LINEAR PROGRAMMING BASED LYAPUNOV FUNCTION COMPUTATION FOR DIFFERENTIAL INCLUSIONS

Robert Baier and Lars Grüne

Chair of Applied Mathematics
University of Bayreuth
95440 Bayreuth, Germany

Sigurður Freyr Hafstein

School of Science and Engineering
Reykjavík University
Menntavegur 1, 101 Reykjavík, Iceland

(Communicated by the associate editor name)

Abstract. We present a numerical algorithm for computing Lyapunov functions for a class of strongly asymptotically stable nonlinear differential inclusions which includes spatially switched systems and systems with uncertain parameters. The method relies on techniques from nonsmooth analysis and linear programming and constructs a piecewise affine Lyapunov function. We provide necessary background material from nonsmooth analysis and a thorough analysis of the method which in particular shows that whenever a Lyapunov function exists then the algorithm is in principle able to compute it. Two numerical examples illustrate our method.

1. **Introduction.** Differential inclusions are a versatile tool to model various dynamical phenomena. They can be used, e.g., in order to describe systems under parametric uncertainties which are ubiquitous in many applications. Via the Filippov regularization they also provide a mathematially rigorous way to handle systems with discontinuities, like spatially switched systems. When analyzing the dynamical behavior of the solutions of differential inclusions, the determination of the stability properties of an equilibrium and — in case of asymptotic stability — its domain of attraction is one of the fundamental problems. In this paper we will investigate this problem for the case of robust or strong asymptotic stability for nonlinear differential inclusions, i.e., when all solutions of the inclusion are asymptotically stable.

Lyapunov functions play an important role in this analysis since their knowledge allows to verify asymptotic stability of an equilibrium and at the same time to estimate its domain of attraction. However, Lyapunov functions are often difficult if not impossible to obtain analytically. Hence, numerical methods may be the only feasible way for computing such functions.

Numerical computations of Lyapunov functions have been extensively studied in recent years. In the literature, two main approaches can be identified. The first approach uses the fact that Lyapunov functions can be characterized by partial differential equations which can then be solved numerically. For nonlinear control systems, which can be seen as a parametrized version of the differential inclusions considered in this paper, such a numerical approach has been presented in [2] using the Zubov equation, a particular Hamilton-Jacobi-Bellman equation. However, this method computes a numerical approximation of a Lyapunov function rather than a Lyapunov function itself. A related method for numerically computing true Lyapunov functions — even smooth ones — has been presented in detail in [8]. However, this method is designed for differential equations and does not directly extend to differential inclusions.

The second main approach uses numerical optimization techniques for computing Lyapunov functions. In [13], a convex optimization approach using linear or quadratic programming has been presented which is, however, only applicable to differential equations. In [15] the authors develop a linear programming method which is based on piecewise linear approximations of the original nonlinear vector field. This approach extends to nonlinear inclusions provided they are generated by piecewise linear and sector bounded uncertainties. Finally, LMI (linear matrix inequalities) optimization techniques have been succesfully applied to the problem, see, e.g., [3] and the references therein, however, this approach is restricted to differential inclusions with polynomial right hand sides.

The contribution of the present paper is the extension of the linear programming based algorithm for computing Lyapunov functions first presented in [19] for ordinary differential equations and further developed in [11] for systems with switching in time. We extend the method to nonlinear differential inclusions defined by polytopes of general nonlinear vector fields on different — overlapping or non-overlapping — domains. This class includes, e.g., Filippov regularizations of discontinuous nonlinear systems like spatially switched systems as well as nonlinear differential equations with polytopic parametric uncertainty. Like in [8] we directly work with the nonlinear vector fields (i.e., we do not use piecewise affine or polynomial approximations) and by means of a thorough analysis of the discretization error we can guarantee that the resulting numerically computed function is a true Lyapunov function of the system, except possibly for a small neighborhood of the origin. Apart from the fact that we use a different type of discretization, the central difference to [8] is that instead of solving the linear partial differential equation $\langle \nabla V(x), f(x) \rangle = -\alpha(\|x\|)$ for one vector field $f(x)$, here a feasible solution to the linear partial differential inequality $\langle \nabla V(x), f_\mu(x) \rangle \leq -\alpha(\|x\|)$ is found for all vectors $f_\mu(x)$ defining our differential inclusion. For a fixed V this inequality may be fulfilled for several different functions f_μ, whereas the equation is in general not. Proceeding this way, we are in particular able to prove that for sufficiently fine and regular discretization our algorithm is always able to compute a Lyapunov function if one exists.

The Lyapunov functions computed by our algorithm are piecewise affine and thus nonsmooth, hence we exploit methods from nonsmooth analysis. Since the results needed for a rigorous treatment of such functions are scattered in different areas in the literature, a second contribution of our paper is a rigorous and self contained presentation of the necessary background results for nonsmooth Lyapunov functions.

The paper is organized as follows. After introducing the setting and several definitions in Section 2, in Section 3 we provide the necessary background results from nonsmooth analysis and precisely define the concept of nonsmooth Lyapunov functions needed for our method. The algorithm along with its detailed analysis can be found in Section 4. Finally, we illustrate the algorithm by two numerical examples in Section 5.

2. **Notation and preliminaries.** In order to introduce the class of differential inclusions to be investigated in this paper, we consider a compact set $G \subset \mathbb{R}^n$ which is divided into M closed subregions $\mathcal{G} = \{G_\mu \,|\, \mu = 1, \ldots, M\}$ with $\bigcup_{\mu=1,\ldots,M} G_\mu = G$. For each $x \in G$ we define the *active index set* $I_\mathcal{G}(x) := \{\mu \in \{1, \ldots, M\} \,|\, x \in G_\mu\}$.

On each subregion G_μ we consider a Lipschitz continuous vector field $f_\mu : G_\mu \to \mathbb{R}^n$. Our differential inclusion on G is then given by

$$\dot{x} \in F(x) := \text{co}\,\{f_\mu(x) \,|\, \mu \in I_\mathcal{G}(x)\}, \tag{1}$$

where "co" denotes the convex hull. A solution of (1) is an absolutely continuous function $x : I \to G$ satisfying $\dot{x}(t) \in F(x(t))$ for almost all $t \in I$, where I is the maximal existence interval. This interval I is of the form $I = [0, T]$ or $I = [0, \infty)$. Since G is compact and $x(t)$ is continuous in t, the maximal existence interval is of the form $I = [0, \infty)$ if and only if $x(t) \in G$ for all $t \geq 0$. Note that we do not impose any invariance properties of G.

To guarantee the existence of a solution of the differential inclusion (1), upper semicontinuity of the right-hand side is an essential assumption, see [7, § 2.7].

Definition 2.1. *A set valued map $F : G \Rightarrow \mathbb{R}^n$ is called upper semicontinuous if for any $x \in G$ and any $\epsilon > 0$ there exists $\delta > 0$ such that*

$$x' \in B_\delta(x) \cap G \quad \text{implies} \quad F(x') \subseteq F(x) + B_\epsilon(0).$$

The following lemma shows that F from (1) is upper semicontinuous. For pairwise disjoint subregions the proof follows from [7, Lemma 3 in § 2.6]. Here we provide an alternative proof idea which also covers overlapping regions.

Lemma 2.2. *The set valued map $F(x) = \text{co}\,\{f_\mu(x) | \mu \in I_\mathcal{G}(x)\}$ from (1) is upper semicontinuous in the sense of Definition 2.1.*

Proof. Pick $x \in G$. We have to show that for every $y' \in F(x')$ there exists $y \in F(x)$ such that $\|y' - y\| < \varepsilon$.

To this end let $A = \bigcup_{\mu \notin I_\mathcal{G}(x)} G_\mu$. This set is compact as a finite union of compact sets. Since $x \notin A$, x has a positive distance from A, i.e., there exists an open ball $B_{\delta_1}(x)$ with $B_{\delta_1}(x) \cap A = \emptyset$ and by definition of A we get $I_\mathcal{G}(x') \subseteq I_\mathcal{G}(x)$ for all $x' \in B_{\delta_1}(x)$.

Since each f_μ is continuous, for any $\epsilon > 0$ we find a positive $\delta \leq \delta_1$ such that $\|f_\mu(x') - f_\mu(x)\| < \epsilon$ holds for all $\mu \in I_\mathcal{G}(x') \subseteq I_\mathcal{G}(x)$ and all $x' \in B_\delta(x)$. Now each $y' \in F(x')$ can be written as a convex combination $y' = \sum_{\mu \in I_\mathcal{G}(x')} \lambda_\mu f_\mu(x')$. Since $I_\mathcal{G}(x') \subseteq I_\mathcal{G}(x)$ we can define $y = \sum_{\mu \in I_\mathcal{G}(x')} \lambda_\mu f_\mu(x) \in F(x)$ in order to obtain

$$\|y' - y\| = \left\| \sum_{\mu \in I_\mathcal{G}(x')} \lambda_\mu (f_\mu(x') - f_\mu(x)) \right\| \leq \underbrace{\sum_{\mu \in I_\mathcal{G}(x')} \lambda_\mu}_{=1} \underbrace{\|f_\mu(x') - f_\mu(x)\|}_{<\epsilon} < \epsilon.$$

This shows the assertion. $\qquad\square$

Note that the differential inclusion (1) is upper semicontinuous due to Lemma 2.2. However, weaker conditions are available in the literature, e.g. almost upper semi-continuity for one-sided Lipschitz differential inclusions in [6]. Two important special cases of (1) are outlined in the following examples.

Example 2.3 (**switched ordinary differential equations**). *We consider a partition of G into pairwise disjoint but not necessarily closed sets H_μ and a piecewise defined ordinary differential equations of the form*

$$\dot{x}(t) = f_\mu(x(t)), \quad x(t) \in H_\mu \tag{2}$$

in which $f_\mu : H_\mu \to \mathbb{R}^n$ is continuous and can be continuously extended to the closures $\operatorname{cl} H_\mu$.

If the ordinary differential equation $\dot{x}(t) = f(x(t))$ with $f : G \to \mathbb{R}^n$ defined by $f(x) := f_\mu(x)$ for $x \in G_\mu$ is discontinuous, then in order to obtain well defined solutions the concept of Filippov solutions, cf. [7, § 2.7], is often used. To this end (2) is replaced by its Filippov regularization, i.e. by the differential inclusion

$$\dot{x}(t) \in F(x(t)) = \bigcap_{\delta > 0} \bigcap_{\mu(N)=0} \overline{\operatorname{co}}\{f((B_\delta(x(t)) \cap G) \setminus N)\} \tag{3}$$

where μ is the Lebesgue measure, $N \subset \mathbb{R}^n$ an arbitrary set of measure zero and $\overline{\operatorname{co}}$ denotes the closure of the convex hull. A straightforward computation shows that if the number of the sets H_μ is finite and each H_μ satisfies $\operatorname{cl} H_\mu = \operatorname{cl} \operatorname{int} H_\mu$, then the inclusion (3) coincides with (1) if we define $G_\mu := \operatorname{cl} H_\mu$ and extend each f_μ continuously to G_μ. This fact is collected, e.g. in [7, § 2.7] and [23].

