Computing Lyapunov functions for strongly asymptotically stable differential inclusions

R. Baier * **L. Grüne** * **S. F. Hafstein** **

* *Chair of Applied Mathematics, University of Bayreuth, D-95440 Bayreuth, Germany, (e-mail: robert.baier, lars.gruene@uni-bayreuth.de).*
** *School of Science and Engineering, Reykjavík University, Menntavegur 1, 101 Reykjavík, Iceland, (e-mail: sigurdurh@ru.is).*

Abstract: We present a numerical algorithm for computing Lyapunov functions for a class of strongly asymptotically stable nonlinear differential inclusions which includes switched systems and systems with uncertain parameters. The method relies on techniques from nonsmooth analysis and linear programming and leads to a piecewise affine Lyapunov function. We provide a thorough analysis of the method and present two numerical examples.

Keywords: Lyapunov methods; stability; numerical methods

1. INTRODUCTION

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

For nonlinear control systems, which can be seen as a parametrized version of the differential inclusions considered in this paper, a numerical approach for computing Lyapunov functions characterizing robust or strong stability has been presented in Camilli et al. [2001] using Hamilton-Jacobi-Bellman equations. However, this method computes a numerical approximation of a Lyapunov function rather than a Lyapunov function itself. A method for numerically computing real Lyapunov functions — even smooth ones — has been presented in detail in Giesl [2007], however, this method is designed for differential equations and it is not clear whether it can be extended to control systems or differential inclusions. Inclusions can be addressed by LMI optimization techniques as in Chesi [2004], however, this approach is restricted to systems with polynomial right hand sides.

In this paper we extend a linear programming based algorithm for computing Lyapunov functions from Marinósson [2002] and Hafstein [2007] to general nonlinear differential inclusions with polytopic right hand sides. This class of inclusions includes switched systems as well as nonlinear differential equations with uncertain parameters. The resulting piecewise linear functions are true Lyapunov functions in a suitable nonsmooth sense.

The paper is organized as follows. After giving necessary background results in the ensuing Sections 2 and 3, we present and rigorously analyze our algorithm in Section 4 and illustrate it by two numerical examples in Section 5.

2. NOTATION AND PRELIMINARIES

We consider a compact set $G \subset \mathbb{R}^n$ which is divided into M closed subregions $\mathcal{G} = \{G_\mu \,|\, \mu = 1, \dots, M\}$ with $\bigcup_{\mu=1,\dots,M} G_\mu = G$. For each $x \in G$ we define the *active index set* $I_\mathcal{G}(x) := \{\mu \in \{1, \dots, M\} \,|\, x \in G_\mu\}$.

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

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

where "co" denotes the convex hull. A solution of (1) is an absolutely continuous functions $x : I \to G$ satisfying $\dot{x}(t) \in F(x(t))$ for almost all $t \in I$, where I is an interval of the form $I = [0, T]$ or $I = [0, \infty)$.

To guarantee the existence of solutions of (1), upper semicontinuity is an essential assumption.

Definition 1. A set-valued map $F : G \rightrightarrows \mathbb{R}^n$ is called upper semicontinuous if for any $x \in G$ and any $\epsilon > 0$ there exists $\delta > 0$ such that $x' \in B_\delta(x) \cap G$ implies $F(x') \subseteq F(x) + B_\epsilon(0)$.

Lemma 3 in § 2.6 in Filippov [1988] shows upper semicontinuity of $F(\cdot)$ in (1) for pairwise disjoint subregions. The proof is based on the closedness of the graph and can be generalized to the overlapping regions in our setting.

Two important special cases of (1) are outlined in the following examples.

Example 2. **(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

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

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

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. Filippov [1988], are often used. To this end (2) is replaced by its Filippov regularization, i.e. by the differential inclusion

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

where μ is the Lebesgue measure and $N \subset \mathbb{R}^n$ an arbitrary set of measure zero. It is well-known (see e.g. § 2.7 in Filippov [1988] and Stewart [1990]) that if the number of the sets H_μ is finite and each H_μ satisfies $\overline{H}_\mu = \overline{\mathrm{int}\, H_\mu}$, then the inclusion (3) coincides with (1) if we define $G_\mu := \overline{H}_\mu$ and extend each f_μ continuously to G_μ.

