

Unified Framework for the Propagation of Continuous-Time Enclosures for Parametric Nonlinear ODEs

Mario E. Villanueva^a · Boris Houska^{a,b} ·
Benoît Chachuat^{a,c}

Received: date / Accepted: date

Abstract This paper presents a framework for constructing and analyzing enclosures of the reachable set of nonlinear ordinary differential equations (ODEs) using continuous-time set-propagation methods. The focus is on convex enclosures that can be characterized in terms of their support functions. A generalized differential inequality is introduced, whose solutions describe such support functions for a convex enclosure of the reachable set under mild conditions. It is shown that existing continuous-time bounding methods that are based on standard differential inequalities or ellipsoidal set propagation techniques can be recovered as special cases of this generalized differential inequality. A way of extending this approach for the construction of nonconvex enclosures is also described, which relies on Taylor models with convex remainder bounds. This unifying framework provides a means for analyzing the convergence properties of continuous-time enclosure methods. The enclosure techniques and convergence results are illustrated with numerical case studies throughout the paper, including a six-state dynamic model of anaerobic digestion.

Keywords interval analysis · ellipsoidal calculus · Taylor models · ordinary differential equations · differential inequalities · convergence analysis · dynamic optimization · global optimization

1 Introduction

A great variety of natural, engineered and industrial systems are inherently dynamic, including batch processes and processes operating at a cyclic steady state. Other such systems are operated in a transient manner on purpose, for instance in order to improve their versatility or robustness. All these systems have their evolution governed by differential equations and are often uncertain due to the presence of time-varying disturbances or due to incomplete/inaccurate knowledge of certain parameters or initial conditions.

The main focus of this paper is on methods for computing time-varying enclosures of the reachable set of parametric nonlinear ODEs. A principal motivation is for applications in the

^aCentre for Process Systems Engineering, Department of Chemical Engineering, Imperial College London, London SW7 2AZ, UK.

^bDepartment of Automation, Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai 200240, China.

^cCorresponding Author. E-mail: b.chachuat@imperial.ac.uk.

field of dynamic optimization. In the context of global optimization, an over-approximation of the reachable set can be used to construct relaxations of the feasible region and the objective function, which are pivotal in branch-and-bound search and its variants; [see, e.g., 15, 34, 49, 54]. In robust optimization likewise, the enclosure of a system's reachable set can be used to ensure feasibility of that system's response despite the presence of uncertainty [26]. Other applications of reachability analysis are for problems in the fields of robust MPC [32], guaranteed state and parameter estimation [2, 27, 28], and system verification and fault detection [36, 63, 64].

Advances in the field of reachability analysis originate from various fields. In viability theory [5], uncertain dynamics are expressed as differential inclusions. In contrast, comparison principles consider auxiliary sets of ODEs that do not depend on the uncertain variables to enclose the reachable set of an uncertain dynamic system. One such comparison principle is provided by the theory of differential inequalities [31, 67], where the right-hand side of the auxiliary ODEs under/overestimate the minimal/maximal value of the right-hand side of the uncertain dynamic system. Other comparison principles relate the reachable sets of uncertain ODEs to the solutions of Hamilton-Jacobi-Bellman (HJB) and Hamilton-Jacobi-Isaacs (HJI) equations [29, 38, 42].

Contributions to reachability analysis are also found in the field of reliable computing. Verified ODE integration techniques were first developed to compute guaranteed solutions for ODEs with rounding errors as the sole source of uncertainty. In the classical approach, the integration horizon is discretized into finite steps and a two-phase algorithm is applied at each step [21, 45]. Phase I is concerned with the computation of an *a priori* enclosure and a step-size, such that existence and uniqueness of the solution can be ascertained on that domain. This typically relies on a high-order Taylor series expansion of the ODE solutions, in combination with interval analysis [19, 46]. Phase II propagates a tightened enclosure until the end of a given step, for instance based on an interval Taylor series [37] or a more involved interval Hermite-Obreschkoff method [44]. Further developments aiming to reduce conservatism in Phase II include wrapping mitigation strategies [37] as well as the use of Taylor models [39, 48] to mitigate the dependency problem [8, 9, 35, 47]. A hybrid approach combining differential inequalities with the two-phase approach is proposed in [51], which provides a verified implementation of the method of differential inequalities. Another verified implementation featuring a simple Runge-Kutta method in combination with a Piccard iteration scheme is described in [52].

