Computational Optimization and Applications (2023) 86:885–923
https://doi.org/10.1007/s10589-023-00517-9

A Filippov approximation theorem for strengthened one-sided Lipschitz differential inclusions

Robert Baier[1] · **Elza Farkhi**[2]

Received: 10 July 2022 / Accepted: 19 July 2023 / Published online: 11 September 2023

Abstract
We consider differential inclusions with strengthened one-sided Lipschitz (SOSL) right-hand sides. The class of SOSL multivalued maps is wider than the class of Lipschitz ones and a subclass of the class of one-sided Lipschitz maps. We prove a Filippov approximation theorem for the solutions of such differential inclusions with perturbations in the right-hand side, both of the set of the velocities (outer perturbations) and of the state (inner perturbations). The obtained estimate of the distance between the approximate and exact solution extends the known Filippov estimate for Lipschitz maps to SOSL ones and improves the order of approximation with respect to the inner perturbation known for one-sided Lipschitz (OSL) right-hand sides from $\frac{1}{2}$ to 1.

Keywords Differential inclusions · Filippov theorem · (Strengthened) one-sided Lipschitz condition · Monotonicity · Set-valued Euler method · Reachable sets

Mathematics Subject Classification 47H05 · 47H06 · 54C60 · 26E25 · 34A60 · 34A36 · 49M25

Dedicated to the memory of Asen L. Dontchev

Robert Baier: Partially supported by Tel Aviv University, Mathematical Institute at Tel Aviv "MINT", and by Bavarian Research and Innovation Agency "BayFor". Elza Farkhi: Partially supported by University of Bayreuth, Mathematical Institute at Tel Aviv "MINT", and by Bavarian Research and Innovation Agency "BayFor".

✉ Robert Baier
robert.baier@uni-bayreuth.de

Elza Farkhi
elza@tauex.tau.ac.il

[1] Chair of Applied Mathematics, University of Bayreuth, 95440 Bayreuth, Germany

[2] School of Mathematical Sciences, Sackler Faculty of Exact Sciences, Tel Aviv University, 69978 Tel Aviv, Israel

1 Introduction

We consider the *differential inclusion*

$$\dot{x}(t) \in F(t, x(t)) \qquad \text{for a.e. } t \in [t_0, T] \quad \text{and} \quad x(t_0) = x^0, \tag{1}$$

where F is a *set-valued map* defined in $\mathbb{R}^{n+1}$ with nonempty compact (possibly convex) sets in $\mathbb{R}^n$ as values, measurable in the time t for all x and upper semi-continuous (not necessarily continuous) in the state x for almost all $t \in I = [t_0, T]$.
The solutions of the inclusion are *absolutely continuous (AC)* functions $x : I \to \mathbb{R}^n$ satisfying (1) almost everywhere.

Filippov-type approximation theorems for differential inclusions follow the original theorem of Filippov [38] and provide approximation estimates for the solutions of (1) in the presence of perturbations by solutions of the original inclusion (1). The perturbations appear in the right-hand side $F(t, x)$ and in the initial set and the approximation estimates are given in terms of the norms of the perturbations. The theorem of Filippov extends classical results on Lipschitz continuity of the (unique) solution of an ODE with respect to perturbations in the right-hand side and the initial point, to Lipschitz stability of the solution set of a differential inclusion. We next recall the classical theorem of Filippov [38] in a slightly simplified form with a fixed Lipschitz constant instead of a time-depending one.

Theorem 1.1 (Filippov [38]) *Let $F : I \times \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ have closed, nonempty sets as values and consider an approximate solution $y : I \to \mathbb{R}^n$ with perturbed initial value $y(t_0) = y^0$ and*

$$\operatorname{dist}(\dot{y}(t), F(t, y(t))) \le \varepsilon(t), \tag{2}$$

where $\varepsilon(\cdot)$ is integrable. For $\rho > 0$ define the tube $\Omega(t) = y(t) + \rho B_1(0)$ for $t \in I$ with $x^0 \in \Omega(t_0)$ and let F be continuous (in the Hausdorff metric in Sect. 2.1) for all $x \in \Omega(t)$, $t \in I$ as well as let $F(t, \cdot)$ be Lipschitz continuous with a constant $L \ge 0$, i.e.

$$\mathrm{d_H}(F(t, x), F(t, y)) \le L|x - y|_2 \quad \text{for } x, y \in \Omega(t), t \in I. \tag{3}$$

Then there exists a (neighboring) solution $x(\cdot)$ of (1) on a subinterval of I such that

$$|y(t) - x(t)|_2 \le \xi(t) = e^{L(t-t_0)}|y^0 - x^0|_2 + \int_{t_0}^{t} e^{L(t-s)}\varepsilon(s)\,ds \tag{4}$$

$$|\dot{y}(t) - \dot{x}(t)|_2 \le L\xi(t) + \varepsilon(t) \qquad \text{a.e.}$$

for $t \in I$ with $\xi(t) \le \rho$.

In other words, an approximate solution satisfying (3) with a (time-dependent) $\varepsilon(\cdot)$-violation of the velocity from the right-hand side $F(t, y(t))$ is close to a solution $x(\cdot)$ of the unperturbed system (1) with a distance proportional to the norm of the violation

$\varepsilon(\cdot)$. The importance of the theorem is confirmed by its numerous applications related to discrete or other approximations of differential inclusions (e.g., [28–31, 34–36, 64]), relaxation theorems (called also Filippov-Ważewski theorems) on the density of the solutions set of (1) in the set of relaxed solutions (e.g., [2, Sec. 2.4], [3, Sec. 10.4], [6, 9, 24, 42, 58]), results on the asymptotic behavior of the solutions and others (e.g., [31–33, 36]). That is why extending the scope of the Filippov theorem beyond the family of Lipschitzian, and beyond the one of continuous maps, is an attractive field of investigation. For more information we refer to [3, 31, 36, 42].

In this respect see also the discussion in [31] of the theorem of Pliś which states the existence of a neighboring trajectory for differential inclusions without assuming uniqueness. It is obtained in [56] for right-hand sides with closed values and integrable Lipschitz modulus and also includes an error estimate with a maximal solution of a corresponding ODE.

In this paper, for any given solution $y(\cdot)$ of the system with *inner and outer vector perturbations* in $\mathbb{R}^n$,

$$\dot{y}(t) \in F\big(t, y(t) + \overline{\delta}(t)\big) + \overline{\varepsilon}(t) \qquad \text{for a.e. } t \in [t_0, T] \quad \text{and} \quad y(t_0) = y^0 = x^0 + \overline{\rho}^0, \tag{5}$$

we want to obtain the existence of a solution of the original system (1) such that the distance between these two solutions is estimated by some norms of the measurable perturbations $\overline{\delta}(\cdot)$, $\overline{\varepsilon}(\cdot)$, $\overline{\rho}^0$ and is small if the perturbations are small.

Our motivation for representing the perturbed system in form (5) and the importance of inner perturbations $\overline{\delta}(\cdot)$, which is essential when F is not continuous with respect to the state variable, are discussed in details in Sect. 2.2.

Removing the continuity of F with respect to the state variable may be problematic, since then the continuous dependence of the solutions with respect to perturbations in the initial condition or the right-hand side may be lost. Fortunately, in some cases the continuous dependence is preserved, possibly in a Hölder form, as in the case of *one-sided Lipschitz (OSL) mapping* F.

The OSL property of *single-valued maps* in $\mathbb{R}^n$ or in Hilbert spaces is known in numerical analysis as a generalization of the Lipschitz continuity ([4, 22], [43, Sec. IV.12], [15]).

An early and restrictive extension of the OSL condition to *set-valued maps* is defined in [37] and [45, 49]. This condition, equivalent to the monotonicity of the map $\mu I - F$ for some $\mu \in \mathbb{R}$, may be satisfied only by maps that are a.e. single-valued [67].

A weaker abstract version of the OSL condition in Banach spaces is formulated in [23], and its most popular form for multimaps in $\mathbb{R}^n$ and Hilbert spaces is coined in [29]. More details on OSL maps may also be found in [25, 27].

Definition 1.2 ([29]) The set-valued map F defined from a domain $[t_0, T] \times D$ in $\mathbb{R}^{n+1}$ to $\mathbb{R}^n$ is called *One-Sided Lipschitz (OSL)* in D with constant $\mu \in \mathbb{R}$ if for **a.e.** $t \in [t_0, T]$, every $x, y \in D$ and every $v \in F(t, x)$ there is $w \in F(t, y)$ such that

$$\langle x - y, v - w\rangle \leq \mu |x - y|_2^2, \tag{6}$$

where $|\cdot|_2$ denotes the *Euclidean norm* in $\mathbb{R}^n$.

The OSL property describes a large family of mappings which contains both Lipschitz and dissipative maps (see also Sect. 2 for examples and a comparison to other known classes of Lipschitz-like maps).

One should note that the constant μ may be zero or even negative, in contrast to the case of Lipschitz continuity. The OSL systems with negative OSL constant have a strongly invariant set which is asymptotically stable and attracts every trajectory [32]. In addition, OSL maps are not necessarily continuous as is shown in Sect. 2: easy examples of discontinuous OSL single-valued functions in $\mathbb{R}^1$ with OSL constant $\mu = 0$ are monotone decreasing functions.

In the case of OSL map F (even in the presence of discontinuities) a Filippov type approximation theorem is proved in [29] for inclusions with OSL and convex-valued right-hand sides with outer perturbations and first order of approximation of the solutions with respect to these perturbation is established. This theorem is applied there to the Euler approximation of differential inclusions and error estimates are derived implying convergence for right-hand sides which may be not Lipschitz in the state variable (this is easy to see for autonomous inclusions). Effective estimates for the Euler scheme providing convergence for OSL mappings being discontinuous in the state variable follow from a Filippov-type theorem for OSL mappings with inner perturbations [30], where order $\frac{1}{2}$ of approximation with respect to the inner perturbations is obtained. This leads to the order $\mathcal{O}(\sqrt{h})$ of the Euler method for differential inclusions with (discontinuous) OSL right-hand sides.

The Strengthened One-Sided Lipschitz (SOSL) condition we define next is intermediate between the Lipschitz and the OSL condition, i.e. weaker than the Lipschitz condition and stronger than the OSL one.

Definition 1.3 ([53, p. 171]) The set-valued map F from a domain $[t_0, T] \times D$ in $\mathbb{R}^{n+1}$ to $\mathbb{R}^n$ is called *Strengthened One-Sided Lipschitz (SOSL)* in D with constant $\mu \in \mathbb{R}$ if for a.e. $t \in [t_0, T]$, every $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n) \in D$ and every $v = (v_1, \ldots, v_n) \in F(t, x)$ there is $w = (w_1, \ldots, w_n) \in F(t, y)$ such that for all $i \in \{1, \ldots, n\}$ we have the implications

$$x_i > y_i \quad \Rightarrow \quad v_i - w_i \leq \mu |x - y|_\infty \tag{7}$$

and

$$x_i < y_i \quad \Rightarrow \quad w_i - v_i \leq \mu |x - y|_\infty, \tag{8}$$

where $|\cdot|_\infty$ denotes the *maximum norm* in $\mathbb{R}^n$.

The two cases in the definition above can be unified with the trivial case $x_i = y_i$ as follows:
For a.e. $t \in [t_0, T]$, every $x, y \in D$ and every $w \in F(t, y)$ there is $v \in F(t, x)$ such that for all $i \in \{1, \ldots, n\}$ the implications (7)–(8) hold, or equivalently

$$(x_i - y_i)(v_i - w_i) \leq \mu |x_i - y_i| \cdot |x - y|_\infty \qquad (i = 1, \ldots, n). \tag{9}$$

Note that also the SOSL constant μ may be negative and F is not necessarily a.e. single-valued. For maps with values in $\mathbb{R}^1$, the SOSL condition is equivalent to the OSL one. Also, the set-valued map F is SOSL iff co F (with convexified values) is SOSL.

A somewhat stronger (uniform) version of the SOSL condition appears earlier in [50, 51] (see remarks, e.g., in [8, 9]). First order convergence of the Euler scheme is derived in [49] for 1d case and in [50, 51] for higher dimensions for the unique solution of a differential inclusion satisfying this condition. Later, first order of convergence of the solution set for the explicit/implicit Euler method is derived in [9, 53] also for the wider class of SOSL maps as defined here.

The following hierarchy between the classes of OSL, SOSL and Lipschitz (in the Hausdorff metric) mappings with compact values in $\mathbb{R}^n$ is not hard to verify (see e.g., Example 2.8 and [8, Example 5.1]):

$$\text{OSL} \supset \text{SOSL} \supset \text{Lip}, \tag{10}$$

and there is no equality between any two classes.

Although the SOSL condition is weaker than the Lipschitz continuity, it is strong enough to provide approximation results for differential inclusions (see [9, 53]), better than for OSL maps and often the same as for Lipschitz maps. This is exactly the case with the Filippov approximation theorem (Theorem 1.1) proved here for SOSL maps in the right-hand side.

As a main result in this paper we prove a Filippov-type theorem for a SOSL right-hand side F with inner and outer perturbations. The obtained estimate of the distance between the perturbed and non-perturbed solutions is of first order, as in the classical Filippov theorem for the Lipschitz case, and improves the corresponding approximation estimate for OSL right-hand side of [30], removing the square root on the norm of the inner perturbation. Thus we prove the correctness of the conjecture in [30, Remark 3.2] stating that, under a suitably defined SOSL condition, one may obtain first order convergence with respect to the inner perturbation.

The paper is organized as follows: In the next section general definitions and known facts as well as examples and properties of OSL and SOSL maps are presented. In Section 3 the main theorem of the paper together with stability results for reachable sets are presented. In Section 4 an application of this theorem to approximations of dynamical systems with numerical experiments are presented.

2 Preliminaries and examples

2.1 Notation

We denote *vectors* in $\mathbb{R}^n$ by $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$. The *(closed) Euclidean unit ball* in $\mathbb{R}^n$ is denoted by $B_1(0)$, the ball around the center x^0 with radius $r > 0$ by $B_r(x^0)$. The *maximum norm* of the vector $x \in \mathbb{R}^n$ is denoted by $|x|_\infty = \max_{1 \le i \le n} |x_i|$, its *Euclidean norm* is denoted by $|x|_2$ or simply as $|x|$. The *norm* of an L_∞-function $f : I \to \mathbb{R}^n$ for a bounded, nonempty interval $I = [t_0, T] \subset \mathbb{R}$ is $\|f\|_{L_\infty} =$

$\operatorname{esssup}_{t\in I}|f(t)|$, for f being an L_1-function we denote the corresponding norm as $\|f\|_{L_1} = \int_I |f(t)|\,dt$. For a real number μ we denote $\mu_+ = \max\{0,\mu\}$, $\mu_- = \min\{0,\mu\}$.

We denote by $\mathcal{K}(\mathbb{R}^n)$ the *set of compact, nonempty subsets* of $\mathbb{R}^n$ and by $\mathcal{C}(\mathbb{R}^n)$ the *set of convex, compact, nonempty subsets* of $\mathbb{R}^n$. To measure distances of bounded, nonempty sets $A, B \subset \mathbb{R}^n$ we introduce the *one-sided Hausdorff distance* $\mathrm{d}(A,B) = \sup_{a\in A}\operatorname{dist}(a,B)$ and the *(two-sided) Hausdorff distance* as $\mathrm{d_H}(A,B) = \max\left\{\mathrm{d}(A,B), \mathrm{d}(B,A)\right\}$, where $\operatorname{dist}(z,B) = \inf_{b\in B}|z-b|_2$ is the *distance of a vector* $z \in \mathbb{R}^n$ *to the set* B. The *norm of a set* is defined by $\|A\|_2 = \mathrm{d_H}(A,\{0\}) = \sup\{|a|_2 : a \in A\}$. Recall that the Hausdorff distance in $\mathcal{K}(\mathbb{R}^n)$ is also obtained via $\mathrm{d_H}(A,B) = \min\left\{\varepsilon > 0 \,|\, A \subset B + \varepsilon B_1(0),\ B \subset A + \varepsilon B_1(0)\right\}$. The *interior*, the *boundary* and the *closure* of a set $A \subset R^n$ are denoted by $\operatorname{int}(A)$, $\operatorname{bd}(A)$ and $\overline{A}$, respectively.

We fix the time interval $I = [t_0, T]$ and denote $F : D \Rightarrow \mathbb{R}^n$ for a *set-valued map* with *domain* $D \subset \mathbb{R}^m$ (usually $m \in \{n, n+1\}$) and which has subsets of $\mathbb{R}^n$ as images. The *graph of the set-valued map* F is defined as

$$\operatorname{Graph} F = \{(x,y) \in D \times \mathbb{R}^n \,:\, y \in F(x)\}.$$

F is *(Lebesgue) measurable* if the *pre-image* $F^{-1}(U) = \{t \in I \,:\, F(t) \cap U \neq \emptyset\}$ is a (Lebesgue) measurable set for each open set $U \subset \mathbb{R}^n$ [3, Sec. 8.1]. For a single-valued map $F(t) = \{f(t)\}$ this corresponds to the usual criterion for *(Lebesgue) measurable functions* $f : I \to \mathbb{R}^n$ that the *pre-image* $f^{-1}(U) = \{t \in I \,:\, f(t) \in U\}$ of an open set $U \subset \mathbb{R}^n$ is (Lebesgue) measurable. F with compact, nonempty images is *upper-semicontinuous (usc)* (in the ε-sense) [2, Sec. 1.1, Definition 5], [3, Sec. 1.4, below Definition 1.4.1] if for all $x \in D$, $\varepsilon > 0$ there exists $\delta > 0$ such that for all $y \in \mathbb{R}^n$ with $|y-x|_2 < \delta$ the inclusion $F(y) \subset F(x) + \varepsilon B_1(0)$, i.e. $\mathrm{d}(F(y), F(x)) \le \varepsilon$ holds (in contrary to set-valued continuity where $\mathrm{d}(F(x), F(y)) \le \varepsilon$ would also hold).