An important subclass of switched systems are piecewise affine systems in which each f_μ in (2) is affine, i.e., $f_\mu(x) = A_\mu x + b_\mu$, see, e.g., Johansson [2003], Liberzon [2003].

Example 3. **(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 $F(x) = \mathrm{co}\,\{f_\mu(x) \,|\, \mu = 1, \ldots, M\}$ with a bounded 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

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

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 4. The 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 $t \to \infty$

If these properties hold, then the *domain of attraction* is defined as the maximal subset of $\mathbb{R}^n$ for which convergence holds, i.e. $D := \{x_0 \,|\, \lim_{t\to\infty} x(t) = 0$ for every solution with $x(0) = x_0\}$.

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 $\mathcal{T} = \{T_\nu \,|\, \nu = 1, \ldots, N\}$, 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 a face of T_{ν_2}. For each $x \in G$ we define the *active index set* $I_{\mathcal{T}}(x) := \{\nu \in \{1, \ldots, N\} \,|\, x \in T_\nu\}$. Let us denote by $\mathrm{diam}(T_\nu) = \max_{x,y\in T_\nu} \|x - y\|$ 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., $\nabla V_\nu := \nabla V|_{\mathrm{int}\, T_\nu} \equiv const$ 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 either $G_\mu \cap T_\nu$ is empty or of the form $\mathrm{co}\,\{x_{j_0}, x_{j_1}, \ldots, x_{j_k}\}$, where $x_{j_0}, x_{j_1}, \ldots, x_{j_k}$ are pairwise disjoint vertices of T_ν and $0 \le k \le n$, i.e. $G_\mu \cap T_\nu$ is a (lower dimensional) 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, cf. Section 2.1 in Clarke [1990], which we introduce for arbitrary Lipschitz continuous functions. We first introduce the corresponding directional derivative.

Definition 5. For a given function $W : \mathbb{R}^n \to \mathbb{R}$ and $l, x \in \mathbb{R}^n$, Clarke's directional derivative of W at x in direction l is defined as

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

Using Clarke's directional derivative as support function, we can state the definition of Clarke's subdifferential

Definition 6. For a locally Lipschitz function $W : \mathbb{R}^n \to \mathbb{R}$ and $x \in \mathbb{R}^n$ Clarke's subdifferential is defined as

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

Theorem 2.5.1 in Clarke [1990] yields the following alternative representation of ∂_{Cl} via limits of gradients.

Proposition 7. Clarke's subdifferential for a Lipschitz continuous function $W : G \to \mathbb{R}$ satisfies

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

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

$$\max \langle \partial_{\mathrm{Cl}} V(x), F(x)\rangle \le -\alpha(\|x\|) \quad (4)$$

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

$$\langle \partial_{\mathrm{Cl}} V(x), F(x)\rangle := \{\langle d, v\rangle \,|\, d \in \partial_{\mathrm{Cl}} V(x),\ v \in F(x)\}. \quad (5)$$

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

$$\max \langle \partial_{\mathrm{Cl}} V(x), F(x)\rangle \le -\|x\| \quad (6)$$

holds for all $x \in G$ with $\|x\| \ge \varepsilon$. Note, however, that even with a nonlinear rescaling of V it may not be possible to obtain (6) for all $x \in G$.

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

It is well known that the existence of a Lyapunov function in the sense of Definition 8 guarantees asymptotic stability of (1) as shown in the following Theorem, cf. also Ryan [1998].

Theorem 9. Consider a Lipschitz continuous function $V : G \to \mathbb{R}$ and F from (1) satisfying (4) and let $x(t)$ be a solution of (1). Then the inequality

$$V(x(t)) \le V(x(0)) - \int_0^t \alpha(\|x(\tau)\|)d\tau \qquad (7)$$

holds for all $t \ge 0$ 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

$$D = \{x_0 \,|\, \lim_{t\to\infty} x(t) = 0 \text{ for a solution with } x(0) = x_0\}$$

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 $0 \in \operatorname{int} C$ and $C \subset \operatorname{int} G$.

