i. Subtracting the two expansions for $t = Mh$ we obtain

$$|x^h(Mh) - x(Mh)| \le C(h^2 + M^2h^3 + M^4h^4), \quad (10)$$

where the constant C is independent on $u \in \hat{\mathcal{U}}$ and h. This estimation is not straightforward and is obtained by separately comparing each term in the above representation of $x(t)$ with that for x^h. The term h^2 comes from $\int_0^{Mh}(u^h(s) - u(s))\,ds$ and $\int_0^{Mh} s(u^h(s) - u(s))\,ds$, due to the definition of u^h using Lemma 2. The term M^4h^4 is the truncation error. We shall present the estimate for some of the other terms, denoting $\Delta(s) = u^h(s) - u(s)$. For the term multiplying f_3 we have integrating by parts

$$\left|\int_0^{Mh} s^2\Delta(s)\,ds\right| = \left|(Mh)^2\int_0^{Mh}\Delta(s)\,ds\right|$$
$$+ \left|\int_0^{Mh} 2s\int_0^s \Delta(\tau)\,d\tau\,ds\right| \le 2M^2h^3.$$

For the term multiplying f_4,

$$\int_0^{Mh} u(s)\int_0^s u(\tau)\,d\tau\,ds = \frac{1}{2}\left(\int_0^{Mh} u(s)\,ds\right)^2,$$

hence

$$\left|\int_0^{Mh} u^h(s)\int_0^s u^h(\tau)\,d\tau\,ds - \int_0^{Mh} u(s)\int_0^s u(\tau)\,d\tau\,ds\right|$$
$$= \frac{1}{2}\left|\left(\int_0^{Mh} u^h(s)\,ds\right)^2 - \left(\int_0^{Mh} u(s)\,ds\right)^2\right|$$
$$= \frac{1}{2}\left|\int_0^{Mh}\Delta(s)\,ds\right|\int_0^{Mh}(u^h(s) + u(s))\,ds \le \frac{1}{2}h^2.2Mh \le h^2.$$

Let us consider also the term multiplying f_5. We estimate the difference of the expression for u^h and u by

$$\left|\int_0^{Mh}\int_0^s u^h(\tau)\int_0^\tau \Delta(\theta)\,\mathrm{d}\theta\,\mathrm{d}\tau\,\mathrm{d}s\right| + \left|\int_0^{Mh}\int_0^s \Delta(\tau)\int_0^\tau u(\theta)\,\mathrm{d}\theta\,\mathrm{d}\tau\,\mathrm{d}s\right|.$$

The first term can be estimated by $(Mh)^2h$ since $\left|\int_0^\tau \Delta(\theta)\,\mathrm{d}\theta\right| \le h$. For the second we obtain the same estimation after an integration by parts and using the last inequality. The rest of the terms can be estimated similarly.

Since $M \le \sqrt{N}$, (10) implies that $|x^h(Mh) - x(Mh)| \le 3Ch^2$. A similar estimation holds for the error created on every interval $[\tau_k, \tau_{k+1}]$, $k = 0, \dots, K$. Then a standard propagation of errors argument implies that

$$|x^h(1) - x(1)| \le \mathrm{e}^L 3Ch^2(K+1) \le C_1 h^{1.5},$$

where $L = \|A_1\| + \|A_2\|$ and $\|A\|$ is the operator norm of A. The theorem is proved. □

We mention that the choice of $K \approx \sqrt{N}$ in the above proof is, in a sense, optimal. With any choice of K the total error (which is the sum of the local errors $h^2 + K^2h^3 + K^4h^4$) is $C(h^2 + K^2h^3 + K^4h^4)N/K \approx C(h/K + Kh^2 + K^3h^3)$, which is minimal at $K = h^{-1/2}$.

One can question the sharpness of the order of convergence 1.5. Indeed, it is proved in [14,7] that in the case $\mathcal{N}(A_1, A_2) = 3$ every point of the reachable set $\hat{R}$ can be reached by a piecewise constant control having not more than $m = 4$ switches. Under the conditions of the above theorem one can obtain by a standard analysis that $H(\hat{R}^h, R) \le Cmh^2$. However, in [7] it is shown by an example (making use of the Fuller phenomenon) with $n = 7$ and $\mathcal{N}(A_1, A_2) = 5$ that in general finite numbers of switches are not enough to generate $\hat{R}$ (thus formally $m = \infty$). The question about a sharp estimation is open.

References

[1] V.M. Veliov, On the time-discretization of control systems, SIAM J. Control Optim. 35 (5) (1997) 1470–1486.
[2] F. Lempio, V.M. Veliov, Discrete approximations to differential inclusions, GAMM Mitt. Ges. Angew. Math. Mech. 21 (1998) 101–135.
[3] T. Donchev, Approximation of lower semicontinuous differential inclusions, Numer. Funct. Anal. Optim. 22 (1–2) (2001) 55–67.
[4] G. Grammel, Towards fully discretized differential inclusions, Set-Valued Anal. 11 (1) (2003) 1–8.
[5] A.S. Morse (Ed.), Control Using Logic-based Switching, Springer, London, 1997.
[6] D. Liberzon, Switchings in Systems and Control, Birghäuser, Boston, 2003.
[7] M. Margaliot, A counterexample to a conjecture of Gurvits on switched systems, IEEE Trans. Automat. Control 52 (6) (2007) 1123–1126.
[8] L. Grüne, P.E. Kloeden, Higher order numerical approximation of switching systems, Systems Control Lett. 55 (9) (2006) 746–754.
[9] J.P. Aubin, A. Cellina, Differential Inclusions, Springer, 1984.
[10] A. Dontchev, E. Farkhi, Error estimates for discretized differential inclusion, Computing 41 (4) (1989) 349–358.
[11] V.M. Veliov, Second order discrete approximations to linear differential inclusions, SIAM J. Numer. Anal. 29 (2) (1992) 439–451.
[12] B.D. Doitchinov, V.M. Veliov, Parametrisations of integrals of set-valued mappings and applications, J. Math. Anal. Appl. 179 (2) (1993) 483–499.
[13] L. Gurvitz, Stability of discrete linear inclusions, Linear Algebra Appl. 231 (1995) 47–85.
[14] Y. Sharon, M. Margaliot, Third-order nilpotency, finite switchings and asymptotic stability, J. Differential Equations 233 (2007) 135–150.