Deutsche Forschungsgemeinschaft
Priority Programme 1962

Improved Regularity Assumptions for Partial Outer Convexification of MIPDECOs

Paul Manns, Christian Kirches

Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization

Preprint Number SPP1962-079

received on September 24, 2018

Edited by
SPP1962 at Weierstrass Institute for Applied Analysis and Stochastics (WIAS)
Leibniz Institute in the Forschungsverbund Berlin e.V.
Mohrenstraße 39, 10117 Berlin, Germany
E-Mail: spp1962@wias-berlin.de

World Wide Web: http://spp1962.wias-berlin.de/

IMPROVED REGULARITY ASSUMPTIONS FOR PARTIAL OUTER CONVEXIFICATION OF MIPDECOS

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, all requirements of differentiability and uniformly bounded derivatives on the involved functions from previous work can be omitted.

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 [11] over oil-spill response planning [20] to optimal control of large-scale gas networks [21]. 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 [15–17] and was transferred to semilinear PDE-constrained MIOCPs by Hante and Sager [9, 10]. 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 regularity assumptions such that they are fulfilled for a broader class of problems and can be checked more easily.

2010 *Mathematics Subject Classification.* 49M20, 49N60, 90C11, 49J20.

Key words and phrases. Mixed-Integer Optimal Control with PDEs, Relaxations of Mixed-Integer Optimal Control, Regularity.

P. Manns and C. Kirches acknowledge funding by Deutsche Forschungsgemeinschaft through Priority Programme 1962.

C. Kirches acknowledges financial support by the German Federal Ministry of Education and Research, program "Mathematics for Innovations in Industry and Service", grants 05M2016-MoPhaPro, 05M17MBA-MOReNet, and program "IKT 2020: Software Engineering", grant 61210304-ODINE..

The authors would like to thank Robert Haller-Dintelmann, TU Darmstadt, and Dirk Lorenz, TU Braunschweig, for helpful discussions on the topic.

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

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

where we assume A to be the generator of a C_0-semigroup on a real Banach space X and $J \in C^1(C([0,T], X) \times L^1((0,T), U), \mathbb{R})$, $x \in C([0,T], X)$ (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 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 do not assume that the integer control is distributed in space.

Problems of the type (MIOCP) can be equivalently reformulated by means of *partial outer convexification*, see the publications by Berkovitz [2], Cesari [3], Sager [15–17]. 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

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

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

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

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

Definition 0.1 (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 0.2 (Sum-Up-Rounding Algorithm, [15, 17])**.** *Let* $0 = t_0 < \ldots < t_N = T$ *be a discretization grid of* $[0,T]$*. with maximum discretization width* $\Delta t := \max_{i\in\{0,N-1\}} t_{i+1} - t_i$*. 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*

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

In case, the maximum is ambiguous, exactly one of the maximizing indices has to be chosen by arg max*.*

The following proposition is due to Sager and states that Sum-Up-Rounding indeed yields sequences of vanishing integrality gap.

Proposition 0.3 (Sum-Up-Rounding yields Vanishing Integrality Gap, [17])**.** *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* $\phi_n := \alpha - \beta_n$ *fulfills*

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

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

Remark 0.4. *Please note that it is necessary to relax the algebraic constraint by an arbitrarily small* $\delta > 0$ *in* (BC_δ) *to approximate feasible points properly.*

Remark 0.5. *In work under review* [12, 13]*, the authors are extending 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 to the PDE setting here without any problems. We don't want to elaborate on that here and just note that the reformulation there shows vanishing constraints* $0 \le \beta_{n,i}(t)c(x(t),u(t),v_i)$ *which are taken care of by the Sum-Up-Rounding variant used in* [12, 13] *and the claim of Proposition 0.3 still holds.*

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

$$\dot{x}(t) = Ax(t) + \sum_{i=1}^{|V|} \alpha_i(t) f(t,x(t),u(t),v_i),\ x(0) = x_0 \tag{0.1}$$

$$\dot{y}_n(t) = Ay_n(t) + \sum_{i=1}^{|V|} \beta_{n,i}(t) f(t,y_n(t),u(t),v_i),\ y(0) = x_0 \tag{0.2}$$

where x solves (0.1) in (RC), and y_n solves (0.2) 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 are going to show

$$\sup_{t\in[0,T]} \|y_n(t) - x(t)\|_X \underset{n\to\infty}{\to} 0 \tag{0.3}$$

under mild regularity assumptions and with regard to Definition 0.1. In particular, Lipschitz continuity of f in x and u and the availability of mild solutions will do. A proof for this result has been given in the presence of ODEs by Sager in [17] and in the presence of semilinear PDEs by Hante and Sager in [10, Thm 1]. Both proofs assume significantly more regularity on f, x and u to show (0.3).