Proof. Using an argument similar to Filippov [1988] (Chapter 3, § 15, (8)), see also Baier et al. [2010], we obtain that that $t \mapsto (V \circ x)(t)$ is absolutely continuous and satisfies

$$\frac{d}{dt}(V \circ x)(t) \le \max\langle \partial_{\mathrm{Cl}} V(x(t)), F(x(t))\rangle \le -\alpha(\|x(t)\|)$$

for almost all $t \ge 0$ 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 (7). Asymptotic stability, i.e., properties (i) and (ii) of Definition 4 can now be concluded by classical Lyapunov function arguments as in Theorem 3.2.7 from Hinrichsen and Pritchard [2005]. The full proof is included in Baier et al. [2010].

Remark 10. A different concept of nonsmooth Lyapunov functions was presented in Bacciotti and Ceragioli [1999]. In this reference, in addition to Lipschitz continuity, the function V is also assumed to be regular in the sense of Definition 2.3.4 in Clarke [1990], i.e. the usual directional derivative in Definition 5 exists for every direction l and coincides with Clarke's directional derivative. Under this additional condition, inequality (4) can be relaxed to

$$\max \dot{\overline{V}}(x) \le -\alpha(\|x\|) \qquad (8)$$

with

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

Note that this is indeed a relaxation of (4), cf. Example 18 below, however, both the relaxed inequality (8) as well as the regularity assumption on V are difficult to be implemented algorithmically, which is why we use (4).

The sufficient condition for (4) involves Clarke's subdifferential of a piecewise linear function. The following Lemma is proved in Kummer [1988], Proposition 4.

Lemma 11. Clarke's generalized gradient of $V \in PL(\mathcal{T})$ is given by $\partial_{\mathrm{Cl}} V(x) = \operatorname{co}\{\nabla V_\nu \,|\, \nu \in I_{\mathcal{T}}(x)\}$.

Now we can simplify the sufficient condition (4) for the particular structure of F in (1).

Proposition 12. Consider $V \in PL(\mathcal{T})$ and F from (1). Then for any $x \in G$ the inequality

$$\langle \nabla V_\nu, f_\mu(x)\rangle \le -\alpha(\|x\|) \qquad (9)$$

for all $\mu \in I_{\mathcal{G}}(x)$ and $\nu \in I_{\mathcal{T}}(x)$ implies (4).

Proof. From Lemma 11 we know that each $d \in \partial_{\mathrm{Cl}} V(x)$ can be written as a convex combination of active gradients ∇V_ν, $\nu \in I_{\mathcal{T}}(x)$.

Moreover, by the definition of F in (1) each $v \in F(x)$ can be written as a convex combination of active function values $f_\mu(x)$, $\mu \in I_{\mathcal{G}}(x)$. The scalar product $\langle d, v\rangle$ with the convex combinations can now be estimated by (9).

We end this section by stating a theorem which ensures that Lyapunov functions — even smooth ones — always exist for asymptotically stable inclusions. It follows immediately from Theorem 1 in Teel and Praly [2000] setting $\alpha(r) := \min\{V(x) \,|\, \|x\| = r\}$.

Theorem 13. If the differential inclusion (1) is asymptotically stable with domain of attraction D, then there exists a C^∞-Lyapunov function $V : D \to \mathbb{R}$.

The theorem in particular implies that if we choose our computational domain (which will again be denoted by G in what follows) as a subset of D, then we can expect to find a function V defined on the whole set G.