A well-developed theory for enclosing the solution set of linear and hybrid uncertain systems has emanated from the field of control theory. These methods commonly proceed by propagating sets forward in time too [10]. One example is the ellipsoidal bounding approach [30], first developed for linear systems under ellipsoidal uncertainty sets, and later extended to address nonlinear systems as well [26]. This method involves constructing vector- and matrix-valued ODEs, whose solutions are, respectively, the center and the shape matrix of ellipsoids that enclose the reachable set of the uncertain dynamic system, pointwise in time [65]. Other set representations used for bounding reachable sets include polytopes [16] and zonotopes [3].

In the field of global optimization, the construction of convex lower-bounding problems for nonlinear dynamic optimization problems is pivotal to the development of efficient complete search methods. Considerable effort has thus been devoted to obtaining convex estimators for the solution of parametric nonlinear ODEs with respect to the uncertain parameters. For linear dynamic systems, a method was developed in [61] which provides affine bounds in the parameter space via the solution of an auxiliary set of ODEs. This method was later extended to propagate affine bounds [62] as well as a pair of convex and concave bounds

[58, 60] for the solutions of parametric nonlinear ODEs based on differential inequalities and (generalized) McCormick relaxations [41, 59]. The refinement of reachable set enclosures by accounting for *a priori* or physical information about a given system was also investigated in [57, 61]. Another way of computing pointwise-in-time, convex under-estimators and concave over-estimators using the α -BB approach [1, 40] has been described in [13, 49, 50]. Discretization-based methods adapted from the field of reliable computing too have been considered in the context of global optimization. An extension of the two-phase approach based on Taylor models for handling parametric uncertainty was considered in [34, 35], and similar extensions for the propagation of McCormick relaxations and McCormick-Taylor models [12] are described in [55, 56].

An important property in comparing the tightness of the reachable set enclosures constructed with various approaches is how quickly these enclosures converge to the exact reachable set as the size of the uncertainty set vanishes, namely their convergence rate in the Hausdorff sense. This is especially relevant in global optimization, where it has been shown that the convergence order can be related to the so-called cluster effect [20, 68]. The convergence properties of McCormick relaxations and Taylor models as applied to factorable functions have been analyzed in [11, 12]. Regarding parametric differential equations, it is not hard to contrive examples whereby the enclosures computed based on differential inequalities would only enjoy first-order Hausdorff convergence, thereby calling for the development of higher-order techniques.

In sum, existing methods for constructing enclosures of the reachable set of nonlinear ODEs are scattered across different disciplines. A possible classification for all of these methods is whether the enclosures are propagated continuously in time or on a time grid after discretization of the integration horizon. The focus in this paper is on the former class, which relies on the formulation and solution of auxiliary ODEs. Our main contribution is two-fold:

- (i) We introduce a generalized differential inequality (GDI), which provides sufficient conditions for a time-varying support function to describe a convex enclosure of the reachable set (Theorem 3). We show that the GDI encompasses results from both the classical theory of differential inequalities and ellipsoidal set-propagation techniques as special cases (Propositions 2 and 3), in addition to providing a means of constructing auxiliary ODEs that propagate other kinds of convex sets in the state space. We also investigate Taylor models for the propagation of more versatile, nonconvex enclosures whereby the GDI is used to describe a convex enclosure of the Taylor remainder term (Propositions 4 and 5).
- (ii) We analyze the convergence properties of continuous-time enclosure methods based on the proposed GDI (Theorem 4). Here again, we recover the classical, linear and quadratic convergence results for differential inequalities and ellipsoidal set propagation, respectively, as special cases (Corollaries 1 and 2). Moreover, we analyze the convergence rate of those nonconvex set-propagation methods based on Taylor models, showing that the Hausdorff convergence order of these enclosures is $q + 1$ when q th-order Taylor models are considered (Corollary 3).

