IMPROVED REGULARITY ASSUMPTIONS FOR PARTIAL OUTER CONVEXIFICATION OF MIXED-INTEGER PDE-CONSTRAINED OPTIMIZATION PROBLEMS[*],[**],[***]

Paul Manns[****] and Christian Kirches

Abstract. Partial outer convexification is a relaxation technique for MIOCPs being constrained by time-dependent differential equations. Sum-Up-Rounding algorithms allow to approximate feasible points of the relaxed, convexified continuous problem with binary ones that are feasible up to an arbitrarily small $\delta > 0$. We show that this approximation property holds for ODEs and semilinear PDEs under mild regularity assumptions on the nonlinearity and the solution trajectory of the PDE. In particular, requirements of differentiability and uniformly bounded derivatives on the involved functions from previous work are not necessary to show convergence of the method.

Mathematics Subject Classification. 49M20, 49N60, 90C11, 49J20.

1. Introduction

Mixed-integer PDE-constrained optimization problems (MIPDECOs) form a broad class of mixed-integer optimal control problems (MIOCPs). They can serve as a powerful modeling tool for a large variety of real-world problems from topology optimization [12] over oil-spill response planning [22] to optimal control of large-scale gas networks [23]. Unfortunately, they combine the linear/quadratic/cubic increase of some variables due to the distribution in the spatial domain with the curse of dimensionality of the branch-and-bound tree for the integer control variable trajectories. Therefore, techniques are necessary to be able to approximate feasible points fastly.

Sum-Up-Rounding is such a technique that computes approximately feasible points of the mixed integer problem from feasible points of a relaxed continuous problem in linear time. It was elaborated for ODE-constrained MIOCPs by Sager [17–19] and was transferred to semilinear PDE-constrained MIOCPs by Hante and Sager [10, 11]. While the aforementioned publications show the power of this approach, they impose regularity assumptions on the problem that are quite restrictive in the PDE-case. We are going to weaken the

Institut für Mathematische Optimierung, Technische Universität Carolo-Wilhelmina zu Braunschweig, Universitätsplatz 2, 38106 Braunschweig, Germany.

[****] Corresponding author: paul.manns@tu-bs.de

regularity assumptions such that they are fulfilled for a broader class of problems and can be checked more easily. However, this prevents us from making *a priori* estimates available, which currently require those regularity assumptions.

In particular, we are dealing with the following MIOCPs which include a potentially unbounded operator A:

$$\begin{aligned}
\min_{x,u,v} \quad & J(x,u) \\
\text{s.t.} \quad & \dot{x}(t) = Ax(t) + f(t, x(t), u(t), v(t)) \text{ a.e. } t \in [0, T] \\
& x(0) = x_0 \\
& v(t) \in V && \text{a.e. } t \in [0, T] \\
& 0 \leq c(x(t), u(t)) && \text{a.e. } t \in [0, T],
\end{aligned} \quad \text{(MIOCP)}$$

where we assume A to be the generator of a C_0-semigroup on a real Banach space X, $J \in C(C([0, T], X) \times L^1((0, T), U), \mathbb{R})$, $x \in C([0, T], X)$ (*i.e.* being a mild solution of the semilinear equation), $u \in L^1((0, T), U)$ for a real Banach space U, $v \in L^\infty((0, T), \mathbb{R}^{n_v})$ with $v(t) \in V$ a.e. where $V \subset \mathbb{R}^{n_v}$ and $|V| < \infty$, and we assume the function $f : [0, T] \times X \times U \times V \to X$ being uniformly continuous in the first and Lipschitz continuous in the second and third argument. In particular, we assume that the integer control is not distributed in space. We assume the constraint function $c : X \times U \to Y$ for some Banach space Y to be Lipschitz continuous in the first argument.

Problems of the type (MIOCP) can be equivalently reformulated by means of *partial outer convexification*, see the publications by Berkovitz [2], Cesari [3], and Sager [17–19]. These proofs were developed for ODEs, but can be applied in the presence of semilinear PDEs as in (MIOCP) without any modification. The *partial outer convexification* of (MIOCP) reads

$$\begin{aligned}
\min_{x,u,\beta} \quad & J(x,u) \\
\text{s.t.} \quad & \dot{x}(t) = Ax(t) + \sum_{i=1}^{|V|} \beta_i(t) f(t, x(t), u(t), v_i) \text{ a.e. } t \in [0, T] \\
& x(0) = x_0 \\
& \beta(t) \in \{0, 1\}^{|V|} && \text{a.e. } t \in [0, T] \\
& 1 = \sum_{i=1}^{|V|} \beta_i(t) && \text{a.e. } t \in [0, T] \\
& -\delta \leq c(x(t), u(t)) && \text{a.e. } t \in [0, T],
\end{aligned} \quad \text{(BC}_\delta\text{)}$$

with the choice $\delta = 0$. Now, one can relax (MIOCP), or equivalently (BC$_\delta$) with $\delta = 0$, by weakening the SOS-1 property of β to convex combinations,

$$\begin{aligned}
\min_{x,u,\alpha} \quad & J(x,u) \\
\text{s.t.} \quad & \dot{x}(t) = Ax(t) + \sum_{i=1}^{|V|} \alpha_i(t) f(t, x(t), u(t), v_i) \text{ a.e. } t \in [0, T] \\
& x(0) = x_0 \\
& \alpha(t) \in [0, 1]^{|V|} && \text{a.e. } t \in [0, T] \\
& 1 = \sum_{i=1}^{|V|} \alpha_i(t) && \text{a.e. } t \in [0, T] \\
& 0 \leq c(x(t), u(t)) && \text{a.e. } t \in [0, T].
\end{aligned} \quad \text{(RC)}$$

To describe the relationship between feasible points of (RC) and feasible points of (BC$_\delta$) for a small $\delta > 0$ and constructed by rounding, we introduce the following definition.

Definition 1.1 (Vanishing integrality gap). Let $(\phi_n)_n \subset L^\infty((0, T), \mathbb{R})$ be a bounded sequence such that $\Phi_n(t) := \int_0^t \phi_n(s) \mathrm{d}s$ satisfies

$$\|\Phi_n\|_{L^\infty} \to 0.$$

Then, we call $(\phi_n)_n$ a sequence of *vanishing integrality gap*.

The mentioned Sum-Up-Rounding algorithm is given below.

Definition 1.2 (Sum-Up-Rounding algorithm, [17, 19]). Let $0 = t_0 < \ldots < t_N = T$ be a discretization grid of $[0, T]$ with maximum discretization width $\Delta t \coloneqq \max\limits_{i \in \{0, N-1\}} t_{i+1} - t_i$. For $\alpha \in L^\infty((0, T), \mathbb{R}^{|V|})$, we define a binary-valued piecewise-constant function $\beta(\alpha) : [0, T] \to \{0, 1\}^{|V|}$ iteratively for $i = 0, \ldots, N-1$ as

$$\beta(\alpha)_j(t)|_{[t_i, t_{i+1}]} \coloneqq \begin{cases} 1 & : & j = \arg\max\limits_{k \in \{1, \ldots, |V|\}} \int_0^{t_{i+1}} \alpha_k(t) \mathrm{d}t - \int_0^{t_i} \beta(\alpha)_k(t) \mathrm{d}t \\ 0 & : & \text{else} \end{cases}.$$

If the maximum is ambiguous, exactly one of the maximizing indices has to be chosen by $\arg\max$.