4. THE ALGORITHM

In this section we present an algorithm for computing Lyapunov functions in the sense of Definition 8 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 Marinósson [2002] and further developed in Hafstein [2007]. 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 $x \in G$, $x \ne x_i$, 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

$$V(x_i) \ge \|x_i\|. \qquad (10)$$

In order to ensure (4), we demand that for every k-face $T = \operatorname{co}\{x_{j_0}, x_{j_1}, \dots, x_{j_k}\}$, $0 \le k \le n$, of a simplex $T_\nu = \operatorname{co}\{x_0, x_1, \dots, x_n\} \in \mathcal{T}$ and every vector field f_μ that is defined on this k-face, the inequalities

$$\langle \nabla V_\nu, f_\mu(x_{j_i})\rangle + A_{\nu\mu} \le -\|x_{j_i}\| \quad \text{for } i = 0, 1, \dots, k. \quad (11)$$

Here, $A_{\nu\mu}$ is an appropriate constant which is chosen in order to compensate for the interpolation error in the points $x \in T$ with $x \ne x_{j_i}$, $i = 0, \dots, k$. being not a vertex of a simplex. Proposition 14, below, will show that the constants $A_{\nu\mu}$ can be chosen such that the condition (11) for $x_{j_0}, x_{j_1}, \dots, x_{j_k}$ ensures

$$\langle \nabla V_\nu, f_\mu(x)\rangle \le -\|x\| \qquad (12)$$

for every $x \in T = \operatorname{co}\{x_{j_0}, x_{j_1}, \dots, x_{j_k}\}$.

Let us illustrate the condition (11) with the 2D-example in Fig. 1, where for simplicity of notation we set $A_{\nu\mu} = 0$. Assume that $T_1 = \operatorname{co}\{x_1, x_2, x_3\}$ and $T_2 = \operatorname{co}\{x_2, x_3, x_4\}$ as well as $T_\nu \subset G_\nu$ and $T_\nu \ne G_\nu$, $\nu = 1, 2$.

Since T_1 and T_2 have the common 1-face $T_1 \cap T_2 = \operatorname{co}\{x_2, x_3\}$, (11) leads to the following inequalities:

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

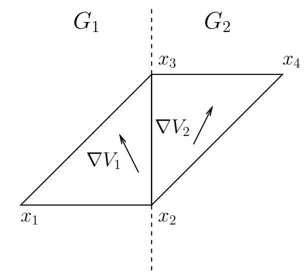

Fig. 1. Gradient conditions (11) for two adjacent simplices

Now we turn to the investigation of the interpolation error on our simplicid grids. In the following proposition we state bounds for the interpolation error for the linear interpolation of C^2-vector fields which follow immediately from the Taylor expansion. Assertion (i) is standard but is stated here in a form which is suitable for (ii), in which we derive an expression for $A_{\nu\mu}$ in (11) which ensures that (12) holds.

Proposition 14. (i) Let $x_0, x_1, \dots, x_k \in \mathbb{R}^n$ be affinely independent vectors and define $T := \mathrm{co}\,\{x_0, x_1, \dots, x_k\}$. Let $\mathcal{U} \subseteq \mathbb{R}^n$ be an open set, $T \subset \mathcal{U}$, and let $f \in \mathcal{C}^2(\mathcal{U})$. Define B as a bound

$$\max_{\substack{z\in T\\ r,s=1,2,\dots,n}} \left|\frac{\partial^2 f}{\partial x_r \partial x_s}(z)\right| \leq B. \qquad (13)$$

for the Hessian of every $f = f_i$, $i = 1, \dots, n$ and $h := \mathrm{diam}(T)$. Then

$$\left| f(\sum_{i=0}^{k} \lambda_i x_i) - \sum_{i=0}^{k} \lambda_i f(x_i) \right| \leq nBh^2$$

for every convex combination $\sum_{i=0}^{k} \lambda_i x_i \in T$.
(ii) If (11) holds with $f_\mu = f$ and $nBh^2 \|\nabla V_\nu\|_1 \leq A_{\nu\mu}$, then (12) holds.