Throughout the paper, several numerical examples are used to illustrate the construction and convergence rate of various types of enclosures. The underlying implementation is based on the library MC++ [43], a C++ library employing operator overloading for bounding factorable functions. Moreover, the numerical integrator available through the GNU Scientific Library (GSL) is used. The C++ class implementing the various bounding techniques based on intervals, ellipsoids and Taylor models with interval

and ellipsoidal remainders is made freely available at: <http://www3.imperial.ac.uk/environmentenergyoptimisation/software>.

The rest of the paper is organized as follows. The problem formulation and the notation used throughout the paper are defined in Sect. 2, and existing continuous-time set-propagation approaches are reviewed in Sect. 3. The GDI for continuous-time propagation of convex enclosures is introduced in Sect. 4, and this framework is shown to encompass existing continuous-time propagation methods in Sect. 5. This latter section also describes an approach based on Taylor models with convex remainder in order for the GDI to produce nonconvex enclosures. The main convergence results exploiting the proposed GDI framework are presented in Sect. 6. Details about the numerical implementation of the bounding techniques are given in Sect. 7, before presenting the numerical case study of an anaerobic digester. Finally, Sect. 8 concludes the paper.

2 Problem Statement

We consider parametric dynamic systems in the form of nonlinear ODEs

$$\forall t \in [0, T] \quad \dot{x}(t, p) = f(t, x(t, p), p) \quad \text{with} \quad x(0, p) = x_0(p). \quad (1)$$

The state $x : [0, T] \times P \rightarrow \mathbb{R}^{n_x}$ is regarded as a function of the uncertain parameter vector $p \in P \subseteq \mathbb{R}^{n_p}$ along the time horizon $[0, T]$. The reachable set of the initial value problem (1) is denoted by

$$X(t, P) := \{x(t, p) \mid p \in P\}. \quad (2)$$

The problem addressed in this paper is the computation of time-varying enclosures $Y(t, P) \supseteq X(t, P)$, for $t \in [0, T]$, using continuous-time set-propagation techniques. Specifically, these enclosures are constructed from the solutions of auxiliary ODEs. By a slight abuse of notation, we shall denote the reachable set and its enclosure by $X(t)$ and $Y(t)$, respectively, when it is clear from the context what the corresponding parameter host set P is.

Throughout the paper, we consider solutions to these auxiliary ODEs in the *extended sense*, whereby a solution can be any absolutely continuous function satisfying the initial condition and the ODE almost everywhere on $[0, T]$; see, e.g., [18, Theorem 1.1]. More precisely, we constrain ourselves to the case of locally Lipschitz-continuous functions, which are sufficient for most practical purposes.

2.1 Notation and Background

The set of compact subsets of \mathbb{R}^n is denoted by \mathbb{K}^n , and the subset of compact convex subsets of \mathbb{K}^n , by \mathbb{K}_C^n . We define the unit ball $\mathcal{B}^n \in \mathbb{K}_C^n$ as $\mathcal{B}^n := \{z \in \mathbb{R}^n \mid \|z\|_2 \leq 1\}$. The diameter $\text{diam}(Z)$ of a set $Z \in \mathbb{K}^n$ is defined as

$$\text{diam}(Z) := \max_{z_1, z_2 \in Z} \|z_1 - z_2\|,$$

and the support function $V[Z] : \mathbb{R}^n \rightarrow \mathbb{R}$ of Z as

$$\forall c \in \mathbb{R}^n, \quad V[Z](c) := \max_z \{c^T z \mid z \in Z\}. \quad (3)$$

In particular, there exists positive constants $C_1 \leq C_2 < \infty$ such that

$$\forall Z \in \mathbb{K}_{\mathbb{C}}^n, \quad C_1 \text{diam}(Z) \leq \max_{c \in \mathbb{R}^n, \|c\| \leq 1} V[Z](c) + V[Z](-c) \leq C_2 \text{diam}(Z). \quad (4)$$

This result follows from the definition of the support function and the diameter of a set, using the equivalence of norms in finite dimensional vector spaces.