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 1. Furthermore, we point out its consequences for the existing theory of Sum-Up-Rounding and partial outer convexification. In Section 2, we prove the aforementioned approximation result. Finally in Section 3, we summarize our results in relation to the literature discussed above and show that our result is a true generalization of them, in particular those in [10, 17].

1. 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 1.1 (Chap. 4, Def. 2.3 in [14], 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 1.2 (Existence and Uniqueness)**.** *The mild solution from Definition 1.1 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 [14, Chap. 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 2.6 in Section 2.

Proposition 1.3 (Generalization of Theorem 2 in [17])**.** *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 (0.1) and (0.2). Furthermore, let $f_i(s) := f(s, x(s), u(s), v_i)$ be in $L^1((0,T),X)$ for $i \in \{1,\dots,|V|\}$.*

Then,

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

We point out the achievement of proving Proposition 1.3 below.

Remark 1.4. *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 [17] 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* [10, Thm 1] *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 1.6.*

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 1.5 (Corollary 6 and 8 in [17])**.** *Let* $(\bar{x},\bar{\alpha},\bar{u})$ *be feasible for* (RC) *such that* $\bar{x}$ *is the unique mild solution of* (0.1) *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* $\phi_n := \bar{\alpha} - \beta_n$ *is 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, 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 1.3, there exists $C_r > 0$ and $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$

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

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

Now, we establish a broad setting where (0.3) holds and which can be checked more easily.

Corollary 1.6. *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* $\phi_n := \bar{\alpha} - \beta_n$ *are of vanishing integrality gap,* $u \in L^1((0,T),U)$ *and* $f : [0,T] \times X \times U \times V \to X$ *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 a $L^1((0,T))$-function into a uniformly Lipschitz continuous function yields another $L^1((0,T))$-function. We observe that

$$x(t) = T(t)x_0 + \int_0^t T(t-s) \sum_{i=1}^{|V|} \alpha_i(s) f(s,x(s),u(s),v_i)\,\mathrm{d}s \tag{1.1}$$

is the mild solution of (0.1) and

$$y_n(t) = T(t)x_0 + \int_0^t T(t-s) \sum_{i=1}^{|V|} \beta_{n,i}(s) f(s, y_n(s), u(s), v_i)\,\mathrm{d}s \tag{1.2}$$

are the mild solutions of (0.2). Then, we apply Proposition 1.3. □

2. Proof of Proposition 1.3

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 (0.3) for a broad class of semilinear PDEs under mild regularity assumptions. Finally, we generalize the result from continuous functions to piecewise continuous ones.

2.1. **Vanishing Integrality Gap for $L^1((0,T),X)$-functions.** By means of an approximation argument, we are going to show the following result which will enable us to relax previous results that rely on the direct applicability of an integration by parts formula.

Lemma 2.1. *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 $n \geq n_0$.*

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

Proof. Let $C_f := \|f\|_{L^\infty((0,T),X)}$ and $C_\phi := \sup_{n\in\mathbb{N}} \|\phi_n\|_{L^\infty}$ which exist 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 Proposition 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

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

Plugging the estimates together, we arrive at

$$\begin{aligned}\sup_{t\in[0,T]}\left\|\int_0^t f(s)\phi_n(s)\,\mathrm{d}s\right\|_X &\le \sup_{t\in[0,T]}\left\|\int_0^t g(s)\phi_n(s)\,\mathrm{d}s\right\|_X + C_\phi\|f-g\|_{L^1((0,T),X)}\\ &< C_g\frac{\varepsilon}{2C_g}+C_\phi\frac{\varepsilon}{2C_\phi}<\varepsilon\end{aligned}$$

for all $n\ge n_0$. □

Example 2.2. *Now, we got rid of differentiability claims. To demonstrate the result of Lemma 2.1 in the absence of differentiability, we consider the following Weierstraß function*

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

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.