Proof. We only prove (ii). If (11) holds for $f_\mu = f$ and $nBh^2 \|\nabla V_\nu\|_1 \leq A_{\nu\mu}$, then we obtain with Hölder's inequality and (i)

$$\begin{aligned} &\langle \nabla V_\nu, f_\mu(x)\rangle = \left\langle \nabla V_\nu, f(\textstyle\sum_{i=1}^{k} \lambda_i x_i)\right\rangle \\ &\leq \textstyle\sum_{i=1}^{k} \lambda_i \langle \nabla V_\nu, f(x_i)\rangle \\ &\quad + \|\nabla V_\nu\|_1 \left\| f(\textstyle\sum_{i=1}^{k} \lambda_i x_i) - \sum_{i=1}^{k} \lambda_i f(x_i)\right\|_\infty \leq -\|x\|. \end{aligned}$$

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 $f(x)$ and the same convex combination of the values of $f(x_i)$ at the vertices is bounded by the corresponding convex combination of error terms, which are small if the simplex is small.

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 (12) and thus (11) may not be feasible near the origin, cf. also the discussion on $\alpha(\|x\|)$ after Definition 8. This is also reflected in the proof of Theorem 16, below, in which we will need a positive distance to the equilibrium at zero.

To accomplish this fact, for $\varepsilon > 0$ we define the subset $\mathcal{T}^\varepsilon := \{T_\nu \in \mathcal{T} \,|\, T_\nu \cap B_\varepsilon(0) = \emptyset\} \subset \mathcal{T}$. Furthermore, if f_μ is defined on a simplex T with $T := \mathrm{co}\,\{x_0, x_1, \dots, x_k\}$, we assume that f_μ possesses a $\mathcal{C}^2$-extension $\overline{f}_\mu : \mathcal{U} \to \mathbb{R}^n$ on an open set $\mathcal{U} \supset T$ which preserves the bound (13).

Algorithm 1.

(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* $C_{\nu,i}$, $i = 1, \dots, n$ and demand that for the i-th component $\nabla V_{\nu,i}$ of ∇V_ν we have
$$|\nabla V_{\nu,i}| \leq C_{\nu,i}, \quad i = 1, \dots, n.$$

(iii) For every $T_\nu := \mathrm{co}\,\{x_0, x_1, \dots, x_n\} \in \mathcal{T}^\varepsilon$, every k-face $T = \mathrm{co}\,\{x_{j_0}, x_{j_1}, \dots, x_{j_k}\}$, $k = 0, \dots, n$, and every μ with $T \subseteq G_\mu$ (recall that by assumption this implies that f_μ is defined on an open set $\mathcal{U} \supset T$) we demand
$$\langle \nabla V_\nu, f_\mu(x_{j_i})\rangle + nB_{\mu,T} h_\nu^2 \sum_{j=1}^{n} C_{\nu,j} \leq -\|x_{j_i}\| \quad (14)$$
for each $i = 0, \dots, k$ with $h_\nu := \mathrm{diam}(T_\nu)$ and $\max_{i,r,s=1,2,\dots,n} \sup_{z\in T} \left|\frac{\partial^2 \overline{f}_{\mu,i}}{\partial x_r \partial x_s}(z)\right| \leq B_{\mu,T}$.
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 (14) 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 resulting linear optimization problem has $nN + P$ variables, where P is the number of different nodes x_j in the triangulation (see (i)) and N is the number of simplices (see (ii)). Due to (i)–(iii), it possesses $P + (N(M+1)+M)n$ inequalities. Here, P is proportional to $1/h^n$ with h being the maximal diameter of the triangulation.

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

Theorem 15. Assume that 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 (9) holds with $\alpha(r) = r$, i.e., $\langle \nabla V_\nu, f_\mu(x)\rangle \leq -\|x\|$ for all $\mu \in I_{\mathcal{G}}(x)$ 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 $x_{j_0}, x_{j_1}, \dots, x_{j_k}$, $0 \leq k \leq n$. 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 (11) holds on T with
$$A_{\nu\mu} = nB_{\mu,T} h_\nu^2 \textstyle\sum_{j=1}^{n} C_{\nu,j} \geq B_{\mu,T} h_\nu^2 \|\nabla V_\nu\|_1.$$
Thus, Proposition 14(ii) yields the assertion.

The next theorem will show, that if (1) possesses a Lyapunov function then Algorithm 1 can compute a Lyapunov function $V \in PL(\mathcal{T}^\varepsilon)$ for a suitable triangulation $\mathcal{T}^\varepsilon$.

