Computing Lyapunov functions for strongly asymptotically stable differential inclusions

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

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.

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

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

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

where “co” denotes the convex hull. A solution of (1) is an absolutely continuous functions $x : I \rightarrow G$ 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 $F : G \Rightarrow \mathbb{R}^n$ is called upper semicontinuous if for any $x \in G$ and any $\epsilon > 0$ there exists $\delta > 0$ such that $x \in B_\delta(x) \cap G$ implies $F(x') \subseteq F(x) + B_\delta(0)$.

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 $H_\mu$ and a piecewise defined ordinary differential equations of the form

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

in which $f_\mu : H_\mu \rightarrow \mathbb{R}^n$ 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 $f : G \rightarrow \mathbb{R}^n$ defined by $f(x) := f_\mu(x)$ for $x \in G_\mu$,

2. NOTATION AND PRELIMINARIES

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

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

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

where “co” denotes the convex hull. A solution of (1) is an absolutely continuous functions $x : I \rightarrow G$ 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 $F : G \Rightarrow \mathbb{R}^n$ is called upper semicontinuous if for any $x \in G$ and any $\epsilon > 0$ there exists $\delta > 0$ such that $x \in B_\delta(x) \cap G$ implies $F(x') \subseteq F(x) + B_\delta(0)$.

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 $H_\mu$ and a piecewise defined ordinary differential equations of the form

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

in which $f_\mu : H_\mu \rightarrow \mathbb{R}^n$ 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 $f : G \rightarrow \mathbb{R}^n$ 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
\[ \dot{x}(t) \in F(x(t)) = \bigcap_{\delta > 0} \bigcap_{\mu(X) = 0} \{ f \in (B_\delta(x(t)) \cap G) \setminus N \} \] (3)
where \( \mu \) 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 \( H_\mu \) is finite and each \( H_\mu \) satisfies \( \overline{H}_\mu \subset \bigcap \overline{H}_\mu \), then the inclusion (3) coincides with (1) if we define \( G_\mu := \overline{H}_\mu \) and extend each \( f_\mu \) continuously to \( G_\mu \).

An important subclass of switched systems are piecewise affine systems in which each \( f_\mu \) 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 \( F(x) = \{ f_\mu(x) | \mu = 1, \ldots, M \} \) 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
\[ \dot{x}(t) \in F(x(t)) = \{ f_\mu(x(t)) | \mu = 1, \ldots, M \} \]
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 \( ||x(0)|| \leq \delta \) satisfies \( ||x(t)|| \leq \varepsilon \) for all \( t \geq 0 \)

(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 \( T = \{ T_\nu | \nu = 1, \ldots, N \} \), i.e. each \( T_\nu \) 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 \( T_{\nu_1} \) and \( T_{\nu_2} \). For each \( x \in G \) we define the active index set \( I_T(x) := \{ \nu | \nu \in \{ 1, \ldots, N \} \text{ and } x \in T_\nu \} \). Let us denote by \( \text{diam}(T_\nu) = \max_{x,y \in T_\nu} ||x-y|| \) the diameter of a simplex.

Then, by \( PL(T) \) we denote the space of continuous functions \( V : G \to \mathbb{R} \) which are affine on each simplex, i.e., \( \nabla V := \{ \nabla V \vert_{\text{int} T_\nu} \equiv \text{const} \} \) for all \( T_\nu \in T \).

For the algorithm to work properly we need the following compatibility between the subregions \( G_\mu \) and the simplices \( T_\nu \): for every \( \mu \) and every \( \nu \) that either \( G_\mu \cap T_\nu \) is empty or of the form \( \{ x_{j_0}, x_{j_1}, \ldots, x_{j_k} \} \), where \( x_{j_0}, x_{j_1}, \ldots, x_{j_k} \) are pairwise disjoint vertices of \( T_\nu \) and \( 0 \leq k \leq n \), i.e. \( G_\mu \cap T_\nu \) is a \( (n-k) \)-face of \( T_\nu \).