2.2 Inner and outer perturbations

We use the term "*inner perturbation*" for the state perturbation $\overline{\delta}(\cdot)$ in the inclusion (5) and "*outer perturbation*" for the perturbation of the set of velocities $\overline{\varepsilon}(\cdot)$ as it is done in the classial book of Filippov [39, Chap. 2, § 7] and e.g., in [19, Definition 2], [44, Sec. 2], [21, Sec. A.4, (2)], [5, Sec. 5], [12, (14)].

The lack of continuity of $F(t,\cdot)$ is the main reason to consider separately perturbations of the state variable (the inner perturbations) and perturbations of the set of velocities (the outer perturbations) as in [39, Chap. 2, § 7]. Indeed, if $F(t,\cdot)$ is Lipschitz continuous with constant L, we have for small $|\overline{\delta}(t)|_2$ the inclusion

$$F(t, x + \overline{\delta}(t)) \subseteq F(t,x) + L|\overline{\delta}(t)|_2 B_1(0).$$

Then any solution $y(\cdot)$ of the perturbed inclusion (5) fulfills the inclusion

$$\dot{y}(t) \in F\big(t, y(t)\big) + \overline{\xi}(t) \qquad \text{for a.e. } t \in [t_0, T] \quad \text{and} \quad y(t_0) = y^0 \in B_r(x^0), \tag{11}$$

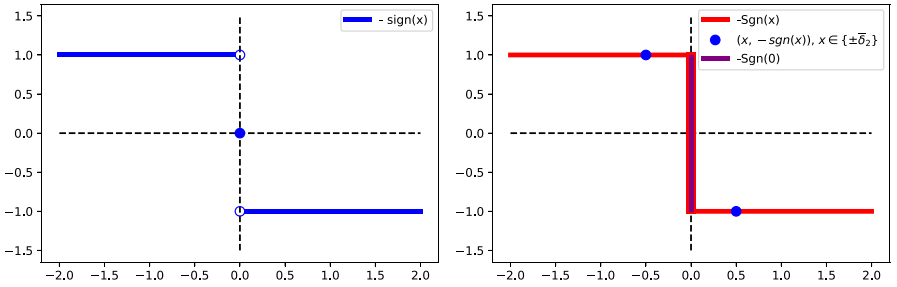

Fig. 1 The function $-\operatorname{sign}(x)$ and its (unperturbed) set-valued variant $-\operatorname{Sign}(x)$ (Color figure online)

where $|\overline{\xi}(t)|_2 \leq L|\overline{\delta}(t)|_2 + |\overline{\varepsilon}(t)|_2$. In the latter inclusion only **a small outer perturbation** is present. In this case it is sufficient to consider only outer perturbations in the Filippov-type theorems.

Yet, *without continuity* of $F(t,\cdot)$, an element of the set $F(t, x+\overline{\delta}(t))$ may be far away from the set $F(t,x)$ for small $|\overline{\delta}(t)|_2$ so that the approximation bound for the outer perturbation $|\overline{\xi}(t)|_2$ in (11) may be large, while the inner perturbations tend to zero.

The following simple example of Filippov illustrates this observation.
Let $F : \mathbb{R} \Rightarrow \mathbb{R}$ be defined by

$$F(x) = -\operatorname{Sign}(x) = \begin{cases} \{1\} & \text{if x} < 0, \\ [-1,1] & \text{if x} = 0, \\ \{-1\} & \text{if x} > 0. \end{cases} \tag{12}$$

The *set-valued map* in Fig. 1 (right plot) is the convex-valued usc "regularization" of $-\operatorname{sign}(x)$ (see (15), left plot) and is discontinuous, only upper semicontinuous, at $x = 0$.

On the graph of $F(x) = -\operatorname{Sign}(x)$ we consider a sequence of points $(x_k, y_k) = (\overline{\delta}_k, -1)$ and $(-x_k, -y_k) = (-\overline{\delta}_k, 1)$ with $\overline{\delta}_k = \frac{1}{k}$ for $k \in \mathbb{N}$ on its graph. In Fig. 1 (right plot) the red graph and the blue points for $k = 2$ are shown.

Due to the upper semi-continuity of F for $x = 0$, i.e. for all $\varepsilon > 0$ there exists $\delta > 0$ such that

$$F(z) \subset F(0) + \varepsilon B_1(0) \quad \text{for all } z \in \mathbb{R} \text{ with } |z| \leq \delta,$$

the sequence $((x_k, y_k))_k$ with $y_k \in F(x_k)$ converges to $(0,-1) \in F(0)$. Similarly, the sequence $((-x_k, -y_k))_k$ converges to $(0,1) \in F(0)$. The missing lower semi-continuity of F at $x = 0$ implies that the inclusion

$$F(0) \subset F(z) + \varepsilon B_1(0) \quad \text{for all } z \in \mathbb{R} \text{ with } |z| \leq \delta,$$

holds only with $\varepsilon \geq 2$ for any small $\delta > 0$, and not for smaller $\varepsilon > 0$.

Thus, replacing an inner perturbation by an outer one may yield too coarse estimates in the Filippov-type theorem. Considering inner perturbations separately from the outer

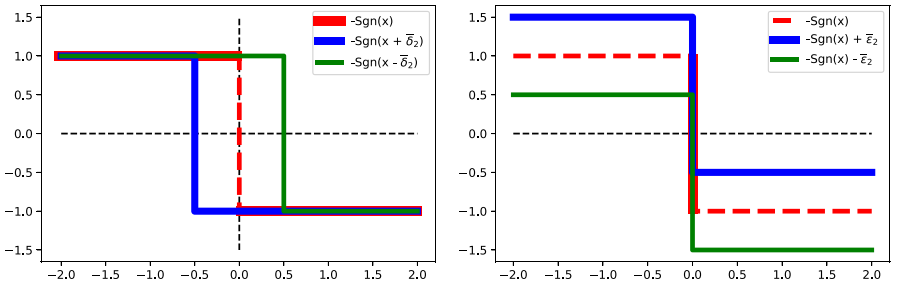

Fig. 2 Inner (on the left) resp. outer (on the right) vector perturbations of $-\operatorname{Sign}(x)$ (Color figure online)

ones refines the estimates and allows to extend the approximation estimates to the case of set-valued maps F which are discontinuous with respect to the state variable.

In Fig. 2 the graphs of two inner vector perturbations $F(x+\overline{\delta}_k)$ (in blue) and $F(x-\overline{\delta}_k)$ (in green) for $\overline{\delta}_k = \frac{1}{k}$ are shown for $k=2$ in the left plot, while the right plot shows two outer vector perturbations $F(x)+\overline{\varepsilon}_k$ (in blue) and $F(x)-\overline{\varepsilon}_k$ (in green) for $\overline{\varepsilon}_k = \frac{1}{k}$ and $k=2$. In both plots the graph of the original mapping $F(x) = -\operatorname{Sign}(x)$ (dashed lines in red color in both plots) is also present.

On Fig. 2 one checks visually that the Hausdorff distance between **the graphs** of $F(\cdot)$ and $F(\cdot+\overline{\delta}_k)$ is bounded by $\overline{\delta}_k$. The same estimate for the graphs hold for the outer vector perturbation $F(\cdot)+\overline{\varepsilon}_k$. Nevertheless, the Hausdorff distance between **the values** of F and the perturbed mapping $F(\cdot+\overline{\delta}_k)$ at a given point $x=0$ is equal to 2.

Let us sketch two more motivations for the systems (5) with vector and set-valued perturbations, respectively. Theorem 1.1 requires Lipschitz continuity in the state variable with closed, not necessarily convex values and essentially that the approximate solution fulfills the inequality (2). The latter together with $\varepsilon_0 = |y^0 - x^0|_2$ means that $y(\cdot)$ is a solution of the differential inclusion

$$\dot{y}(t) \in F(t, y(t)) + \varepsilon(t)B_1(0) \qquad \text{for a.e. } t \in I \quad \text{and} \quad y(t_0) \in x^0 + \varepsilon_0 B_1(0)$$

with set-valued outer perturbation $\varepsilon(t)B_1(0)$. In this case we can rewrite the inclusion in the form (5) with $\overline{\delta}(t) = 0$ by [21, Proposition 3.5].

The second motivation we would like to sketch comes from set-valued discretization methods for solving the differential inclusion (1) as the set-valued Euler method [10, 17, 35, 66]. A discrete solution for the step size $h = \frac{T-t_0}{N}$ with a given $N \in \mathbb{N}$ taking values on the grid points $t_j = t_0 + jh$, $j = 0, \ldots, N$, has the form $y^{j+1} = y^j + hw^j$, $w^j \in F(y^j)$, where we have assumed F to be autonomous for simplicity. To prove the convergence for this set-valued method, one essential step is to obtain the existence of a neighboring solution in continuous time. Consider the *piecewise linear interpolant*

$$y(t) = y^j + (t - t_j)w^j \quad \text{for } t \in I_j \tag{13}$$

on the subinterval $I_j = [t_j, t_{j+1}]$, $j = 0, \dots, N-1$. It is absolutely continuous with the derivative

$$\dot{y}(t) = w^j \in F(y^j) \tag{14}$$

in the interior of I_j. The right-hand side in (14) can be seen as an inner vector perturbation of the right-hand side $F(y(t))$ in (1), since

$$F(y^j) = F(y(t) + \overline{\delta}(t)) \quad \text{with} \quad \overline{\delta}(t) = \begin{cases} y^j - y(t) & \text{for } t \in I_j \backslash \{t_{j+1}\},\ j = 0, \dots, N-2, \\ y^{N-1} - y(t) & \text{for } t \in I_{N-1},\ j = N-1. \end{cases}$$

Thus, $y(\cdot)$ is a solution of the perturbed differential inclusion (14) and the Filippov Theorem 3.7 guarantees the existence of a neighboring solution of (1) at a distance $\mathcal{O}(h)$ for SOSL right-hand sides, if the inner perturbation $\overline{\delta}(t)$ is $\mathcal{O}(h)$ in norm. If the original inclusion (1) has a unique solution, this Filippov theorem already implies error estimates of order 1 for the set-valued Euler's or some Runge–Kutta methods (see [45, 49]).

2.3 Examples for SOSL/OSL set-valued maps

We list some classes of SOSL set-valued maps. An OSL (or SOSL) function in this subsection means a single-valued function taking values in $\mathbb{R}$ or $\mathbb{R}^n$. Since every single-valued map with the values from an OSL function is an OSL set-valued map (see Remark 2.3), we start the discussion with SOSL and OSL (single-valued) functions and the special case of linear functions.

Lemma 2.1 *Let $A \in \mathbb{R}^{n\times n}$ be a matrix and $b(t) \in \mathbb{R}^n$ for $t \in I$. Then the affine function $f(t, x) = Ax + b(t)$ for $x \in \mathbb{R}^n$, $t \in I$ is*

(i) OSL with constant $\mu = \lambda_{max}$, where λ_{max} is the maximal eigenvalue of the symmetrized matrix $A_{sym} = \frac{1}{2}(A + A^\top)$,

(ii) SOSL with constant $\mu = \max\limits_{i=1,\dots,n} \left(\max\{0, a_{ii}\} + \sum\limits_{\substack{j=1,\dots,n \\ j\neq i}} |a_{ij}| \right)$

The SOSL constant can be estimated via $\max\{0, \max\limits_{i=1,\dots,n} a_{ii}\} + \max\limits_{i=1,\dots,n} \sum\limits_{\substack{j=1,\dots,n \\ j\neq i}} |a_{ij}|$.

Proof (i) Let $x, y \in \mathbb{R}^n$, $t \in I$ and $v = f(t, x) = Ax + b(t)$, $w = f(t, y) = Ay + b(t)$. Then,

$$\begin{aligned} \langle x - y, v - w\rangle = \langle x - y, A(x - y)\rangle &= \frac{1}{2}\langle x - y, A(x - y)\rangle \\ &\quad + \frac{1}{2}\langle A^\top(x - y), x - y\rangle \\ &= \langle x - y, \frac{1}{2}(A + A^\top)(x - y)\rangle \\ &= \langle x - y, A_{\mathrm{sym}}(x - y)\rangle \le \lambda_{\max}|x - y|_2^2 \end{aligned}$$

is OSL with the claimed constant by the estimate with the Rayleigh quotient.

(ii) Let $i \in \{1, \dots, n\}$ and consider $v_i - w_i$. By $v = Ax + b(t)$, $w = Ay + b(t)$ we have $v_i - w_i = a_i(x - y)$ with the i-th row vector $a_i^\top \in \mathbb{R}^n$. Hence,

$$\begin{aligned}(x_i - y_i)(v_i - w_i) &= (x_i - y_i) \cdot \langle a_i^\top, x - y\rangle = (x_i - y_i) \sum_{j=1}^{n} a_{ij}(x_j - y_j)\\ &\le a_{ii}(x_i - y_i)^2 + |x_i - y_i| \sum_{\substack{j=1,\dots,n\\ j\neq i}} |a_{ij}| \cdot |x_j - y_j|\\ &\le \underbrace{\big(\max\{0, a_{ii}\} + \sum_{\substack{j=1,\dots,n\\ j\neq i}} |a_{ij}|\big)}_{=:\mu_i} |x_i - y_i| \cdot |x - y|_\infty.\end{aligned}$$

Obviously,

$$\begin{aligned}\mu_i &\le \max_{k=1,\dots,n} \mu_k = \max_{k=1,\dots,n} \left(\max\{0, a_{kk}\} + \sum_{\substack{j=1,\dots,n\\ j\neq k}} |a_{kj}|\right)\\ &\le \max\{0, \max_{k=1,\dots,n} a_{kk}\} + \max_{k=1,\dots,n} \sum_{\substack{j=1,\dots,n\\ j\neq k}} |a_{kj}|.\end{aligned}$$

□

In the previous lemma we could have estimated the SOSL constant by the bigger row-sum norm $\|A\|_\infty = \max\limits_{i=1,\dots,n} \sum\limits_{j=1,\dots,n} |a_{ij}|$, but then the SOSL constant could no longer be zero, e.g., for diagonal matrices with negative diagonal elements. Both constants can be non-positive as it is the case for $f(x) = Ax$ with the matrix $A = \begin{pmatrix} -2 & -1 \\ 1 & -1 \end{pmatrix}$ with eigenvalues $-2, -1$ for the symmetrized matrix A_{sym} (the OSL constant is $\mu = -1$) or for $f(x) = Bx$ with the diagonal matrix $B = \operatorname{diag}(\{-2, -1\})$ and the SOSL constant $\mu = 0$. It is easy to see with Lemma 2.1 that in the first case $f(x) = Ax$ is also SOSL but with positive constant $\mu = 1$.

Remark 2.2 Each real-valued monotone decreasing function with domain in $\mathbb{R}$ is SOSL (hence OSL) with constant $\mu = 0$ and every *dissipative* function from $\mathbb{R}^n$ to $\mathbb{R}^n$ (see [20, Chap. 3, (1)]) is OSL with the same constant.

The negation of the *sign* function

$$f(x) = -\operatorname{sign}(x) \quad \text{with} \quad \operatorname{sign}(x) = \begin{cases} -1 & \text{if } x < 0,\\ 0 & \text{if } x = 0,\\ 1 & \text{if } x > 0 \end{cases} \tag{15}$$

for $x \in \mathbb{R}$ (see in Fig. 1, left picture) is discontinuous at $x = 0$ and SOSL with constant $\mu = 0$. The function $g(x) = -x - \operatorname{sign}(x)$ is OSL with constant -1 and SOSL with constant 0.

We now list some classes of set-valued SOSL and OSL and show connections to previously defined notions in the literature.

Remark 2.3 Let $F : \mathbb{R}^n \Rightarrow \mathbb{R}^n$ be a set-valued map. Each single-valued map with $F(x) = \{f(x)\}$ and an OSL/SOSL function f is an OSL/SOSL set-valued map.

Let F be *dissipative* (see [20, Chap. 3, (1)]), i.e. $G = -F$ is *monotone/accretive* (see [21, Sec. 4.3] so that for all $x, y \in \mathbb{R}^n$ and all $v \in G(x)$, $w \in G(y)$ the inequality $\langle x - y, v - w\rangle \geq 0$ holds. Then F is OSL with constant 0. An important example for dissipative set-valued maps is $F(x) = -\partial g(x)$, the Moreau-Rockafellar subdifferential for a convex function $g : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$, see [21, Chap. 1, Sec. 4, Problems 12].

We state some more examples of OSL and SOSL maps and refer to [29, 30] for similar example classes and discussions on earlier OSL/SOSL concepts. The next result in (iv) generalizes [51, Lemma 3.6] to SOSL maps.