Theorem 16. Assume that 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.

Note 17. The precise conditions on the triangulation are given in the formula (18) of the proof. The triangulation must ensure that each triangle has a sufficiently small diameter $h = \operatorname{diam}(T_\nu)$ and fulfills an angle condition via X^* to prevent too flat triangles. If the simplices $T_\nu \in \mathcal{T}$ are all similar, then it suffices to assume that $\max_{\nu=1,2,\dots,N} \operatorname{diam}(T_\nu)$ is small enough, cf. Theorem 8.2 and Theorem 8.4 in Hafstein [2007].

Proof of Theorem 16.

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

$$c_0 := \max_{x \in G\setminus B_\varepsilon(0)} \frac{\|x\|}{W^*(x)} \text{ and } c_\mu := \max_{G_\mu \setminus B_\varepsilon(0)} \frac{-2\|x\|}{\langle \nabla W^*(x), f_\mu(x)\rangle}$$

for every $\mu = 1,2,\dots,M$. We set $c = \max_{\mu=0,1,\dots,M} c_\mu$ 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 $\mu = 1,2,\dots,M$ 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 $T_\nu = \operatorname{co}\{x_0, x_1, \dots, x_n\} \in \mathcal{T}^\varepsilon$ 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 $x_1 - x_0, x_2 - x_0, \dots, x_n - x_0$ as row vectors consecutively, i.e.

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

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

The matrix $X_{\nu,y}$ is independent of the order of $x_0, x_1, \dots, x_n$ and thus, well-defined, see Baier et al. [2010]. Let us define

$$X^*_\nu = \max_{y \text{ vertex of } T_\nu} \|X^{-1}_{\nu,y}\|_2 \text{ and } X^* = \max_{\nu=1,2,\dots,N} X^*_\nu. \tag{15}$$

(iii) By Whitney's extension theorem in Whitney [1934] we can extend W to an open set containing G. For every $i = 1,2,\dots,n$ we have by Taylor's theorem

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

where H_W is the Hessian of W and $z_i = x_0 + \vartheta_i(x_i - x_0)$ for some $\vartheta_i \in]0,1[$. The following equality holds, if we define $w_\nu := (W(x_1) - W(x_0), \dots, W(x_n) - W(x_0))^T$:

$$w_\nu - X_{\nu,y}\nabla W(x_0) = \frac{1}{2}\begin{pmatrix} \langle x_1 - x_0, H_W(z_1)(x_1 - x_0)\rangle \\ \vdots \\ \langle x_n - x_0, H_W(z_n)(x_n - x_0)\rangle \end{pmatrix} =: \frac{1}{2}\xi_w \tag{16}$$

Setting

$$A := \max_{\substack{z \in T \\ i,j=1,2,\dots,n}} \left| \frac{\partial^2 W}{\partial x_i \partial x_j}(z)\right| \text{ and } h := \max_{\nu=1,2,\dots,N} \operatorname{diam}(T_\nu),$$

we apply Proposition 14 and get the bound $\|\xi_{w,i}\| \leq nAh^2$ for $i = 1,2,\dots,n$. Hence,

$$\|\xi_w\| \leq n^{\frac{3}{2}} A h^2. \tag{17}$$

Furthermore, for every $i,j = 1,2,\dots,n$ there is a $\tilde{z}_i$ on the line segment between x_i and 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, $\|\nabla W(x_i) - \nabla W(x_0)\|_2 \leq nAh$ follows by Proposition 14. From this we obtain the inequality

$$\begin{aligned} &\|X^{-1}_{\nu,y} w_\nu - \nabla W(x_i)\|_2 \\ &\leq \|X^{-1}_{\nu,y} w_\nu - \nabla W(x_0)\|_2 + \|\nabla W(x_i) - \nabla W(x_0)\|_2 \\ &\leq \frac{1}{2}\|X^{-1}_{\nu,y}\|_2 n^{\frac{3}{2}} A h^2 + nAh \leq nAh(\frac{1}{2}X^* n^{\frac{1}{2}} h + 1) \end{aligned}$$

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

(iv) Define $D := \max_{\mu=1,2,\dots,M} \sup_{z \in G_\mu \setminus\{0\}} \|f_\mu(z)\|_2 / \|z\|$. 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, $D \leq \alpha L$ with $\|z\|_2 \leq \alpha\|z\|$.