Since the functions in \( PL(T) \) computed by the proposed algorithm are in general nonsmooth, we need a generalized concept for derivatives. In this paper we use Clarke’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 \( x, t \in \mathbb{R}^n \), Clarke’s directional derivative of \( W \) at \( x \) in direction \( t \) is defined as
\[ W_t'(x; l) = \limsup_{h \to 0} \frac{W(y + hl) - W(y)}{h} \]

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
\[ \partial Cl W(x) = \{ d \in \mathbb{R}^n | \forall l \in \mathbb{R}^n : \langle l, d \rangle \leq W_t'(x; l) \} \]

Theorem 2.5.1 in Clarke [1990] yields the following alternative representation of \( \partial Cl W \) via limits of gradients.

**Proposition 7.** Clarke’s subdifferential for a Lipschitz continuous function \( W : G \to \mathbb{R} \) satisfies
\[ \partial Cl W(x) = \{ \lim_{i \to \infty} \nabla W(x_i) | x_i \to x, \nabla W(x_i) \text{ exists for } i \to \infty \} \]

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
\[ \max \langle \partial Cl V(x), F(x) \rangle \leq -\alpha(||x||) \]
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
\[ \langle \partial Cl V(x), F(x) \rangle := \{ \langle d, v \rangle | d \in \partial Cl V(x), v \in F(x) \} \]

Given \( \varepsilon > 0 \), since \( G \) is compact, changing \( V \) to \( \gamma V \) for \( 0 < \gamma \in \mathbb{R} \) sufficiently large we can always assume without loss of generality that
\[ \max \langle \partial Cl V(x), F(x) \rangle \leq -||x|| \]
holds for all \( x \in G \) with \( ||x|| \leq \varepsilon \). 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

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

(7)

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

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

contains every connected component $C \subseteq V^{-1}([0, c])$ of a sublevel set $V^{-1}([0, c]) := \{x \in G \mid V(x) \in [0, c]\}$ for some $c > 0$ which satisfies $0 \in \text{int } C$ and $C \subseteq \text{int } G$.

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

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

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 (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 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

$$\max \tilde{V}(x) \leq -\alpha(\|x\|)$$

(8)

with

$$\tilde{V}(x) := \{a \in \mathbb{R} \mid \text{there exists } v \in F(x) \text{ with } (p, v) = a \text{ for all } p \in \partial_C V(x)\}.$$

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 $V \in PL(T)$ is given by $\partial_C V(x) = \text{co} \{\nabla V_\nu \mid \nu \in I_T(x)\}$.

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

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

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

(9)

for all $\mu \in I_G(x)$ and $\nu \in I_T(x)$ implies (4).

**Proof.** From Lemma 11 we know that each $d \in \partial_C V(x)$ can be written as a convex combination of active gradients $\nabla V_\nu, \nu \in I_T(x)$.

Moreover, by the definition of $F$ in (1) each $v \in F(x)$ can be written as a convex combination of active function values $f_\mu(x), \mu \in I_G(x)$. The scalar product $(d, v)$ 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 $\alpha(r) := \min \{V(x) \mid \|x\| = r\}$.

**Theorem 13.** If the differential inclusion (1) is asymptotically stable with domain of attraction $D$, then there exists a $C^\infty$-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 $T_\nu \in T$ 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

$$V(x_i) \geq \|x_i\|.$$ (10)

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

$$\langle \nabla V_\nu, f_\mu(x_{j_i}) \rangle + A_{\nu, \mu} \leq -\|x_{j_i}\|$$

for $i = 0, 1, \ldots, k$. (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 $x \neq x_{j_i}$, $i = 0, \ldots, k$, 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 $x_{j_0}, x_{j_1}, \ldots, x_{j_k}$ ensures

$$\langle \nabla V_\nu, f_\mu(x) \rangle \leq -\|x\|$$

(12)

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

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 $T_1 = \{x_1, x_2, x_3\}$ and $T_2 = \{x_2, x_3, x_4\}$ as well as $T_\nu \subseteq G_\nu$ and $T_\nu \neq G_\nu, \nu = 1, 2$.