An important subclass of switched systems are piecewise affine systems in which each f_μ in (2) is given by an affine map, i.e.,

$$f_\mu(x) = A_\mu x + b_\mu,$$

see, e.g., [14, 18].

Example 2.4 (**polytopic inclusions**). *Consider a differential inclusion $\dot{x}(t) \in F(x(t))$ in which $F(x) \subset \mathbb{R}^n$ is a closed polytope $F(x) = \operatorname{co}\{f_\mu(x) \,|\, \mu = 1, \ldots, M\}$ with a finite number of vertices $f_\mu(x)$ for each $x \in G$. If the vertex maps $f_\mu : G \to \mathbb{R}^n$ are Lipschitz continuous, then the resulting inclusion*

$$\dot{x}(t) \in F(x(t)) = \operatorname{co}\{f_\mu(x(t)) \,|\, \mu = 1, \ldots, M\}$$

is of type (1) with $G_\mu = G$ for all $\mu = 1, \ldots, M$.

The aim of this paper is to present an algorithm for the computation of Lyapunov functions for asymptotically stable differential inclusions of the type (1). Here asymptotic stability is defined in the following strong sense.

Definition 2.5. *The differential inclusion (1) is called (strongly) asymptotically stable (at the origin) if the following two properties hold:*

(i) *For each $\varepsilon > 0$ there exists $\delta > 0$ such that each solution $x(t)$ of (1) with $\|x(0)\| \le \delta$ satisfies $\|x(t)\| \le \varepsilon$ for all $t \ge 0$.*

(ii) *There exists a neighborhood N of the origin such that for each solution $x(t)$ of (1) with $x(0) \in N$ the convergence $x(t) \to 0$ holds as $t \to \infty$.*

Assuming the properties (i) and (ii), the domain of attraction w.r.t. G *is defined as the maximal subset of $\mathbb{R}^n$ for which convergence holds, i.e.*

$$D := \{x_0 \in \mathbb{R}^n \mid \text{every solution with } x(0) = x_0 \text{ is defined on } [0, \infty),$$
$$\text{i.e., it stays in } G, \text{ and satisfies } \lim_{t \to \infty} x(t) = 0\}. \tag{4}$$

Note that if a solution $x(\cdot)$ leaves the region G for some t, then its starting value will not be contained in D. The numerical algorithm we propose will compute a continuous and piecewise affine function $V : G \to \mathbb{R}$. In order to formally introduce this class of functions, we divide G into N n-simplices $\mathcal{T} = \{T_\nu \mid \nu = 1, \ldots, N\}$, i.e. each T_ν is the convex hull of $n + 1$ affinely independent vectors with $\bigcup_{\nu=1,\ldots,N} T_\nu = G$. The intersection $T_{\nu_1} \cap T_{\nu_2}$ is either empty or a common face of T_{ν_1} and T_{ν_2}, i.e. $T_{\nu_1} \cap T_{\nu_2} = \text{co}\{y \mid y \text{ is a vertex of } T_{\nu_i}, \text{ i=1,2}\}$. For each $x \in G$ we define the *active index set* $I_{\mathcal{T}}(x) := \{\nu \in \{1, \ldots, N\} \mid x \in T_\nu\}$. Let us denote by $\text{diam}(T_\nu) := \max_{x,y \in T_\nu} \|x - y\|$ the diameter of a simplex.

Then, by $PL(\mathcal{T})$ we denote the space of continuous functions $V : G \to \mathbb{R}$ which are affine on each simplex, i.e.

$$\nabla V_\nu := \nabla V|_{\text{int } T_\nu} \equiv const \quad \text{for all } T_\nu \in \mathcal{T}.$$

For the algorithm to work properly we need the following compatibility between the subregions G_μ and the simplices T_ν: for every μ and every ν that

$$\text{either } G_\mu \cap T_\nu \text{ is empty or of the form co}\{x_{j_0}, x_{j_1}, \ldots, x_{j_k}\}, \tag{5}$$

where $x_{j_0}, x_{j_1}, \ldots, x_{j_k}$ are pairwise disjoint vertices of T_ν and $0 \leq k \leq n$, i.e., $G_\mu \cap T_\nu$ is a k-face of T_ν.

Since the functions in $PL(\mathcal{T})$ computed by the proposed algorithm are in general nonsmooth, we need a generalized concept for derivatives. In this paper we use Clarks's generalized gradient which we introduce for arbitrary Lipschitz continuous functions. Following [4] we first introduce the corresponding directional derivative.

Definition 2.6. *(i) For a given function $W : \mathbb{R}^n \to \mathbb{R}$ and $l, x \in \mathbb{R}^n$, we will denote the directional derivative*

$$W'(x; l) = \lim_{h \downarrow 0} \frac{W(x + hl) - W(x)}{h}$$

as directional derivative of W at x in direction l (if the limit exists).
(ii) Clarke's directional derivative (cf. [4, Section 2.1]) is defined as

$$W'_{Cl}(x; l) = \limsup_{\substack{y \to x \\ h \downarrow 0}} \frac{W(y + hl) - W(y)}{h}.$$

Using Clarke's directional derivative as support function, we can state the definition of Clarke's subdifferential (see [4, Section 2.1]).

Definition 2.7. *For a locally Lipschitz function $W : \mathbb{R}^n \to \mathbb{R}$ and $x \in \mathbb{R}^n$ Clarke's subdifferential is defined as*

$$\partial_{Cl} W(x) = \{d \in \mathbb{R}^n \mid \forall l \in \mathbb{R}^n : \langle d, l \rangle \leq W'_{Cl}(x; l)\}.$$

In [4, Theorem 2.5.1] the following alternative representation of ∂_{Cl} via limits of gradients is shown.

Proposition 2.8. *For a Lipschitz continuous function $W : G \to \mathbb{R}$ Clarke's sub-differential satisfies*

$$\partial_{Cl}W(x) = \text{co} \left\{ \lim_{i \to \infty} \nabla W(x_i) \,|\, x_i \to x, \, \nabla W(x_i) \text{ exists} \right.$$

$$\left. \text{and} \lim_{i \to \infty} \nabla W(x_i) \text{ exists} \right\}.$$

3. **Lyapunov functions.** There is a variety of possibilities of defining Lyapunov functions for differential inclusions. While it is known that asymptotic stability of (1) with domain of attraction D implies the existence of a smooth Lyapunov function defined on D, see Theorem 3.7, below, for our computational purpose we make use of piecewise affine and thus in general nonsmooth functions. Hence, we need a definition of a Lyapunov function which does not require smoothness. It turns out that Clarke's subgradient introduced above is just the right tool for this purpose.

Definition 3.1. *A positive definite[1] and Lipschitz continuous function $V : G \to \mathbb{R}$ is called a Lyapunov function of (1) if the inequality*

$$\max \langle \partial_{Cl}V(x), F(x) \rangle \leq -\alpha(\|x\|) \tag{6}$$

holds for all $x \in G$, where $\alpha : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ is continuous with $\alpha(0) = 0$ and $\alpha(r) > 0$ for $r > 0$ and we define the set valued scalar product as

$$\langle \partial_{Cl}V(x), F(x) \rangle := \{ \langle d, v \rangle \,|\, d \in \partial_{Cl}V(x), \; v \in F(x) \}. \tag{7}$$

Given $\varepsilon > 0$, since G is compact, changing V to γV for $\gamma \in \mathbb{R}$ sufficiently large we can always assume without loss of generality that

$$\max \langle \partial_{Cl}V(x), F(x) \rangle \leq -\|x\| \tag{8}$$

holds for all $x \in G$ with $\|x\| \geq \varepsilon$. Note, however, that even with a nonlinear rescaling of V it may not be possible to obtain (8) for all $x \in G$.

It is well known that the existence of a Lyapunov function in the sense of Definition 3.1 guarantees asymptotic stability of (1), see, e.g., [21]. For the convenience of the reader we include a proof of this fact. To this end, we first need the following preparatory proposition.

Proposition 3.2. *Let $x(t)$ be a solution of (1) and $V : G \to \mathbb{R}$ be a Lipschitz continuous function. Then the mapping $t \mapsto (V \circ x)(t)$ is absolutely continuous and satisfies*

$$\frac{d}{dt}(V \circ x)(t) \leq \max \langle \partial_{Cl}V(x(t)), F(x(t)) \rangle$$

for almost all $t \geq 0$ with $x(\tau) \in G$ for all $\tau \in [0, t]$.

Proof. We will start with the proof as in [7, Chapter 3, § 15, (8)]. The complete proof is included for the reader's convenience.

The functions $t \mapsto x(t)$ and $t \mapsto (V \circ x)(t)$ are absolutely continuous as a composition, see [17, remarks after Corollary 3.52].

Let us consider a set N of measure zero such that for every $t \notin N$:

- The derivative $\frac{d}{dt}(V \circ x)$ exists at time t.
- The derivative $\dot{x}$ exists at time t and $\dot{x}(t) \in F(x(t))$.

---

[1]i.e., $V(0) = 0$ and $V(x) > 0$ for all $x \in G \setminus \{0\}$

- t is a Lebesgue point of $\dot{x}$, i.e.

$$\lim_{h\to 0}\frac{1}{h}\int_t^{t+h}\|\dot{x}(s)-\dot{x}(t)\|ds = 0$$

(see [20, Chapter IX, § 4, Theorem 5]).

Hence,

$$\lim_{h\to 0}\|\frac{x(t+h)-x(t)}{h}-\dot{x}(t)\| = \lim_{h\to 0}\|\frac{1}{h}\int_t^{t+h}\dot{x}(s)ds-\dot{x}(t)\|$$

$$\leq \lim_{h\to 0}\frac{1}{h}\int_t^{t+h}\|\dot{x}(s)-\dot{x}(t)\|ds = 0$$

and we have proved the following error estimate of the abbreviated Taylor expansion for $x(\cdot)$ as stated in [7, Chapter 3, § 15, (8)]:

$$x(t+h) = x(t) + h\dot{x}(t) + o(h),$$

$$\|V(x(t+h))-V(x(t)+h\dot{x}(t))\| \leq L\cdot\|x(t+h)-x(t)-h\dot{x}(t)\| = o(h).$$

We will use this to prove that the time derivative coincides with the usual (right) directional derivative:

$$\frac{d}{dt}(V\circ x)(t) = \lim_{h\to 0}\frac{V(x(t+h))-V(x(t))}{h} = \lim_{h\downarrow 0}\frac{V(x(t+h))-V(x(t))}{h}$$

$$= \lim_{h\downarrow 0}\frac{V(x(t)+h\dot{x}(t))-V(x(t))}{h} = V'(x(t);\dot{x}(t))$$

By considering the sequence $y_n = x(t)$ in the definition of Clarke's directional derivative, it is clear that

$$V'(x(t);\dot{x}(t)) \leq V'_{\mathrm{Cl}}(x(t);\dot{x}(t)) = \max_{d\in\partial_{\mathrm{Cl}}V(x(t))}\langle d,\dot{x}(t)\rangle$$

$$\leq \max_{d\in\partial_{\mathrm{Cl}}V(x(t))}\max_{v\in F(x(t))}\langle d,v\rangle = \max\langle\partial_{\mathrm{Cl}}V(x(t)),F(x(t))\rangle,$$

where we used Definition 2.6, $\dot{x}(t)\in F(x(t))$ and (7).   □

Now we can prove asymptotic stability.