(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 variable $C_{\nu,i}$ in the linear programming problem from the algorithm set $C_{\nu,i} = C := \max_{z \in G} \|\nabla W(z)\|_2$ and for every $T_\nu \in \mathcal{T}^\varepsilon$ and every vertex x_i of T_ν set $V(x_i) = W(x_i)$. By doing this, we have assigned values to all variables of the linear programming problem.

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 and just as clearly $|\nabla V_{\nu,i}| \leq C_{\nu,i}$ for all $T_\nu \in \mathcal{T}$ and all $i = 1,2,\dots,n$.

Pick an arbitrary $T_\nu \in \mathcal{T}^\varepsilon$. Then, by the definition of w_ν and $X_{\nu,y}$, we have $\nabla V_\nu = X^{-1}_{\nu,y} w_\nu$. 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 $\operatorname{co}\{x_0, x_1, \dots, x_k\}$, $0 \leq k \leq n$, where the x_i are vertices of T_ν. Then, by (ii) and (16)–(17), we have for every $i = 0,1,\dots,k$ that

$$\begin{aligned} &\langle \nabla V_\nu, f_\mu(x_i)\rangle \\ &= \langle \nabla W(x_i), f_\mu(x_i)\rangle + \langle X^{-1}_{\nu,y} w_\nu - \nabla W(x_i), f_\mu(x_i)\rangle \\ &\leq -2\|x_i\| + \|X^{-1}_{\nu,y} w_\nu - \nabla W(x_i)\|_2 \, \|f_\mu(x_i)\|_2 \\ &\leq -2\|x_i\| + nAh(\frac{1}{2}X^* n^{\frac{1}{2}} h + 1) \cdot D\|x_i\|. \end{aligned}$$

The constraints $\langle \nabla V_\nu, f_\mu(x_i)\rangle + nB_{\nu\mu} h_\nu^2 \sum_{j=1}^n C_{\nu,j} \leq -\|x_i\|$ are therefore fulfilled whenever h is so small that

$$-2\|x_i\| + n^2 B h^2 C + nAh(\tfrac{1}{2}X^* n^{\frac{1}{2}} h + 1) D\|x_i\| \leq -\|x_i\|,$$

with X^* given in (15) and $B := \max_{\substack{\mu=1,2,\dots,M \\ \nu=1,2,\dots,N}} B_{\mu\nu}$. Because $\|x_i\| \geq \varepsilon$, this inequality is satisfied if

$$n^2 B \tfrac{h^2}{\varepsilon} C + nAh(X^* n^{\frac{1}{2}} h + 1) D \leq 1. \tag{18}$$

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

5. EXAMPLES

We illustrate our algorithm by two examples, the first one is taken from Bacciotti and Ceragioli [1999].

Example 18. **(Nonsmooth harmonic oscillator with nonsmooth friction)** Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be given by

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

with $\operatorname{sgn} x_i = 1$, $x_i \geq 0$ and $\operatorname{sgn} x_i = -1$, $x_i < 0$. 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 regularization is of type (1). In Bacciotti and Ceragioli [1999] 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 10. It is, however, not a Lyapunov function in the sense of our Definition 8. 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 $F(x) = \operatorname{co}\{\binom{-3/2}{1}, \binom{-1/2}{-1}\}$ and for $\partial_{\mathrm{Cl}} V$ we get $\partial_{\mathrm{Cl}} V(x) = \operatorname{co}\{\binom{1}{1}, \binom{-1}{1}\}$. This implies $\max \langle \partial_{\mathrm{Cl}} V(x), F(x) \rangle \geq \langle \binom{-1}{1}, \binom{-3/2}{1} \rangle = 5/2 > 0$ which shows that (4) 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.