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

$$\langle \nabla V_1, f_1(x) \rangle \leq -\|x\| \quad \text{for every } x \in \{x_1, x_2, x_3\} \subseteq T_1,$$

$$\langle \nabla V_2, f_1(x) \rangle \leq -\|x\| \quad \text{for every } x \in \{x_2, x_3, x_4\} \subseteq T_2,$$

$$\langle \nabla V_1, f_2(x) \rangle \leq -\|x\| \quad \text{for every } x \in \{x_2, x_3\} \subseteq T_1 \cap T_2,$$

$$\langle \nabla V_2, f_2(x) \rangle \leq -\|x\| \quad \text{for every } x \in \{x_2, x_3\} \subseteq T_1 \cap T_2.$$
Now we turn to the investigation of the interpolation error on our simplicial 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 $A_{\nu m}$ in (11) which ensures that (12) holds.

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

$$\max_{r,s=1,2,\ldots,n} \left| \frac{\partial^2 f}{\partial x_r \partial x_s}(z) \right| \leq B.$$  

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

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

for every convex combination $\sum_{i=0}^{k} \lambda_i x_i \in T$.

(ii) If (11) holds with $f_\mu = f$ and $nBh^2 \| \nabla V_\mu \|_1 \leq A_{\nu m}$, then (12) holds.

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

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

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 $T_\nu \in T$ close to the equilibrium at zero from $T$. 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 $T^\varepsilon := \{ T_\nu \in T \mid T_\nu \cap B(0, \varepsilon) = \emptyset \} \subset T$. Furthermore, if $f_\mu$ is defined on a simplex $T$ with $T := \text{co} \{x_0, x_1, \ldots, x_k\}$, we assume that $f_\mu$ possesses a $C^2$-extension $\overline{f}_\mu \colon U \to \mathbb{R}^n$ on an open set $U \supset T$ which preserves the bound (13).

**Algorithm 1.**

(i) For all vertices $x_i$ of the simplices $T_\nu \in T^\varepsilon$ 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 $T_\nu \in T^\varepsilon$ we introduce the variables $C_{\nu,i}$, $i = 1, \ldots, n$ and demand that for the $i$-th component $\nabla V_\mu$ of $\nabla V_\nu$ we have

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

for each $i = 0, 1, \ldots, k$ with $h_\nu := \text{diam}(T_\nu)$ and

$$\max_{r,s=1,2,\ldots,n} \sup_{T \in T^\varepsilon} \left| \frac{\partial^2 f}{\partial x_r \partial x_s}(z) \right| \leq B_\nu.$$  

Note, that if $f_\mu$ is defined on the face $T \subset T^\varepsilon$, then $f_\mu$ 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 $T_\nu \in T^\varepsilon$ and the condition $V \in PL(T^\varepsilon)$ uniquely define the function $V : \bigcup_{T_\nu \in T^\varepsilon} T_\nu \to \mathbb{R}$.

The resulting linear optimization problem has $nN + P$ variables, where $P$ is the number of different nodes $x_j$ in the triangulation (see (ii) 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 $1/h^2$ with $h$ being the maximal diameter of the triangulation.

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

**Theorem 15.** Assume that the linear program constructed by the algorithm has a feasible solution. Then, on each $T_\nu \in T^\varepsilon$ the function $V$ from (iv) is positive definite and for every $x \in T_\nu \in T^\varepsilon$ inequality (9) holds with $\alpha(r) = r$, i.e., $\langle \nabla V_\nu, f_\mu(x) \rangle \leq -\| x \|$ for all $\mu \in I_G(x)$ and $\nu \in I_T(x)$.

**Proof.** Let $f_\mu$ be defined on the $k$-face $T = T_\nu \cap G_\mu$ with vertices $x_0, x_1, \ldots, x_k$, $0 \leq k \leq n$. Then every $x \in T$ is a convex combination $x = \sum_{i=0}^{k} \lambda_i x_i$. Conditions (ii) and (iii) of the algorithm imply that (11) holds on $T$ with

$$A_{\nu m} = nB_\nu \| h_\varepsilon \| \sum_{j=1}^{n} C_{\nu,j} \geq B_\nu \| h_\varepsilon \| \| \nabla V_\nu \|_1.$$  

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 $V \in PL(T^\varepsilon)$ for a suitable triangulation $T^\varepsilon$.