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

ϕ_n was chosen such that

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

If $k\ge n$, the sin *terms oscillate inside the constant segments of f_n and cancel each other there and if $k\le n-2$, 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

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

Remark 2.3. *Having Lemma 2.1 at hand, the mentioned improvement from Remark 1.4 (1) can now be proven quite easily similar to the reasoning in* [17]. *However, as we have promised a more general result working for semilinear PDEs as well, we are going to invest some extra effort.*

2.2. **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 2.4. *Let X be a Banach space, $(T(t))_{t\ge0}$ be a C_0-semigroup on X, $f\in L^1((0,T),X)$. Then,*

$$\sup_{t\in[0,T]}\int_0^t\|(T(t+h-s)-T(t-s))f(s)\|_X\,\mathrm{d}s \underset{h\downarrow0}{\to} 0$$

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 [14, Chap. 1, Thm 2.2] or Arendt et al. 's monograph [1, Thm 3.1.7], and set $C := \sup_{t\in[0,T]} \|T(t)\|_{op}$. A simple estimation using the semigroup property of T 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. □

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

Lemma 2.5. *Let X be a Banach space, $(T(t))_{t\geq 0}$ be a C_0-semigroup on X, $f \in L^1((0,T),X)$, $(\phi_n)_n \subset L^\infty((0,T),\mathbb{R})$ with $\phi_n(t) \in [-1,1]$ for a.e. $t \in [0,T]$ such that*

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

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 continutity of the Bochner integral, we know $(\nu_n)_n \subset C([0,T],X)$. Note that the uniform boundedness of $(\phi_n)_n$, the boundedness of $T(t)$ on compact intervals already used in the proof of Lemma 2.4 imply the uniform boundedness of $(\nu_n)_n$. We prove the claim by employing Theorem 1 in [18] by Simon which is a practical application and extension of the Arzela-Ascoli theorem.

Hence, using [18, Thm 1], we have to verify the following two conditions.

$$B_{t_1,t_2} := \left\{ \int_{t_1}^{t_2} \nu_n(t) \,\mathrm{d}t : n \in \mathbb{N} \right\} \subset\subset X \text{ for all } 0 < t_1 < t_2 < T \tag{2.1}$$

and

$$\sup_{n\in\mathbb{N}} \|\nu_n(\cdot + h) - \nu_n(\cdot)\|_{L^p((0,T-h),X)} \underset{h\downarrow 0}{\to} 0. \tag{2.2}$$

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

Regarding (2.1), 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 (2.2), 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

$$\int_0^t \|(T(t+h-s)-T(t-s))f(s)\|_X\,\mathrm{d}s \underset{h\downarrow 0}{\to} 0 \tag{2.3}$$

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 2.4 to (2.3). This was the last step necessary to show that (2.2) 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 $p=\infty$). □

Equipped with Lemma 2.1 and Lemma 2.5 we are enabled to generalize the approximation result (0.3) from the settings in [12] and [10] for mild solutions of semilinear PDEs whose differential operators generate C_0-semigroups. This is the statement of Theorem 2.6 below which implies Proposition 1.3.

Theorem 2.6. *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 (0.1) and y_n are the unique mild solutions of (0.2) 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 (1.1) and (1.2) 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 [17], we insert a zero and obtain

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

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. [14, 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 2.1, $\nu_{n,i}(t) \to 0$ 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 finite, regular, σ-additive measures $\mu : \mathcal{B} \to X^*$ 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]:

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

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 2.5, 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. □

3. 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 2.1 allowed us to shift the integration by parts to a smooth approximation of the L^1-function.

Our findings can be interpreted as a constructive complement to the Filippov-Važewski theorem, [7, 19], 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 compactness argument in Lemma 2.5 allowed us to deduce strong convergence from weak convergence. This is in particular valuable because the demand for continuously differentiable solution trajectories might not be very restrictive for ODEs but can be quite restrictive for PDEs. Finally, we would like to mention that the approximation argument in Lemma 2.1 allows to extend our proof without the previous differentiability assumptions but prevents us from finding priori estimates on the approximation error as they are available in [9, 10, 12].

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 2.6 and phrase them for our needs which is of course 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 $|\langle x, e\rangle| \le \|x\|_X \|e\|_E$ 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 $\|f_n(t)\|_X \le g(t)$ 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 $\mathbb{R}\backslash(0, T)$ allows the application of the convolution. □

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

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

is in $L^1((0, t), X)$. ■

References

[1] Wolfgang Arendt, Charles JK Batty, Matthias Hieber, and Frank Neubrander, *Vector-valued Laplace transforms and Cauchy problems*, Vol. 96, Springer Science & Business Media, 2011. doi:10.1007/978-3-0348-0087-7.

[2] Leondard D. Berkovitz, *Optimal Control Theory*, Springer-Verlag, 1974. doi:10.1007/978-1-4757-6097-2.