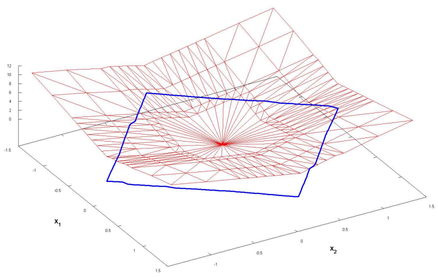

Fig. 2. Lyapunov function and level set for Example 18

There are three facts worth noting. First, we can set the error terms $B_{\mu,T} = 0$ for any triangulation fulfilling the conditions of Theorem 16 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 (6) at the origin. This is because $F(0,0)$ 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. Note that this does not happen if $F(0) = \{0\}$. Finally, 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 Bacciotti and Ceragioli [1999]. 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.

Example 19. **(pendulum with uncertain friction)** Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be given with

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

g is the earth gravitation and equals approximately $9.81 m/s^2$, 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

$$\dot{x}(t) \in F(x(t)) = \operatorname{co}\{f_\mu(x(t)) \mid \mu = 1, 2\}. \tag{19}$$

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 3. Algorithm 1 succeeds in computing a Lyapunov function, see Fig. 3. In fact, here the algorithm works even for $\varepsilon = 0$, cf. Baier et al. [2010].

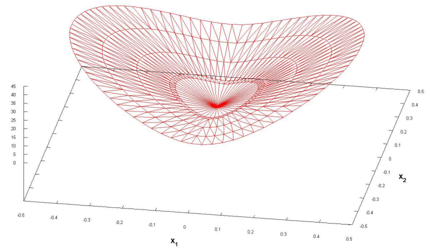

Fig. 3. Lyapunov function for Example 19

REFERENCES

A. Bacciotti and F. Ceragioli. Stability and stabilization of discontinuous systems and nonsmooth Lyapunov functions. *ESAIM Control Optim. Calc. Var.*, 4:361–376, 1999.

R. Baier, Grüne, and S. F. Hafstein. Linear programming based lyapunov function computation for differential inclusions, 2010. submitted.

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*, volume 258 of *LNCIS*, pages 277–289, London, 2001. Springer.

G. Chesi. Estimating the domain of attraction for uncertain polynomial systems. *Automatica*, 40(11):1981–1986, 2004.

F. H. Clarke. *Optimization and nonsmooth analysis.* SIAM, Philadelphia, PA, 1990.

A. F. Filippov. *Differential equations with discontinuous righthand sides.* Kluwer Academic Publishers Group, Dordrecht, 1988.

P. Giesl. *Construction of global Lyapunov functions using radial basis functions*, volume 1904 of *Lecture Notes in Mathematics.* Springer, Berlin, 2007.

S. F. Hafstein. *An algorithm for constructing Lyapunov functions*, volume 8 of *Electron. J. Differential Equ. Monogr.* Texas State Univ., Dep. of Mathematics, San Marcos, 2007.

D. Hinrichsen and A. J. Pritchard. *Mathematical systems theory I. Modelling, state space analysis, stability and robustness.* Springer-Verlag, Berlin, 2005.

M. Johansson. *Piecewise linear control systems.* Lecture Notes in Control and Inform. Sci. Springer-Verlag, Berlin, 2003.

B. Kummer. Newton's method for nondifferentiable functions. In *Advances in mathematical optimization*, pages 114–125. Akademie-Verlag, Berlin, 1988.

D. Liberzon. *Switching in systems and control.* Birkhäuser Boston Inc., Boston, MA, 2003.

S. Marinósson. Lyapunov function construction for ordinary differential equations with linear programming. *Dyn. Syst.*, 17(2): 137–150, 2002.

E. P. Ryan. An integral invariance principle for differential inclusions with applications in adaptive control. *SIAM J. Control Optim.*, 36(3):960–980, 1998.

D. Stewart. A high accuracy method for solving ODEs with discontinuous right-hand side. *Numer. Math.*, 58(3):299–328, 1990.

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.*, 5:313–367, 2000.

H. Whitney. Analytic extensions of differentiable functions defined in closed sets. *Trans. Amer. Math. Soc.*, 36(1):63–89, 1934.