The algorithm can be summarized as follows. For the time intervals indexed by $i = 0, \ldots, N-1$, rounded controls are computed one after another. The index $j \in \{1, \ldots, |V|\}$ identifies the discrete control choices. On the first interval $[0, t_1]$, the component of ω corresponding to the largest component of $\int_0^{t_1} \alpha$ is set to one; the others are set to zero. For all subsequent intervals indexed by i, the algorithm computes the integrated difference between α and ω up to the time points t_i, the so-called *integrated control deviation*. The obtained vector is added to $\int_{t_i}^{t_{i+1}} \alpha$ and the rounding is computed afterwards using the maximizing index of this sum of vectors. In this way, the outcomes of previous rounding decisions are taken into account, which enables the following proposition that is due to Sager and states that Sum-Up-Rounding indeed yields sequences of vanishing integrality gap.

Proposition 1.3 (Sum-Up-Rounding yields Vanishing Integrality Gap, [19]). *Let $\alpha \in L^\infty((0, T), \mathbb{R}^{|V|})$ solve* (RC) *and β_n denote the binary control be computed from α with maximum discretization width $\frac{1}{n}$ by means of Sum-Up-Rounding. Then, the sequence of control deviations $\phi_n \coloneqq \alpha - \beta_n$ satisfies*

$$\sup_{t \in [0, T]} \left\| \int_0^t \phi_n(s) \mathrm{d}s \right\|_\infty \le C \frac{1}{n}$$

for a constant $C > 0$. That is, each coordinate sequence of $(\phi_n)_n$ is of vanishing integrality gap.

Remark 1.4. Note that it is necessary to relax the algebraic constraint by an arbitrarily small $\delta > 0$ in (BC_δ) to avoid the situation of a degenerate feasible set, *cf.* ([3], Chap. 18.7).

Remark 1.5. In work under review [13, 15], the authors extend the theory for additional "combinatorial" constraints of the form $0 \le c(x(t), u(t), v(t))$. Some of the results presented there can be included in the PDE setting here without any problems. We do not want to elaborate on that here and just note that the reformulation there shows that vanishing constraints $0 \le \beta_{n,i}(t) c(x(t), u(t), v_i)$ can be taken care of by a Sum-Up-Rounding variant described in [13, 15], and that the claim of Proposition 1.3 still holds.

If the involved initial value problem (IVP) provides enough regularity, this implies $y_n \to x$ and one obtains a result of the following form.

Proposition 1.6. *Let α, u and the corresponding solution x of the IVP be feasible for* (RC), *$(\beta_n)_n$ be constructed such that $(\phi_n)_n$ is of vanishing integrality gap. Furthermore, let additional regularity assumptions on the IVP hold (see below). Then, the sequence of state vectors $(y_n)_n$ corresponding to $(\beta_n)_n$ and u satisfies*

$$y_n \to x.$$

Furthermore, by continuity of J and c

$$J(y_n, u) \to J(x, u)$$

and

$$c(y_n, u) \to c(x, u).$$

In absence of the constraint c, Proposition 1.6 implies that the infimal value of (BC_0) coincides with the minimal value of (RC) and a sequence of feasible points exists such that the infimal value is approached. This holds true regardless of whether a minimizer of (BC_0) exists or not. In the presence of the constraint, Proposition 1.6 implies that a sequence of points exists such that for all $\delta > 0$ there exists n_δ such that all subsequent elements of the sequence are feasible for (BC_δ) and the objective converges to the minimizing value of (RC). In this case, convergence to the infimal value of (BC_0) cannot be guaranteed as the constraint can lead to a degenerate feasible set, see an example by Cesari ([3], Chap. 18.7), ([13], Sect. 4).

As indicated by the requirements of Proposition 1.6, some regularity assumptions on the PDE are needed to obtain the convergence $y_n \to x$. Sufficient regularity assumptions have been provided in the presence of ODEs by Sager in [19] and in the presence of semilinear PDEs by Hante and Sager in [11] Theorem 1. In particular, they require

$$\left\| \frac{\mathrm{d}}{\mathrm{d}s} T(t - s) f(x(s), u(s), v_i) \right\|_X \leq C$$

for some $C > 0$ a.e. on $0 < s < t < T$ and $i \in \{1, \dots, |V|\}$ where $(T(t))_{t \geq 0}$ denotes the semigroup generated by A. In [10], the results are extended to a class of hyperbolic PDEs where regularity conditions involving differentiability of the mapping $x \mapsto f(x, u, v_i)$ and piecewise smooth controls are required to prove the result, see Hypothesis 3 and the results thereafter in [10] for the details. Such requirements may be difficult to check, especially if classical solutions of the IVP cannot be expected to exist. This can happen, for example, if the initial value x_0 is not in the domain of A, but just in X, or the trajectory $s \mapsto f(x(s), u(s), v_i)$ can only be assumed to be an L^1-function, but not to be differentiable w.r.t s, *cf.* ([1], Prop. 3.1.16, Cor. 3.1.17).

1.1. Contribution

We generalize the existing results on the ability to approximate solution trajectories for (RC) with binary-valued ones feasible for (BC_δ), computed with Sum-Up-Rounding, to a class of semilinear PDEs. In particular, we consider the following IVPs,

$$\dot{x}(t) = Ax(t) + \sum_{i=1}^{|V|} \alpha_i(t) f(t, x(t), u(t), v_i), \ x(0) = x_0 \tag{1.1}$$

$$\dot{y}_n(t) = Ay_n(t) + \sum_{i=1}^{|V|} \beta_{n,i}(t) f(t, y_n(t), u(t), v_i), \ y(0) = x_0 \tag{1.2}$$

where x solves (1.1) in (RC), and y_n solves (1.2) in (BC_δ) with β_n being computed by Sum-Up-Rounding from α on a partition of $[0, T]$ into n equidistant intervals. We show

$$\sup_{t \in [0,T]} \|y_n(t) - x(t)\|_X \underset{n \to \infty}{\to} 0 \tag{1.3}$$

under the regularity assumption $f(x(s), u(s), v_i) \in L^1((0,T), X)$ for $i \in \{1, \dots, |V|\}$ and with regard to Definition 1.1. In particular, Lipschitz continuity of f in x and u and the availability of mild solutions will do. Furthermore, we characterize the convergence of $(\beta_n)_n$ to α by means of weak(*) topologies in L^p-spaces in Theorem 3.4.

1.2. Structure of the remainder

We state our main statement and a setup comprising a broad class of PDEs and corresponding control problems for which it holds in Section 2. Furthermore, we point out its consequences for the existing theory of Sum-Up-Rounding and partial outer convexification. In Section 3, we prove the aforementioned approximation result. Therefore, we combine of the convergence of the β_n to α in a weak sense with a compactness result provided by semigroup theory and the findings in [20]. We demonstrate the results on a computational example in Section 4. Finally in Section 5, we summarize our results in relation to the literature discussed above. Furthermore, we put the results in context of the Filippov–Ważewski theorem where related questions have been studied outside the mixed-integer optimization context several decades ago.