[3] Lamberto Cesari, *Optimization — Theory and Applications*, Springer Verlag, 1983. doi:10.1007/978-1-4613-8165-5.

[4] Francesco Saverio De Blasi and Giulio Pianigiani, *Evolution inclusions in non separable Banach spaces*, Comment. Math. Univ. Carolin **40** (1999), no. 2, 227–250. doi:10338.dmlcz/119079.

[5] Nicolae Dinculeanu, *Vector measures*, VEB Deutscher Verlag der Wissenschaften, Berlin, 1967. doi:10.1016/B978-1-4831-9762-3.50003-2.

[6] Ivan Dobrakov, *On representation of linear operators on* $C_0(T, X)$, Czechoslovak Mathematical Journal **21** (1971), no. 1, 13–30. https://dml.cz/bitstream/handle/10338.dmlcz/101000/CzechMathJ_21-1971-1_3.pdf.

[7] Alekseı 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, Vol. 1 (1962), no. (1) doi:10.1137/0301006.

[8] Halina Frankowska, *A priori estimates for operational differential inclusions*, Journal of Differential Equations **84** (1990), no. 1, 100–128. doi:10.1016/0022-0396(90)90129-D.

[9] Falk M Hante, *Relaxation methods for hyperbolic pde mixed-integer optimal control problems*, Optimal Control Applications and Methods **38** (2017), no. 6, 1103–1110. doi:10.1002/oca.2315.

[10] Falk M Hante and Sebastian Sager, *Relaxation methods for mixed-integer optimal control of partial differential equations*, Computational Optimization and Applications **55** (2013), no. 1, 197–225. doi:10.1007/s10589-012-9518-3.

[11] Jaroslav Haslinger and Raino AE Mäkinen, *On a topology optimization problem governed by two-dimensional Helmholtz equation*, Computational Optimization and Applications **62** (2015), no. 2, 517–544. doi:doi.org/10.1007/s10589-015-9746-4.

[12] Christian Kirches, Felix Lenders, and Paul Manns, *Approximation properties and tight bounds for constrained mixed-integer optimal control*, Optimization Online Preprint **5404** (2016). submitted http://www.optimization-online.org/DB_FILE/2016/04/5404.pdf.

[13] Paul Manns, Christian Kirches, and Felix Lenders, *A linear bound on the integrality gap for sum-up rounding in the presence of vanishing constraints* (2017). submitted http://www.optimization-online.org/DB_FILE/2018/04/6580.pdf.

[14] Amnon Pazy, *Semigroups of linear operators and applications to partial differential equations*, Vol. 44, Springer Science & Business Media, 1983. doi:10.1007/978-1-4612-5561-1.

[15] Sebastian Sager, *Numerical methods for mixed-integer optimal control problems*, Der andere Verlag Tönning, Lübeck, Marburg, 2005. doi:https://mathopt.de/PUBLICATIONS/Sager2005.pdf.

[16] ______, *Reformulations and Algorithms for the Optimization of Switching Decisions in Nonlinear Optimal Control*, Journal of Process Control **19** (2009), no. 8, 1238–1247. doi:10.1016/j.jprocont.2009.03.008.

[17] Sebastian Sager, Hans-Georg Bock, and Moritz Diehl, *The Integer Approximation Error in Mixed-Integer Optimal Control*, Mathematical Programming, Series A **133** (2012), no. 1–2, 1–23. doi:10.1007/s10107-010-0405-3.

[18] Jacques Simon, *Compact sets in the space* $L^p((0, T), B)$, Annali di Matematica pura ed applicata **146** (1986), no. 1, 65–96. doi:10.1007/BF01762360.

[19] Tadeusz Ważewski, *On an optimal control problem*, Differential equations and their applications, 1963, pp. 229–242. doi:10338.dmlcz/702189.

[20] Fengqi You and Sven Leyffer, *Oil spill response planning with minlp*, SIAG/OPT Views-and-News **21** (2010), no. 2, 1–8. doi:10.1002/aic.12536.

[21] Victor M Zavala, *Stochastic optimal control model for natural gas networks*, Computers & Chemical Engineering **64** (2014), 103–113. doi:10.1016/j.compchemeng.2014.02.002.

Institut für Mathematische Optimierung, Technische Universität Carolo-Wilhelmina zu Braunschweig, Universitätsplatz 2, 38106 Braunschweig, GERMANY; paul.manns@tu-bs.de\&\c.kirches@tu-bs.de