**Theorem 16.** Assume that the system (1) possesses a $C^2$-Lyapunov function $W^* : G \to \mathbb{R}$ and let $\varepsilon > 0$. Then, there exists a triangulation $T^\varepsilon$ such that the linear
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 \( b = \text{diam}(T_v) \) and fulfills an angle condition via \( X^* \) to prevent too flat triangles. If the simplices \( T_v \in T \) are all similar, then it suffices to assume that \( \max_{v=1,\ldots,N} \text{diam}(T_v) \) is small enough, cf. Theorem 8.2 and Theorem 8.4 in Haifast [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

\[
c_0 := \max_{x \in G \setminus B_{\varepsilon}(0)} \| W(x) \| \quad \text{and} \quad c_\mu := \max_{G \setminus B_{\varepsilon}(0)} \langle \nabla W(x), f_\mu(x) \rangle
\]

for every \( \mu = 1, \ldots, M \). We set \( c := \max_{\nu=0,1,\ldots,M} c_\mu \) 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 \( \mu = 1, \ldots, M \). We have \( \langle \nabla W(x), f_\mu(x) \rangle \leq -2\| x \| \) for every \( x \in G \setminus B_{\varepsilon}(0) \).

(ii) For every \( T_v \in T \), we have \( \text{co} \{x_0, x_1, \ldots, x_n\} \in T^c \) 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.

\[
X_{\nu,y} = (x_1 - x_0, x_2 - x_0, \ldots, x_n - x_0)^T.
\]

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

The matrix \( X_{\nu,y} \) is independent of the order of \( 0, x_1, \ldots, x_n \) and thus, well-defined, see Baier et al. [2010]. Let us define

\[
X_{\nu}^* = \max_{y \in \text{vertex of } T_v} \| X_{\nu,y}^{-1} \|_2 \quad \text{and} \quad X^* = \max_{\nu=1,2,\ldots,N} X_{\nu}^*.
\]

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

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

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

\[
w_\nu = X_{\nu,y} \nabla W(x_0)
\]

\[
= \frac{1}{2} \begin{pmatrix}
(x_1 - x_0, H_W(z_1)(x_1 - x_0)) \\
\vdots \\
(x_n - x_0, H_W(z_n)(x_n - x_0))
\end{pmatrix} = \frac{1}{2} \xi_{\nu,w}
\]

Setting

\[
A := \max_{i,j=1,2,\ldots,n} \left| \frac{\partial^2 W}{\partial x_i \partial x_j}(z) \right| \quad \text{and} \quad h := \max_{\nu=1,2,\ldots,N} \text{diam}(T_v),
\]

we apply Proposition 14 and get the bound \( \| \xi_{\nu,w} \| \leq n Ah^2 \) for \( i = 1, 2, \ldots, n \). Hence,

\[
\| \xi_{\nu,w} \| \leq n^{\frac{3}{2}} Ah^2.
\]

Furthermore, for every \( i, j = 1, 2, \ldots, n \) there is a \( z_i \) on the line segment between \( x_i \) and \( x_0 \), such that

\[
\partial_j W(x_i) - \partial_j W(x_0) = \langle \nabla \partial_j W(z_i), x_i - x_0 \rangle,
\]

where \( \partial_j W \) denotes the \( j \)-th component of \( \nabla W \). Hence,

\[
\| \nabla W(x_i) - \nabla W(x_0) \|_2 \leq n Ah^2
\]

for every \( i = 0, 1, \ldots, n \). This last inequality is independent of the simplex \( T_v = \{x_0, x_1, \ldots, x_n\} \).

(iv) Define \( D := \max_{i=1,2,\ldots,M} \sup_{x \in G \setminus B_{\varepsilon}(0)} \| \nabla W(x) \|_2 / \| z \| \| z \| \).