Proposition 2.4 *Let $F : \mathbb{R}^n \Rightarrow \mathbb{R}^n$ be a set-valued map and let one of the following assumptions hold:*

(i) *F is* Lipschitz *with constant $L \geq 0$, i.e. $\mathrm{d_H}(F(x), F(y)) \leq L|x - y|_2$, and set $\mu_F = \sqrt{n}L$.*

(ii) *$G : \mathbb{R}^n \Rightarrow \mathbb{R}^n$ is OSL/SOSL with constant $\mu_G \in \mathbb{R}$, $U, V \subset \mathbb{R}^n$ are nonempty and set $F(x) = G(x + U) + V$, $\mu_F = \mu_G$.*

(iii) *$G : \mathbb{R}^n \Rightarrow \mathbb{R}^n$ is OSL/SOSL with constant $\mu_G \in \mathbb{R}$, $\lambda \geq 0$ and set $F = \lambda G$, $\mu_F = \lambda\mu_G$.*

(iv) *$G, H : \mathbb{R}^n \Rightarrow \mathbb{R}^n$ are OSL/SOSL maps with constants $\mu_G, \mu_H \in \mathbb{R}$ and set $F = G + H$, $\mu_F = \mu_G + \mu_H$.*

(v) *$F_i : \mathbb{R} \Rightarrow \mathbb{R}$ are OSL maps with constants $\mu_i \in \mathbb{R}$, $i = 1, \ldots, n$ and set $F(x) = \sum_{i=1}^n F_i(x_i)e^i$ for $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ with the standard unit vectors $e^i \in \mathbb{R}^n$, $i = 1, \ldots, n$, the notation*

$$F_i(x_i)e^i = \{v \in \mathbb{R}^n : v_i \in F_i(x_i) \text{ and } v_j = 0 \text{ for } j \in \{1, \ldots, n\}, j \neq i\} \quad (16)$$

and $\mu_F = \max\{0, \max_{i=1,\ldots,n} \mu_i\}$.

If G, H are SOSL in (ii)–(iv), then F is also SOSL in (i)–(v) with the stated constant μ_F.
If G, H are OSL in (ii)–(iv), then F is OSL in (i)–(v) with constant μ_F (with $\mu_F = L$ in (i)).

Proof (i) is simple for the OSL or SOSL case and follows for $x, y \in \mathbb{R}^n$, $v \in F(x)$, $i = 1, \ldots, n$ from

$$\begin{aligned} \langle x - y, v - w\rangle &\leq |x - y|_2 \cdot |v - w|_2, \\ \text{resp. } (x_i - y_i)(v_i - w_i) &\leq |x_i - y_i| \cdot |v_i - w_i| \leq |x_i - y_i| \cdot |v - w|_2 \\ &\leq |x_i - y_i| \cdot L|x - y|_2 \leq L\sqrt{n} \cdot |x_i - y_i| \cdot |x - y|_\infty \end{aligned}$$

with $w \in F(y)$ such that $v \in w + L|x-y|_2\, B_1(0)$, see without proofs [30, Remark 2.1], [29, Remark 2.2], [26, Remark 1].
For (ii) see [30, Lemma 3.1] for OSL maps, for SOSL maps let $i = 1, \dots, n, x, y \in \mathbb{R}^n$, $z \in F(x)$ with $z = w + v$ and $w \in G(x+u)$, $u \in U$. Choose $\widetilde{w} \in G(y+u)$ such that the SOSL condition holds for G and set $\widetilde{z} = \widetilde{w} + v \in G(y+U) + V$. Then,

$$(x_i - y_i)(z_i - \widetilde{z}_i) = (x_i - y_i) \cdot \big((w_i + v_i) - (\widetilde{w}_i + v_i)\big) \le \mu_G |x_i - y_i| \cdot |x - y|_\infty.$$

The proofs of (iii)-(iv) are standard and left to the reader.
(v) Let $x, y \in \mathbb{R}^n$ and $v \in F(x)$ so that $v_i \in F_i(x_i)$ for $i = 1, \dots, n$. By the OSL condition there exists $w_i \in F_i(t, y) \subset \mathbb{R}$ with $(x_i - y_i)(v_i - w_i) \le \mu_i |x_i - y_i|^2$. We set $w = (w_1, \dots, w_n)$ so that $w \in F(y)$ and

$$(x_i - y_i)(v_i - w_i) \le \max\{\mu_i, 0\} |x_i - y_i| \cdot |x - y|_\infty$$

which proves the SOSL condition with constant μ_F. □

It is remarkable that many well-known functions (or their negation) in machine learning, electrical engineering, control theory or physics are SOSL (see e.g., [14, Sec. 2], [11, Sec. 2.4], [47, 57]), some functions are listed in the following example.

Example 2.5 All functions $f_i : \mathbb{R} \to \mathbb{R}$, $i = 1, 2$, below have the real numbers as domain and range and belong to the class of SOSL functions.

(i) The negation of the *sigmoidal* function

$$f_1(x) = -\sigma(x, \alpha) \quad \text{with} \quad \sigma(x, \alpha) = \frac{2}{1 + \exp(-\frac{x}{\alpha})} - 1 \tag{17}$$

for $x \in \mathbb{R}$ and some fixed $\alpha > 0$ is SOSL with constant $\mu = 0$, since $f_1(\cdot)$ is monotone decreasing, and is $C^\infty(\mathbb{R})$, in particular is Lipschitz with constant $L_1 = \frac{1}{2\alpha}$.

(ii) The negation of the *saturation* function

$$f_2(x) = -\operatorname{sat}(\beta x) \quad \text{with} \quad \operatorname{sat}(x) = \begin{cases} \operatorname{sign}(x) & \text{if } |x| > 1, \\ x & \text{if } |x| \le 1 \end{cases} \tag{18}$$

for $x \in \mathbb{R}$ and some fixed $\beta > 0$ is Lipschitz with constant $L_2 = \beta$. $f_2(\cdot)$ is SOSL with constant $\mu = 0$, since $f_2(\cdot)$ is also monotone decreasing.

The sigmoidal or the saturation function are used in practical realization (approximation) of the discontinuous sign function from Remark 2.2 (see e.g., [47, Sec. 3.1]), or in the theoretical analysis of discontinuous differential equations. This approximation is usually performed by the choice of small values $\alpha > 0$ for the sigmoidal function $f_1(x)$ in (17) or for the saturation function $f_2(x)$ by large values for $\beta > 0$ in (18).

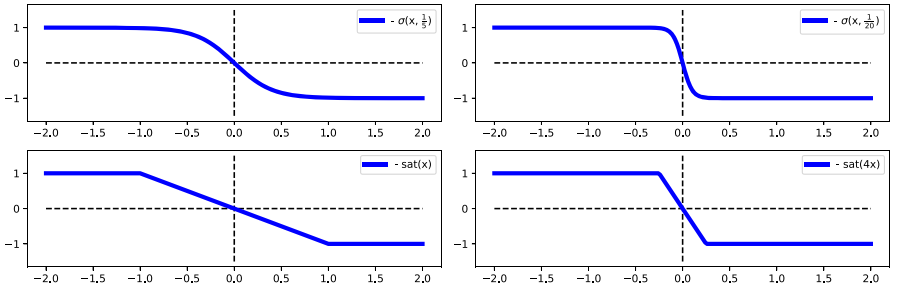

Fig. 3 Sigmoidal (first row) and saturation function (second row) for $\alpha \in \{\frac{1}{5}, \frac{1}{20}\}$, $\beta \in \{1, 4\}$(*Colorfigureonline*)

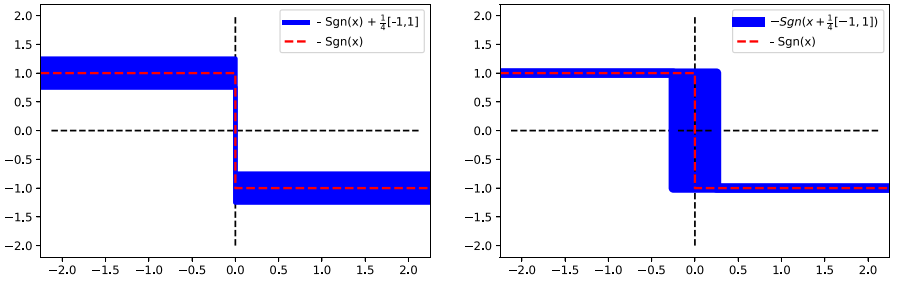

Fig. 4 The two SOSL maps F_1 and F_2 (outer and inner set perturbation of G on the left/right) (Color figure online)

In (i) $L_1 = \max_{x\in\mathbb{R}} |\dot{f}_1(x)| = |\dot{f}_1(0)| = \frac{1}{2\alpha}$ tends to ∞ for $\alpha \to 0 + 0$ and in (ii) the Lipschitz constant $L_2 = \beta$ explodes if the non-saturation zone $[-\frac{1}{\beta}, \frac{1}{\beta}]$ is narrowed for $\beta \to \infty$. This behavior can be observed in Fig. 3.

Further examples of SOSL (monotone decreasing) functions used in machine learning are the negation of the Heaviside and the ReLU/ramp function (see e.g., [14, Chap. 2]).

Next we present examples of OSL and SOSL set-valued maps which are not single-valued.

Example 2.6 We study examples of SOSL set-valued maps $F_i : \mathbb{R} \rightrightarrows \mathbb{R}$, $i = 1, 2$, with convex, compact, nonempty images which are set perturbations of the OSL map $G(x) = -\operatorname{Sign}(x)$ in the sense of Proposition 2.4(ii). Compare both perturbations with the original set-valued map G in Fig. 1 (right plot).

(i) $F_1(x) = -\operatorname{Sign}(x) + \frac{1}{4}[-1, 1]$ (outer perturbation of OSL set-valued map G) in Fig. 4 (left) is OSL (and SOSL) with constant $\mu = 0$ due to Proposition 2.4(ii) (apply with $U = \{0\}$ and $V = \frac{1}{4}[-1, 1]$). F_1 is discontinuous (only usc) and not dissipative.

(ii) $F_2(x) = -\operatorname{Sign}(x + \frac{1}{4}[-1, 1])$ (inner perturbation of OSL set-valued map G) in Fig. 4 (right) is OSL (and SOSL) with constant $\mu = 0$ due to Proposition 2.4(ii) (apply with $U = \frac{1}{4}[-1, 1]$ and $V = \{0\}$). F_2 has the same properties as F_1 in (i).

Example 2.7 An example of a discontinuous SOSL set-valued map defined in $\mathbb{R}$ with non-degenerate intervals as values, a negative SOSL constant $\mu = 0$ and which cannot be represented as the sum of a Lipschitz multifunction and a dissipative (SOSL) single-valued function is $F(x) = \text{co}\{-\text{sign}(x), -(\text{sign}(x) + x^{1/3})\}$.

We end this section by one example which is OSL but not SOSL.

Example 2.8 ([8, Example 5.6]) Consider the set-valued map $F(x) = -\partial g(x)$ for $x \in \mathbb{R}^n$ which is the convex subdifferential of Rockafellar/Moreau of the real-valued function $g(x) = |x|_2$.

Then F is OSL and dissipative by Remark 2.3 (i.e. $-F$ is monotone) but not SOSL for $n \geq 2$.

Another example would be the Hölder continuous function of degree $\frac{1}{3}$ from [30, Example 5.4] which is OSL with constant $\mu = \frac{1}{2}$ but not SOSL. More variants of Lipschitz-type or OSL-type set-valued maps and corresponding examples can be found in [9] and [7, 8].

3 Filippov-type theorems for SOSL maps

3.1 Existence and boundednes of solutions

For the proof of Filippov theorems under weaker conditions than Lipschitz continuity we need an existence result for differential inclusions under weak assumptions.

Theorem 3.1 ([54, Corollary 6 of Theorem 1]) *Consider $F : I \times \mathbb{R}^n \Rightarrow \mathbb{R}^n$, $x^0 \in \mathbb{R}^n$ such that*

(i) $F(t, x)$ is an orientor field, i.e. $F(t, x)$ is closed and nonempty,
(ii) $F(\cdot, x)$ is measurable in $t \in I$ for all $x \in \mathbb{R}^n$,
(iii) for almost all $t \in I$
– either $F(t, \cdot)$ is upper semi-continuous (= usc) *at $x \in \mathbb{R}^n$ and $F(t, x)$ is convex*
– or $F(t, \cdot)$ is lower-semicontinuous (= lsc) *at some neighborhood of $x \in \mathbb{R}^n$*
(iv) $F(\cdot, \cdot)$ is (weakly) locally integrably bounded*, i.e. for every bounded set $\widetilde{S} \subset I \times \mathbb{R}^n$ the distance function* $\text{dist}(0, F(t, x))$ *is bounded by an integrable function $k : I \to \mathbb{R}$ for all $(t, x) \in \widetilde{S}$.*

Then, there exists a solution of the differential inclusion (1).

Remark 3.2 The assumptions of the theorem provide two options to guarantee the existence: convex images with upper semi-continuity only or lower semi-continuity with nonconvex closed images (similar to the discussion in [2, Sec. 2.1, p. 94]). From now on we mainly follow the first option for the rest of the paper, since the set-valued map $-\text{Sign}(x)$ which appears in most of our applications is only usc.

A similar local existence result can be found in [63] in Theorem 8.13, where (ii) is replaced by the weaker existence of a (strongly) measurable selector of $F(\cdot, x)$. The global existence follows from Theorem 8.15 together with Example 8.17.

We now summarize our basic assumptions on the right-hand side $F : I \times \mathbb{R}^n \Rightarrow \mathbb{R}^n$ of the differential inclusion. Here, the boundedness condition in **(A1)** is slightly stronger than (iv) in the previous existing result (that guarantees the boundedness of at least one solution), since we consider a sub-inclusion of $F(t, x)$ and also need the boundedness of all solutions.

(A1) $F(t, x) \subset \mathbb{R}^n$ is compact and nonempty and is *integrably bounded on bounded sets*, i.e. for every constant C and for every compact $S \subset \mathbb{R}^n$ with $||S||_2 \leq C$ there is an L_1-function $K_F(\cdot; C)$ such that

$$||F(t, S)||_2 \leq K_F(t; C). \tag{19}$$

(A2) $F(\cdot, x)$ is Lebesgue measurable in $t \in I$ for all $x \in \mathbb{R}^n$.
(A3) $F(t, \cdot)$ is upper semi-continuous at $x \in \mathbb{R}^n$ for almost all $t \in I$.
(A4) $F(t, x)$ is convex.
(A5) F is SOSL with a constant $\mu \in \mathbb{R}$.

In the case that (A2)–(A3) hold, F is called *upper Carathéodory* in [1, Sec. 4].

We first state a version of Gronwall's lemma in differential form which does not require the usual non-negativity of functions defining the right-hand side (20) of the inequality. It is inspired by the proofs of [65, Lemma 2.4.4] and [21, Sec. 8.5].

Lemma 3.3 *Let $I = [t_0, T]$, $k(\cdot)$ and $p(\cdot)$ are in $L_1(I)$, $\psi : I \to \mathbb{R}$ absolutely continuous and*

$$\dot{\psi}(t) \leq k(t)\psi(t) + p(t) \quad \text{for a.e. } t \in I. \tag{20}$$

Then,

$$\psi(t) \leq \varphi(t) = e^{K(t)}\left(\varphi(t_0) + \int_{t_0}^{t} e^{-K(s)} p(s)\, ds\right), \tag{21}$$

where $K(t) = \int_{t_0}^{t} k(s)\, ds$ and $\varphi(\cdot)$ solves the initial value problem

$$\dot{\varphi}(t) = k(t)\varphi(t) + p(t) \quad \text{for a.e. } t \in I, \quad \varphi(t_0) = \psi(t_0)$$

Proof Define the AC function $\eta(t) = e^{-K(t)}\psi(t)$ for $t \in I$. Then, via (20)

$$\dot{\eta}(t) = e^{-K(t)}(-k(t)\psi(t) + \dot{\psi}(t)) \leq e^{-K(t)} p(t). \tag{22}$$

$\eta(t) = \eta(t_0) + \int_{t_0}^{t} \dot{\eta}(s)\, ds$ by absolute continuity together with (22) and $\psi(t) = e^{K(t)}\eta(t)$ yields

$$\psi(t) \leq e^{K(t)}\left(\varphi(t_0) + \int_{t_0}^{t} e^{-K(s)} p(s)\, ds\right) = \varphi(t).$$

□

We prove a technical lemma for the boundedness of solutions with inner vector and outer set-valued perturbations similar to [29, Lemma 3.1] and [30, Lemma 3.2]. Note that the integrable boundedness condition in (A1) (see e.g., [31]), is weaker than simply boundedness on bounded sets [30] and the linear growth condition, $||F(t,x)||_2 \le c(t)(1+|x|_2)$ with $c(\cdot) \in L_1(I)$ [21, Chap. 2, § 6], but stronger than the condition (iv) in Theorem 3.1. The assumption (A1) allows the estimates for all perturbed solutions in the next lemma below.