2. MAIN STATEMENT AND CONSEQUENCES

As mentioned above, mild solutions are the solution concept of semilinear PDEs with which we will work in the remainder. Therefore, we recall its definition and existence and uniqueness.

Definition 2.1 (Chap. 4, Def. 2.3 in [16], Prop. 3.1.16 in [1]). Let A generate a C_0-semigroup $(T(t))_{t \geq 0}$ on X, $x_0 \in X$ and $f \in L^1((0,T), X)$. Then, the function $x \in C([0,T], X)$ defined by means of the variation of constants formula

$$x(t) := T(t)x_0 + \int_0^t T(t-s)f(s)\mathrm{d}s$$

for $t \in [0,T]$ is called a *mild solution* of the IVP

$$\dot{x}(t) = Ax(t) + f(t), \; x(0) = x_0.$$

Corollary 2.2 (Existence and uniqueness). *The mild solution from Definition 2.1 is uniquely defined.*

Defining solutions like this makes sense as it gives a uniquely defined term which coincides with the classical solution where available, see *e.g.* the results in Pazy's monograph ([16], Chaps. 4, 5, 6) or Arendt *et al.*'s monograph ([1], Chap. 3.1). Now, we state our main result, which will be proven as Theorem 3.7 in Section 3.

Proposition 2.3 (Extension of Thm. 2 in [19]). *Let $\alpha \in L^\infty((0,T), \mathbb{R}^{|V|})$ such that $\|\alpha\|_{L^\infty} \leq 1$, $(\beta_n)_n$ be binary-valued functions such that the coordinate sequences of $(\phi_n)_n$ defined by $\phi_n := \alpha - \beta_n$ are of vanishing integrality gap. Let x, y_n for $n \in \mathbb{N}$ be the unique mild solutions of (1.1) and (1.2). Furthermore, let $f_i(s) := f(s, x(s), u(s), v_i)$ be in $L^1((0,T), X)$ for $i \in \{1, \ldots, |V|\}$.*
Then,

$$\|x - y_n\|_{C([0,T],X)} \underset{n \to \infty}{\to} 0.$$

We point out the achievement of proving Proposition 2.3 below.

Remark 2.4. In particular, we have strengthened the results from the literature as follows.

(1) For the ODE-case, the regularity assumptions (6c) in Theorem 2 and (17) in Corollary 6 in [19] that $s \mapsto f(s, y(s), u(s), v_i) \in C^1([0,T], \mathbb{R}^n)$ with

$$\|f(\cdot, x(\cdot), u(\cdot), v_i)\|_{L^\infty} \leq M, \; \|f(\cdot, y(\cdot), u(\cdot), v_i)'\|_{L^\infty} \leq C$$

can be weakened to $[0,T] \ni s \mapsto f(s, x(s), u(s), v_i) \in \mathbb{R}^n$ being in $L^1((0,T), \mathbb{R}^n)$ which is a trivial corollary with the choice $A := 0$ and $X = \mathbb{R}^n$.

(2) For semilinear PDEs whose differential operator generates a C_0-semigroup $(T(t))_{t\geq 0}$, the prerequisite H_2 in (Thm. 1 of [11]) that for all $t \in [0, T]$, the function $s \mapsto T(t-s)f(s, y(s), u(s))$ is a piecewise H^1-function and

$$\left\| \frac{\mathrm{d}}{\mathrm{d}s} T(t-s)f(s, x(s), u(s)) \right\|_X \leq C \text{ for a.e. } 0 < s < t < T$$

can be weakened to $[0, T] \ni s \mapsto f(s, x(s), u(s)) \in X$ being in $L^1((0, T), X)$. Feasible setups for the IVPs can be validated by checking the prerequisites of Corollary 2.6.

To provide a self-contained article, we state and prove the following proposition summarizing the relationship between (RC) and (BC_δ). It follows from a continuity argument.

Proposition 2.5 (Cors. 6 and 8 in [19]). *Let $(\bar{x}, \bar{\alpha}, \bar{u})$ be feasible for (RC) such that $\bar{x}$ is the unique mild solution of* (1.1) *in the setting $u = \bar{u}$, $\alpha = \bar{\alpha}$. Let $[0, T] \ni s \mapsto f(s, x(s), u(s), v_i) \in X$ be in $L^1((0, T), X)$ for $1 \leq i \leq |V|$. Let $(\beta_n)_n$ be binary-valued functions such that the coordinate sequences of $(\phi_n)_n$ defined by $\phi_n := \bar{\alpha} - \beta_n$ are of vanishing integrality gap. Then, for every $\delta > 0$, there exists $(y^\delta, \bar{u}, \beta^\delta)$ being feasible for (BC_δ) such that*

$$|J(\bar{x}, \bar{u}) - J(y^\delta, \bar{u})| < \delta.$$

Proof. By continuity of J and c, the fact that there exists $\varepsilon > 0$ such that $\|\bar{x} - y\|_{C([0,T],X)} < \varepsilon$ implies

$$|J(\bar{x}, \bar{u}) - J(y, \bar{u})| < \delta \text{ and } \|c(\bar{x}(t), \bar{u}(t)) - c(y(t), \bar{u}(t))\|_Y < \delta.$$

By Proposition 2.3, there exist $C_r > 0$ and $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$

$$\|\bar{x} - y_n\|_{C([0,T],X)} < \min\{\delta, \varepsilon\}$$

holds. We choose $\beta^\delta := \beta_{n_0}$ and $y^\delta := y_{n_0}$ and the claim follows. $\square$

Now, we establish a broad setting where (1.3) holds and which can be checked more easily.

Corollary 2.6. *Let $\alpha \in L^\infty((0, T), \mathbb{R}^{|V|})$, $(\beta_n)_n$ be binary-valued functions such that the coordinate sequences of $(\phi_n)_n$ defined by $\phi_n := \alpha - \beta_n$ are of vanishing integrality gap, let $u \in L^1((0, T), U)$, and let $f : [0, T] \times X \times U \times V \to X$ be continuous in the first and uniformly Lipschitz continuous in the second and third argument. Then,*

$$\|x - y_n\|_{C([0,T],X)} \underset{n \to \infty}{\to} 0.$$

Proof. First, we note that plugging an $L^1((0, T))$-function into a uniformly Lipschitz continuous function yields another $L^1((0, T))$-function. We observe that

$$x(t) = T(t)x_0 + \int_0^t T(t-s) \sum_{i=1}^{|V|} \alpha_i(s) f(s, x(s), u(s), v_i) \mathrm{d}s \tag{2.1}$$

is the mild solution of (1.1) and

$$y_n(t) = T(t)x_0 + \int_0^t T(t-s) \sum_{i=1}^{|V|} \beta_{n,i}(s) f(s, y_n(s), u(s), v_i) \mathrm{d}s \tag{2.2}$$

are the mild solutions of (1.2). Then, we apply Proposition 2.3. $\square$

3. Proof of Proposition 2.3

We approach the main statement in several steps. First, we show that $(\phi_n)_n$ being of vanishing integrality gap implies $\int_0^t \phi_n f \to 0$ uniformly for $f \in L^1((0,T), X)$. Teaming this insight up with some compactness arguments, we show (1.3) for a broad class of semilinear PDEs under mild regularity assumptions. Finally, we generalize the result from continuous functions to piecewise continuous ones.