Theorem 3.3. *Consider a Lipschitz continuous function $V : G \to \mathbb{R}$ and F from* (1) *satisfying* (6) *and let $x(t)$ be a solution of* (1). *Then the inequality*

$$V(x(t)) \leq V(x(0)) - \int_0^t \alpha(\|x(\tau)\|)d\tau \tag{9}$$

holds for all $t \geq 0$ satisfying $x(\tau)\in G$ for all $\tau\in[0,t]$.

In particular, if V is positive definite then (1) *is asymptotically stable and its domain of attraction w.r.t. G defined in* (4) *contains every connected component $C \subseteq V^{-1}([0,c])$ of a sublevel set*

$$V^{-1}([0,c]) := \{x\in G\,|\,V(x)\in[0,c]\}$$

for some $c > 0$ which satisfies $0\in\mathrm{int}\,C$ and $C\subset\mathrm{int}\,G$.

Proof. Proposition 3.2 shows that that $t\mapsto(V\circ x)(t)$ is absolutely continuous and satisfies

$$\frac{d}{dt}(V\circ x)(t) \leq -\alpha(\|x(t)\|)$$

for almost all $t \geq 0$ with $x(t)\in G$. Under the assumption that $x(\tau)\in G$ for all $\tau\in[0,t]$ we can integrate this inequality from 0 to t which yields (9).

By the following classical arguments for Lyapunov functions (see also [5, Theorem 1.2] and [12, Theorem 3.2.7]), the asymptotic stability, i.e., properties (i) and (ii) of Definition 2.5, can now be concluded.

step 1: Before showing (i) and (ii), we prove by contradiction that every solution starting in a connected component $C \subseteq V^{-1}([0,c])$ for some $c > 0$ with $0 \in \operatorname{int} C$ and $C \subset \operatorname{int} G$ stays in C for all $t \geq 0$ and is hence defined on $I = [0,\infty)$. To this end, pick any solution $x(t)$ with $x(0) \in C$ and assume that $x(t_1) \notin C$ holds for some $t_1 > 0$. Then by continuity of the solution there exists a time $t_2 \geq 0$ such that $x(t_2) \in \partial C$ and $x(\tau) \in C$ for all $\tau \in [0,t_2]$. Note that this implies $x(\tau) \in G$ for all $\tau \in [0,t_2]$. Hence, the integral inequality (9) is valid for $t = t_2$ and implies $V(x(t_2)) \leq V(x(0))$ where equality holds if and only if $x(\tau) = 0$ for all $\tau \in [0,t_2]$. In this case we get $x(t_2) = 0$ which contradicts $x(t_2) \in \partial C$ because $0 \in \operatorname{int} C$. If $x(t_2) \neq 0$ we get the strict inequality $V(x(t_2)) < V(x(0)) \leq c$ which again contradicts $x(t_2) \in \partial C$ because by definition of C we have $V(x) = c$ for all $x \in \partial C$.

step 2: Now we prove Definition 2.5 (i) and (ii). In order to show (i), first observe that it is sufficient to prove (i) for all sufficiently small $\varepsilon > 0$. Hence, we can restrict ourselves to those $\varepsilon > 0$ for which the closed ball $\operatorname{cl} B_\varepsilon(0)$ satisfies $\operatorname{cl} B_\varepsilon(0) \subset \operatorname{int} G$. Since $V : G \to \mathbb{R}$ is continuous and positive definite, for each such $\varepsilon > 0$ we get $c_\varepsilon := \min\{V(x) \mid \|x\| \geq \varepsilon\} > 0$. The corresponding sublevel set $V^{-1}([0,c_\varepsilon])$ is contained in the closed ball $\operatorname{cl} B_\varepsilon(0)$. Since V is continuous with $V(0) = 0$ we can furthermore conclude that $V^{-1}([0,c_\varepsilon])$ contains a ball $\operatorname{cl} B_\delta(0)$ for some $\delta > 0$. Clearly, this ball must be contained in the connected component $C \subseteq V^{-1}([0,c_\varepsilon])$ with $0 \in \operatorname{int} C$. By our choice of sufficiently small ε we get $C \subseteq \operatorname{cl} B_\varepsilon(0) \subset \operatorname{int} G$. Thus, any solution with $\|x(0)\| \leq \delta$ starts in C and hence satisfies $x(t) \in C \subseteq V^{-1}([0,c_\varepsilon])$ for all $t \geq 0$. By choice of c_ε we obtain $\|x(t)\| \leq \varepsilon$ and thus (i).

In order to show (ii), pick an arbitrary solution $x(t)$ with $x(0) \in C$ with C from the assumption. Then the solution remains in C for all $t \geq 0$ and we can thus use the integral inequality (9) for all $t \geq 0$. We claim that this implies $V(x(t)) \to 0$. Indeed, since $V(x(t))$ is monotone decreasing and bounded from below by 0 we obtain $V(x(t)) \searrow c^* \geq 0$. Assuming $c^* > 0$ yields $x(t) \notin V^{-1}([0,c^*])$. Then, since V is continuous with $V(0) = 0$ and α is continuous with $\alpha(r) > 0$ for $r > 0$ this implies the existence of $\delta > 0$ with $\alpha(\|x(t)\|) \geq \delta$ for all $t \geq 0$. Thus, the right hand side of (9) and consequently also $V(x(t))$ decreases unboundedly which contradicts $V(x(t)) \searrow c^* > 0$. Thus, $V(x(t)) \searrow 0$ as $t \to \infty$.

Now, the positive definiteness of V implies that $V(x(t)) \to 0$ is only possible if $x(t) \to 0$. This shows (ii) and hence finishes the proof. $\square$

In step 1 of the proof, the condition "$C \subset \operatorname{int} G$" on C and the property $x(t) \in C$ guarantees that the values of $x(\cdot)$ remain in G.

Remark 3.4. *A different concept of nonsmooth Lyapunov functions was presented in* [1]. *In this reference, in addition to Lipschitz continuity, the function V is also assumed to be regular in the sense of* [4, *Definition 2.3.4*], *i.e. the usual directional derivative in Definition 2.6 exists for every direction l and coincides with Clarke's directional derivative. Under this additional condition, inequality* (6) *can be relaxed to*

$$\max \overline{\dot{V}}(x) \leq -\alpha(\|x\|) \tag{10}$$

with

$$\dot{\overline{V}}(x) := \{a \in \mathbb{R} \mid \text{there exists } v \in F(x) \text{ with } \langle p, v \rangle = a \text{ for all } p \in \partial_{Cl}V(x)\}.$$

Here the right hand side $-\alpha(\|x\|)$ *in* (10) *could be replaced by "0" in case of a LaSalle type invariance principle as in* [1]. *Note that this is indeed a relaxation of* (6), *cf. Example* 5.1, *below. While for theoretical constructions this variant is appealing, both the relaxed inequality* (10) *as well as the regularity assumption on V are difficult to be implemented algorithmically, which is why we use* (6). *Note, however, that this does not limit the applicability of our algorithm because asymptotic stability of* (1) *implies the existence of a smooth Lyapunov function, cf. Theorem* 3.7 *below. This in turn implies that both a regular Lyapunov function satisfying* (10) *and a not necessarily regular Lyapunov function satisfying* (6) *exist. Thus, in terms of existence, neither concept is stronger or weaker than the other.*

For computational purposes in our algorithm we now derive a simpler sufficient condition for (6) using the particular structure of $F(x)$ in (1). This sufficient condition requires the evaluation of Clarke's subdifferential of a piecewise linear function. To this end we first need the following lemma which is proved in [16, Proposition 4] and [22, Proposition A.4.1]. We again provide an independent proof in order to keep this paper self contained.

Lemma 3.5. *Clarke's generalized gradient of* $V \in PL(\mathcal{T})$ *is given by*

$$\partial_{Cl}V(x) = \text{co}\{\nabla V_\nu \mid \nu \in I_{\mathcal{T}}(x)\}.$$

Proof. Fix $x \in G$. Since the simplices $T_\nu \in \mathcal{T}$ are closed we have

$$d(x, T_\nu) = \inf_{y \in T_\nu} \|x - y\| = 0$$

if and only if $x \in T_\nu$, i.e., if and only if $\nu \in I_{\mathcal{T}}(x)$. Hence, since there are only finitely many T_ν we find $\varepsilon > 0$ such that $d(x, T_\nu) > \varepsilon$ for all $\nu \notin I_{\mathcal{T}}(x)$.

Now consider an arbitrary sequence $x_i \to x$ with $x_i \in G$ such that $\nabla V(x_i)$ exists for all i and $\lim_{i \to \infty} \nabla V(x_i)$ exists. Since $x_i \to x$ we know $\|x - x_i\| < \varepsilon$ for all sufficiently large i which implies $\nabla V(x_i) = \nabla V_\nu$ for some $\nu \in I_{\mathcal{T}}(x)$. Since there are only finitely many different indices $\nu \in I_{\mathcal{T}}(x)$,

$$\lim_{i \to \infty} \nabla V(x_i) = \nabla V_\nu \in \text{co}\{\nabla V_\nu \mid \nu \in I_{\mathcal{T}}(x)\}$$

follows. By definition of $\partial_{Cl}V(x)$ as the convex hull of all such limits this implies

$$\partial_{Cl}V(x) \subseteq \text{co}\{\nabla V_\nu \mid \nu \in I_{\mathcal{T}}(x)\}.$$

In order to prove the converse inclusion, let $\nu \in I_{\mathcal{T}}(x)$. Then, since cl int $T_\nu = T_\nu$, we find a sequence $x_i \to x$ with $x_i \in \text{int } T_\nu$ implying $\nabla V_\nu \in \partial_{Cl}V(x)$. Now convexity of $\partial_{Cl}V(x)$ implies

$$\text{co}\{\nabla V_\nu \mid \nu \in I_{\mathcal{T}}(x)\} \subseteq \partial_{Cl}V(x)$$

and thus the assertion. $\qquad\square$

Now we can simplify the sufficient condition (6) for the particular structure of F in (1).

Proposition 3.6. *Consider* $V \in PL(\mathcal{T})$ *and* F *from* (1). *Then for any* $x \in G$ *the inequality*

$$\langle \nabla V_\nu, f_\mu(x) \rangle \le -\alpha(\|x\|) \quad \text{for all } \mu \in I_{\mathcal{G}}(x) \text{ and } \nu \in I_{\mathcal{T}}(x) \tag{11}$$

implies (6).

Proof. From Lemma 3.5 we know that each $d \in \partial_{\mathrm{Cl}} V(x)$ can be written as a convex combination

$$d = \sum_{\nu \in I_{\mathcal{T}}(x)} \alpha_\nu \nabla V_\nu$$

for coefficients $\alpha_\nu \geq 0$ with $\sum_{\nu \in I_{\mathcal{T}}(x)} \alpha_\nu = 1$.