Lemma 3.4 *Let $F : I \times \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ fulfill assumptions (A1) and be OSL with constant $\mu \in \mathbb{R}$.*

Then for all $K_\delta, K_\varepsilon, K_0 \ge 0$ there exist constants $C_B, C_F \ge 0$ such that for all measurable vector perturbations $\overline{\delta}(\cdot) \in L_\infty(I)$, $\overline{\varepsilon}(\cdot) \in L_1(I)$ and all initial values $y^0 \in \mathbb{R}^n$ with

$$\|\overline{\delta}(\cdot)\|_{L_\infty} \le K_\delta, \quad \|\overline{\varepsilon}(\cdot)\|_{L_1} \le K_\varepsilon, \quad |y^0 - x^0|_2 \le K_0, \tag{23}$$

the solutions $y(\cdot)$ of the perturbed inclusion (5) satisfy

$$\|y(\cdot)\|_{L_\infty} \le C_B, \qquad \|\dot{y}(\cdot)\|_{L_1} \le C_F, \tag{24}$$

where

$$C_B = e^{\mu_+(T-t_0)}\big(|x^0|_2 + K_0\big) + e^{\mu_+(T-t_0)}\big(\|K_F(\cdot; K_\delta)\|_{L_1} + K_\varepsilon\big), \tag{25}$$

$$C_F = \|K_F(\cdot; C_B + K_\delta)\|_{L_1} + K_\varepsilon. \tag{26}$$

Proof F is OSL so that for all $x, \widetilde{x} \in \mathbb{R}^n$ and for a.e. $t \in I$ (see (6) and [29])

$$\begin{aligned}\delta^*(x-\widetilde{x}, F(t,x)) &= \max_{v \in F(t,x)} \langle x-\widetilde{x}, v\rangle \le \max_{\widetilde{v} \in F(t,\widetilde{x})} \big(\langle x-\widetilde{x}, \widetilde{v}\rangle + \mu|x-\widetilde{x}|_2^2\big)\\ &= \delta^*(x-\widetilde{x}, F(t,\widetilde{x})) + \mu|x-\widetilde{x}|_2^2\end{aligned}$$

For a.e. $t \in I$, $\dot{y}(t) \in F(t, y(t) + \overline{\delta}(t)) + \overline{\varepsilon}(t)$. Then using the above inequality for support functions,

$$\begin{aligned}\langle y(t), \dot{y}(t) - \overline{\varepsilon}(t)\rangle &\le \delta^*(y(t), F(t, y(t)+\overline{\delta}(t))) \le \delta^*(y(t), F(t, \overline{\delta}(t))) + \mu|y(t)|_2^2\\ &\le \delta^*(y(t), F(t, \|\overline{\delta}(\cdot)\|_{L_\infty} B_1(0)) + \mu \cdot |y(t)|_2^2\\ &\le |y(t)|_2 \cdot \|F(t, K_\delta B_1(0))\|_2 + \mu \cdot |y(t)|_2^2\\ &\le \mu \cdot |y(t)|_2^2 + K_F(t; K_\delta) \cdot |y(t)|_2\end{aligned}$$

holds for a.e. $t \in I$ by a suitable L_1-function $K_F(\cdot; K_\delta)$, since the set $K_\delta B_1(0)$ is bounded. Hence,

$$\begin{aligned}\langle y(t), \dot{y}(t)\rangle &= \langle y(t), \dot{y}(t) - \overline{\varepsilon}(t)\rangle + \langle y(t), \overline{\varepsilon}(t)\rangle\\ &\le \mu \cdot |y(t)|_2^2 + K_F(t; K_\delta) \cdot |y(t)|_2 + |\overline{\varepsilon}(t)|_2 \cdot |y(t)|_2.\end{aligned}$$

Introducing the function $p(t) = |y(t)|_2$, it is trivial to show by definition that $p(\cdot)$ is AC and that $p^2(\cdot)$ is differentiable at each point of differentiability of $y(\cdot)$, i.e. almost everywhere in I. Since $p(t)^2$ is a composition of an (outer) locally Lipschitz function $g(s) = s^2$ and an AC function $p(t)$, the (extended) chain rule holds for a.e. $t \in I$ by [59, Theorem 2] yielding

$$p(t)\dot{p}(t) = \frac{1}{2}\frac{d}{dt}p(t)^2 = \frac{1}{2}\frac{d}{dt}\langle y(t), y(t)\rangle = \langle y(t), \dot{y}(t)\rangle \leq \mu \cdot p(t)^2 + (K_F(t; K_\delta) + |\overline{\varepsilon}(t)|_2) \cdot p(t) \tag{27}$$

Next we want to prove (28) for almost every $t \in I$.

$$\dot{p}(t) \leq \mu \cdot p(t) + K_F(t; K_\delta) + |\overline{\varepsilon}(t)|_2 \tag{28}$$

Case 1: Consider the points t where $p(t) \neq 0$ and $\dot{p}(t)$ exists.
In the (measurable) set of points $t \in I$ where $p(t) \neq 0$ we can cancel $p(t)$ on both sides of the estimate (27) and get (28).

Case 2: If t lies in the (measurable) set $\mathcal{N} = \{\tau \in I : p(\tau) = 0\}$, we can consider only its subset of the points of density (which is of full measure by the Lebesgue density theorem, see [13, Chap. II, Theorem 5.1]), at which also the derivative $\dot{p}(t)$ exists, since $p(\cdot)$ is absolutely continuous. Consider an arbitrary sequence $\{t_k\}_k$ in $\mathcal{N}$ converging to such a density point t and calculate

$$\dot{p}(t) = \lim_{k\to\infty} \frac{p(t) - p(t_k)}{t - t_k} = 0$$

since $p(t) = 0$. Then (28) is trivially fulfilled.

In both cases (28) holds for a.e. $t \in I$ and it follows from the Gronwall inequality (Lemma 3.3) that

$$\begin{aligned} |y(t)|_2 = p(t) &\leq e^{\mu(t-t_0)}|y^0|_2 + \int_{t_0}^{t} e^{\mu(t-s)}\big(K_F(s; K_\delta) + |\overline{\varepsilon}(s)|_2\big)\,ds \\ &\leq e^{\mu(T-t_0)}\big(|x^0|_2 + K_0\big) + e^{\mu(T-t_0)}\Bigg(\|K_F(\cdot; K_\delta)\|_{L_1} + \underbrace{\int_{t_0}^{T} |\overline{\varepsilon}(s)|_2\,ds}_{\leq K_\varepsilon}\Bigg) = C_B \end{aligned} \tag{29}$$

which proves the first inequality in (24). Furthermore, we have for a.e. $t \in I$:

$$\begin{aligned} |\dot{y}(t)|_2 &\leq \|F(t, y(t) + \overline{\delta}(t)) + \overline{\varepsilon}(t)\|_2 \leq \|F(t, y(t) + \overline{\delta}(t))\|_2 + |\overline{\varepsilon}(t)|_2 \\ &\leq \|F(t, (C_B + K_\delta)B_1(0)))\|_2 + |\overline{\varepsilon}(t)|_2 \leq K_F(t; C_B + K_\delta) + |\overline{\varepsilon}(t)|_2, \\ \|\dot{y}(\cdot)\|_{L_1} &= \int_{t_0}^{T} |\dot{y}(t)|_2\,dt \leq \|K_F(\cdot; C_B + K_\delta)\|_{L_1} \\ &\quad + \int_{t_0}^{T} |\overline{\varepsilon}(t)|_2\,dt \leq \|K_F(\cdot; C_B + K_\delta)\|_{L_1} + K_\varepsilon = C_F \end{aligned}$$

which proves the second inequality in (24). □

To prove a Filippov-type theorem for the SOSL case, we state an equivalent condition to the SOSL property which refines the working condition in [53, Sec. 2, (31)] and is applied in the proofs in this section.

Lemma 3.5 *Let $F : I \times \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ have nonempty images. The following condition is equivalent to the SOSL condition for F:*
For a.e. $t \in I$ and every $x, y, \tilde{y} \in \mathbb{R}^n$, $w \in F(t, y)$ there is $v \in F(t, x)$ such that

$$(x_i - \tilde{y}_i)(v_i - w_i) \le \mu |x_i - \tilde{y}_i| \cdot |x - \tilde{y}|_\infty + |\mu| \cdot |x_i - \tilde{y}_i| \cdot |y - \tilde{y}|_\infty \tag{30}$$

for every index $i \in \{1, \dots, n\}$ satisfying

$$|x_i - y_i| > |y - \tilde{y}|_\infty. \tag{31}$$

Proof For given $t, x, y, \tilde{y}$ we denote by J the set of indices satisfying (31).

First, we assume that (30) holds for any given $t, x, y, \tilde{y}$ and indices $i \in J$. Choosing $\tilde{y} = y$, we get from (30) the SOSL condition in the form (9) for a.e. $t \in I$.

Conversely, let F be SOSL. Then there exists a subset $\widetilde{I} \subset I$ of full measure such that the inequalities in Definition 1.3 hold for given $x, y, \tilde{y} \in \mathbb{R}^n$. Let $t \in \widetilde{I}$ and $i \in J$. Without loss of generality suppose $x_i > y_i$. Then it follows from (31) that $x_i > \tilde{y}_i$. We obtain from the SOSL condition that for the given $x, y \in \mathbb{R}^n$, $w \in F(t, y)$ there is $v \in F(t, x)$ such that for $i \in J$

$$v_i - w_i \le \mu |x - y|_\infty.$$

We multiply this inequality by the positive number $x_i - \tilde{y}_i = |x_i - \tilde{y}_i|$ and obtain

$$(x_i - \tilde{y}_i)(v_i - w_i) \le \mu |x_i - \tilde{y}_i| \cdot |x - y|_\infty. \tag{32}$$

Then, for $\mu \ge 0$ we apply the triangle inequality $|x - y|_\infty \le |x - \tilde{y}|_\infty + |y - \tilde{y}|_\infty$ and get

$$(x_i - \tilde{y}_i)(v_i - w_i) \le \mu |x_i - \tilde{y}_i| \cdot |x - \tilde{y}|_\infty + \mu |x_i - \tilde{y}_i| \cdot |y - \tilde{y}|_\infty$$

which obviously implies the claim for $t \in \widetilde{I}$.

In the case $\mu < 0$ we use the inverse inequality $|x - y|_\infty \ge |x - \tilde{y}|_\infty - |y - \tilde{y}|_\infty$ and get from (32)

$$(x_i - \tilde{y}_i)(v_i - w_i) \le \mu |x_i - \tilde{y}_i| \cdot |x - \tilde{y}|_\infty - \mu |x_i - \tilde{y}_i| \cdot |y - \tilde{y}|_\infty$$

which also implies (30). □

Remark 3.6 The working condition for SOSL maps in Lemma 3.5 plays a key role in the definition of an auxiliary differential sub-inclusion in the proof of the Filippov theorem for SOSL maps. The corresponding one for OSL maps

$$\langle x - \widetilde{y}, v - w\rangle \le \mu|x - y|_2^2 + |y - \tilde{y}|_2 \cdot |v - w|_2$$

which is equivalent to the OSL condition is used in [30] for the same purpose in the proof of the Filippov theorem in the OSL case.

We recall the working condition for SOSL maps in [53, Sec. 2, (31)]:

For (a.e.) $t \in I$ and all $x, y, \tilde{y} \in \mathbb{R}^n$, $v \in F(t, x)$ there is $w \in F(t, y)$ such that

$$(x_i - \tilde{y}_i)(v_i - w_i) \le \mu|x - \tilde{y}|_\infty^2 + \mu \cdot |x - \tilde{y}|_\infty \cdot |y - \tilde{y}|_\infty \tag{33}$$

for indices $i \in \{1, \dots, n\}$ satisfying

$$|x_i - y_i| > \kappa|y - \tilde{y}|_\infty. \tag{34}$$

Both working conditions (30)–(31) and (33)–(34) are equivalent to the SOSL condition for $\mu \ge 0$, $\kappa = 1$, but only (30)–(31) is equivalent to the SOSL property if $\mu < 0$.

3.2 Filippov approximation theorem for the SOSL case

We now state the main result of this paper, the Filippov theorem for inclusions with SOSL right-hand sides.

Theorem 3.7 (Filippov-type theorem for the SOSL case with inner perturbations)

Let $F : I \times \mathbb{R}^n \Rightarrow \mathbb{R}^n$ satisfy (A1)–(A5), consider the inner vector perturbation $\overline{\delta}(\cdot) \in L_\infty(I)$ and let $\widetilde{y}(\cdot)$ be a solution of the inclusion

$$\dot{\widetilde{y}}(t) \in F(t, \widetilde{y}(t) + \overline{\delta}(t)) \quad \text{for a.e. } t \in I \quad \text{and} \quad \widetilde{y}(t_0) = y^0. \tag{35}$$

Then there exists a solution $x(\cdot)$ of the inclusion (1) such that for all $t \in [t_0, T]$

$$\begin{aligned} |\widetilde{y}(t) - x(t)|_\infty \le\ & \max\left\{e^{\mu(t-t_0)}|y^0 - x^0|_\infty,\ 2e^{\mu_+(t-t_0)} \cdot \|\overline{\delta}(\cdot)\|_{L_\infty}\right\} \\ & + |\mu| \cdot \int_{t_0}^{t} e^{\mu(t-s)}|\overline{\delta}(s)|_2\, ds. \end{aligned} \tag{36}$$

Proof The proof is done in several steps. Denote by Ω the measurable set of points t in I in which all $\widetilde{y}_i(\cdot)$, $i = 1, \dots, n$, are differentiable in t as well as (9), (30)–(31), (35) and the upper-semicontinuity of $F(t, \cdot)$ hold. Since $\widetilde{y}(t)$ is absolutely continuous, Ω has full measure in I.

Step 1: definition of an auxiliary differential inclusion involving the criterion of Lemma 3.5

For the given functions $\widetilde{y}(\cdot), \overline{\delta}(\cdot)$ we set $y(t) = \widetilde{y}(t) + \overline{\delta}(t)$ and for any $x \in \mathbb{R}^n$, $t \in I = [t_0, T]$ we denote by $J(t,x)$ the set of indices $i \in \{1, 2, \ldots, n\}$ satisfying the condition

$$|x_i - y_i(t)| > \|\overline{\delta}(\cdot)\|_{L_\infty}. \tag{37}$$

Clearly, for the given $t, x, y(t), \widetilde{y}(t)$, we have $J(t,x) \subset J$, where J is the set of indices for which (31) holds (see the proof of Lemma 3.5). For $(t,x) \in \Omega \times \mathbb{R}^n$ let us introduce the set-valued mapping

$$G(t,x) = \Big\{ v \in F(t,x) \mid \forall i \in J(t,x) : \ (x_i - \widetilde{y}_i(t))(v_i - \hat{\widetilde{y}}_i(t)) \\ \leq \mu |x_i - \widetilde{y}_i(t)| \cdot |x - \widetilde{y}(t)|_\infty + |\mu| \cdot |x_i - \widetilde{y}_i(t)| \cdot |\overline{\delta}(t)|_\infty \Big\}. \tag{38}$$

Note that $G(t,x)$ is well-defined by (38) for all $x \in \mathbb{R}^n$ and $t \in \Omega$. For $t \in I \backslash \Omega$, $x \in \mathbb{R}^n$ we define $G(t,x) = F(t,x)$ and consider the auxiliary differential inclusion

$$\dot{x}(t) \in G(t, x(t)) \qquad \text{for a.e. } t \in I \quad \text{and} \quad x(t_0) = x^0. \tag{39}$$

Step 2: verification of the conditions in Theorem 3.1 ensuring the existence of a solution of (39)
(i), (iii) The values of $G(t,x)$ are convex, compact, nonempty.
For $t \in I \backslash \Omega$, $x \in \mathbb{R}^n$, all three conditions in (i) hold by the assumptions on F, since $G(t,x) = F(t,x)$.
For $t \in \Omega$, the above mentioned inclusion $J(t,x) \subset J$ and Lemma 3.5 imply that $G(t,x) \neq \emptyset$ for all $x \in \mathbb{R}^n$, since $\hat{\widetilde{y}}(t) \in F(t, \widetilde{y}(t) + \overline{\delta}(t)) = F(t, y(t))$ for $t \in \Omega$. The convexity and closedness follow directly from (38). For the upper semi-continuity we now rewrite the definition of $G(t,x)$ for $t \in I$, $x \in \mathbb{R}^n$. We introduce for $i \in \{1, \ldots, n\}$:

- The set-valued map $H_i : \mathbb{R}^n \Rightarrow I$

$$H_i(x) = \{t \in I \ : \ \xi_i(t,x) > \|\overline{\delta}(\cdot)\|_{L_\infty}\} \quad \text{with} \quad \xi_i(t,x) = |y_i(t) - x_i|$$

  collects the times t for which (37) holds.
- The set-valued maps $\widetilde{D}_i, D_i, D : I \times \mathbb{R}^n \Rightarrow \mathbb{R}^n$, the functions $\eta_i, \beta_i : I \times \mathbb{R}^n \to \mathbb{R}$ by

$$\eta_i(t,x) = \widetilde{y}_i(t) - x_i,$$
$$\beta_i(t,x) = \begin{cases} (\widetilde{y}_i(t) - x_i)\hat{\widetilde{y}}_i(t) - \mu|x_i - \widetilde{y}_i(t)| \cdot |x - \widetilde{y}(t)|_\infty - |\mu| \cdot |x_i - \widetilde{y}_i(t)| \cdot |\overline{\delta}(t)|_\infty \\ \quad \text{for } t \in \Omega, x \in \mathbb{R}^n, \\ -|\widetilde{y}_i(t) - x_i| \cdot K_F(t, |x|_2) \quad \text{for } t \in I \backslash \Omega, x \in \mathbb{R}^n, \end{cases} \tag{40}$$

$$\widetilde{D}_i(t,x) = \{v \in \mathbb{R}^n \ : \ \eta_i(t,x) v_i \geq \beta_i(t,x)\} \cap F(t,x). \tag{41}$$

Note that for $t \in I \backslash \Omega$ the inequality in (41) is trivially satisfied for every $v \in F(t,x)$ by (19).

$$D_i(t,x) = \chi_{H_i(x)}(t)\widetilde{D}_i(t,x) + (1 - \chi_{H_i(x)}(t))F(t,x), \tag{42}$$

$$G(t,x) = \bigcap_{i=1}^{n} D_i(t,x). \tag{43}$$

It is easy to verify by (43) that G has closed values, since the values of F and D_i are closed for $t \in I$, $x \in \mathbb{R}^n$.