3.1. Vanishing integrality gap for $L^1((0,T), X)$-functions

By means of an approximation argument, we show the following result which enables us to relax assumptions made for the proofs of previous results that relied on the direct applicability of an integration by parts formula.

Lemma 3.1. *Let X be a Banach space, $f \in L^1((0,T), X)$, $(\phi_n)_n \subset L^\infty((0,T), \mathbb{R})$ be bounded and of vanishing integrality gap. Furthermore, let $\Phi_n(t) := \int_0^t \phi_n(s)\mathrm{d}s$ and $\varepsilon > 0$. Then, there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ we have*

$$\sup_{t \in [0,T]} \left\| \int_0^t f(s)\phi_n(s)\mathrm{d}s \right\|_X < \varepsilon.$$

Proof. Let $C_\phi := \sup_{n \in \mathbb{N}} \|\phi_n\|_{L^\infty}$, which exists by assumption. Let $\varepsilon > 0$.

We use that fact that $\overline{C^\infty([0,T], X)}^{\|\cdot\|_{L^1}} = L^1((0,T), X)$ (see Prop. B.1). Hence, there exists $g \in C^\infty([0,T], X)$ such that

$$\|f - g\|_{L^1((0,T),X)} < \frac{\varepsilon}{2C_\phi}$$

with $C_g := \|g\|_{L^\infty((0,T),X)} + T\|g'\|_{L^\infty((0,T),X)}$. We insert a zero

$$\int_0^t f(s)\phi_n(s)\mathrm{d}s = \int_0^t g(s)\phi_n(s)\mathrm{d}s + \int_0^t \phi_n(s)(f(s) - g(s))\mathrm{d}s$$

and apply integration by parts for the first summand which then reads

$$\int_0^t g(s)\phi_n(s)\mathrm{d}s = g(t)\Phi_n(t) - \int_0^t g'(s)\Phi_n(s)\mathrm{d}s$$

and, consequently,

$$\left\| \int_0^t g(s)\phi_n(s)\mathrm{d}s \right\|_X \leq C_g \|\Phi_n\|_{L^\infty}.$$

Due to the convergence of Φ_n, there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, we have

$$\|\Phi_n\|_{L^\infty} < \frac{\varepsilon}{C_g 2}.$$

Putting the estimates together, we arrive at

$$\sup_{t \in [0,T]} \left\| \int_0^t f(s)\phi_n(s)\mathrm{d}s \right\|_X \leq \sup_{t \in [0,T]} \left\| \int_0^t g(s)\phi_n(s)\mathrm{d}s \right\|_X + C_\phi \|f - g\|_{L^1((0,T),X)}$$

$$< C_g \frac{\varepsilon}{2C_g} + C_\phi \frac{\varepsilon}{2C_\phi} = \varepsilon$$

for all $n \geq n_0$. $\qquad \square$

Example 3.2. Now, we got rid of differentiability assumptions. To demonstrate the result of Lemma 3.1 in the absence of differentiability, we consider the following Weierstraß function

$$f : [0, 2\pi] \to \mathbb{R}$$
$$f(x) := \lim_{n \to \infty} f_n(x)$$
$$f_n(x) := \sum_{k=0}^{n-1} \frac{2^k \sin(2^k x)}{3^k}$$

which is nowhere differentiable. Furthermore, we consider the following sequence of functions $\phi_n : [0, 2\pi] \to [-1, 1]$ for which we have an equidistant discretization step width $\frac{2\pi}{2^n}$ which makes this example straightforward.

$$\phi_n(x) := \begin{cases} 1 & : \quad x \in 2\pi \cdot \left[\frac{2i}{2^n}, \frac{2i+1}{2^n}\right] \quad i \in \{0, \ldots, 2^{n-1} - 1\} \\ -1 & : \quad x \in 2\pi \cdot \left[\frac{2i+1}{2^n}, \frac{2i+2}{2^n}\right] \quad i \in \{0, \ldots, 2^{n-1} - 1\} \end{cases}.$$

The sequence ϕ_n was chosen such that

$$\int_0^{2\pi} \frac{2^k \sin(2^k x)}{3^k} \phi_n(x) \mathrm{d}x = \frac{2^k}{3^k} \begin{cases} \int_0^{2\pi} |\sin(2^k x)| \mathrm{d}x & : \quad k+1 = n \\ 0 & : \quad k+1 \neq n \end{cases}.$$

If $k \geq n$, the sin terms oscillate inside the constant segments of f_n and cancel each other there. If $k \leq n - 2$, f_n oscillates and cancels itself within segments where sin has the same sign and is symmetric with respect to the extreme point in this segment.

By means of Lebesgue's dominated convergence theorem, we obtain

$$\int_0^{2\pi} f(x)\phi_n(x)\mathrm{d}x = \lim_{m \to \infty} \int_0^{2\pi} \sum_{k=0}^{m} \frac{2^k \sin(2^k x)}{3^k} \phi_n(x)\mathrm{d}x = \frac{2^{n-1}}{3^{n-1}} \int_0^{2\pi} |\sin(2^{n-1}x)|\mathrm{d}x \leq \frac{2^{n-1}}{3^{n-1}} 2\pi \underset{n \to \infty}{\to} 0.$$

Remark 3.3. Having Lemma 3.1 at hand, the improvement mentioned in Remark 2.4 (1) can now be proven quite easily similar to the reasoning in [19]. However, as we have promised a more general result that works for semilinear PDEs as well, we are going to invest some extra effort.

Using the proof of Lemma 3.1, we can characterize the convergence of $(\beta_n)_n$ and $(\phi_n)_n$ by means of weak topologies, which is done in Theorem 3.4 below.

Theorem 3.4. *Let* $\alpha \in L^\infty((0, T), \mathbb{R}^{|V|})$ *be given such that* $\alpha \geq 0$ *and* $\sum_{i=1}^{|V|} \alpha_i = 1$ *a.e. Let* $(\beta_n)_n \subset L^\infty((0, T), \mathbb{R}^{|V|})$ *be a sequence of* $\{0, 1\}^{|V|}$-*valued functions such that* $(\phi_n)_n$ *is of vanishing integrality gap where* $\phi_n := \alpha - \beta_n$ *for all n.*
Then,

$$\beta_n \rightharpoonup \alpha \text{ in } L^p((0, T), \mathbb{R}^{|V|}) \text{ for } 1 \leq p < \infty$$

and

$$\beta_n \rightharpoonup^* \alpha \text{ in } L^p((0, T), \mathbb{R}^{|V|}) \text{ for } 1 < p \leq \infty.$$

Proof. We employ Lemma 3.1 with $X = \mathbb{R}$ to obtain $\phi_n \rightharpoonup^* 0$ in $L^\infty((0, T), \mathbb{R}^{|V|})$. This implies $\beta_n \rightharpoonup^* \alpha$ in $L^\infty((0, T), \mathbb{R}^{|V|})$. The other claims follow immediately as we have tested with L^1-functions and $L^p \subset L^1$ for $p > 1$ on finite measure spaces. $\square$

3.2. Approximation error of binary controls generated by Sum-Up-Rounding

Before we can prove our result, we need the following two preparatory lemmata. The first transforms a pointwise convergence into a uniform one.