Moreover, by the definition of F in (1) each $v \in F(x)$ can be written as a convex combination

$$v = \sum_{\mu \in I_{\mathcal{G}}(x)} \lambda_\mu f_\mu(x)$$

for coefficients $\lambda_\mu \geq 0$ with $\sum_{\mu \in I_{\mathcal{G}}(x)} \lambda_\mu = 1$. Thus from (11) we get

$$\langle d, v \rangle = \left\langle \sum_{\nu \in I_{\mathcal{T}}(x)} \alpha_\nu \nabla V_\nu, \sum_{\mu \in I_{\mathcal{G}}(x)} \lambda_\mu f_\mu(x) \right\rangle$$
$$= \underbrace{\sum_{\nu \in I_{\mathcal{T}}(x)} \alpha_\nu}_{=1} \underbrace{\sum_{\mu \in I_{\mathcal{G}}(x)} \lambda_\mu}_{=1} \underbrace{\langle \nabla V_\nu, f_\mu(x) \rangle}_{\leq -\alpha(\|x\|)} \leq -\alpha(\|x\|).$$

$\square$

We end this section by stating a theorem which ensures that Lyapunov functions — even smooth ones — always exist for asymptotically stable inclusions. Its proof relies on [5, Theorem 1.2] or [24, Theorem 1].

Theorem 3.7. *If the differential inclusion* (1) *is asymptotically stable with domain of attraction D w.r.t. G, then there exists a $\mathcal{C}^\infty$-Lyapunov function $V : D \to \mathbb{R}$.*

Proof. From [24, Theorem 1] applied with $\mathcal{G} = D$ we obtain the existence of a positive definite $\mathcal{C}^\infty$ Lyapunov function $V : D \to \mathbb{R}$ satisfying $\max \langle \nabla V(x), F(x) \rangle \leq -V(x)$ for all $x \in D$. Setting $\alpha(r) := \min\{V(x) \,|\, \|x\| = r\}$ yields the assertion. $\square$

Often we can expect the existence of a Lyapunov function on a larger set than D. The reason for this is that the set G on which we consider (1) is typically a computational domain for our algorithm which is a subset $G \subset \widetilde{G}$ of a larger domain $\widetilde{G}$ on which (1) is defined. In this case, the domain of attraction $\widetilde{D}$ for (1) considered on $\widetilde{G}$ may be strictly larger than the domain of attraction D for the restriction of (1) to G. Thus, Theorem 3.7 ensures the existence of a Lyapunov function on $\widetilde{D}$ whose restriction to $G \cap \widetilde{D}$ is still a Lyapunov function in our sense. In particular, if $G \subseteq \widetilde{D}$, then V is defined on the whole set G. We will use this observation in Corollary 4.8, below. Note, however, that the domain of attraction D of (1) with respect to G is in general smaller than $\widetilde{D} \cap G$.

4. **The algorithm.** In this section we present an algorithm for computing Lyapunov functions in the sense of Definition 3.1 on $G \setminus B_\varepsilon(0)$, where $\varepsilon > 0$ is an arbitrary small positive parameter. To this end, we use an extension of an algorithm first presented in [19] and further developed in [11]. The basic idea of this algorithm is to impose suitable conditions on V on the vertices x_i of the simplices $T_\nu \in \mathcal{T}$ which together with suitable error bounds in the points $x \in G$, $x \neq x_i$, ensures that the resulting V has the desired properties for all $x \in G \setminus B_\varepsilon(0)$.

In order to ensure positive definiteness of V, for every vertex x_i of our simplices we demand

$$V(x_i) \geq \|x_i\|. \tag{12}$$

In order to ensure (6), we demand that for every k-face $T = \text{co}\,\{x_{j_0}, x_{j_1}, \ldots, x_{j_k}\}$, $0 \leq k \leq n$, of a simplex $T_\nu = \text{co}\,\{x_0, x_1, \ldots, x_n\} \in \mathcal{T}$ and every vector field f_μ that is defined on this k-face, the inequalities

$$\langle \nabla V_\nu, f_\mu(x_{j_i}) \rangle + A_{\nu\mu}\|\nabla V_\nu\|_1 \leq -\|x_{j_i}\| \quad \text{for } i = 0, 1, \ldots, k, \tag{13}$$

hold true. Here, $A_{\nu\mu} \geq 0$ is an appropriate constant which is chosen in order to compensate for the interpolation error in the points $x \in T$ with $x \neq x_{j_i}$, $i = 0, \ldots, k$. Corollary 4.3, below, will show that the constants $A_{\nu\mu}$ can be chosen such that the condition (13) for $x_{j_0}, x_{j_1}, \ldots, x_{j_k}$ ensures

$$\langle \nabla V_\nu, f_\mu(x) \rangle \leq -\|x\| \quad \text{for every } x \in T = \text{co}\,\{x_{j_0}, x_{j_1}, \ldots, x_{j_k}\}. \tag{14}$$

Let us illustrate the condition (13) with the 2D-example in Figure 1, where for simplicity of notation we set $A_{\nu\mu} = 0$. Assume that $T_1 = \text{co}\,\{x_1, x_2, x_3\}$ and $T_2 = \text{co}\,\{x_2, x_3, x_4\}$ as well as $T_\nu \subset G_\nu$ and $T_\nu \neq G_\nu$, $\nu = 1, 2$.
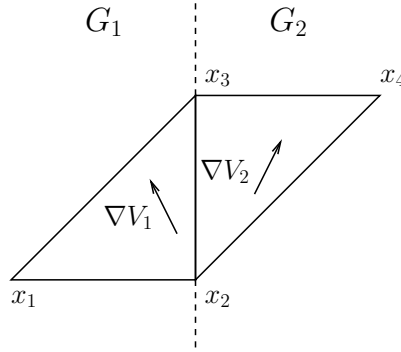

FIGURE 1. Gradient conditions (13) for two adjacent simplices

Since T_1 and T_2 have the common 1-face $T_1 \cap T_2 = \text{co}\,\{x_2, x_3\}$, (13) leads to the following inequalities:

$$\begin{aligned}
-\|x\| &\geq \langle \nabla V_1, f_1(x) \rangle &&\text{for every } x \in \{x_1, x_2, x_3\} \subset T_1, \\
-\|x\| &\geq \langle \nabla V_2, f_2(x) \rangle &&\text{for every } x \in \{x_2, x_3, x_4\} \subset T_2, \\
-\|x\| &\geq \langle \nabla V_1, f_2(x) \rangle &&\text{for every } x \in \{x_2, x_3\} \subset T_1 \cap T_2, \\
-\|x\| &\geq \langle \nabla V_2, f_1(x) \rangle &&\text{for every } x \in \{x_2, x_3\} \subset T_1 \cap T_2.
\end{aligned}$$

Now we turn to the investigation of the interpolation error on our simplicid grids. In the following proposition and lemma we derive bounds for the interpolation error for the linear interpolation of $\mathcal{C}^2$-vector fields which follow immediately from the Taylor expansion. These are standard but are provided here in a form which is suitable for Corollary 4.3, in which we derive an expression for $A_{\nu\mu}$ in (13) which ensures that (14) holds.

Proposition 4.1. Let $x_0, x_1, \ldots, x_k \in \mathbb{R}^n$ be affinely independent vectors, define $T := \mathrm{co}\{x_0, x_1, \ldots, x_k\}$, $h := \mathrm{diam}(T)$ and consider a convex combination $\sum_{i=0}^k \lambda_i x_i \in T$.
a) If $g : G \to \mathbb{R}$ is Lipschitz with constant L, then

$$\left| g\left(\sum_{i=0}^k \lambda_i x_i \right) - \sum_{i=0}^k \lambda_i g(x_i) \right| \le Lh.$$

b) If $g \in \mathcal{C}^2(\mathcal{U}, \mathbb{R})$ with $\mathcal{U} \subseteq \mathbb{R}^n$ is an open set with $T \subset \mathcal{U}$, then

$$\left| g\left(\sum_{i=0}^k \lambda_i x_i \right) - \sum_{i=0}^k \lambda_i g(x_i) \right|$$
$$\le \frac{1}{2} \sum_{i=0}^k \lambda_i B_H \|x_i - x_0\|_2 \left(\max_{z \in T} \|z - x_0\|_2 + \|x_i - x_0\|_2 \right) \le B_H h^2,$$

where $B_H := \max_{z \in T} \|H(z)\|_2$ and $H(z)$ is the Hessian of g at z.

Proof. a) The Lipschitz continuity of g and the convex combination yield the immediate estimate

$$|g\left(\sum_{i=0}^k \lambda_i x_i \right) - g(x_0)| \le L\| \sum_{i=0}^k \lambda_i x_i - x_0\| \le L \sum_{i=0}^k \lambda_i \|x_i - x_0\| \le Lh.$$

b) By Taylor's theorem

$$g\left(\sum_{i=0}^k \lambda_i x_i \right)$$
$$= g(x_0) + \nabla g(x_0) \cdot \sum_{i=0}^k \lambda_i (x_i - x_0) + \frac{1}{2} \sum_{i=0}^k \lambda_i (x_i - x_0)^T H(z) \sum_{j=0}^k \lambda_j (x_j - x_0)$$
$$= \sum_{i=0}^k \lambda_i \left(g(x_0) + \nabla g(x_0) \cdot (x_i - x_0) + \frac{1}{2}(x_i - x_0)^T H(z) \sum_{j=0}^k \lambda_j (x_j - x_0) \right)$$

for some z on the line segment between x_0 and $\sum_{i=0}^k \lambda_i x_i$. Further, again by Taylor's theorem, we have for every $i = 0, 1, \ldots, k$ that

$$g(x_i) = g(x_0) + \nabla g(x_0) \cdot (x_i - x_0) + \frac{1}{2}(x_i - x_0)^T H(z_i)(x_i - x_0)$$

for some z_i on the line segment between x_0 and x_i. Hence,

$$
\left| g\left(\sum_{i=0}^{k} \lambda_i x_i \right) - \sum_{i=0}^{k} \lambda_i g(x_i) \right|
$$

$$
= \frac{1}{2} \left| \sum_{i=0}^{k} \lambda_i (x_i - x_0)^T \left(H(z) \sum_{j=0}^{k} \lambda_j (x_j - x_0) - H(z_i)(x_i - x_0) \right) \right|
$$

$$
\leq \frac{1}{2} \| \sum_{i=0}^{k} \lambda_i (x_i - x_0) \|_2 \left(\|H(z)\|_2 \| \sum_{j=0}^{k} \lambda_j x_j - x_0 \|_2 + \|H(z_i)\|_2 \|x_i - x_0\|_2 \right)
$$

$$
\leq \frac{1}{2} \sum_{i=0}^{k} \lambda_i B_H \|x_i - x_0\|_2 \left(\max_{z \in T} \|z - x_0\|_2 + \|x_i - x_0\|_2 \right).
$$

Since each norm difference $\|z - x_0\|$ for $z \in T$ and $\|x_i - x_0\|$ for $i = 0, 1, \ldots, k$ is bounded by $h = \mathrm{diam}(T)$, this finishes the proof. $\qquad\square$

This proposition shows that when a point $x \in T$ is written as a convex combination of the vertices x_i of the simplex T, then the difference between $g(x)$ and the same convex combination of the function values $g(x_i)$ of g at the vertices x_i is bounded by the corresponding convex combination of error terms, which are small if the simplex is small. In the following lemma we prove an observation which allows us in case b) to derive a simpler expression for the error term in the subsequent corollary. The proof uses standard estimates of the operator norm of $H(z)$ and the bound B on the second derivatives.