(ii) $G(\cdot, x)$ is measurable for any $x \in \mathbb{R}^n$

Let us first mention that all functions $\eta_i(t,x)v_i$ and $\beta_i(t,x)$ for a fixed $x \in \mathbb{R}^n$ are *Carathéodory* in $(t,v) \in I \times \mathbb{R}^n$, i.e. measurable in t for fixed v and continuous with respect to v for fixed t.

For a fixed $x \in \mathbb{R}^n$ the set $H_i(x)$ is measurable as the pre-image of the open interval $U = (\|\overline{\delta}(\cdot)\|_{L_\infty}, \infty)$ for the measurable function $\varphi(\cdot) = |y_i(\cdot) - x_i|$. For fixed $(t,x) \in \Omega \times \mathbb{R}^n$ the first operand in the intersection of set-valued map $\widetilde{D}_i(t,x)$ is measurable in t respectively by [16, Théorème 3.5]. The measurability of $\widetilde{D}_i(\cdot,x)$ follows from the intersection with the measurable set-valued map $F(\cdot,x)$, the one of $D_i(\cdot,x)$ follows from (42), since $H_i(x)$ is a measurable set and therefore, the characteristic function $\chi_{H_i(x)}(\cdot)$ is measurable in t by [18, Example 2.1.2] as well as the product $\chi_{H_i(x)}(\cdot)\widetilde{D}_i(\cdot,x)$ by [16, Corollaire 1]. As a finite intersection in (43) the measurability of $G(\cdot,x)$ is granted on I by [3, Theorem 8.2.4].

(iii) $G(t,\cdot)$ is usc for $t \in \Omega$

For this we show that for a fixed $t \in \Omega$ the graph of $D_i(t,\cdot)$ is closed for every $i = 1, \ldots, n$.

For sequences with $\lim_{k\to\infty} x^k = x^*$ and $\lim_{k\to\infty} v^k = v^*$ with $v^k \in D_i(t,x^k)$ we show that $v^* \in D_i(t,x^*)$.

case a: $t \in H_i(x^*)$

The continuity of $\xi(t,\cdot)$ yields that $t \in H_i(x^k)$ and $v^k \in \widetilde{D}_i(t,x^k)$ from (42) for large k.

The left- and right-hand sides $\eta_i(t,x)v_i$ and $\beta_i(t,x)$ in the inequality (41) are continuous in (x,v), so that the convergence of both sequences $\{x^k\}_k$, $\{v^k\}_k$ yield the inequality (41) in the first set of the intersection also for (x^*, v^*). Since the graph of $F(t,\cdot)$ is closed, $v^* \in \widetilde{D}_i(t,x^*)$ is valid and $v^* \in D_i(t,x^*)$ from $t \in H_i(x^*)$.

case b: $t \notin H_i(x^*)$

By definition in (42) $D_i(t,x^*) = F(t,x^*)$ and $v^* \in D_i(t,x^*)$ holds trivially.

Therefore in all cases the graphs of $D_i(t,\cdot)$ and $G(t,\cdot)$ are closed and $G(t,\cdot)$ is usc due to [2, Sec. 1.1, Theorem 1] (see also [3, Propositions 1.4.8-−1.4.9]), since $F(t,x)$ is compact and $F(t,\cdot)$ is usc in x.

(iv) G is locally integrably bounded as a subset of F, which is integrably bounded on bounded sets by (A1).

Hence, we have checked all assumptions of the Existence Theorem 3.1.

Step 3: solution of the auxiliary differential inclusion

By Theorem 3.1, there exists a solution $x(t)$ of the auxiliary inclusion (39). We set $z(t) = x(t) - \widetilde{y}(t)$ for the next two steps. Clearly, $z(\cdot)$ is AC and we can assume

without loss of generality (possibly after removing a set of measure zero from Ω) that $x(\cdot)$ and $z(\cdot)$ are differentiable for $t \in \Omega$.
In the next steps we prove the estimate (36).

Step 4: local SOSL estimate for $z_i(\cdot)$ on open subsets of Ω
For $i = 1, \dots, n$ we define the sets

$$\begin{aligned} \theta_i &= \left\{t \in I \,:\, |y_i(t) - x_i(t)| > \|\overline{\delta}(\cdot)\|_{L_\infty}\right\}, \\ T^i_{\max} &= \left\{t \in I \,:\, |z_i(t)| = |z(t)|_\infty\right\}, \\ \Theta_i &= \operatorname{int}\theta_i \cap \operatorname{int} T^i_{\max}. \end{aligned}$$

By the continuity of $z(\cdot)$ and the measurability of $y_i(\cdot) - x_i(\cdot)$, θ_i and $T^i_{\max}$ are measurable sets so that Θ_i is measurable and open. Define the open set $\Theta = \bigcup_{i=1}^n \Theta_i$. Then

$$I\backslash\Theta = \bigcap_{i=1}^n \left(I\backslash(\operatorname{int}(\theta_i) \cap \operatorname{int}(T^i_{\max}))\right) = \bigcap_{i=1}^n \left((I\backslash \operatorname{int}(\theta_i) \cup (I\backslash \operatorname{int}(T^i_{\max})\right) \tag{44}$$

is a closed set. Then clearly $I = \Theta \cup \operatorname{int}(I\backslash\Theta) \cup \operatorname{bd}(I\backslash\Theta)$. It is well-known that every open set $V \subset \mathbb{R}$ is a countable union of disjoint open intervals (see e.g., [60, Theorem 1.3] or [41, Proposition 0.21]). Every such disjoint open interval is the maximal interval (with respect to set inclusion) containing a given point of V. We will call these disjoint open intervals *(maximal) components* of V.

Step 4a: We now show that for any $i \in \{1, \dots, n\}$ and any (maximal) component of Θ_i, $\Delta = (t', t'')$ and every $t \in \overline{\Delta} = [t', t'']$ the following estimate holds:

$$|z(t)|_\infty \le e^{\mu(t-t')}|z(t')|_\infty + |\mu| \cdot \int_{t'}^{t} e^{\mu(t-s)}|\overline{\delta}(s)|_2 \, ds \tag{45}$$

Note that if (45) holds on the open interval Δ, then it is also true on its closure by the continuity of $z(\cdot)$ and of the function in the right-hand side of (45). For (45) we show the following estimate for a.e. $t \in \Theta_i$:

$$\frac{d}{dt}|z_i(t)| \le \mu|z(t)|_\infty + |\mu| \cdot |\overline{\delta}(t)|_2 \tag{46}$$

We use the definition of $G(t, x)$ (see (38)) for $t \in \Theta_i \cap \Omega$, since $\dot{x}(t) \in F(t, x(t))$ and (31) holds for $t \in \Delta \subset \Theta_i$ Hence, for $t \in \Theta_i \cap \Omega$

$$\begin{aligned} z_i(t)\dot{z}_i(t) &= (x_i(t) - \widetilde{y}_i(t))(\dot{x}_i(t) - \dot{\widetilde{y}}_i(t)) \\ &\le \mu|x_i(t) - \widetilde{y}_i(t)| \cdot |x(t) - \widetilde{y}(t)|_\infty + |\mu| \cdot |x_i(t) - \widetilde{y}_i(t)| \cdot |y(t) - \widetilde{y}(t)|_\infty \\ &= \mu|z_i(t)| \cdot |z(t)|_\infty + |\mu| \cdot |z_i(t)| \cdot |\overline{\delta}(t)|_2, \end{aligned} \tag{47}$$

since $y(t) = \widetilde{y}(t) + \overline{\delta}(t)$.

For the absolutely continuous function $p(\tau) = |z_i(\tau)|$ and $\tau \in I$ we can argue as in the proof of Lemma 3.4 to get that $p(\cdot)^2$ and $z_i(\cdot)^2$ are differentiable at the points where $|z_i(\cdot)|$ is differentiable that is w.l.o.g. in Ω (eventually removing a set of zero measure from Ω). Furthermore, the (extended) chain rule holds for $p(\tau)^2 = z_i(\tau)^2$ and a.e. $\tau \in I$ (w.l.o.g. we may assume that $\tau \in \Omega$) yielding together with (47)

$$\begin{aligned} p(\tau)\dot{p}(\tau) &= \frac{1}{2}\frac{d}{d\tau}p(\tau)^2 = \frac{1}{2}\frac{d}{d\tau}z_i(\tau)^2 = z_i(\tau)\dot{z}_i(\tau) \\ &\leq \mu p(\tau)\cdot|z(\tau)|_\infty + |\mu|\cdot p(\tau)\cdot|\overline{\delta}(\tau)|_2 \quad \text{for } \tau \in \Omega. \end{aligned} \tag{48}$$

We can repeat the arguments of cases 1 and 2 in the proof of Lemma 3.4 to show that (46) holds for $t \in \Theta_i \cap \Omega$. We can apply the Gronwall inequality (Lemma 3.3) together with $p(t) = |z_i(t)| = |z(t)|_\infty$ for $t \in \Delta \subset \Theta_i \subset T^i_{\max}$ and it follows from (46) that (45) holds.

Step 4b: We show that the inequality (45) proved in step 4a for a (maximal) component of Θ_i also holds for $t \in \overline{\Delta} = [t', t'']$ for any (maximal, possibly larger) component $\Delta = (t', t'')$ of $\Theta = \bigcup_{i=1}^n \Theta_i$.

Indeed, take an arbitary (maximal) component $\Delta_i = (t_i', t_i'')$ of Θ_i. If it does not intersect any (maximal) component $\Delta_j = (t_j', t_j'')$ of Θ_j for $j \neq i$, then Δ_i is also a (maximal) component of Θ and we can apply the result of step 4a.

If $\Delta_i \cap \Delta_j \neq \emptyset$ for some $j \neq i$, we now show that (45) holds in the closure of the interval $\Delta_i \cup \Delta_j = (t', t'')$.
There are two possibilities:

a) the inclusions $\Delta_i \subset \Delta_j$ or $\Delta_j \subset \Delta_i$ hold
   In this case we simply apply step 4a on the larger interval.
b) Δ_i and Δ_j overlap partially, i.e. either $t_j' \leq t_i' < t_j'' \leq t_i''$ or $t_i' \leq t_j' < t_i'' \leq t_j''$
   Assume for instance the first sub-case (the second one is similar to prove). Writing (45) for the interval $[t_j', t_i']$, we get

$$|z(t_i')|_\infty \leq e^{\mu(t_i'-t_j')}|z(t_j')|_\infty + |\mu|\int_{t_j'}^{t_i'} e^{\mu(t_i'-s)}|\overline{\delta}(s)|_2\, ds. \tag{49}$$

Let $t \in [t_j', t_i'']$. If $t \in [t_j', t_j'']$, then (45) holds by Claim 1 for this interval. In the other case $t \in [t_i', t_i'']$. Then we apply (45) in (t_i', t_i''), and for $|z(t_i')|_\infty$ we use (49) and get

$$\begin{aligned} |z(t)|_\infty &\leq e^{\mu(t-t_i')}|z(t_i')|_\infty + |\mu|\int_{t_i'}^{t} e^{\mu(t-s)}|\overline{\delta}(s)|_2\, ds \\ &\leq e^{\mu(t-t_i')}\left(e^{\mu(t_i'-t_j')}|z(t_j')|_\infty + |\mu|\int_{t_j'}^{t_i'} e^{\mu(t_i'-s)}|\overline{\delta}(s)|_2\, ds\right) \\ &\quad + |\mu|\int_{t_i'}^{t} e^{\mu(t-s)}|\overline{\delta}(s)|_2\, ds \end{aligned}$$

$$= e^{\mu(t-t'_j)}|z(t'_j)|_\infty + |\mu| \int_{t'_j}^{t'_i} e^{\mu(t-s)}|\overline{\delta}(s)|_2 \, ds$$
$$+ |\mu| \int_{t'_i}^{t} e^{\mu(t-s)}|\overline{\delta}(s)|_2 \, ds,$$

where we have used (49) in the second estimate. The estimate above implies that (45) holds in the closure of the union (t', t'') of any two intersecting (maximal) components Δ_i, Δ_j of Θ_i and Θ_j, respectively.

Since every (maximal) component of Θ is a union of countably many intersecting components of Θ_i, $i = 1, \ldots, n$, using the above argument and induction, we obtain that (45) holds in the closure of any (maximal) component of Θ.

In the next step we derive an error estimate in $I\backslash\Theta$ representing an error *reset* in the estimate, since errors at previous times are not accumulated in this case.

Step 4c (SOSL error reset): We now prove that for all $t \in \mathrm{int}(I)\backslash\Theta$ we have

$$|z(t)|_\infty \le 2\|\overline{\delta}(\cdot)\|_{L_\infty}. \tag{50}$$

Fix $t \in \mathrm{int}(I)\backslash\Theta$, and define $J_{\max}(t) = \{i \in \{1, \ldots, n\} : |z_i(t)| = |z(t)|_\infty\}$ as set of "maximal" indices. Obviously, $J_{\max}(t) \neq \emptyset$ and $t \in T^i_{\max}$ for all $i \in J_{\max}(t)$. Consider the possible cases:

1) there exists $i_0 \in J_{\max}(t)$ with $t \in \mathrm{int}(T^{i_0}_{\max})$
Since $t \in \mathrm{int}(I)\backslash\Theta$, it follows from a similar representation as in (44) that $t \in \mathrm{int}(I)\backslash\mathrm{int}(\Theta_{i_0})$. Hence, there are two sub-cases:
$\alpha)$ $t \in \mathrm{int}(I)\backslash\overline{\Theta}_{i_0}$, i.e. $t \notin \mathrm{int}(\Theta_{i_0})$ and $t \notin \mathrm{bd}(\Theta_{i_0})$
Then $t \notin \Theta_{i_0}$ and

$$|x_{i_0}(t) - y_{i_0}(t)| \le \|\overline{\delta}(\cdot)\|_{L_\infty}, \quad |z_{i_0}(t)| = |z(t)|_\infty,$$

thus by the triangle inequality and $t \in T^{i_0}_{\max}$

$$|z_{i_0}(t)| - |\overline{\delta}_{i_0}(t)| \le |z_{i_0}(t) - \overline{\delta}_{i_0}(t)| = |x_{i_0}(t) - \big(\widetilde{y}_{i_0}(t) + \overline{\delta}_{i_0}(t)\big)|$$
$$= |x_{i_0}(t) - y_{i_0}(t)| \le \|\overline{\delta}(\cdot)\|_{L_\infty}$$
so that $|z(t)|_\infty = |z_{i_0}(t)| \le \|\overline{\delta}(\cdot)\|_{L_\infty} + |\overline{\delta}_{i_0}(t)| \le \|\overline{\delta}(\cdot)\|_{L_\infty} + |\overline{\delta}(t)|_2 \le 2\|\overline{\delta}(\cdot)\|_{L_\infty}$.

$\beta)$ $t \in \mathrm{bd}(\Theta_{i_0})$
Since $t \in \mathrm{bd}(\Theta_{i_0})\backslash\,\mathrm{bd}(I)$, there exists a sequence $\{\tau_k\}_k \subset (\mathrm{int}(I)\backslash\overline{\Theta}_{i_0}) \cap \mathrm{int}(T^{i_0}_{\max})$ converging to t. Hence, by the definition of Θ_{i_0} and $T^{i_0}_{\max}$,

$$|x_{i_0}(\tau_k) - y_{i_0}(\tau_k)| \le \|\overline{\delta}(\cdot)\|_{L_\infty}, \quad |z_{i_0}(\tau_k)| = |z(\tau_k)|_\infty$$

for $k \in \mathbb{N}$. Thus, by the triangle inequality and $\tau_k \in T^{i_0}_{\max}$

$$
\begin{aligned}
|z_{i_0}(\tau_k)| - |\overline{\delta}_{i_0}(\tau_k)| &\le |z_{i_0}(\tau_k) - \overline{\delta}_{i_0}(\tau_k)| \\
&= |x_{i_0}(\tau_k) - \big(\widetilde{y}_{i_0}(\tau_k) + \overline{\delta}_{i_0}(\tau_k)\big)| \\
&= |x_{i_0}(\tau_k) - y_{i_0}(\tau_k)| \le \|\overline{\delta}(\cdot)\|_{L_\infty}, \\
\text{so that} \quad |z(\tau_k)|_\infty = |z_{i_0}(\tau_k)| &\le \|\overline{\delta}(\cdot)\|_{L_\infty} + |\overline{\delta}_{i_0}(\tau_k)| \\
&\le \|\overline{\delta}(\cdot)\|_{L_\infty} + |\overline{\delta}(\tau_k)|_2 \le 2\|\overline{\delta}(\cdot)\|_{L_\infty}.
\end{aligned}
$$

The continuity of $z(\cdot)$ yields $|z(t)|_\infty \le 2\|\overline{\delta}(\cdot)\|_{L_\infty}$ and also (50).

2) for all $i \in J_{\max}(t)$, $t \notin \operatorname{int}(T^i_{\max})$
Then, since $t \in \bigcap_{i\in J_{\max}(t)} T^i_{\max}$, it follows that there exists $i_0 \in J_{\max}(t)$ with $t \in \operatorname{bd}(T^{i_0}_{\max})$. $T^i_{\max}$ is closed by the continuity of $z_i(\cdot)$, $z(\cdot)$ so that $\operatorname{int}(I)\backslash T^i_{\max}$ is open and $\operatorname{bd}(T^i_{\max}) = \operatorname{bd}(\operatorname{int}(I)\backslash T^i_{\max})$ is contained in a union of countable many points which has measure 0. Thus we obtain (50) for a.e. $t \in \operatorname{int}(I)\backslash\Theta$. By the continuity of $|z(\cdot)|_\infty$ we get that (50) holds for every $t \in \operatorname{int}(I)\backslash\Theta$.