Lemma 3.5. *Let X be a Banach space, $(T(t))_{t\geq 0}$ be a C_0-semigroup on X, $f \in L^1((0,T), X)$. Then,*

$$\sup_{t\in[0,T]} \int_0^t \|(T(t+h-s) - T(t-s))f(s)\|_X \, \mathrm{d}s \underset{h\downarrow 0}{\to} 0.$$

Proof. We note that $t \mapsto \|T(t)\|_{op}$ is dominated by an exponential function on compact intervals, a standard result *e.g.* from Pazy's monograph ([16], Chap. 1, Thm. 2.2) or Arendt *et al.*'s monograph ([1], Thm. 3.1.7), and let $C := \sup_{t\in[0,T]} \|T(t)\|_{op}$. A simple estimation using the semigroup property of $(T(t))_{t\geq 0}$ and submultiplicativity of the norm gives

$$\sup_{t\in[0,T]} \int_0^t \|(T(t+h-s) - T(t-s))f(s)\|_X \, \mathrm{d}s \leq \sup_{t\in[0,T]} \int_0^t \|T(t-s)\|_{op} \|(T(h) - I)f(s)\|_X \, \mathrm{d}s$$

$$\leq C \int_0^T \|(T(h) - I)f(s)\|_X \, \mathrm{d}s.$$

An application of Lebesgue's dominated convergence theorem finishes the proof. $\square$

The second shows that a certain sequence of functions in $C([0,T], X)$ is relatively compact.

Lemma 3.6. *Let X be a Banach space, $(T(t))_{t\geq 0}$ be a C_0-semigroup on X, $f \in L^1((0,T), X)$, $(\phi_n)_n \subset L^\infty((0,T), \mathbb{R})$ with $\phi_n(t) \in [-1,1]$ for a.e. $t \in [0,T]$ such that*

$$\nu_n(t) \underset{n\to\infty}{\to} 0, \quad \text{for all } t \in [0,T]$$

with the setting

$$\nu_n(t) := \int_0^t \phi_n(s) T(t-s) f(s) \mathrm{d}s.$$

Then, the set $\{\nu_n : n \in \mathbb{N}\}$ is relatively compact in $L^p((0,T), X)$ for $p \in [1,\infty)$ and $C([0,T], X)$ in the norm-topology.

Proof. Again, we set $C := \sup_{t\in[0,T]} \|T(t)\|_{op}$. Due to the absolute continuity of the Bochner integral, we know $(\nu_n)_n \subset C([0,T], X)$. Note that the uniform boundedness of $(\phi_n)_n$ and the boundedness of $T(t)$ on compact intervals already used in the proof of Lemma 3.5 imply the uniform boundedness of $(\nu_n)_n$. We prove the claim by employing Theorem 1 in [20] by Simon, which is a practical application and extension of the Arzelà–Ascoli theorem.

Hence, using Theorem 1 of [20], we have to verify the following two conditions.

$$B_{t_1,t_2} := \left\{ \int_{t_1}^{t_2} \nu_n(t) \mathrm{d}t : n \in \mathbb{N} \right\} \subset\subset X \text{ for all } 0 < t_1 < t_2 < T \tag{3.1}$$

and

$$\sup_{n\in\mathbb{N}} \|\nu_n(\cdot + h) - \nu_n(\cdot)\|_{L^p((0,T-h),X)} \underset{h\downarrow 0}{\to} 0 \tag{3.2}$$

to show convergence in $L^p((0,T),X)$, and in $C([0,T],X)$ in the case $p = \infty$.

Regarding (3.1), we use that $\nu_n(t) \to 0$ pointwise and $(\nu_n)_n$ is uniformly bounded. Thus, we can employ Lebesgue's dominated convergence theorem, which yields

$$\left\| \int_{t_1}^{t_2} \nu_n(t) \mathrm{d}t \right\|_X \to 0$$

for all $0 < t_1 < t_2 < T$. Hence, B_{t_1,t_2} consists of the elements of a Cauchy sequence and is therefore relatively compact in X.

To show (3.2), we observe

$$\|\nu_n(t+h) - \nu_n(t)\|_X = \left\| \int_0^{t+h} T(t+h-s)f(s)\phi_n(s)\mathrm{d}s - \int_0^t T(t-s)f(s)\phi_n(s)\mathrm{d}s \right\|_X$$

$$\leq \left\| \int_t^{t+h} T(t+h-s)f(s)\phi_n(s)\mathrm{d}s \right\|_X + \left\| \int_0^t (T(t+h-s) - T(t-s))f(s)\phi_n(s)\mathrm{d}s \right\|_X.$$

For the integrand of the first term, we get

$$\left\| \int_t^{t+h} T(t+h-s)f(s)\phi_n(s)\mathrm{d}s \right\|_X \leq C \int_0^T \|f(s)\|_X \chi_{[t,t+h]}(s)\mathrm{d}s$$

and convergence to zero for $h \downarrow 0$ by Lebesgue's dominated convergence theorem independent of the specific choice of ϕ_n. For the second term, we estimate

$$\left\| \int_0^t (T(t+h-s) - T(t-s))f(s)\phi_n(s)\mathrm{d}s \right\|_X \leq \int_0^t \|(T(t+h-s) - T(t-s))f(s)\|_X |\phi_n(s)|\mathrm{d}s$$

$$\leq \int_0^t \|(T(t+h-s) - T(t-s))f(s)\|_X \mathrm{d}s.$$

By means of Lebesgue's dominated convergence theorem, we get

$$\int_0^t \|(T(t+h-s) - T(t-s))f(s)\|_X \mathrm{d}s \underset{h \downarrow 0}{\to} 0 \tag{3.3}$$

for all $t \in [0, T-h]$. Another application of Lebesgue's dominated convergence theorem gives

$$\int_0^{T-h} \left\| \int_0^t (T(t+h-s) - T(t-s))f(s)\mathrm{d}s \right\|_X^p \mathrm{d}t \underset{h \downarrow 0}{\to} 0$$

for all $p \in [1, \infty)$. For the case $p = \infty$, *i.e.* convergence in $C([0,T],X)$, we apply Lemma 3.5 to (3.3). This was the last step necessary to show that (3.2) holds for $\{\nu_n : n \in \mathbb{N}\}$. Now, we infer that $\{\nu_n : n \in \mathbb{N}\}$ is relatively compact in the norm-topology of $L^p((0,T),X)$ for $p \in [1, \infty)$, and of $C([0,T],X)$ in the case $p = \infty$. $\square$

Equipped with Lemmas 3.1 and 3.6, we are enabled to generalize the approximation result (1.3) from the settings in [13] and [11] for mild solutions of semilinear PDEs whose differential operators generate C_0-semigroups. This is the statement of Theorem 3.7 below, which implies Proposition 2.3.