Lemma 4.2. *Let $T \subset \mathcal{U} \subset \mathbb{R}^n$, where $\mathcal{U}$ is open and T is compact, and let $g \in \mathcal{C}^2(\mathcal{U}, \mathbb{R})$. Denote the Hessian of g by H and let B be a constant, such that*

$$
\max_{\substack{z \in T \\ r,s=1,2,\ldots,n}} \left| \frac{\partial^2 g}{\partial x_r \partial x_s}(z) \right| \leq B. \tag{15}
$$

Then

$$
\max_{z \in T} \|H(z)\|_2 \leq nB.
$$

Proof. The proof follows from the simple calculation

$$
\max_{z \in T} \|H(z)\|_2 = \max_{\substack{z \in T \\ \|u\|_2=1}} \|H(z)u\|_2 = \max_{\substack{z \in T \\ \|u\|_2=1}} \sqrt{\sum_{i=1}^{n} \left(\sum_{j=1}^{n} h_{ij}(z)u_j \right)^2}
$$

$$
\leq \max_{\|u\|_2=1} \sqrt{\sum_{i=1}^{n} \left(\sum_{j=1}^{n} B|u_j| \right)^2} \leq \max_{\|u\|_2=1} \sqrt{\sum_{i=1}^{n} nB^2 \sum_{j=1}^{n} |u_j|^2}
$$

$$
= \sqrt{n^2 B^2} = nB.
$$

$\qquad\square$

Using Proposition 4.1 and Lemma 4.2 we arrive at the following corollary.

Corollary 4.3. *Let $x_0, x_1, \ldots, x_k \in \mathbb{R}^n$ be affinely independent vectors, define $T := \mathrm{co}\{x_0, x_1, \ldots, x_k\}$, $h := \mathrm{diam}(T)$ and consider a convex combination $\sum_{i=0}^{k} \lambda_i x_i \in T$.*

Consider a function $g : G \to \mathbb{R}^n$ with components $g = (g_1, \ldots, g_n)$.
(i) If $g \in \mathcal{C}^2(\mathcal{U}, \mathbb{R}^n)$ with $\mathcal{U} \subseteq \mathbb{R}^n$ is an open set with $T \subset \mathcal{U}$. Let B be a constant satisfying (15) for every $g = g_i$, $i = 1, \ldots, n$, i.e.

$$\max_{\substack{z \in T \\ i,r,s=1,2,\ldots,n}} \left| \frac{\partial^2 g_i}{\partial x_r \partial x_s}(z) \right| \le B \,.$$

Then

$$\left\| g \left(\sum_{i=0}^k \lambda_i x_i \right) - \sum_{i=0}^k \lambda_i g(x_i) \right\|_\infty \le nBh^2 \,.$$

(ii) Let L be the common Lipschitz constant of g_i, $i = 1, \ldots, n$, in case a) of Proposition 4.1 and B the common bound (15) for the second derivatives g_i in case b). If (13) holds and h satisfies

$$Lh \le A_{\nu\mu} \text{ resp. } nBh^2 \le A_{\nu\mu} \tag{16}$$

in case a) resp. b), then (14) holds.

Proof. (i) For every convex combination $z = \sum_{i=0}^k \lambda_i x_i$ with $z \in T$ and $z = (z_1, \ldots, z_n)$, there is an $m \in \{1, \ldots, n\}$ with $\|z\|_\infty = |z_m|$ such that

$$\left\| g \left(\sum_{i=0}^k \lambda_i x_i \right) - \sum_{i=0}^k \lambda_i g(x_i) \right\|_\infty \le B_{H^m} h^2,$$

where we used Proposition 4.1 b) and defined

$$B_{H^m} := \max_{z \in T} \| H^m(z) \|_2.$$

Here, $H^m(z) = \left(h_{ij}^m(z) \right)_{i,j=1,2,\ldots,n}$ is the Hessian of the m-th component g_m of the vector field g at point z. Then, by Lemma 4.2 and the assumption on B, B_{H^m} is bounded by nB.

(ii) If (13) holds and h satisfies (16), then we obtain with Hölder's inequality and (i) in case b)

$$\langle \nabla V_\nu, g(x) \rangle = \left\langle \nabla V_\nu, g \left(\sum_{i=0}^k \lambda_i x_i \right) \right\rangle$$

$$= \left\langle \nabla V_\nu, \sum_{i=0}^k \lambda_i g(x_i) \right\rangle + \left\langle \nabla V_\nu, g \left(\sum_{i=0}^k \lambda_i x_i \right) - \sum_{i=0}^k \lambda_i g(x_i) \right\rangle$$

$$\le \sum_{i=0}^k \lambda_i \langle \nabla V_\nu, g(x_i) \rangle + \| \nabla V_\nu \|_1 \left\| g \left(\sum_{i=0}^k \lambda_i x_i \right) - \sum_{i=0}^k \lambda_i g(x_i) \right\|_\infty$$

$$\le \sum_{i=0}^k \lambda_i (-\|x_i\| - A_{\nu\mu} \| \nabla V_\nu \|_1) + \| \nabla V_\nu \|_1 nBh^2$$

$$= \sum_{i=0}^k \lambda_i (-\|x_i\|) - A_{\nu\mu} \| \nabla V_\nu \|_1 + \| \nabla V_\nu \|_1 nBh^2$$

$$\le -\sum_{i=0}^k \lambda_i \|x_i\| \le -\left\| \sum_{i=0}^k \lambda_i x_i \right\| = -\|x\|.$$

Case a) is similar to prove.   □

Before running the algorithm, one might want to remove some of the $T_\nu \in \mathcal{T}$ close to the equilibrium at zero from $\mathcal{T}$. The reason for this is that inequality (14) and thus (13) may not be feasible near the origin, cf. also the discussion on $\alpha(\|x\|)$ after Definition 3.1. This is also reflected in the proof of Theorem 4.6, below, in which we will need a positive distance to the equilibrium at zero.

To accomplish this fact, we define the subset

$$\mathcal{T}^\varepsilon := \{T_\nu \in \mathcal{T} \mid T_\nu \cap B_\varepsilon(0) = \emptyset\} \subset \mathcal{T}$$

for $\varepsilon > 0$.

Furthermore, if f_μ is defined on a simplex $T := \mathrm{co}\,\{x_0, x_1, \ldots, x_k\}$, we assume in case b) that f_μ possesses a $\mathcal{C}^2$-extension $\overline{f}_\mu : \mathcal{U} \to \mathbb{R}^n$ on an open set $\mathcal{U} \supset T$. If T is an n-simplex and f_μ is $\mathcal{C}^2$ on T, then this follows by Whitney's extension theorem [25] and we have

$$\max_{\substack{z \in T \\ i,r,s=1,2,\ldots,n}} \left| \frac{\partial^2 \overline{f}_{\mu,i}}{\partial x_r \partial x_s}(z) \right| = \max_{i,r,s=1,2,\ldots,n} \sup_{z \in \mathrm{int}\,T} \left| \frac{\partial^2 f_{\mu,i}}{\partial x_r \partial x_s}(z) \right|,$$

where $\overline{f}_{\mu,i}$ and $f_{\mu,i}$ are the i-th components of the vector fields $\overline{f}_\mu$ and f_μ respectively.

Algorithm 4.4.

(i) *For all vertices x_i of the simplices $T_\nu \in \mathcal{T}^\varepsilon$ we introduce $V(x_i)$ as the variables and $\|x_i\|$ as lower bounds in the constraints of the linear program and demand $V(x_i) \geq \|x_i\|$. Note that every vertex x_i only appears once here.*

(ii) *For every simplex $T_\nu \in \mathcal{T}^\varepsilon$ we introduce the variables $C_{\nu,i}$, $i = 1, \ldots, n$ and demand that for the i-th component $\nabla V_{\nu,i}$ of ∇V_ν we have*

$$|\nabla V_{\nu,i}| \leq C_{\nu,i}, \quad i = 1, \ldots, n.$$

(iii) *For every $T_\nu := \mathrm{co}\,\{x_0, x_1, \ldots, x_n\} \in \mathcal{T}^\varepsilon$, every k-face $T = \mathrm{co}\,\{x_{j_0}, x_{j_1}, \ldots, x_{j_k}\}$ of T_ν, $0 \leq k \leq n$, and every μ with $T \subseteq G_\mu$ we demand one of the two inequalities*

$$a) \qquad \langle \nabla V_\nu, f_\mu(x_{j_i}) \rangle + L h_\nu \sum_{j=1}^n C_{\nu,j} \leq -\|x_{j_i}\|, \quad \text{if } f_\mu \text{ L-Lipschitz on } T_\nu, \quad (17)$$

$$b)\ \langle \nabla V_\nu, f_\mu(x_{j_i}) \rangle + n B_{\mu,T} h_\nu^2 \sum_{j=1}^n C_{\nu,j} \leq -\|x_{j_i}\|, \quad \text{if } f_\mu \in C^2(\mathcal{U}), \mathcal{U} \supset T_\nu, \quad (18)$$

for $i = 0, 1, \ldots, k$ with $h_\nu := \mathrm{diam}(T_\nu)$, $B_{\mu,T} \geq \max\limits_{i,r,s=1,2,\ldots,n} \sup\limits_{z \in T} \left| \dfrac{\partial^2 \overline{f}_{\mu,i}}{\partial x_r \partial x_s}(z) \right|$.

Note, that if f_μ is defined on the face $T \subset T_\nu$, then f_μ is also defined on any face $S \subset T$ of T. However, it is easily seen that the constraints (17) resp. (18) for the simplex S are redundant, for they are automatically fulfilled if the constraints for T are valid.

(iv) *If the linear program with the constraints (i)–(iii) has a feasible solution, then the values $V(x_i)$ from this feasible solution at all the vertices x_i of all the simplices $T_\nu \in \mathcal{T}^\varepsilon$ and the condition $V \in PL(\mathcal{T}^\varepsilon)$ uniquely define the function*

$$V : \bigcup_{T_\nu \in \mathcal{T}^\varepsilon} T_\nu \to \mathbb{R}.$$

The following theorem shows that V from (iv) defines a Lyapunov function on the simplices $T_\nu \in \mathcal{T}^\varepsilon$.

Theorem 4.5. *Assume that each f_μ is Lipschitz on G_μ and the linear program constructed by the algorithm has a feasible solution. Then, on each $T_\nu \in \mathcal{T}^\varepsilon$ the function V from (iv) is positive definite and for every $x \in T_\nu \in \mathcal{T}^\varepsilon$ inequality (11) holds with $\alpha(r) = r$, i.e.,*

$$\langle \nabla V_\nu, f_\mu(x) \rangle \leq -\|x\| \quad \text{for all } \mu \in I_\mathcal{G}(x) \text{ and } \nu \in I_\mathcal{T}(x).$$

Proof. Let f_μ be defined on the k-face $T = T_\nu \cap G_\mu$ with vertices $x_{j_0}, x_{j_1}, \ldots, x_{j_k}$, $0 \leq k \leq n$. Then every $x \in T$ is a convex combination $x = \sum_{i=0}^{k} \lambda_i x_{j_i}$. Conditions (ii) and (iii) of the algorithm imply that (13) holds on T with $A_{\nu\mu} = L h_\nu$ resp. $A_{\nu\mu} = n B_{\mu,T} h_\nu^2$, because

$$A_{\nu\mu} \sum_{j=1}^{n} C_{\nu,j} = n B_{\mu,T} h_\nu^2 \sum_{j=1}^{n} C_{\nu,j} \geq n B_{\mu,T} h_\nu^2 \sum_{j=1}^{n} |\nabla V_{\nu,i}| = n B_{\mu,T} h_\nu^2 \|\nabla V_\nu\|_1$$

in case b) (case a) is similar). Thus, Corollary 4.3(ii) yields the assertion. $\qquad\square$

Clearly, the Lipschitz assumption on f_μ is weaker than the $\mathcal{C}^2$-assumption in case b) of Proposition 4.1, but the triangulation in case a) must be finer to fulfil the more demanding condition (17) in comparison with (18).