Step 4d: $t \in \operatorname{bd}(I) = \{t_0, T\}$
If $t = t_0$, then

$$
|z(t_0)|_\infty = |x(t_0) - \widetilde{y}(t_0)|_\infty = |x^0 - y^0|_\infty. \tag{51}
$$

Otherwise, $t = T$ which follows either from step 4b) and the continuity of $z(\cdot)$ (if T is at the boundary of Θ) with

$$
|z(T)|_\infty \le e^{\mu(T-t')}|z(t')|_\infty + |\mu| \int_{t'}^{T} e^{\mu(T-s)}|\overline{\delta}(s)|_2 \, ds \tag{52}
$$

or from step 4c) and the continuity of $z(\cdot)$ (if T is at the boundary of $\operatorname{int}(I)\backslash\Theta$) with

$$
|z(T)|_\infty \le 2\|\overline{\delta}(\cdot)\|_{L_\infty}. \tag{53}
$$

Step 5: We show that if $\Delta = (t', t'')$ is a (maximal) component of Θ with smallest value $t' \in I$, then either $t' = t_0$ or $|z(t')|_\infty \le 2\|\overline{\delta}(\cdot)\|_{L_\infty}$.

Indeed, if $t' > t_0$, then in each left neighborhood $(t' - \varepsilon, t')$ there is a point $\tau_\varepsilon \notin \Theta$, since otherwise one can extend Δ to the left in Θ and it will not be maximal in Θ. Thus for every $\varepsilon = \frac{1}{k}$, $k \in \mathbb{N}$, there is a $\tau_k \in (t' - \frac{1}{k}, t')\backslash\Theta$. As in step 4c we get $|z(\tau_k)|_\infty \le 2\|\overline{\delta}(\cdot)\|_{L_\infty}$, $k \in \mathbb{N}$. By the continuity of $z(\cdot)$ and its norm, we get $|z(t')|_\infty \le 2\|\overline{\delta}(\cdot)\|_{L_\infty}$.

Step 6: We show that the inequality

$$
\begin{aligned}
|z(t)|_\infty \le \max\{&e^{\mu(t-t_0)}|y^0 - x^0|_\infty, \; 2e^{\mu_+(t-t_0)} \cdot \|\overline{\delta}(\cdot)\|_{L_\infty}\} \\
&+ |\mu| \cdot \int_{t_0}^{t} e^{\mu(t-s)}|\overline{\delta}(s)|_2 \, ds
\end{aligned} \tag{54}
$$

holds for all $t \in I$.

Step 6a) We prove that (54) holds for $t \in \overline{\Theta}$.
Take a (maximal) component $\Delta = (t', t'')$ of Θ. By step 5, either $t' = t_0$ or $|z(t)|_\infty \le 2\|\overline{\delta}(\cdot)\|_{L_\infty}$. If $t' = t_0$, then by step 4b (or (52) for $t'' = T$) we have for $t \in \overline{\Delta} = [t_0, t'']$

$$\begin{aligned}|z(t)|_\infty &\le e^{\mu(t-t_0)}|z(t_0)|_\infty + |\mu| \cdot \int_{t_0}^{t} e^{\mu(t-s)}|\overline{\delta}(s)|_2 \, ds \\ &\le \max\{e^{\mu(t-t_0)}|z(t_0)|_\infty, \ 2e^{\mu_+(t-t_0)} \cdot \|\overline{\delta}(\cdot)\|_{L_\infty}\} \\ &\quad + |\mu| \cdot \int_{t_0}^{t} e^{\mu(t-s)}|\overline{\delta}(s)|_2 \, ds\end{aligned}$$

which proves (54) in this case together with $|z(t_0)|_\infty = |x^0 - y^0|_\infty$.
Let $t' > t_0$. Then by step 5, $|z(t')|_\infty \le 2\|\overline{\delta}(\cdot)\|_{L_\infty}$ and by this inequality and (45) we have for $t \in \overline{\Delta}$

$$\begin{aligned}|z(t)|_\infty &\le e^{\mu(t-t')} \cdot 2\|\overline{\delta}(\cdot)\|_{L_\infty} + |\mu| \cdot \int_{t'}^{t} e^{\mu(t-s)}|\overline{\delta}(s)|_2 \, ds \\ &\le \max\{e^{\mu(t-t_0)}|z(t_0)|_\infty, \ 2e^{\mu_+(t-t_0)} \cdot \|\overline{\delta}(\cdot)\|_{L_\infty}\} \\ &\quad + |\mu| \cdot \int_{t_0}^{t} e^{\mu(t-s)}|\overline{\delta}(s)|_2 \, ds.\end{aligned}$$

Trivially, (54) holds by (51) for $t = t_0$. Thus we have shown that (54) holds in each component of Θ, hence in the closure of Θ (by the continuity of $z(\cdot)$).

Step 6b) We prove that (54) holds for $t \in I \backslash \overline{\Theta}$.
By step 4c or (53) (if $T \in \mathrm{bd}(\Theta)$), we have to distinguish three cases, namely $t = t_0$, $t \in (t_0, T)$ and $t = T$. In the first two cases we have $|z(t_0)|_\infty = e^{\mu(t_0-t_0)}|z(t_0)|_\infty$ by (51) or $|z(t)|_\infty \le 2\|\overline{\delta}(\cdot)\|_{L_\infty}$ by (50) so that (54) holds. For the third case we use the inequality (53). Therefore, the estimate (54) follows immediately in all cases. □

We now prove a version of Filippov's Theorem for SOSL maps with inner and outer perturbations similar to the OSL case in [29, Theorem 3.2] and [30, Theorem 3.1] with a new proof idea.

Corollary 3.8 *(Filippov-type theorem for SOSL maps with inner and outer perturbations) Let $F : I \times \mathbb{R}^n \Rightarrow \mathbb{R}^n$ satisfy the assumptions (A1)–(A5) and let $y(\cdot)$ satisfying the perturbed inclusion (5) with vector perturbations $\varepsilon(\cdot) \in L_1(I)$, $\overline{\delta}(\cdot) \in L_\infty(I)$.*

Then, there exists a solution $x(\cdot)$ of (1) such that for all $t \in I$

$$\begin{aligned}&|y(t) - x(t)|_\infty \le \max\left\{ e^{\mu(t-t_0)}|y^0 - x^0|_\infty, \ 2e^{\mu_+(t-t_0)} \cdot \left(\|\overline{\delta}(\cdot)\|_{L_\infty} + \|\overline{\varepsilon}(\cdot)\|_{L_1}\right)\right\} \\ &\quad + |\mu| \cdot \int_{t_0}^{t} e^{\mu(t-s)}|\overline{\delta}(s)|_2 \, ds + \left(|\mu| \int_{t_0}^{t} e^{\mu(t-s)} \, ds + 1\right) \cdot \int_{t_0}^{t} |\overline{\varepsilon}(s)|_2 \, ds \qquad (55) \\ &\le \max\left\{e^{\mu(t-t_0)}|y^0 - x^0|_\infty, \ 2e^{\mu_+(t-t_0)} \cdot \left(\|\overline{\delta}(\cdot)\|_{L_\infty} + \|\overline{\varepsilon}(\cdot)\|_{L_1}\right)\right\} \\ &\quad + C_1(\mu)\|\overline{\delta}(\cdot)\|_{L_\infty} + C_2(\mu)\|\overline{\varepsilon}(\cdot)\|_{L_1} \qquad (56)\end{aligned}$$

with $C_1(\mu) = |\mu| \cdot \max_{t\in I} \int_{t_0}^{t} e^{\mu(t-s)}\, ds$, $C_2(\mu) = C_1(\mu) + 1$.

Proof The function $z(t) = y(t) - \int_{t_0}^{t} \overline{\varepsilon}(s)\, ds$ is AC with

$$\dot{z}(t) = \dot{y}(t) - \overline{\varepsilon}(t) \in F(t, y(t) + \overline{\delta}(t)), \quad z(t_0) = y(t_0) = y^0$$

satisfies the differential inclusion (35) with right-hand side $F(t, z(t)+\widetilde{\delta}(t))$ and a new inner vector perturbation $\widetilde{\delta}(t) = \overline{\delta}(t) + \int_{t_0}^{t} \overline{\varepsilon}(s)\, ds$. $\widetilde{\delta}(\cdot)$ is also an L_∞-function with

$$|\widetilde{\delta}(t)|_2 \le |\overline{\delta}(t)|_2 + \int_{t_0}^{t} |\overline{\varepsilon}(r)|_2\, dr, \quad \|\widetilde{\delta}(\cdot)\|_{L_\infty} \le \|\overline{\delta}(\cdot)\|_{L_\infty} + \|\overline{\varepsilon}(\cdot)\|_{L_1}.$$

Theorem 3.7 guarantees the existence of a solution $x(\cdot)$ of the original differential inclusion (1) with the estimate (36). Then,

$$\begin{aligned}
|y(t) - x(t)|_\infty &\le |y(t) - z(t)|_\infty + |z(t) - x(t)|_\infty \\
&\le \Big| \int_{t_0}^{t} \overline{\varepsilon}(s)\, ds \Big|_\infty + \max\left\{ e^{\mu(t-t_0)} |y^0 - x^0|_\infty,\ 2e^{\mu_+(t-t_0)} \cdot (\|\overline{\delta}(\cdot)\|_{L_\infty} + \|\overline{\varepsilon}(\cdot)\|_{L_1}) \right\} \\
&\quad + |\mu| \cdot \int_{t_0}^{t} e^{\mu(t-s)} \left(|\overline{\delta}(s)|_2 + \int_{t_0}^{s} |\overline{\varepsilon}(r)|_2\, dr \right) ds \\
&\le \max\left\{ e^{\mu(t-t_0)} |y^0 - x^0|_\infty,\ 2e^{\mu_+(t-t_0)} \cdot \left(\|\overline{\delta}(\cdot)\|_{L_\infty} + \|\overline{\varepsilon}(\cdot)\|_{L_1} \right) \right\} \\
&\quad + |\mu| \cdot \int_{t_0}^{t} e^{\mu(t-s)} |\overline{\delta}(s)|_2\, ds + \left(|\mu| \int_{t_0}^{t} e^{\mu(t-s)}\, ds + 1 \right) \cdot \int_{t_0}^{t} |\overline{\varepsilon}(s)|_2\, ds.
\end{aligned}$$

□

Note that $C_1(\mu)$ in Corollary 3.8 can be calculated as 0 for $\mu = 0$ and estimated by 1 for $\mu < 0$.

Remark 3.9 Note that the estimate (56) in the SOSL case proves the conjecture of [30, Remark 3.2] and provides order 1 with respect to the norm of the inner perturbation $\|\overline{\delta}(\cdot)\|_{L_\infty}$ and of the outer perturbation $\|\overline{\varepsilon}(\cdot)\|_{L_1}$. In the OSL case in [30, Theorem 3.1] the corresponding estimate

$$|y(t) - x(t)|_2 \le e^{\mu(t-t_0)} |y^0 - x^0|_2 + \int_{t_0}^{t} e^{\mu(t-s)} |\overline{\varepsilon}(s)|_2\, ds + C \sqrt{\int_{t_0}^{t} e^{2\mu(t-s)} |\overline{\delta}(s)|_2\, ds} \tag{57}$$

(with a constant C depending only on μ, C_B, C_F) is of order 1 in the outer perturbation but of order $\frac{1}{2}$ in the inner perturbation. Hence, the SOSL case provides a better order of the estimates which is also visible in the second motivation of Subsec. 2.2. Under the boundedness assumption (A1) not only the solutions of the perturbed system (5) by Lemma 3.4 are bounded but also the states y^j and velocities w^j of the Euler's method uniformly in the step-size h (see the reasoning in [30] for the OSL case). Then $\|\overline{\delta}(\cdot)\|_\infty = \mathcal{O}(h)$ holds for the SOSL and the OSL case, but only for the SOSL case the estimate for the Euler polygons in (13) would be $\mathcal{O}(h)$.

A direct proof of Corollary 3.8 following the lines of the proof in [30, Theorem 3.1] in the OSL case may improve the constants $C_1(\mu)$ and $C_2(\mu)$. On the other hand, the measurability of $F(\cdot, x+\overline{\delta}(\cdot))$ is a subtle issue (see for results in [21, Proposition 3.5] for continuous $\overline{\delta}(\cdot)$) and would need either an additional upper Scorza-Dragoni property [1, Sec. 5] or another existence result requiring only a strong measurable selection of $F(\cdot, x)$ plus assumptions on its boundedness (see [63, Chap. 3, Theorem 8.13 and following results]).

3.3 Stability and approximation results

From the presented results we can easily derive stability results for reachable sets with respect to the initial sets or the vector perturbations.

Definition 3.10 Let $X^0 \subset \mathbb{R}^n$ be a nonempty initial set. The *reachable set* $\mathcal{R}(t, t_0, X^0)$, sometimes denoted as $\mathcal{R}_F(t, t_0, X^0)$, of the differential inclusion (1) *at a given time* $t \in I$ with initial condition $x(t_0) \in X^0$ and right-hand side F is defined as the set of all end points of solutions at this time, i.e.

$$\mathcal{R}(t, t_0, X^0) = \left\{ x(t) \in \mathbb{R}^n \,:\, x(\cdot) \text{ is a feasible solution on } [t_0, t] \textit{of } (1) \text{ with } x(t_0) \in X^0 \right\}.$$

Corollary 3.11 *For reachable sets of* (1) *starting from two compact, nonempty initial sets* $X^0, Y^0 \subset \mathbb{R}^n$ *and* $F : I \times \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *satisfying the assumptions (A1)–(A5) we have the estimate*

$$\mathrm{d_H}(\mathcal{R}(t, t_0, X^0), \mathcal{R}(t, t_0, Y^0)) \le e^{\mu(t-t_0)}\, \mathrm{d_H}(X^0, Y^0) \quad \textit{for } t \in I$$

and weak (set-valued) exponential stability *holds if the SOSL constant* μ *is negative and* $t \to \infty$.

The same estimate is stated in [29, Theorem 3.2] for the OSL case. Note that the OSL and SOSL estimate do not differ, since the error terms with respect to the initial condition coincide.

Corollary 3.12 *Let* $X^0 \subset \mathbb{R}^n$ *be a compact, nonempty set and let the assumptions of Corollary 3.8 be satisfied. If* $\mathcal{R}_{\delta,\varepsilon}(t, t_0, X^0)$ *denotes the reachable set of the perturbed inclusion* (5) *at time* $t \in I$ *with initial set* X^0*, then*

$$\begin{aligned}\mathrm{d_H}(\mathcal{R}(t, t_0, X^0), \mathcal{R}_{\delta,\varepsilon}(t, t_0, X^0)) \le\ & 2e^{\mu_+(t-t_0)}\big(\|\overline{\delta}(\cdot)\|_{L_\infty} + \|\overline{\varepsilon}(\cdot)\|_{L_1}\big) \\ & + C_1(\mu)\|\overline{\delta}(\cdot)\|_{L_\infty} + C_2(\mu)\|\overline{\varepsilon}(\cdot)\|_{L_1}\end{aligned}$$

with $C_1(\mu), C_2(\mu)$ *as in Corollary 3.8.*

This result is a direct result of Corollary 3.8. The next approximation result is formulated in the spirit of the classical Filippov Theorem 1.1 and focuses on distances of graphs of the two right-hand sides.

Proposition 3.13 *Let* $F : I \times \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *satisfy the assumptions (A1)–(A5), and let* Graph $F(t, \cdot)$ *be measurable (w.r.t. to t).*

(i) Let $y : I \to \mathbb{R}^n$ be AC such that $y(t_0) = y^0$ and

$$\operatorname{dist}((y(t), \dot{y}(t)), \operatorname{Graph} F(t, \cdot)) \le \gamma(t) \quad \text{for a.e. } t \in I \tag{58}$$

with $\gamma(\cdot) \in L_\infty(I)$. Then there exists a solution $x(\cdot)$ of (1) satisfying

$$|y(t) - x(t)|_2 \le \max\left\{e^{\mu(t-t_0)}|y^0 - x^0|_\infty,\ 2e^{\mu_+(t-t_0)} \cdot \left(\|\gamma(\cdot)\|_{L_\infty} + \|\gamma(\cdot)\|_{L_1}\right)\right\} + C_1(\mu)\|\gamma(\cdot)\|_{L_\infty} + C_2(\mu)\|\gamma(\cdot)\|_{L_1} \quad \text{for } t \in I \tag{59}$$

with $C_1(\mu), C_2(\mu)$ as in Corollary 3.8 and μ the SOSL constant of F.