Theorem 3.7. *Let X be a real Banach space and A be the generator of a C_0-semigroup $(T(t))_{t\geq 0}$. Let $\alpha \in L^\infty((0,T),\mathbb{R})$ with $0 \leq \alpha \leq 1$ a.e., $(\beta_n)_n \subset L^\infty((0,T),\mathbb{R})$ be binary-valued functions and $u \in L^1((0,T),U)$ be such that x is the unique mild solution of* (1.1) *and y_n are the unique mild solutions of* (1.2) *for $n \in \mathbb{N}$ and that $(\phi_n)_n$ with $\phi_n := \alpha - \beta_n$ is of vanishing integrality gap and $f_i(s) := f(s,x(s),u(s),v_i)$ is in $L^1((0,T),X)$.*

Furthermore, let $\varepsilon > 0$. Then, there exist $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, we obtain:

$$\|x(t) - y_n(t)\|_X \leq \varepsilon \exp(C_r t)$$

with $C_r > 0$ independent of ε.

Proof. Let $t \in [0,T]$. As the mild solutions x, y_n are continuous, we can evaluate them and use the variation of constants formulas (2.1) and (2.2) to compute their difference

$$\|x(t) - y_n(t)\|_X = \left\| \sum_{i=1}^{|V|} \int_0^t T(t-s)(\alpha_i(s)f_i(s) - \beta_{n,i}(s)f_i(s,y_n(s),u(s)))\mathrm{d}s \right\|_X.$$

As done in [19], we insert a zero and obtain

$$\|x(t) - y_n(t)\|_X \leq \left\| \sum_{i=1}^{|V|} \int_0^t T(t-s)(\alpha_i(s)f_i(s) - \beta_{n,i}(s)f_i(s)\mathrm{d}s \right\|_X$$

$$+ \left\| \sum_{i=1}^{|V|} \int_0^t \beta_{n,i}(s)T(t-s)(f_i(s) - f_i(s,y_n(s),u(s)))\mathrm{d}s \right\|_X$$

$$\leq \sum_{i=1}^{|V|} \left\| \int_0^t \phi_{n,i}(s)T(t-s)f_i(s)\mathrm{d}s \right\|_X + |V|L \sup_{t\in[0,T]} \|T(t)\|_{op} \int_0^t \|x(s) - y_n(s)\|_X \mathrm{d}s,$$

where L denotes the Lipschitz constant of f in the second argument and we have used that $|\beta_{n,i}(t)| \leq 1$ for all $t \in [0,T]$. Noting that $\|T(t)\|_{op}$ is dominated by an exponential function on compact intervals, see *e.g.* ([16], Chap. 1, Thm. 2.2) or ([1], Thm. 3.1.7), we set $C_r := |V|L \sup_{t\in[0,T]} \|T(t)\|_{op}$. Let $i \in \{1,\ldots,|V|\}$ be fixed. Now, we have to handle the sequence

$$\nu_{n,i}(t) := \int_0^t \phi_{n,i}(s)T(t-s)f_i(s)\mathrm{d}s.$$

We are going to show

$$\sup_{t\in[0,T]} \|\nu_{n,i}(t)\|_X \to 0.$$

The function $s \mapsto T(t-s)f_i(s)$ is in $L^1((0,t),X)$, see Proposition B.2. Hence, by means of Lemma 3.1, $\nu_{n,i}(t) \to 0$ for all $t \in [0,T]$. Furthermore, $(\nu_{n,i})_n \subset C([0,T],X)$ and

$$\sup_{n\in\mathbb{N}} \sup_{t\in[0,T]} \|\nu_{n,i}(t)\|_X < \infty$$

because $(\phi_{n,i})_n \subset L^\infty((0,T),\mathbb{R})$ with $\phi_{i,n}(t) \in [-1,1]$ for a.e. $t \in [0,T]$ and the $\nu_{n,i}$ are continuous due to the absolute continuity of the Bochner integral. Hence, $(\nu_{n,i})_n$ is a bounded sequence in $C([0,T],X)$ that converges to

0 pointwise. By means of Dinculeanu and Singer's extension of the Riesz–Markov–Kakutani theorem, elements ψ of the topological dual of $C([0,T], X)$ can be identified with regular Borel measures $\mu : \mathcal{B} \to X^*$ of finite variation where $\mathcal{B}$ is the Borel σ-field on $[0,T]$, see Proposition A.1 *e.g.* from Dinculeanu's monograph ([5], Chap. III.19, Cor. 2) or Dobrakov's article [6]:

$$\psi(\nu_{n,i}) \coloneqq \int_0^T \nu_{n,i} \mathrm{d}\mu.$$

This setting allows to apply Lebesgue's dominated convergence theorem, see Proposition A.2 *e.g.* from Dinculeanu's monograph ([5], Chap. 8, Thm. 3), from which we obtain

$$\nu_{n,i} \rightharpoonup 0.$$

Now, it suffices to check that $\{\nu_{n,i} : n \in \mathbb{N}\}$ is relatively compact w.r.t the $\|\cdot\|_{C([0,T],X)}$-topology because weakly convergent sequences contained in $\|\cdot\|$-compact sets converge in $\|\cdot\|$ to the same limit and sequential compactness and compactness are equivalent on metric spaces. Employing Lemma 3.6, we infer that $\{\nu_{n,i} : n \in \mathbb{N}\}$ is relatively compact in the $\|\cdot\|_{L^\infty((0,T),X)} = \|\cdot\|_{C([0,T],X)}$-topology and due to $\nu_{n,i} \rightharpoonup 0$

$$\lim_{n\to\infty} \|\nu_{n,i}\|_{C([0,T],X)} = 0.$$

An $\frac{\varepsilon}{|V|}$-argument gives

$$\sup_{t\in[0,T]} \sum_{i=1}^{|V|} \left\| \int_0^t \phi_{n,i}(s) T(t-s) f_i(s) \mathrm{d}s \right\|_X < \varepsilon$$

for all $n \geq n_0$ and some $n_0 \in \mathbb{N}$. The application of Grönwall's inequality finishes the proof. $\square$

4. COMPUTATIONAL EXAMPLE

We provide an example to demonstrate our findings computationally, and consider a problem where differentiability cannot be assumed. In more detail, we consider the IVP

$$\begin{aligned} \dot{x}(t) &= -\partial_s x(t) + \alpha_1(t) f_1(t) + \alpha_2(t) f_2(t) \\ x(0) &\equiv 0.5 \end{aligned} \tag{4.1}$$