The Minkowski sum of two compact sets $W, Z \in \mathbb{K}^n$ is denoted by

$$W \oplus Z := \{w + z \mid w \in W, z \in Z\},$$

and the Hausdorff distance between W and Z is given by

$$d_H(W, Z) := \max \left\{ \max_{w \in W} \min_{z \in Z} \|w - z\|, \max_{z \in Z} \min_{w \in W} \|w - z\| \right\}. \quad (5)$$

In particular, the Hausdorff distance of any two convex compact sets $W, Z \in \mathbb{K}_{\mathbb{C}}^n$ is bounded by the maximum difference of their support functions

$$d_H(W, Z) \leq \max_{c \in \mathbb{R}^n, \|c\|=1} |V[W](c) - V[Z](c)|. \quad (6)$$

This result follows readily from (5), by noting that the Hausdorff distance $d_H(W, Z) = \|w^* - z^*\|$ at a min-max point [resp. a max-min point] $(w^*, z^*) \in W \times Z$ is bounded by the difference $|V[W](c) - V[Z](c)|$ in the direction $c = w^* - z^*$ [resp. in the direction $c = z^* - w^*$].

The set of n -dimensional interval vectors is denoted by \mathbb{IR}^n . The midpoint and radius of an interval vector $P := [p^L, p^U] \in \mathbb{IR}^n$ are defined as $\text{mid}(P) := \frac{1}{2}(p^U + p^L)$ and $\text{rad}(P) := \frac{1}{2}(p^U - p^L)$, respectively. The n -by- n matrix $\text{diag rad}(P) \in \mathbb{R}^{n \times n}$ is a diagonal matrix whose elements are the components of $\text{rad}(P)$.

The set of n -dimensional positive semi-definite symmetric matrices is denoted by \mathbb{S}_+^n . An ellipsoid with center $c \in \mathbb{R}^n$ and shape matrix $M \in \mathbb{S}_+^n$ is denoted by

$$\mathcal{E}(c, M) := \left\{ c + M^{\frac{1}{2}} v \mid v \in \mathbb{R}^n : v^T v \leq 1 \right\}$$

or, simply, $\mathcal{E}(M)$ if the ellipsoid is centered at the origin.

A q th-order Taylor model of a $(q+1)$ -times continuously-differentiable function $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ on the set $P \in \mathbb{IR}^n$ and at $\hat{p} \in P$ is the pair $(\mathcal{P}_f^q, \mathcal{R}_f^q)$ with $\mathcal{P}_f^q: \mathbb{R}^n \rightarrow \mathbb{R}^m$ a q th-order multivariate polynomial and $\mathcal{R}_f^q \in \mathbb{K}^m$ satisfying

$$\forall p \in P, \quad \mathcal{P}_f^q(p) := \sum_{\substack{\gamma \in \mathbb{N}^n, \\ |\gamma| \leq q}} \frac{\partial^\gamma f(\hat{p})}{\gamma!} (p - \hat{p})^\gamma \quad \text{and} \quad f(p) - \mathcal{P}_f^q(p) \in \mathcal{R}_f^q.$$

This last statement uses multi-index notation. A multi-index γ is a vector in \mathbb{N}^n , $n > 0$. The order of γ is $|\gamma| := \sum_{i=1}^n \gamma_i$. Given a point $p \in \mathbb{R}^n$, p^γ is a shorthand notation for the expression $\prod_{i=1}^n p_i^{\gamma_i}$. Moreover, $\partial^\gamma f$ is a shorthand notation for the partial derivative $\frac{\partial^{|\gamma|} f}{\partial p_1^{\gamma_1} \dots \partial p_n^{\gamma_n}}$.

3 Review of Existing Convex Set Propagation Methods

3.1 Differential Inequalities

The theory of differential inequalities provides sufficient conditions for time-varying interval bounds $Y(t) := [y^L(t), y^U(t)]$ to yield an enclosure of a parametric ODE's reachable set $X(t)$. These conditions can be used for propagating an interval enclosure $Y(t)$ of $X(t)$ along the integration horizon.