In the next theorem we will prove, that if (1) possesses a $\mathcal{C}^2$-Lyapunov function, then Algorithm 4.4 succeeds in computing a Lyapunov function $V \in PL(\mathcal{T}^\varepsilon)$ for a suitable triangulation $\mathcal{T}^\varepsilon$. In the following Corollary 4.8, we will give a sufficient condition for the existence of such a $\mathcal{C}^2$-Lyapunov function.

Theorem 4.6. *Assume that each f_μ is Lipschitz on G_μ and the system (1) possesses a $\mathcal{C}^2$-Lyapunov function $W^* : G \to \mathbb{R}$ and let $\varepsilon > 0$.*

Then, there exists a triangulation $\mathcal{T}^\varepsilon$ such that the linear programming problem constructed by the algorithm has a feasible solution and thus delivers a Lyapunov function $V \in PL(\mathcal{T}^\varepsilon)$ for the system.

Remark 4.7. *The precise conditions on the triangulation are given in the formulas (25) resp. (24) of the proof for the two cases f_μ being Lipschitz continuous resp. twice continuously differentiable. The triangulation must ensure that each triangle has a sufficiently small diameter and fulfills an angle condition to prevent too flat triangles. If the simplices $T_\nu \in \mathcal{T}$ are all similar as in [11], then it suffices to assume that $\max_{\nu=1,2,\ldots,N} \operatorname{diam}(T_\nu)$ is small enough, cf. [11, Theorem 8.2 and Theorem 8.4]. Here we are using more general triangulations $\mathcal{T}$ and therefore, we have to compromise for triangulations that can lead to problems. Essentially, we still assume that $\max_{\nu=1,2,\ldots,N} \operatorname{diam}(T_\nu)$ is small enough, but additionally we have to assume that the simplices $T_\nu \in \mathcal{T}^\varepsilon$ are regular in the sense that e.g. $X_\nu^* \cdot \operatorname{diam}(T_\nu) \leq X^* h \leq R$, for some constant $R > 0$ (cf. parts (ii),(v) and equation (19) of the proof). This is a similar condition as in FEM methods. Note that for any triangulation $\mathcal{T}$ such that $\mathcal{T}^\varepsilon$ satisfies assumption (24) resp. (25), this inequality will also be satisfied for the scaled down triangulation*

$$(c\mathcal{T})^\varepsilon := \{c T_\nu = \operatorname{co}\{c x_0, \ldots, c x_n\} \mid T_\nu = \operatorname{co}\{x_0, \ldots, x_n\} \in \mathcal{T}, \, c T_\nu \cap B_\varepsilon(0) = \emptyset\}$$

for any $c \in (0, 1]$, cf. also Remark 4.9, below.

Proof of Theorem 4.6: We will split the proof into several steps.

(i) Since continuous functions take their maximum on compact sets and $G \setminus B_\varepsilon(0)$ is compact, we can define

$$c_0 := \max_{x \in G \setminus B_\varepsilon(0)} \frac{\|x\|}{W^*(x)}$$

and for every $\mu = 1, 2, \ldots, M$

$$c_\mu := \max_{G_\mu \setminus B_\varepsilon(0)} \frac{-2\|x\|}{\langle \nabla W^*(x), f_\mu(x) \rangle}.$$

We set $c = \max_{\mu=0,1,\ldots,M} c_\mu$ and define $W(x) := c \cdot W^*(x)$. Then, by construction, W is a Lyapunov function for the system, $W(x) \geq \|x\|$ for every $x \in G \setminus B_\varepsilon(0)$, and for every $\mu = 1, 2, \ldots, M$ we have $\langle \nabla W(x), f_\mu(x) \rangle \leq -2\|x\|$ for every $x \in G_\mu \setminus B_\varepsilon(0)$.

(ii) For every $T_\nu = \mathrm{co}\,\{x_0, x_1, \ldots, x_n\} \in \mathcal{T}^\varepsilon$ pick out one of the vertices, say $y = x_0$, and define the $n \times n$ matrix $X_{\nu,y}$ by writing the components of the vectors $x_1 - x_0, x_2 - x_0, \ldots, x_n - x_0$ as row vectors consecutively, i.e.

$$X_{\nu,y} = \left(x_1 - x_0, x_2 - x_0, \ldots x_n - x_0\right)^T.$$

$X_{\nu,y}$ is invertible, since its rows are linearly independent. We are interested in the quantity $X_{\nu,y}^* = \|X_{\nu,y}^{-1}\|_2 = \lambda_{\min}^{-\frac{1}{2}}$, where $\lambda_{\min}$ is the smallest eigenvalue of $X_{\nu,y}^T X_{\nu,y}$.

First, we show that $X_{\nu,y}^*$ is properly defined, i.e. is independent of the order of the $x_0, x_1, \ldots, x_n$. Denote by S_n the permutations σ of $\{0, 1, \ldots, n\}$. Then the row permutating matrices by left multiplication are the matrices $E_\sigma = (\delta_{\sigma(i),\sigma(j)})_{i,j=0,1,\ldots,n}$, $\sigma \in S_n$. If we show that $\|(E_\sigma X_{\nu,y})^{-1}\|_2 = \|X_{\nu,y}^{-1}\|_2 = \lambda_{\min}^{-\frac{1}{2}}$ for all $\sigma \in S_n$, then we have showed that $X_{\nu,y}^*$ is independent of the order of $x_0, x_1, \ldots, x_n$. For every $\sigma \in S_n$

$$E_\sigma^T E_\sigma = \left(\sum_{k=0}^n \delta_{\sigma(k),\sigma(i)} \cdot \delta_{\sigma(k),\sigma(j)}\right)_{i,j=0,1,\ldots,n} = (\delta_{i,j})_{i,j=0,1,\ldots,n} = I.$$

Hence,

$$(E_\sigma X_{\nu,y})^T E_\sigma X_{\nu,y} = X_{\nu,y}^T \underbrace{E_\sigma^T E_\sigma}_{=I} X_{\nu,y} = X_{\nu,y}^T X_{\nu,y}.$$

This proves that $X_{\nu,y}^*$ is properly defined. Let us denote by

$$X_\nu^* = \max_{y \text{ vertex of } T_\nu} \|X_{\nu,y}^{-1}\|_2 \quad \text{and} \quad X^* = \max_{\nu=1,2,\ldots,N} X_\nu^*. \tag{19}$$

(iii) We consider an arbitrary but fixed $T_\nu = \mathrm{co}\,\{x_0, x_1, \ldots, x_n\} \in \mathcal{T}^\varepsilon$ and set $y = x_0$. By Whitney's extension theorem [25] we can extend W to an open set containing G so W is defined on an open set containing $T_\nu \subset G$. For every $i = 1, 2, \ldots, n$ we have by Taylor's theorem

$$W(x_i) = W(x_0) + \langle \nabla W(x_0), x_i - x_0 \rangle + \frac{1}{2}\langle x_i - x_0, H_W(z_i)(x_i - x_0) \rangle,$$

where H_W is the Hessian of W and $z_i = x_0 + \vartheta_i(x_i - x_0)$ for some $\vartheta_i \in]0, 1[$. We define

$$w_{\nu,y} := \begin{pmatrix} W(x_1) - W(x_0) \\ W(x_2) - W(x_0) \\ \vdots \\ W(x_n) - W(x_0) \end{pmatrix}$$

so that the following equality holds:

$$w_{\nu,y} - X_{\nu,y}\nabla W(x_0) = \frac{1}{2}\begin{pmatrix} \langle x_1 - x_0, H_W(z_1)(x_1 - x_0)\rangle \\ \langle x_2 - x_0, H_W(z_2)(x_2 - x_0)\rangle \\ \vdots \\ \langle x_n - x_0, H_W(z_n)(x_n - x_0)\rangle \end{pmatrix} =: \frac{1}{2}\xi_{\nu,y} \qquad (20)$$

Setting

$$A := \max_{\substack{z \in G \\ i,j=1,2,\ldots,n}} \left| \frac{\partial^2 W}{\partial x_i \partial x_j}(z) \right|$$

and

$$h := \max_{\nu=1,2,\ldots,N} \operatorname{diam}(T_\nu)$$

we have by Lemma 4.2 that

$$\|(x_i - x_0)^T H_W(z_i)(x_i - x_0)\|_2 \le h^2 \|H_W(z_i)\|_2 \le nAh^2$$

for $i = 1, 2, \ldots, n$. Hence,

$$\|\xi_{\nu,y}\|_2 = \left\| \begin{pmatrix} \langle x_1 - x_0, H_W(z_1)(x_1 - x_0)\rangle \\ \langle x_2 - x_0, H_W(z_2)(x_2 - x_0)\rangle \\ \vdots \\ \langle x_n - x_0, H_W(z_n)(x_n - x_0)\rangle \end{pmatrix} \right\|_2 \le n^{\frac{3}{2}}Ah^2. \qquad (21)$$

Furthermore, for every $i, j = 1, 2, \ldots, n$ there is a $\tilde{z}_i$ on the line segment between x_i and $y = x_0$, such that

$$\partial_j W(x_i) - \partial_j W(x_0) = \langle \nabla \partial_j W(\tilde{z}_i), x_i - x_0\rangle,$$

where $\partial_j W$ denotes the j-th component of ∇W. Hence, by Lemma 4.2

$$\|\nabla W(x_i) - \nabla W(x_0)\|_2 \le nAh.$$

From this we obtain the inequality

$$\|X_{\nu,y}^{-1}w_{\nu,y} - \nabla W(x_i)\|_2$$
$$\le \|X_{\nu,y}^{-1}w_{\nu,y} - \nabla W(x_0)\|_2 + \|\nabla W(x_i) - \nabla W(x_0)\|_2$$
$$\le \frac{1}{2}\|X_{\nu,y}^{-1}\|_2 n^{\frac{3}{2}}Ah^2 + nAh \le nAh\left(\frac{1}{2}X^* n^{\frac{1}{2}}h + 1\right) \qquad (22)$$

for every $i = 0, 1, \ldots, n$. This last inequality is independent of the simplex $T_\nu = \operatorname{co}\{x_0, x_1, \ldots, x_n\}$.