in one spatial dimension, *i.e.* $\Omega = [\ell, r]$. We assume a constant influx of 0 on the left side of the domain and do not impose any condition at the right boundary of the domain, which can be interpreted as a free outflow of the domain. It is well known that $-\partial_s$ generates the right translation semigroup, see Example 2.9 of [16]. In particular, the translation semigroup does not provide smoothing properties like the heat semigroup or other semigroups associated with parabolic equations do. We choose f_1 to be a nowhere differentiable Weierstraß function in time, see Example 3.2, multiplied by a constant function in space and $f_2(t) \equiv 0$. As *Sum-Up-Rounding* does not require optimality of the (forward) solution it approximates, we may choose α somewhat arbitrarily for the purpose of demonstration. We use $\alpha_1 = \alpha_2 \equiv 0.5$. Clearly, *Sum-Up-Rounding* produces a chattering that approximates the constant function 0.5 weakly. We have visualized this in Figure 1 for a coarse and a finer rounding grid. We discretize the time horizon $[0,10]$ and Ω into 4096 intervals each and solve with the Lax–Friedrichs scheme for hyperbolic conservation laws, see for example Leveque's monograph [14] for the details. To compute the value of the right-most cell with the Lax–Friedrichs scheme, we add a ghost-cell with zero-order extrapolation, see (Sect. 7.2.1 of [14]). Let $\omega^{(1)}, \ldots, \omega^{(6)}$ denote the sequence of *Sum-Up-Rounding* approximations of α with $N^{(1)} = 128, \ldots, N^{(6)} = 4096$ rounding intervals. Let $y^{(1)}, \ldots, y^{(6)}$ denote
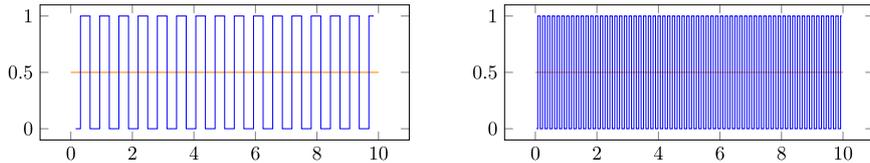
FIGURE 1. *Sum-Up-Rounding* approximation $\beta_1(t)$ (blue) of $\alpha_1(t) = 0.5$ (orange), with $N = 32$ (*left*) and $N = 128$ (*right*).
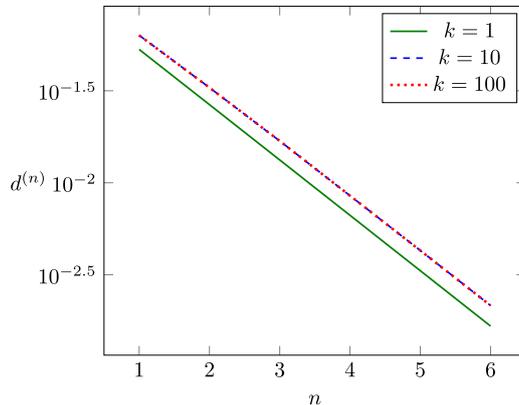

FIGURE 2. Convergence of the relative error of the state vector for (4.1) with f_1 being a Weierstraß function and $\omega^{(n)}$ computed by *Sum-Up-Rounding* on refined grids indexed by $n = 1, \ldots, 6$. The cases $k = 10$ and $k = 100$ coincide within the figure's resolution.

the corresponding solutions of (4.1) with $\omega^{(n)}$ instead of α. We have computed the relative error

$$d^{(n)} := \frac{\sup_{t \in [0,T]} \|y^{(n)}(t) - x(t)\|_{L^1}}{\sup_{t \in [0,T]} \|x(t)\|_{L^1}}$$

for $n = 1, \ldots, 6$. As the Weierstraß function cannot be evaluated exactly, we have approximated it by including $k = 1, 10, 100$ summands of its defining cosine series. The convergence of $d^{(n)}$ is very similar for the three choices of k. When only one summand of the cosine series is included, *i.e.* the smoothness is *highest*, the convergence is a little faster than when more summands are included. In numbers, we have $d^{(6)} = 1.6642 \times 10^{-3}$ for $k = 1$, $d^{(6)} = 2.1519 \times 10^{-3}$ for $k = 10$ and $d^{(6)} = 2.1521 \times 10^{-3}$ for $k = 100$. To see the whole process, we have visualized convergence of $d^{(n)}$ in Figure 2. To visualize the violation of the differentiability in the right hand side, we have plotted the approximants of the Weierstraß function used for our computations in Figure 3.

5. CONCLUSION

As mentioned before, previous proofs employed the integration by parts directly on $\int_0^t \phi_{n,i}(s)T(t-s)f_i(s)\mathrm{d}s$. As differentiability of $\phi_{n,i}$ is not available, the demand of a certain amount of differentiability to $s \mapsto T(t-s)f_i(s)$ was inherent to them. Lemma 3.1 allowed us to shift the integration by parts to a smooth approximation of the L^1-function. However, we would like to stress that the approximation argument in Lemma 3.1, which allows to extend our proof without the previous differentiability assumptions, currently prevents us from finding *a priori* estimates on the approximation error as they are available in [10, 11, 13, 19]. The compactness argument in Lemma 3.6 allowed us to deduce strong convergence from weak convergence. This is in particular valuable
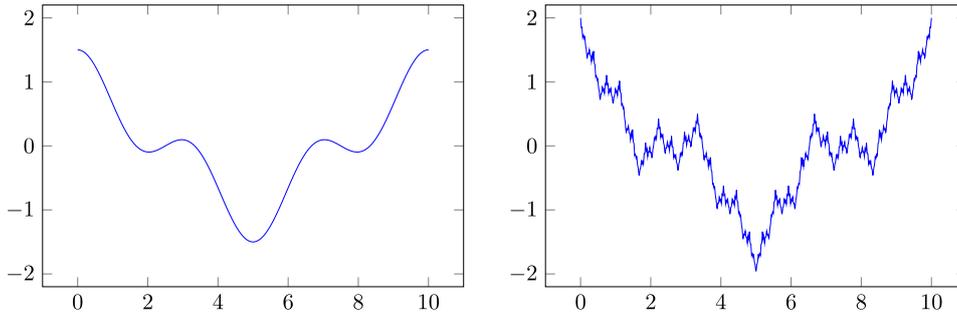
FIGURE 3. Approximants of the Weierstraß function by including one summand (*left*) and ten summands (*right*).

because the requirement of continuously differentiable solution trajectories might not be very restrictive for ODEs, but can be quite restrictive for PDEs.

Our findings can be interpreted as a constructive and algorithmic complement to the Filippov–Ważewski theorem [7, 21], which states that the solutions of a set of differential inclusions with set-valued nonlinear term are dense in the set of differential inclusions with convexified nonlinear term under similar conditions, see [4, 8] for the case of semilinear evolution equations based on C_0-semigroups. The same idea has been pursued by Gamkrelidze in [9] where he discovered that the infimal value of an OCP can be approximated with trajectories emanating from feasible controls even when no feasible limiting control exists. The optimal state trajectory is called *optimal sliding state* by him.

With respect to the numerics, we point out that the fact that refining the grids on which the IVPs are solved can lead to a loss of (piecewise) differentiability in the limit if the IVP does not have a differentiable solution, but the discretizations do. As far as the question of convergence of the MIOCP approximation process as $n \to \infty$ is concerned, our results show that this loss of (piecewise) differentiability is no cause for concern anymore.

APPENDIX A. RESULTS FROM MEASURE THEORY

We state the results from measure theory needed to obtain the weak convergence $\nu_n \rightharpoonup 0$ in Theorem 3.7 and phrase them for our needs which is of course a special case of the very general results in Dinculeanu's monograph [5].