(ii) If $G : I \times \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ satisfies the assumptions (A1)–(A5) such that Graph $G(t, \cdot)$ *is measurable (w.r.t. to t) and*

$$\mathrm{d}(\operatorname{Graph} G(t, \cdot), \operatorname{Graph} F(t, \cdot)) \le \gamma(t) \quad \text{for a.e. } t \in I, \tag{60}$$

then the one-sided Hausdorff distance $\mathrm{d}(\mathcal{R}_G(t, t_0, y^0), \mathcal{R}_F(t, t_0, x^0))$ *can be estimated by the same right-hand side as in (59) for $t \in I$.*

Proof (i) Let $y(\cdot)$ be given. Then

$$\left(y(t), \dot{y}(t)\right) \in \operatorname{Graph} F(t, \cdot) + \gamma(t)\widetilde{B}_1(0) \quad \text{for a.e. } t \in I \tag{61}$$

with $\widetilde{B}_1(0)$ the closed unit ball in $\mathbb{R}^{2n}$. The map $H : I \rightrightarrows \mathbb{R}^{2n}$ with

$$H(t) = \left(\left(y(t), \dot{y}(t)\right) + \gamma(t)\widetilde{B}_1(0)\right) \cap \operatorname{Graph} F(t, \cdot)$$

is measurable by [3, Theorem 8.2.4] and has closed, nonempty images by construction. By [3, Theorem 8.1.4], it has a measurable selection $\left(z(t), w(t)\right) \in H(t)$ for $t \in I$ which satisfies

$$|y(t) - z(t)|_2 \le \gamma(t), \qquad |\dot{y}(t) - w(t)|_2 \le \gamma(t) \quad \text{for a.e. } t \in I. \tag{62}$$

Then for a.e. $t \in I$

$$\dot{y}(t) = w(t) + \left(\dot{y}(t) - w(t)\right) \in F(t, z(t)) + \left(\dot{y}(t) - w(t)\right) = F(t, y(t) + \overline{\delta}(t)) + \overline{\varepsilon}(t),$$

where $\overline{\delta}(\cdot) = z(\cdot) - y(\cdot) \in L_\infty(I), \overline{\varepsilon}(\cdot) = \dot{y}(\cdot) - w(\cdot) \in L_1(I)$. Applying Corollary 3.8 together with (62), there exists a solution $x(\cdot)$ of (1), such that (59) holds for the given function $y(\cdot)$.
(ii) For $y(t) \in \mathcal{R}_G(t, t_0, y^0)$ for $t \in I$, we have

$$\dot{y}(t) \in G(t, y(t)) \quad \text{for a.e. } t \in I \quad \text{and} \quad y(t_0) = y^0 \tag{63}$$

as well as $\big(y(t), \dot{y}(t)\big) \in \operatorname{Graph} G(t, \cdot)$ so that (61) also holds and the proof above continues as before by using $x(t) \in \mathcal{R}_F(t, t_0, x^0)$. □

Remark 3.14 It follows from the last proposition that the estimate (59) also holds for the (two-sided) Hausdorff distance between the reachable sets of the inclusions (1) and (63) under the assumption that (58) holds for the Hausdorff distance between the graphs of F and G. Then $\mu = \max\{\mu_F, \mu_G\}$ and the constants C_1 and C_2 are the maximal corresponding constants.

The last three claims can be considered as both approximation and stability results: if the interval I is finite, the estimates of the Hausdorff distances between the original and "perturbed" reachable sets in all three results are uniform in time. This also implies estimates of the distances between the corresponding *solution funnels*, i.e. the union of the graphs of all solutions. On an infinite time interval, the Hausdorff distances between the reachable sets stay small, if the SOSL constant is non-positive and the Hausdorff distance between the initial sets or the norms of the perturbations $\overline{\delta}(\cdot), \overline{\varepsilon}(\cdot)$ or of the bound $\gamma(\cdot)$ for the graphs are small.

For instance, let us consider the right-hand side of the differential inclusion $\dot{x}(t) \in -\operatorname{Sign}(x(t))$ which is replaced by a sequence of sigmoidal or saturation functions with growing Lipschitz constants. If the stability with respect to the initial value is studied with the help of the classical Filippov Theorem 1.1 for Lipschitz right-hand side, the estimate will explode for increasing time. Applying Theorem 3.7 for SOSL right-hand side, the estimate is uniformly bounded by the Hausdorff distance of the initial sets, since the SOSL constant for all functions of the sequence is 0. The approximation estimates in this case would not suffer on exploding Lipschitz constants (which appear in Example 2.5 if the Filippov theorem for Lipschitz right-hand side would be applied). In contrast to the exploding estimates obtained in the classical Filippov theorem, Proposition 3.13 gives good estimates, since the graphs of the sigmoidal or saturation functions tend to the graph of $-\operatorname{Sign}(\cdot)$ and all SOSL constants are non-positive.

4 Examples of differential inclusions with SOSL right-hand sides

In this section we present examples of dynamical systems with SOSL right-hand sides. In the case of Filippov's regularization of discontinuous ODEs with unique solution, Theorem 3.7 implies first order of convergence of the Euler approximants to this solution, as we have motivated in Subsection 2.2. The numerical experiments presented here confirm this order of convergence. The combination of the discrete and continuous Filippov-type approximation theorems was successfully applied in [17] to obtain error estimates of the Euler method for Lipschitz differential inclusions with state constraints and may also work in the case of SOSL mappings.

We now consider examples from differential equations based on applications.

Example 4.1 We consider the second-order differential equation on the time interval $I = [0, T]$ which was introduced by Flügge-Lotz/Klotter in [40, (1.3a) and (1.5d)]

$$\ddot{y}(\tau) + 2D\dot{y}(\tau) + \frac{b}{\omega^2}\,\mathrm{sign}\left(\rho_1 y(\tau) + \rho_2 \dot{y}(\tau)\right) + y(\tau) = 0 \tag{64}$$

for $b > 0$, $D > 0$, $\omega > 0$, initial value $y^0 = \binom{3}{4}$ and a motion under the influence of bang-bang controls. The example can also be found in [62, Beispiel 1.3] and with a slightly different factor for $y(\tau)$ in [51, Example 5.2]. As mentioned in [40] the control function $u(\tau) = \rho_1 y(\tau) + \rho_2 \dot{y}(\tau)$ anticipates the behavior of the solution component $y(\tau)$, acts as a feedback controller and precedes or follows it in time depending on $\rho = \frac{\rho_2}{\rho_1} > 0$ or $\rho < 0$. In [40, (1.5d)] the value $\frac{b}{\omega^2}$ is set to 1 and the damping factor D to 0.1.

The Filippov regularization is

$$\left.\begin{aligned} \dot{y}_1(t) &= y_2(t), \\ \dot{y}_2(t) &\in -2Dy_2(t) - y_1(t) - \tfrac{b}{\omega^2}\,\mathrm{Sign}(\rho_1 y_1(t) + \rho_2 y_2(t)). \end{aligned}\right\} \tag{65}$$

Let $\rho_1 = 0$ and $\rho_2 > 0$:
In this case the model is similar to [57, (2)] (with the right-hand side 0 in (64) replaced by a driving force $\varphi(\eta\tau)$ with a constant η and the equivalent simpler controller $\mathrm{sign}(\dot{y}(\tau))$) and comprises two important engineering equations. One model originates from an electric circuit with capacitor, coil, resistor (which damps the condesator charging) and rectifier eventually switching the sign of the condensator charging driven by an excitation with a periodic alternating (AC) voltage. The other model describes a mechanical system with a spring driven by forced vibrations with viscous damping as well as combined dry and Coulomb friction. In the latter D and $\mu = \frac{b}{\omega^2}$ are the Coulomb and sliding/dry friction coefficients, respectively.

This equation is also treated in several articles on discontinuous differential equations (e.g., in [62, Beispiel 0.1], [21, Example 13.3] and in [46, (1.4)]. In Fig. 5 (left) the (approximated) solution components $y_1(t)$ (blue) and $y_2(t)$ (red) are shown together with the black dashed switching curve $y_2 = 0$, where $\frac{b}{\omega^2} = 4$, $\eta = \pi$, $\varphi(s) = 2\cos(s)$, $T = 6$. Whenever the solution intersects with this curve, the solution component $y_2(t)$ has a corner due to $-\mathrm{Sign}(y_2)$ in $F(t, y)$.

The right-hand side is SOSL with constant $\mu_F = 1$. To see this, rewrite the right-hand side of (65) as $F(t, y) = Ay + b(t) - \mu\widetilde{S}(y)$ with $A = \begin{pmatrix} 0 & 1 \\ -1 & -2D \end{pmatrix}$, the vector $b(t) = \binom{0}{\varphi(\eta t)}$ and the set-valued map $\widetilde{S}(y) = \{0\} \times \mathrm{Sign}(y_2)$ for $y = (y_1, y_2) \in \mathbb{R}^2$. The affine part $Ay + b(t)$ is estimated by Lemma 2.1 with SOSL constant

$$\mu = \max_{i=1,2}\left(\max\{0, a_{ii}\} + \sum_{\substack{j=1,2 \\ j\neq i}} |a_{ij}|\right) = \max\left\{0 + |1|,\ \max\{0, -2D\} + |-1|\right\} = 1.$$

It is easy to prove that $\widetilde{S}(\cdot)$ is SOSL of constant 0 so that F is SOSL (even uniform SOSL) by Proposition 2.4(iv) with constant $\mu = 1$.

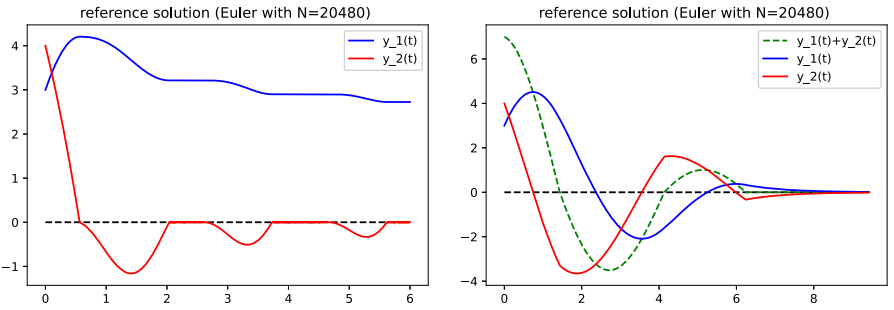

Fig. 5 Solution components and switching curves for Examples 4.1 and 4.2 (Color figure online)

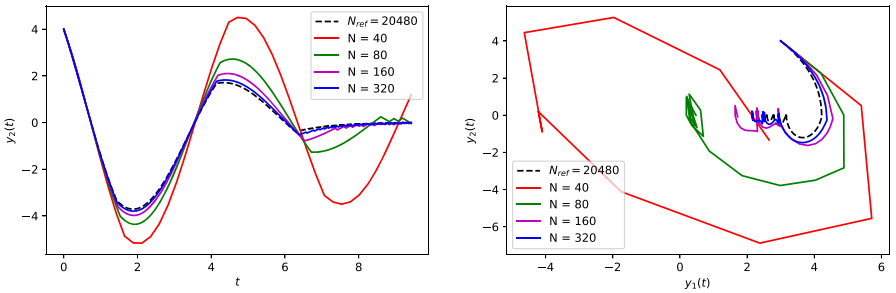

Fig. 6 Euler polygons for Example 4.2, 2nd component (left) and phase portrait (right) (Color figure online)

With Lemma 2.1 and the symmetrized matrix A_{sym} it is straight forward to prove that the right-hand side $F(t,\cdot)$ is even disspative (i.e. uniform OSL with constant $\mu_F = 0$).

Example 4.2 We continue Example 4.1 with the general model in [40, (1.4) and (1.5d)].

case $\rho_1 > 0$ and $\rho_2 > 0$:

The numerical test with the explicit Euler method on the time interval $I = [0, 3\pi]$ and $\rho = 1$ indicates graphically convergence order 1 with respect to the step size. In Fig. 5 (right) the (approximated) solution components $y_1(t)$ (blue) and $y_2(t)$ (red) are shown together with the green dashed function $y_1(t)+y_2(t)$, where $\frac{b}{\omega^2} = 1, \varphi(s) = 0$, $T = 3\pi$. Whenever the green function intersects with the black dashed axis $y_2 = 0$, the solution component $y_2(t)$ has a kink due to $-\operatorname{Sign}(y_1 + y_2)$ in $G(t, y)$. In Fig. 6 the second component of the Euler polygons for $N \in \{40, 80, 160, 320\}$ subintervals are shown together with the reference trajectory calculated with $N_{\text{ref}} = 20480$ (dashed black line). Note that there are corners at the phase portrait for the green trajectory around the points $(-2.5, 2.5)$, $(1.5, -1.5)$ and $(3.5, -4)$ reflecting discontinuities of the velocity when the trajectory crosses the line of discontinuity of the right-hand side $y_1 + y_2 = 0$. All solutions in the left plot show small zig-zagging behavior near the times t with $y_1(t) + y_2(t) = 0$.

In Table 1 the maximum errors (4th column) of the Euler iteration at each grid point are calculated for various step sizes h_k with respect to the reference solution. From this data of subsequent step sizes, the error at the k-th step is compared with the

Table 1 Convergence table for Example 4.2

k	N_k	h_k	Error(h_k)	Estim. order for h_k by k and 6
0	10	0.942478	51.2071	1.4401
1	20	0.471239	12.4716	1.3206
2	40	0.235619	3.3784	1.1797
3	80	0.117810	1.2293	1.0868
4	160	0.058905	0.5251	1.0166
5	320	0.029452	0.2260	0.8169
6	640	0.014726	0.1283	–

sixth step. The order is estimated and shows roughly $\mathcal{O}(h)$. A least squares analysis for matching the true errors with the unknowns C and p in Ch^p yields approximately $C = 36.502$, $p = 1.4350$, whereas $C = 14.397$ for fixed $p = 1$.

The speciality of this variant is the linear combination of components of the solution in the controller deciding on the sign switch in the controller $\operatorname{sign}(y(t) + \rho\dot{y}(t))$ so that it is not clear whether the right-hand side of the differential inclusion is SOSL or not. Nevertheless, the model fits very nicely to the choice of a basis in $\mathbb{R}^n$ for uniform SOSL set-valued maps in [52].

As suggested in [51] we introduce the transformed system with $z_1(t) = y_1(t)$, $z_2(t) = y_1(t) + \rho y_2(t)$ so that we can express $y_2(t) = \frac{1}{\rho}(z_2(t) - y_1(t)))$. Thus, we consider the equivalent differential inclusion $z'(t) \in G(t, z(t))$ with

$$\begin{aligned}
\dot{z}_1(t) &= \frac{1}{\rho}(z_2(t) - z_1(t)),\\
\dot{z}_2(t) &\in -2\widetilde{D}z_2(t) - (\rho + 2\widetilde{D})z_1(t) - \frac{b}{\omega^2}\operatorname{Sign}(z_2(t)),
\end{aligned}$$

where $\widetilde{D} = (D - \frac{1}{2\rho})$. We prove the strengthened OSL condition and consider $z = (z_1, z_2)$, $\widetilde{z} = (\widetilde{z}_1, \widetilde{z}_2) \in \mathbb{R}^2$, $v = (v_1, v_2) \in G(t, z)$, $s(z_2) \in \operatorname{Sign}(z_2)$. $G(t, z)$ is expressed as $Bz - \mu\widetilde{S}(z)$ with the matrix $B = \begin{pmatrix} -\frac{1}{\rho} & \frac{1}{\rho} \\ -(\rho + 2\widetilde{D}) & -2\widetilde{D} \end{pmatrix}$ and μ, $\widetilde{S}(\cdot)$ as in Example 4.1. The linear part $z \mapsto Bz$ is SOSL by Lemma 2.1 with constant

$$\begin{aligned}
\mu_B &= \max_{i=1,2}\left(\max\{0, a_{ii}\} + \sum_{\substack{j=1,\dots,2\\ j\neq i}} |a_{ij}|\right)\\
&= \max\left\{\max\{0, -\frac{1}{\rho}\} + |\frac{1}{\rho}|, \max\{0, -2\widetilde{D}\} + |-(\rho + 2\widetilde{D})|\right\}\\
&= \max\left\{\frac{1}{\rho}, \max\{0, -2\widetilde{D}\} + |\rho + 2\widetilde{D}|\right\}.
\end{aligned}$$

We can argue with Proposition 2.4(iv) as in Example 4.1 to see that the transformed differential inclusion with right-hand side $G(t,z)$ is SOSL (even uniformly) with constant $\mu_G = \mu_B$.

We discuss analytically another higher-dimensional example with three coupled strings with six states.

Example 4.3 ([61], [51, Example 5.3], [55, (16)]) Consider the system of three coupled springs with dry friction of second order

$$y_i'(t) = y_{3+i}(t) \quad (i = 1,2,3), \tag{66}$$
$$y_4'(t) \in -y_1(t) + (y_2(t) - y_1(t)) - y_4(t) - 0.3\,\mathrm{Sign}(y_4(t)), \tag{67}$$
$$y_5'(t) \in -(y_2(t) - y_1(t)) + (y_3(t) - y_2(t)) - y_5(t) - 0.3\,\mathrm{Sign}(y_5(t)), \tag{68}$$
$$y_6'(t) \in -(y_3(t) - y_2(t)) - y_6(t) - 0.3\,\mathrm{Sign}(y_6(t)) + 10\cos(\pi t) \tag{69}$$

on $t \in I = [0,6]$ with initial condition $y(0) = (-1,1,-1,-1,1,1)^\top$. With the matrix $A = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -2 & 1 & 0 & -1 & 0 & 0 \\ 1 & -2 & 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 & 0 & -1 \end{pmatrix}$, the set-valued map for the inhomogenity $B(x) = \sum_{i=1}^6 B_i(x_i)e^i$ and the vector function $c(t) = \sum_{i=1}^6 c_i(t)e^i$ with the notation (16) used in Proposition 2.4 and $B_i(x_i) \subset \mathbb{R}$, $c_i(t) \in \mathbb{R}$ with

$$B_i(x_i) = \begin{cases} \{0\} & (i = 1,2,3), \\ -0.3\,\mathrm{Sign}(x_i) & (i = 4,5,6), \end{cases} \qquad c_i(t) = \begin{cases} 0 & (i = 1,\dots,5), \\ 10\cos(\pi t) & (i = 6). \end{cases}$$

Then $F(t,x) = Ax + B(x) + c(t)$.