(iv) Define

$$D := \max_{\mu=1,2,\ldots,M} \sup_{z \in G_\mu\setminus\{0\}} \frac{\|f_\mu(z)\|_2}{\|z\|}.$$

Note, that $D < +\infty$ because all norms on $\mathbb{R}^n$ are equivalent and for every μ the vector field f_μ is Lipschitz on G_μ and, if defined, $f_\mu(0) = 0$. In this case, $D \le \alpha L$ with $\|z\|_2 \le \alpha\|z\|$.

(v) In the final step we assign values to the variables $V(x_i), C_{\nu,i}$ of the linear programming problem from the algorithm and show that they fulfill the constraints.

For every $T_\nu \in \mathcal{T}^\varepsilon$ and every vertex x_i of T_ν set $V(x_i) = W(x_i)$. Clearly, by the construction of W and of the piecewise linear function V from the variables $V(x_i)$, we have $V(x_i) \geq \|x_i\|$ for every $T_\nu \in \mathcal{T}^\varepsilon$ and every vertex x_i of T_ν.

Pick an arbitrary but fixed $T_\nu = \mathrm{co}\,\{x_0, x_1, \ldots, x_n\} \in \mathcal{T}^\varepsilon$ and set $y = x_0$. Then, by the definition of $w_{\nu,y}$ and $X_{\nu,y}$, cf. part (iii) of the proof, we have

$$\nabla V_\nu = X_{\nu,y}^{-1} w_{\nu,y}\,,$$

since V is piecewise linear and

$$V(x) = V(x_0) + w_{\nu,y}^T \left(X_{\nu,y}^T\right)^{-1} (x - x_0)\,.$$

For every variable $C_{\nu,i}$ in the linear programming problem from the algorithm set

$$C_{\nu,i} = \|\nabla V_\nu\|_2 = \|X_{\nu,y}^{-1} w_{\nu,y}\|_2.$$

Then evidently, $C_{\nu,i} \geq |\nabla V_{\nu,i}|$ for every $T_\nu \in \mathcal{T}^\varepsilon$. The boundedness of ∇W on G assures that there is a constant C such that

$$\|X_{\nu,y}^{-1} w_{\nu,y}\|_2 \leq \|X_{\nu,y}^{-1}\|_2 \max_{z \in G} \|\nabla W(z)\|_2 h \leq R \max_{z \in G} \|\nabla W(z)\|_2 =: C$$

with R from Remark 4.7. Thus, $C_{\nu,i} \leq C$ holds uniformly in ν and i.

Let f_μ be an arbitrary vector field defined on the whole of T_ν or one of its faces, i.e. f_μ is defined on $T := \mathrm{co}\,\{x_{j_0}, x_{j_1}, \ldots, x_{j_k}\}$, $0 \leq k \leq n$, where the x_{j_i} are vertices of T_ν. Then, by (ii) and (20)–(22), we have for every $i = 0, 1, \ldots, k$ that

$$\langle \nabla V_\nu, f_\mu(x_{j_i}) \rangle = \langle \nabla W(x_{j_i}) + \nabla V_\nu - \nabla W(x_{j_i}), f_\mu(x_{j_i}) \rangle$$
$$= \langle \nabla W(x_{j_i}), f_\mu(x_{j_i}) \rangle + \langle X_{\nu,y}^{-1} w_{\nu,y} - \nabla W(x_{j_i}), f_\mu(x_{j_i}) \rangle$$
$$\leq -2\|x_{j_i}\| + \|X_{\nu,y}^{-1} w_{\nu,y} - \nabla W(x_{j_i})\|_2 \|f_\mu(x_{j_i})\|_2$$
$$\leq -2\|x_{j_i}\| + nAh \left(\frac{1}{2} X^* n^{\frac{1}{2}} h + 1\right) \cdot D\|x_{j_i}\|.$$

In case b), i.e. $f_\mu \in C^2(\mathcal{U})$, $\mathcal{U} \supset T_\nu$, the linear constraints

$$\langle \nabla V_\nu, f_\mu(x_{j_i}) \rangle + nB_{\mu,T} h_\nu^2 \sum_{j=1}^n C_{\nu,j} \leq -\|x_{j_i}\|$$

are fulfilled whenever h is so small that

$$-2\|x_{j_i}\| + n^2 B h^2 C + nAh \left(\frac{1}{2} X^* n^{\frac{1}{2}} h + 1\right) D\|x_{j_i}\| \leq -\|x_{j_i}\| \qquad (23)$$

with X^* given by (19) and

$$\max_{\substack{\mu=1,2,\ldots,M \\ T \text{ face of simplex in } \mathcal{T}^\varepsilon}} B_{\mu,T} \leq B\,.$$

Because $\|x_{j_i}\| \geq \varepsilon$ inequality (23) is satisfied if

$$n^2 B \frac{h^2}{\varepsilon} C + nAh \left(\frac{1}{2} X^* n^{\frac{1}{2}} h + 1\right) D \leq 1\,. \qquad (24)$$

Again, case a) follows similarly for f_μ being Lipschitz, if

$$nL\frac{h}{\varepsilon}C + nAh\left(\frac{1}{2}X^*n^{\frac{1}{2}}h + 1\right)D \leq 1\,. \tag{25}$$

Since T_ν and f_μ were arbitrary, this proves the theorem.

$\square$

Corollary 4.8. *Consider a differential inclusion F of type* (1) *defined on a set $\widetilde{G} \subseteq \mathbb{R}^n$. Assume that F is strongly asymptotically stable with domain of attraction $\widetilde{D}$ w.r.t. $\widetilde{G}$ and that each f_μ is Lipschitz. Consider a computational domain $G \subseteq \widetilde{D}$ such that the restriction $F|_G$ is again of the form* (1) *and the assumptions from Section* 2 *hold for $F|_G$ and G.*

Then, for each $\varepsilon > 0$ there exists a triangulation $\mathcal{T}^\varepsilon$ such that the linear programming problem constructed by the algorithm has a feasible solution and thus delivers a Lyapunov function $V \in PL(\mathcal{T}^\varepsilon)$ for the system.

Proof. By Theorem 3.7 there exists a $\mathcal{C}^\infty$ Lyapunov function $W^* : \widetilde{D} \to \mathbb{R}$ whose restriction to G is a $\mathcal{C}^\infty$ Lyapunov function on G. Hence, the assertion follows from Theorem 4.6. $\square$

Remark 4.9. *Note that in assumption* (24) *resp.* (25) *the parameter ε does not only appear explicitly but also implicitly. This is because the value A defined in part* (iii) *depends on W defined in part* (i) *which in turn depends on ε via c_0 and c_μ. In particular, the value A will become larger if ε becomes smaller. Hence, given a triangulation $\mathcal{T}$ such that $\mathcal{T}^\varepsilon$ satisfies assumption* (24) *resp.* (25)*, this inequality may not be satisfied for the triangulation*

$$c(\mathcal{T}^\varepsilon) := \{cT_\nu = \mathrm{co}\,\{cx_0, \ldots, cx_n\} \,|\, T_\nu = \mathrm{co}\,\{x_0, \ldots, x_n\} \in \mathcal{T}^\varepsilon\}$$
$$= \{cT_\nu = \mathrm{co}\,\{cx_0, \ldots, cx_n\} \,|\, T_\nu = \mathrm{co}\,\{x_0, \ldots, x_n\} \in \mathcal{T},\, cT_\nu \cap B_{c\varepsilon}(0) = \emptyset\}$$

for $c \in (0, 1)$, because here we do not only shrink the triangles by the factor c but also the size of the neighborhood $B_{c\varepsilon}(0)$. Hence, the growth of A when passing from ε to $c\varepsilon < \varepsilon$ may make (24) *resp.* (25) *invalid. The corresponding assumption will, however, always be satisfied for the triangulation $(c\mathcal{T})^\varepsilon$ defined in Remark 4.7 because for this triangulation ε remains fixed.*

5. **Examples.** We illustrate our algorithm by two examples, the first one is taken from [1].

Example 5.1 (Nonsmooth harmonic oscillator with nonsmooth friction). *Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be given by*

$$f(x_1, x_2) = \left(-\,\mathrm{sgn}\,x_2 - \frac{1}{2}\,\mathrm{sgn}\,x_1,\, \mathrm{sgn}\,x_1\right)^T$$

with $\mathrm{sgn}\,x_i = 1$, $x_i \geq 0$ and $\mathrm{sgn}\,x_i = -1$, $x_i < 0$. This vector field is piecewise constant on the four regions

$$G_1 = [0, \infty) \times [0, \infty), \qquad G_2 = (-\infty, 0] \times [0, \infty),$$
$$G_3 = (-\infty, 0] \times (-\infty, 0], \qquad G_4 = [0, \infty) \times (-\infty, 0],$$

hence its Filippov regularization is of type (1) *and the triangulation could be chosen such that the compatibility condition* (5) *holds. In* [1] *it is shown that the function $V(x) = |x_1| + |x_2|$ with $x = (x_1, x_2)^T$ is a Lyapunov function in the sense of*

Remark 3.4. *It is, however, not a Lyapunov function in the sense of our Defini-tion 3.1. For instance, if we pick x with $x_1 = 0$ and $x_2 > 0$ then $I_{\mathcal{G}}(x) = \{1, 2\}$ and the Filippov regularization F of f is*

$$F(x) = \text{co} \left\{ \begin{pmatrix} -3/2 \\ 1 \end{pmatrix}, \begin{pmatrix} -1/2 \\ -1 \end{pmatrix} \right\}$$

and for $\partial_{Cl} V$ we get

$$\partial_{Cl} V(x) = \text{co} \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\}.$$

This implies

$$\max \langle \partial_{Cl} V(x), F(x) \rangle \geq \left\langle \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} -3/2 \\ 1 \end{pmatrix} \right\rangle = 5/2 > 0$$

which shows that (6) does not hold.

Despite the fact that $V(x) = |x_1| + |x_2|$ is not a Lyapunov function in our sense, our algorithm produces a Lyapunov function (see Fig. 2) which is — up to rescaling — rather similar to this V.