Proposition A.1 (Riesz–Markov–Kakutani theorem, Chap. 19, Cor. 2 in [5]). *Let X be a Banach space. Then, there exists an isomorphism between continuous linear functionals $\psi \in C([0,T], X)^*$ and regular Borel measures $\mu : \mathcal{B} \to X^*$ with finite variation defined by*

$$\psi(f) = \int_0^T f \mathrm{d}\mu.$$

Proposition A.2 (Lebesgue theorem, Chap. 8, Thm. 3 in [5]). *Let X, E be Banach spaces with a bilinear mapping $X \times E \ni (x, e) \mapsto \langle x, e \rangle \in \mathbb{R}$ such that $|\langle x, e \rangle| \leq \|x\|_X \|e\|_E$ and $\mu : \mathcal{B} \to E$ be a finite measure. Let $(f_n)_n$ be μ-integrable X-valued functions on $[0,T]$ such that $(f_n)_n$ converges μ-almost everywhere to a function $f : [0,T] \to X$. If there exists a positive $\|\mu\|$-integrable function g with $\|f_n(t)\|_X \leq g(t)$ for μ-almost every $t \in [0,T]$ and each $n \in \mathbb{N}$ where $\|\mu\|$ denotes the variation of μ, f is μ-integrable and $f_n \to f$, in particular*

$$\int_0^T f \mathrm{d}\mu = \lim_{n \to \infty} \int_0^T f_n \mathrm{d}\mu.$$

Remark A.3. The existence of the bilinear mapping in Proposition A.2 takes care that the integration of step functions w.r.t μ can be defined properly with sums. Then, μ-integrable functions f are those for which

$$\int_0^T f \mathrm{d}\mu = \lim_{n \to \infty} \int_0^T f_n \mathrm{d}\mu$$

holds with $(f_n)_n$ being a Cauchy sequence of step functions which converges to f μ-almost everywhere.

Appendix B. Results on L^1-functions

We state the following approximation result of L^1-functions by means of smooth functions.

Proposition B.1. *Let X be a Banach space. Then,*

$$\overline{C^\infty([0,T], X)}^{\|\cdot\|_{L^1((0,T),X)}} = L^1((0,T), X).$$

Proof. The scalar case can be found in many analysis textbooks. For the vector-valued case, one can *e.g.* apply Lemma 1.3.3 from [1] to obtain $f * \rho_n \to f$ in $\|\cdot\|_{L^1}$ for $f \in L^1(\mathbb{R}, X)$ and $(\rho_n)_n$ being a mollifier. The choice for the smooth mollifier to have $f * \phi_n \in C^\infty$ can be the same as for the scalar-valued case. Extending $f \in L^1((0,T), X)$ to $L^1(\mathbb{R}, X)$ by setting it to zero on $\mathbb{R} \setminus (0,T)$ allows the application of the convolution. $\quad\square$

Proposition B.2 (Prop. 1.3.4 in [1])**.** *Let X be a Banach space and $(T(t))_{t \geq 0}$ be a C_0-semigroup on X. Let $f \in L^1((0,T), X)$ and $0 < t \leq T$. Then, the function*

$$[0,t] \ni s \mapsto T(t-s)f(s) \in X$$

is in $L^1((0,t), X)$.

References

[1] W. Arendt, C.J.K. Batty, M. Hieber and F. Neubrander, Vector-Valued Laplace Transforms and Cauchy Problems, Vol. 96. Springer Science & Business Media, Basel (2011).

[2] L.D. Berkovitz, Optimal Control Theory. Springer-Verlag, New York (1974).

[3] L. Cesari, Optimization – Theory and Applications. Springer Verlag, New York (1983).

[4] F.S. De Blasi and G. Pianigiani, Evolution inclusions in non separable Banach spaces. *Comment. Math. Univ. Carolin.* **40** (1999) 227–250.

[5] N. Dinculeanu, Vector Measures. VEB Deutscher Verlag der Wissenschaften, Berlin (1967).

[6] I. Dobrakov, On representation of linear operators on $C_0(T, X)$. *Czech. Math. J.* **21** (1971) 13–30.

[7] A.F. Filippov, On some problems of optimal control theory. *Vestnik Moskowskovo Universiteta, Math* **2** (1958) 25–32. [English version: On certain questions in the theory of optimal control. *J. SIAM Ser. A Control* **1** (1962) 76–84].

[8] H. Frankowska, A priori estimates for operational differential inclusions. *J. Differ. Equ.* **84** (1990) 100–128.

[9] R.V. Gamkrelidze, On sliding optimal states. *Dokl. Akad. Nauk SSSR* **143** (1962) 1243–1245. (English translation: *Sov. Math. Dokl.* **3**, 559–562)).

[10] F.M. Hante, Relaxation methods for hyperbolic pde mixed-integer optimal control problems. *Optim. Control Appl. Methods* **38** (2017) 1103–1110.

[11] F.M. Hante and S. Sager, Relaxation methods for mixed-integer optimal control of partial differential equations. *Comput. Optim. Appl.* **55** (2013) 197–225.

[12] J. Haslinger and R.A.E. Mäkinen, On a topology optimization problem governed by two-dimensional Helmholtz equation. *Comput. Optim. Appl.* **62** (2015) 517–544.

[13] C. Kirches, F. Lenders and P. Manns, Approximation properties and tight bounds for constrained mixed-integer optimal control. Preprint Optimization Online n°5404 (2016). Available on: http://www.optimization-online.org/DB_FILE/2016/04/5404.pdf.

[14] R.J. LeVeque, Finite Volume Methods for Hyperbolic Problems. Cambridge University Press, Cambridge (2002).

[15] P. Manns, C. Kirches and F. Lenders, A linear bound on the integrality gap for sum-up rounding in the presence of vanishing constraints. Preprint Optimization Online n°6580 (2017). Available on: http://www.optimization-online.org/DB_FILE/2018/04/6580.pdf.

[16] A. Pazy, Semigroups of Linear Operators and Applications to Partial Differential Equations, Vol. 44. Springer Science & Business Media, Switzerland (1983).

[17] S. Sager, Numerical Methods for Mixed-Integer Optimal Control Problems. Der andere Verlag Tönning, Lübeck, Marburg (2005). Available on: https://mathopt.de/PUBLICATIONS/Sager2005.pdf.

[18] S. Sager, Reformulations and algorithms for the optimization of switching decisions in nonlinear optimal control. *J. Process Control* **19** (2009) 1238–1247.

[19] S. Sager, H.G. Bock and M. Diehl, The integer approximation error in mixed-integer optimal control. *Math. Program. Ser. A* **133** (2012) 1–23.

[20] J. Simon, Compact sets in the space $L^p((0, T), B)$. *Ann. Mat. Pura Appl.* **146** (1986) 65–96.

[21] T. Ważewski, On an optimal control problem, in Differential Equations and Their Applications. Publishing House of the Czechoslovak Academy of Sciences, New York (1963) 229–242.

[22] F. You and S. Leyffer, Oil spill response planning with MINLP. *SIAG/OPT Views-and-News* **21** (2010) 1–8.

[23] V.M. Zavala, Stochastic optimal control model for natural gas networks. *Comput. Chem. Eng.* **64** (2014) 103–113.