Note, that \( D < +\infty \) because all norms on \( \mathbb{R}^n \) are equivalent and for every \( \mu \) the vector field \( f_\mu \) is Lipschitz on \( G \), 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 \( C_{\nu,i} = C := \max_{\nu=1,2,\ldots,M} \| \nabla W(x) \|_2 \) and for every \( T_v \in T^c \) and every vertex \( x_i \) of \( T_v \) 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 \( T_v \in T^c \) and every vertex \( x_i \) of \( T_v \) and just as clearly \( \| V(x_i) \| \leq C_{\nu,i} \) for all \( T_v \in T \) and all \( i = 1, 2, \ldots, n \).

Pick an arbitrary \( T_v \in T^c \). Then, by the definition of \( w_\nu \) and \( X_{\nu,y} \), we have \( \nabla V(x) = X_{\nu,y} w_\nu \). Let \( f_{\nu} \) be an arbitrary vector field defined on the whole of \( T_v \) or one of its faces, i.e. \( f_{\nu} \) is defined on \( \{x_0, x_1, \ldots, x_k\}, 0 \leq k \leq n \), where \( x_i \) are the vertices of \( T_v \). Then, by (ii) and (16)–(17), we have for every \( i = 0, 1, \ldots, k \) that

\[
\langle \nabla V(x), f_{\nu}(x_i) \rangle \leq -\| x_i \| + \| X_{\nu,y}^{-1} w_\nu \|_2 \| \nabla W(x_i) \|_2 \leq -\| x_i \| + n Ah^2 (\frac{1}{2} X^* \| x_i \| + 1) D \| x_i \|.
\]

The constraints \( \langle \nabla V(x), f_{\nu}(x) \rangle + n B_{\alpha_h} h^2 \sum_{i=1}^n C_{\nu,i} \leq \| x_i \| \) are therefore fulfilled whenever \( h \) is so small that

\[
-2\| x_i \| + n^2 B_{\alpha_h} h^2 C + n Ah^2 (\frac{1}{2} X^* \| x_i \| + 1) D \| x_i \| \leq \| x_i \|
\]

with \( X^* \) given in (15) and \( B := \max_{\nu=1,2,\ldots,M} B_{\alpha_h} \).

Because \( \| x_i \| \geq \varepsilon \), this inequality is satisfied if

\[
n^2 B_{\alpha_h} \frac{C}{2} + n Ah \| X^* \| + 1 \leq D \| x_i \| \leq 1.
\]

Since \( T_v \) and \( f_{\nu} \) 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

\[
f(x_1, x_2) = (-\text{sgn} x_2 - \frac{1}{2} \text{sgn} x_1, \text{sgn} x_1, \text{sgn} x_1)^T
\]
with $\sgn x_1 = 1$, $x_1 \geq 0$ and $\sgn x_1 = -1$, $x_1 < 0$. This vector field is piecewise constant on the four regions

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

and 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 $V(x) = \{1,2\}$ and the Filippov regularization $F$ of $f$ is $F(x) = \partial \mathcal{V}(x) = \partial \mathcal{C}(x)$ and for $\partial \mathcal{C}V$ we get $\partial \mathcal{C}V(x) = \partial \mathcal{C}V((\frac{-1}{2}, -\frac{1}{2}))$ and $\partial \mathcal{C}V((\frac{-1}{2}, -\frac{1}{2})) = \mathcal{B} \times \mathcal{B} > 0$ 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 $V$.

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 $f_{\mu}$ vanish in the interiors of the simplexes. 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 $F(0) = \emptyset$. 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

\[
f(x_1,x_2) = (x_2, -kx_2 - g \sin(x_1))
\]

g is the earth gravitation and equals approximately 9.81m/s², $k$ is a nonnegative parameter modelling the friction of the pendulum. It is known that the system is asymptotically 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 $\hat{x}(t) \in F(x(t)) = \{f_\mu(x(t)) | \mu = 1,2\}$, 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 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].