In the x_1, x_2-plane a subset of the domain of attraction secured by the Lyapunov function is depicted in Figure 2.
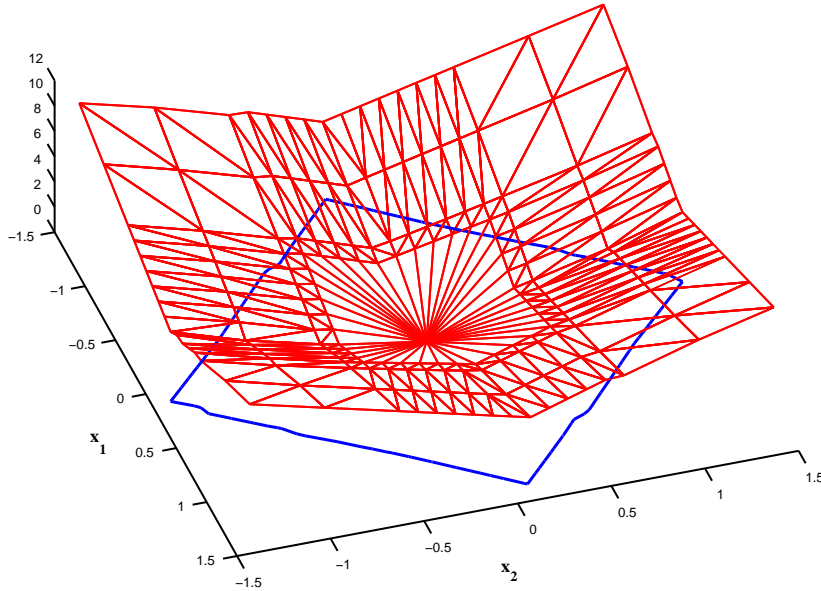

FIGURE 2. Lyapunov function and level set for Example 5.1

There are two facts worth noting. First, we can set the error terms $B_{\mu,T} = 0$ for any triangulation fulfilling the conditions of Theorem 4.6 because the second-order derivatives of the f_μ vanish in the interiors of the simplices. Second, for a sufficiently fine but fixed grid we can take $\varepsilon > 0$ arbitrary small, but we cannot set

$\varepsilon = 0$ *because the Lyapunov function cannot fulfill the inequality* (8) *at the origin. This is because*

$$F(0,0) = \text{co} \left\{ \begin{pmatrix} -3/2 \\ 1 \end{pmatrix}, \begin{pmatrix} -1/2 \\ -1 \end{pmatrix}, \begin{pmatrix} 3/2 \\ -1 \end{pmatrix}, \begin{pmatrix} 1/2 \\ 1 \end{pmatrix} \right\}$$

is a quadrilateral containing $(0,0)$ *as an inner point and thus contains vectors of all directions. Hence, our condition at* 0 *would require* $\nabla V(0,0) = (0,0)^T$ *but this is not possible because of condition (i) of our algorithm and the definition of the Clarke generalized gradient. This is a property of the algorithm for differential inclusions and does not happen if* $F(0) = \{0\}$ *as is the case when we are considering ordinary differential equations (and using less strict bounds, cf. Example* 5.2*). Second, it is interesting to compare the level sets of the Lyapunov function on Fig.* 2 *to the level sets of the Lyapunov function* $V(x) = |x_1| + |x_2|$ *from* [1]*. The fact that the level set in Fig.* 2 *is not a perfect rhombus (as it is for* $V(x) = |x_1| + |x_2|$*) is not due to numerical inaccuracies. Rather, the small deviations are necessary because, as shown above,* $V(x) = |x_1| + |x_2|$ *is not a Lyapunov function in our sense.*

The following example extends the one in [10] by adding the uncertainty in the friction.

Example 5.2 (pendulum with uncertain friction). *Let* $f : \mathbb{R}^2 \to \mathbb{R}^2$ *be given with*

$$f(x_1, x_2) = (x_2, -kx_2 - g\sin(x_1))^T,$$

where g *is the earth gravitation and equals approximately* $9.81 m/s^2$ *and* k *is a nonnegative parameter modelling the friction of the pendulum.*
It is known that the system is asymptotic stable for $k > 0$*, e.g. in the interval* $[0.2, 1]$*.*

If the friction k *is unknown and time-varying, we obtain an inclusion of the type* (1) *with*

$$\dot{x}(t) \in F(x(t)) = \text{co} \{ f_\mu(x(t)) \mid \mu = 1, 2 \}.$$

where $G_1 = G_2$ *and* $f_1(x) = (x_2, -0.2x_2 - g\sin(x_1))^T$*,* $f_2(x) = (x_2, -x_2 - g\sin(x_1))^T$*.*
This is a system of the type of Example 2.4 *where the right-hand side of the differential inclusion is multivalued on the whole domain. Trivially, the subregions* G_μ *satisfy the compatibility condition* (5) *for any triangulation of* G*. Algorithm* 4.4 *succeeds in computing a Lyapunov function (see Fig.* 3*), even with* $\varepsilon = 0$*. This seems contradictory for the constant* $B_{\mu,T}$ *cannot be set to zero. The reason why this is possible is that we took advantage of our system vanishing at the origin and our triangulation of* G *having the origin as a central vertex of a triangle fan, cf.* [9]*.*

The constraint (18) *in (iii) in the algorithm can obviously not be fulfilled for* $x_{j_i} = 0$ *if* $B_{\mu,T} > 0$*. By a more careful analysis and using the special structure of the triangulation around the origin as well as* $F(0) = \{0\}$*, the simple, but conservative estimate from Corollary* 4.3 *can be improved via the inequality*

$$\left| g \left(\sum_{i=0}^k \lambda_i x_i \right) - \sum_{i=0}^k \lambda_i g(x_i) \right|$$

$$\leq \frac{1}{2} \sum_{i=0}^k \lambda_i B_H \|x_i - x_0\|_2 \left(\max_{z \in T} \|z - x_0\|_2 + \|x_i - x_0\|_2 \right)$$

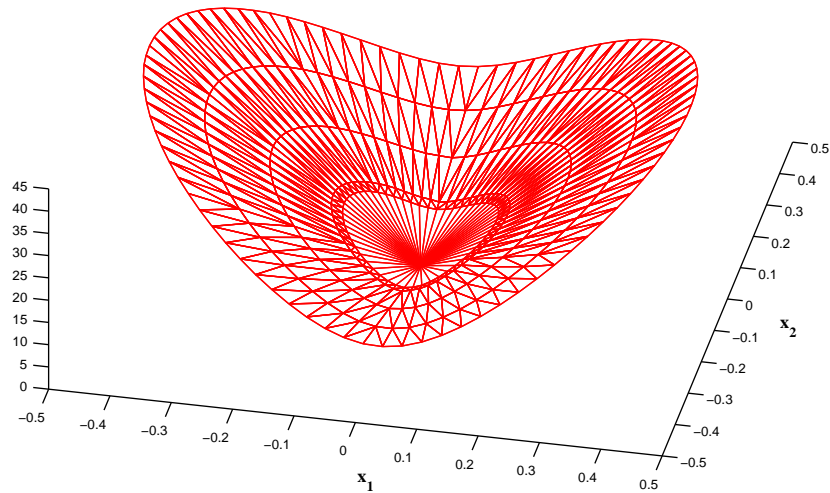*from Proposition* 4.1*, see* [9] *for details.*

FIGURE 3. Lyapunov function for Example 5.2

As a consequence for this particular example the computed Lyapunov function is valid even for a neighborhood of the origin.

REFERENCES

[1] A. Bacciotti and F. Ceragioli, *Stability and stabilization of discontinuous systems and non-smooth Lyapunov functions*, ESAIM Control Optim. Calc. Var., **4** (1999), 361–376 (electronic).

[2] F. Camilli, L. Grüne and F. Wirth, *A regularization of Zubov's equation for robust domains of attraction*, in "Nonlinear Control in the Year 2000, Vol. 1," vol. 258 of Lecture Notes in Control and Inform. Sci., Springer, London, (2001), 277–289.

[3] G. Chesi, *Estimating the domain of attraction for uncertain polynomial systems*, Automatica, **40** (2004), 1981–1986.

[4] F. H. Clarke, "Optimization and Nonsmooth Analysis," vol. 5 of Classics in Applied Mathematics, 2nd edition, SIAM, Philadelphia, PA, 1990. First edition published in John Wiley & Sons, Inc., New York, 1983.

[5] F. H. Clarke, Yu. S. Ledyaev and R. J. Stern, *Asymptotic stability and smooth Lyapunov functions*, J. Differential Equations, **149** (1998), 69–114.

[6] T. Donchev, V. Ríos and P. Wolenski, *Strong invariance and one-sided Lipschitz multifunctions*, Nonlinear Anal., **60** (2005), 849–862.

[7] A. F. Filippov, "Differential Equations with Discontinuous Righthand Sides," vol. 18 of Mathematics and its Applications (Soviet Series), Kluwer Academic Publishers Group, Dordrecht, 1988. Translated from the Russian.

[8] P. Giesl, "Construction of Global Lyapunov Functions Using Radial Basis Functions," vol. 1904 of Lecture Notes in Math., Springer, Berlin, 2007.

[9] P. Giesl and S. F. Hafstein, *Existence of piecewise affine Lyapunov functions in two dimensions*, J. Math. Anal. Appl., **371** (2010), 233–248.

[10] L. Grüne and O. Junge, "Gewöhnliche Differentialgleichungen. Eine Einführung aus der Perspektive der dynamischen Systeme. Bachelorkurs Mathematik," Vieweg Studium, Vieweg+Teubner, Wiesbaden, 2009.

[11] S. F. Hafstein, "An Algorithm for Constructing Lyapunov Functions," vol. 8 of Electron. J. Differential Equ. Monogr., Texas State Univ., Dep. of Mathematics, San Marcos, TX, 2007. Available electronically at `http://ejde.math.txstate.edu/`.

[12] D. Hinrichsen and A. J. Pritchard, "Mathematical Systems Theory I. Modelling, State Space Analysis, Stability and Robustness," Springer-Verlag, Berlin, 2005.

[13] T. A. Johansen, *Computation of Lyapunov functions for smooth nonlinear systems using convex optimization*, Automatica, **36** (2000), 1617–1626.

[14] M. Johansson, "Piecewise Linear Control Systems. A Computational Approach," vol. 284 of Lecture Notes in Control and Inform. Sci., Springer-Verlag, Berlin, 2003.

[15] P. Julián, J. Guivant and A. Desages, *A parametrization of piecewise linear Lyapunov functions via linear programming*, Internat. J. Control, **72** (1999), 702–715. Special issue on Multiple model approaches to modelling and control.

[16] B. Kummer, *Newton's method for nondifferentiable functions*, in "Advances in mathematical optimization," Akademie-Verlag, Berlin, (1988), 114–125.

[17] G. Leoni, "A First Course in Sobolev Spaces," vol. 105 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2009.

[18] D. Liberzon, "Switching in Systems and Control," Systems & Control: Foundations & Applications, Birkhäuser Boston Inc., Boston, MA, 2003.

[19] S. Marinósson, *Lyapunov function construction for ordinary differential equations with linear programming*, Dyn. Syst., **17** (2002), 137–150.

[20] I. P. Natanson, "Theory of Functions of a Real Variable," Frederick Ungar Publishing Co., New York, 1955. Translated by L. F. Boron with the collaboration of E. Hewitt.

[21] E. P. Ryan, *An integral invariance principle for differential inclusions with applications in adaptive control*, SIAM J. Control Optim., **36** (1998), 960–980 (electronic).

[22] S. Scholtes "Introduction to piecewise differentiable equations," habilitation thesis, Universität Karlsruhe, Institut für Statistik und Mathematische Wirtschaftstheorie, Karlsruhe, Germany, May 1994. Preprint no. 53/1994.

[23] D. Stewart, *A high accuracy method for solving ODEs with discontinuous right-hand side*, Numer. Math., **58** (1990), 299–328.

[24] A. R. Teel and L. Praly, *A smooth Lyapunov function from a class-$\mathcal{KL}$ estimate involving two positive semidefinite functions*, ESAIM Control Optim. Calc. Var., **5** (2000), 313–367 (electronic).

[25] H. Whitney, *Analytic extensions of differentiable functions defined in closed sets*, Trans. Amer. Math. Soc., **36** (1934), 63–89.

E-mail address: `robert.baier@uni-bayreuth.de`
E-mail address: `lars.gruene@uni-bayreuth.de`
E-mail address: `sigurdurh@ru.is`