Theorem 1 *Consider the initial value problem (1), and assume that the right-hand side function f is jointly continuous in (t, x, p) and locally Lipschitz-continuous in x uniformly on $[0, T] \times P$, with $P \subset \mathbb{R}^{n_p}$ compact, and the initial value function x_0 is continuous on P . Let the functions $y^L, y^U : [0, T] \rightarrow \mathbb{R}^{n_x}$ be Lipschitz-continuous, with $y^L(t) \leq y^U(t)$, and satisfy*

$$\text{a.e. } t \in [0, T], \quad \dot{y}_i^L(t) \leq \min_{\xi, \rho} \left\{ f_i(t, \xi, \rho) \left| \begin{array}{l} \xi_i = y_i^L(t) \\ \xi \in [y^L(t), y^U(t)] \\ \rho \in P \end{array} \right. \right\} \quad (7)$$

$$\text{and } \dot{y}_i^U(t) \geq \max_{\xi, \rho} \left\{ f_i(t, \xi, \rho) \left| \begin{array}{l} \xi_i = y_i^U(t) \\ \xi \in [y^L(t), y^U(t)] \\ \rho \in P \end{array} \right. \right\}, \quad (8)$$

for each $i = 1, \dots, n_x$, with $[y^L(0), y^U(0)] \supseteq \{x_0(p) \mid p \in P\}$. Then, $X(t) \subseteq [y^L(t), y^U(t)]$ for all $t \in [0, T]$.

This result is well established in the more general setting of a continuous right-hand side function f ; see, e.g., [31, 67] for ODE systems without parameters, where the differential inequalities are stated in terms of left-sided Dini derivatives and with strict inequalities. The corresponding result with weak inequalities follows from a uniqueness argument, such as (local) Lipschitz-continuity of the ODE right-hand side [25, 57]. The parameter-dependent case is also addressed in [57, 62].

Revisiting the classical theory of differential inequalities with stronger assumptions has two principal motivations. The first one is related to well-posedness of the initial value problems, as uniqueness of the ODE solutions is required in most practical applications. The second one is concerned with numerical computation of the bounding trajectories, which usually applies interval analysis to estimate the minimum and maximum values in the right-hand sides of (7,8). From this standpoint, (local) Lipschitz-continuity of the right-hand side function f and of the bounding trajectories y^L and y^U as well as weak differential inequalities are certainly justified. A proof of these results will be obtained in Sect. 5.1 as a corollary of the generalized differential inequalities introduced in Theorem 3.

A practical application of Theorem 1 is in propagating a continuous-time enclosure of the reachable set by considering equalities in (7,8) and possibly overestimating the right-hand sides using interval analysis. This approach presents the advantage that the number of ODEs in the auxiliary bounding system is $2n_x$, and therefore the size of the bounding system is proportional to the number of state variables in the original ODE system. This makes the approach computationally tractable for large-scale systems, at least in principle. Nonetheless, an important limitation with differential inequalities-based bounds is that they often suffer large overestimation, thus providing a poor approximation of the actual reachable set $X(t)$ or even blowing up to infinity in finite time. This instability can even occur in the case of linear ODEs, as illustrated in the following example.

Example 1 Consider the linear differential equations

$$\begin{aligned} \dot{x}_1(t, p) &= x_2(t, p) & \text{with } x_1(0, p) &= p_1 \\ \dot{x}_2(t, p) &= -x_1(t, p) & \text{with } x_2(0, p) &= p_2 \end{aligned} \quad (9)$$

with $p \in P := [-1, 1]^2$. The solutions of this Hamiltonian system remain bounded, although not asymptotically stable. An exact expression for the reachable set is:

$$X(t) = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix} X(0). \quad (10)$$

On the other hand, the tightest possible bounding trajectories y^L, y^U satisfying Theorem (1) are the solutions of the auxiliary ODE system

$$\begin{aligned} \dot{y}_1^L(t) &= y_2^L(t) & \text{with } y_1^L(0) &= p_1^L \\ \dot{y}_1^U(t) &= y_2^U(t) & \text{with } y_1^U(0) &= p_1^U \\ \dot{y}_2^L(t) &= -y_1^U(t) & \text{with } y_2^L(0) &= p_2^L \\ \dot{y}_2^U(t) &= -y_1^L(t) & \text{with } y_2^U(0) &= p_2^U, \end{aligned} \quad (11)$$

which yields an unstable enclosure of the reachable set $X(t)$ of (9) as

$$Y(t) = \exp(t)Y(0). \quad (12)$$

This instability is entailed by the inability of interval vectors to represent rotated boxes exactly, the so-called wrapping effect. \diamond