The diagonal elements of A are either 0 or -1 so that the maximal sum of absolute values of off-diagonal elements is 4 (attained in the fifth row). Hence, the function $(t,x) \mapsto Ax + c(t)$ is SOSL with constant $\mu_A = 4$ by Lemma 2.1. The set-valued map B is strengthened uniform OSL with constant 0 by Proposition 2.4(v). By Proposition 2.4(iv) the set-valued map F is strengthened uniform OSL with constant $\mu_F = \mu_A = 4$.

Example 4.4 Inner set-valued perturbations of the differential inclusion (66)–(69) in Example 4.3 involving $\delta_i > 0$, $i = 4,5,6$, yield the system

$$y_i'(t) = y_{3+i}(t) \quad (i = 1,2,3), \tag{70}$$
$$y_4'(t) \in -y_1(t) + (y_2(t) - y_1(t)) - y_4(t) - 0.3\,\mathrm{Sign}(y_4(t) + \delta_4[-1,1]), \tag{71}$$
$$y_5'(t) \in -(y_2(t) - y_1(t)) + (y_3(t) - y_2(t)) - y_5(t) - 0.3\,\mathrm{Sign}(y_5(t) + \delta_5[-1,1]), \tag{72}$$
$$y_6'(t) \in -(y_3(t) - y_2(t)) - y_6(t) - 0.3\,\mathrm{Sign}(y_6(t) + \delta_6[-1,1]) + 10\cos(\pi t) \tag{73}$$

which is SOSL with constant $\mu = 4$ due to Proposition 2.4(ii) and Example 4.3, but not strengthened uniform OSL.

The new differential inclusion can be seen in the light of a computer implementation of the system of Example 4.3. In practice an algorithm implementing a discrete set-valued Euler's method will not test whether a floating point number y_i, $i = 4, 5, 6$, is exactly zero or not to evaluate $-\operatorname{Sign}(y_i)$. Due to rounding errors one would choose an implementation which returns $-\operatorname{Sign}(0)$ for the argument y_i if the absolute value of y_i is less or equal δ_i close to the floating point precision multiplied by a factor depending on an upper bound of $|y_i|$, i.e. $|y_i| \leq \delta_i$. This is exactly the case when $y_i \in \delta_i[-1, 1]$ so that $-\operatorname{Sign}(y_i + \delta_i[-1, 1]) = [-1, 1]$. Hence, inner set-valued perturbations can incorporate strategies for taking into account rounding errors in floating point arithmetics.

Further examples in the analysis of block designs or cascading state observers [48, (4.2) and below (A.3)] also lead to SOSL systems.

Conclusions

Well-posedness and regularity of solutions in perturbed problems is a topic studied persistently by A. Dontchev. In the paper [35] he and his co-author proved the order of convergence 1 for the set-valued Euler's method in the Lipschitz case and partially repeated the proof of the celebrated Filippov theorem (Theorem 1.1) for convex and compact-valued right-hand sides, since they were not aware of this theorem. The authors of this paper believe that continuing this tradition is an appropriate way to honor the memory of Asen L. Dontchev.

While the form of the perturbed problem in Theorem 3.7 and Corollary 3.8 is different from that in the original theorem of Filippov, the formulation of Proposition 3.13 is more in the spirit of this theorem.

We are currently preparing a follow-up paper that will focus on discrete approximations of differential inclusions for SOSL maps which will benefit from the available Filippov approximation theorems in continuous time (presented here) and in discrete time [9].

The authors would like to thank the reviewers for their particularly careful reading, the valuable suggestions and the encouragement for a better presentation of the material. Their remarks helped us to improve substantially the paper.

Funding Open Access funding enabled and organized by Projekt DEAL. The authors are grateful for the partial support of Tel Aviv University, Mathematical Institute at Tel Aviv "MINT", University of Bayreuth, and Bavarian Research and Innovation Agency "BayFor" which enabled their participation in two bilateral workshops in Tel Aviv and Bayreuth in 2019 and in which the basis of this paper was laid.

Data availability The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Competing Interests The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Appell, J., De Pascale, E., Thái, Nguyên Hôàng., Zabreĭko, P.P.: Multi-valued superpositions. Diss. Math. (Rozprawy Mat.) **345** (1995)
2. Aubin, J.-P., Cellina, A.: Differential Inclusions. Vol. 264. Grundlehren der mathematischen Wissenschaften, Springer, Berlin, pp. xiii+342 (1984)
3. Aubin, J.-P., Frankowska, H.: Set-Valued Analysis. Vol. 2. Systems & Control: Foundations & Applications. Birkhauser Boston Inc., Boston, pp. xx+461 (1990)
4. Auzinger, W., Frank, R., Macsek, F.: Asymptotic error expansions for stiff equations: the implicit Euler scheme. SIAM J. Numer. Anal **27**(1), 67–104 (1990)
5. Bacciotti, A.: On several notions of generalized solutions for discontinuous differential equations and their relationships. Research Report 19. Dipartimento di Matematica del Politecnico di Torino, Torino (2003). https://citeseerx.ist.psu.edu/doc_view/pid/c3c4eadd04721dbbd301b96edac713d5301d9d6b
6. Baier, R., Chahma, I.A., Lempio, F.: Stability and convergence of Euler's method for state-constrained differential inclusions. SIAM J. Optim. **18**(3) (2007). D. Dentcheva, J. Revalski (eds.), special issue on "Variational Analysis and Optimization", 1004-1026 (electronic)
7. Baier, R., Farkhi, E.: Regularity of set-valued maps and their selections through set differences. Part 1: Lipschitz continuity. Serdica Math. J. **39**, 3–4 (2013). Special issue dedicated to the 65th anniversary of Professor Asen L. Dontchev and to the 60th anniversary of Professor Vladimir M. Veliov, pp. 365–390
8. Baier, R., Farkhi, E.: Regularity of set-valued maps and their selections through set differences. Part 2: One-sided Lipschitz properties. Serdica Math. J. **39**, 3–4 (2013). Special issue dedicated to the 65th anniversary of Professor Asen L. Dontchev and to the 60th anniversary of Professor Vladimir M. Veliov, pp. 391–422
9. Baier, R., Farkhi, E.: Discrete Filippov-type stability for one-sided Lipschitzian difference inclusions. In: Feichtinger, G., Kovacevic, R., Tragler, G. (eds.) Control Systems and Mathematical Methods in Economics. Essays in Honor of Vladimir M. Veliov, vol. 687. Lecture Notes in Economy and Math. Systems. Dedicated to Vladimir Veliov's 65-th birthday. Springer, Cham, pp. 27–55 (2018)
10. Beyn, W.J., Rieger, J.: Numerical fixed grid methods for differential inclusions. Computing **81**(1), 91–106 (2007)
11. Blanchini, F., Miani, S.: Set-Theoretic Methods in Control. Systems & Control: Foundations & Applications. Birkhauser Boston Inc., Boston, pp. xvi+481 (2008)
12. Bressan, A.: Singularities of stabilizing feedbacks. In: Vol. 56(4). Control Theory and its Applications (Grado, 1998). 1998, 87–104 (2001)
13. Bruckner, A. M.: Differentiation of Real Functions. Vol. 659. Lecture Notes in Math. Berlin: Springer, pp. x+247 (1978)
14. Calin, O.: Deep Learning Architectures. A Mathematical Approach. Springer Series in the Data Sciences. Springer, Cham, pp. xxx+760 (2020)
15. Cannarsa, P., Da Prato, G., Frankowska, H.: Invariance for quasi-dissipative systems in Banach spaces. J. Math. Anal. Appl **457**(2), 1173–1187 (2018)
16. Castaing, C.: Sur les multi-applications mesurables. Rev. Fr. Inform. Rech. Oper. **1**(1), 91–126 (1967)

17. Chahma, I.A.: Set-valued discrete approximation of state-constrained differential inclusions. Bayreuth. Math. Schr. **67**, 3–162 (2003)
18. Cohn, D.L.: Measure Theory. Second ed. Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks]. Birkhäuser/Springer, New York, pp. xxi +457 (2013)
19. Colombo, G.: Approximate and relaxed solutions of differential inclusions. Rend. Sem. Mat. Univ. Padova **81**, 229–238 (1989)
20. Deimling, K.: Nonlinear Functional Analysis. Springer, Berlin, pp. xiv+450 (1985)
21. Deimling, K.: Multivalued Differential Equations. Vol. 1. de Gruyter Series in Nonlinear Analysis and Applications. Walter de Gruyter, Berlin (1992)
22. Dekker, K., Verwer, J.G.: Stability of Runge–Kutta Methods for Stiff Nonlinear Differential Equations. Vol. 2. CWI Monographs. North-Holland, Amsterdam, pp. ix+307 (1984)
23. Donchev, T.D.: Functional-differential inclusion with monotone right-hand side. Nonlinear Anal. **16**(6), 533–542 (1991)
24. Donchev, T. D.: Qualitative properties of a class differential inclusions. Glas. Mat. Ser. III **31**(51)2, 269–276 (1996)
25. Donchev, T.D.: Properties of one-sided Lipschitz multivalued maps. Nonlinear Anal. **49**(1), 13–20 (2002)
26. Donchev, T.D.: Properties of the reachable set of control systems. Syst. Control Lett. **46**(5), 379–386 (2002)
27. Donchev, T. D.: One sided Lipschitz multifunctions and applications. In: Optimal Control, Stabilization and Nonsmooth Analysis. Vol. 301. Lecture Notes in Control and Inform. Sci. Springer, Berlin, pp. 333–341 (2004)
28. Donchev, T.D., Dontchev, A.L.: Singular perturbations in infinite-dimensional control systems. SIAM J. Control Optim. **42**(5), 1795–1812 (2003)
29. Donchev, T.D., Farkhi, E.: Stability and Euler approximation of one-sided Lipschitz differential inclusions. SIAM J. Control Optim. **36**(2), 780–796 (1998)
30. Donchev, T.D., Farkhi, E.: Approximations of one-sided Lipschitz differential inclusions with discontinuous right-hand sides. In: Calculus of Variations and Differential Equations (Haifa, 1998). Vol. 410. Chapman & Hall/CRC Res. Notes Math. Chapman & Hall/CRC, Boca Raton, FL, pp. 101–118 (2000) isbn: 978-1-584-88024-0
31. Donchev, T.D., Farkhi, E.: On the theorem of Filippov-Pliś. and some applications. Control Cybern. **38**(4A), 1251–1271 (2009)
32. Donchev, T.D., Farkhi, E., Reich, S.: Fixed set iterations for relaxed Lipschitz multimaps. Nonlinear Anal. **53**(7–8), 997–1015 (2003)
33. Donchev, T.D., Farkhi, E., Reich, S.: Discrete approximations and fixed set iterations in Banach spaces. SIAM J. Optim. **18**(3), 895–906 (2007)
34. Dontchev, A.L., Donchev, T.D., Slavov, Ľ: A Tikhonov-type theorem for singularly perturbed differential inclusions. Nonlinear Anal. **26**(9), 1547–1554 (1996)
35. Dontchev, A.L., Farkhi, E.: Error estimates for discretized differential inclusions. Computing **41**(4), 349–358 (1989)
36. Dontchev, A.L., Lempio, F.: Difference methods for differential inclusions: a survey. SIAM Rev. **34**(2), 263–294 (1992)
37. Dunn, J.C.: Iterative construction of fixed points for multivalued operators of the monotone type. J. Funct. Anal. **27**(1), 38–50 (1978)
38. Filippov, A.F.: Classical solutions of differential equations with multi-valued right-hand side. SIAM J. Control **5**, 609–621 (1967)
39. Filippov, A.F.: Differential Equations with Discontinuous Righthand Sides. Vol. 18. Mathematics and its Applications (Soviet Series). English translation of Russian original "Differentsialnye uravneniya s razryvnoi pravoi chastyu", Nauka, Moscow, 1985. Dordrecht: Kluwer Academic Publishers Group, pp. x+304 (1988)
40. Flügge-Lotz, I., Klotter, K.: Über Bewegungen eines Schwingers unter dem Einfluss von Schwarz-Weiss-Regelungen. I. Bewegungen eines Schwingers von einem Freiheitsgrad; Regelung mit Stellungszuordnung ohne Schaltverschiebungen [On movements of an oscillator under the influence of black and white controls. I. Movements of a vibrator of one degree of freedom; control with position assignment without switching displacements]. In: Z. Angew. Math. Mech. 28, pp. 317–337 (1948)
41. Folland, G.B.: Real Analysis. Modern Techniques and their Applications. Pure and Applied Mathematics (New York), 2nd edn. First edition published in 1984. Wiley, New York, pp. xvi+386 (1999)

42. Frankowska, H., Rampazzo, F.: Filippov's and Filippov-Ważewski's theorems on closed domains. J. Differ. Equ. **161**(2), 449–478 (2000)
43. Hairer, E., Wanner, G.: Solving Ordinary Differential Equations. II Stiff and Differential-Algebraic Problems. Second ed. Vol. 14. Springer Series in Computational Mathematics. Springer, Berlin, pp. xvi+614 (1996)
44. Hájek, O.: Discontinuous differential equations. I. J. Differ. Equ. **32**(2), 149–170 (1979)
45. Kastner-Maresch, A.E.: Implicit Runge–Kutta methods for differential inclusions. Numer. Funct. Anal. Optim. **11**(9-10) (1990), 937–958 (1991)
46. Kastner-Maresch, A.E.: The implicit midpoint rule applied to discontinuous differential equations. Computing **49**(1), 45–62 (1992)
47. Krasnova, S.A., Mysik, N.S.: Cascade synthesis of a state observer with nonlinear correcting influences. Autom. Remote Control **75**(2), 263–280 (2014)
48. Krasnova, S.A., Utkin, V.A., Utkin, A.V.: A block approach to the analysis and design of invariant nonlinear tracking systems. Autom. Remote Control **78**(12), 2120–2140 (2017)
49. Lempio, F.: Difference methods for differential inclusions. In: Modern Methods of Optimization. Proceedings of a Summer School at the *Schlos* Thurnau of the University of Bayreuth (Germany), FRG, October 1–6, 1990. Vol. 378. Lecture Notes in Econom. and Math. Systems. Springer, Berlin, pp. 236–273 (1992)
50. Lempio, F.: Modified Euler methods for differential inclusions. In: Set-Valued Analysis and Differential Inclusions. A Collection of Papers resulting from a Workshop held in Pamporovo, Bulgaria, September 17–21, 1990. Vol. 16. Progr. Systems Control Theory. Birkhauser, Boston, pp. 131–148. (1993) isbn: 0-8176-3733-8
51. Lempio, F.: Euler's method revisited. Proc. Steklov Inst. Math. **211**, 429–449 (1995)
52. Lempio, F., Silin, D.B.: Differential inclusions with strongly one-sided-Lipschitz right-hand sides. Differ. Equ. **32**(11), 1485–1491 (1997)
53. Lempio, F., Veliov, V.M.: Discrete approximations of differential inclusions. Bayreuth. Math. Schr. **54**, 149–232 (1998)
54. Łojasiewicz Jr., S.: Some theorems of Scorza-Dragoni type for multifunctions with application to the problem of existence of solutions for differential multivalued equations. In: Mathematical Control Theory. Vol. 14 (1). Banach Center Publ. Warsaw: PWN pp. 625–643 (1985)
55. Marszal, M., Stefański, A.: Synchronization properties in coupled dry friction oscillators. In: Nonlinear Dynamical Systems with Self-Excited and Hidden Attractors. Vol. 133. Stud. Syst. Decis. Control. Springer, Cham, pp. 87–113 (2018)
56. Pliś, A.: On trajectories of orientor fields. Bull. Acad. Polon. Sci. Ser. Sci. Math. Astronom. Phys. **13**, 571–573 (1965)
57. Reissig, R.: Erzwungene Schwingungen mit zäher Dämpfung und starker Gleitreibung [Forced oscillations with viscous damping and strong sliding friction]. Math. Nachr. **11**, 231–238 (1954)
58. Rieger, J.: A proof of the relaxation theorem for differential inclusions based on Euler approximations. Numer. Funct. Anal. Optim. **33**(10), 1244–1249 (2012)
59. Serrin, J., Varberg, D.E.: A general chain rule for derivatives and the change of variables formula for the Lebesgue integral. Am. Math. Mon. **76**, 514–520 (1969)
60. Stein, E.M., Shakarchi, R.: Real Analysis. Measure Theory, Integration, and Hilbert Spaces. Vol. 3. Princeton Lectures in Analysis. Princeton University Press, Princeton, pp. xx+402 (2005)
61. Stewart, D.E.: High-Accuracy Numerical Methods for Ordinary Differential Equations with Discontinuous Right-Hand Side. PhD thesis. The University of Queensland (Australia), Brisbane (1990)
62. Taubert, K.: Differenzverfahren für Schwingungen mit trockener und zäher Reibung und für Reglungsysteme [Difference methods for vibrations with dry and viscous friction and for control systems]. In: Numer. Math. **26**(4), 379–395
63. Tolstonogov, A.: Differential Inclusions in a Banach Space. Vol. 524. Mathematics and its Applications. Translated from the 1986 Russian original and revised by the author. Kluwer Academic Publishers, Dordrecht, pp. xvi+302 (2000)
64. Veliov, V.M.: Differential inclusions with stable subinclusions. Nonlinear Anal. **23**(8), 1027–1038 (1994)
65. Vinter, R.: Optimal Control. Systems & Control: Foundations & Applications. Birkhauser, Boston, pp. xviii+507 (2000)
66. Wolenski, P.R.: The exponential formula for the reachable set of a Lipschitz differential inclusion. SIAM J. Control Optim. **28**(5), 1148–1161 (1990)

67. Zarantonello, E.H.: Dense single-valuedness of monotone operators. Isr. J. Math. **15**, 158–166 (1973)