Remark 1 In the special case of linear ODE systems, it should be noted that general transformations can be applied in order to avoid the foregoing bound instability. In [53], for instance, the original system is first transformed into complex Jordan normal form and an exponential enclosure technique is used based on complex interval analysis and Piccard iterations. One can also shift the trajectories of the original ODEs as given below in (13), and then use differential inequalities to bound the shifted ODEs.

3.2 Ellipsoidal Propagation Approach

In order to improve upon interval enclosures, an alternative approach involves propagating ellipsoidal enclosures of the reachable set $X(t)$. In the presence of uncertain parameters in the ODE right-hand side or initial condition, it is convenient to construct ellipsoidal enclosures after shifting the trajectories so that

$$\forall (t, p) \in [0, T] \times P, \quad x(t, p) - \hat{x}(t) - G(t)(p - \hat{p}) \in \mathcal{E}(Q(t)), \quad (13)$$

with $Q(t) \in \mathbb{S}_+^{n_x}$. Here, $\hat{p} \in \mathbb{R}^{n_p}$ can be any reference point, for instance $\hat{p} := \text{mid}(P)$; the reference trajectory $\hat{x} : [0, T] \rightarrow \mathbb{R}^{n_x}$ is such that $\hat{x}(t) := x(t, \hat{p})$; and $G(t) \in \mathbb{R}^{n_x \times n_p}$ is the solution of the variational differential equation

$$\forall t \in [0, T], \quad \dot{G}(t) = A(t)G(t) + B(t) \quad \text{with } G(0) = B_0,$$

where the following shorthand notation is used:

$$A(t) := \frac{\partial f}{\partial x}(t, \hat{x}(t), \hat{p}), \quad B(t) := \frac{\partial f}{\partial p}(t, \hat{x}(t), \hat{p}), \quad \text{and } B_0 := \frac{\partial x_0}{\partial p}(\hat{p}),$$

assuming that the right-hand side function f is continuously-differentiable in (x, p) , and the initial value function x_0 continuously-differentiable with respect to p . With this transformation, an enclosure $Y(t)$ of the reachable set $X(t)$ of the original ODEs can then be recovered as:

$$Y(t) = \{\hat{x}(t)\} \oplus G(t)[P - \hat{p}] \oplus \mathcal{E}(Q(t)). \quad (14)$$

The following theorem provides sufficient conditions for time-varying ellipsoidal bounds $\mathcal{E}(Q(t))$ to satisfy (13). This construction requires a right-hand side nonlinearity bound $\Omega(t, Q, P, \hat{p}) \in \mathbb{I}\mathbb{R}^{n_x}$ such that

$$\begin{aligned} \forall(\xi, \rho) \in \{\hat{x}(t)\} \oplus G(t)[P - \hat{p}] \oplus \mathcal{E}(Q) \times P, \\ f(t, \xi, \rho) - f(t, \hat{x}(t), \hat{p}) - A(t)(\xi - \hat{x}(t)) - B(t)(\rho - \hat{p}) \in \Omega(t, Q, P, \hat{p}), \end{aligned} \quad (15)$$

as well as an initial value nonlinearity bound $\Omega_0(P, \hat{p}) \in \mathbb{I}\mathbb{R}^{n_x}$ satisfying

$$\forall \rho \in P, \quad x_0(\rho) - \hat{x}(0) - B_0(\rho - \hat{p}) \in \Omega_0(P, \hat{p}). \quad (16)$$

By a slight abuse of notation, the more compact notation $\Omega_t(Q)$ and Ω_0 is used subsequently to denote the nonlinearity bounds as P and \hat{p} are fixed.

Theorem 2 *Consider the initial value problem (1), and assume that the right-hand side function f is jointly continuous in (t, x, p) and continuously-differentiable in (x, p) for $t \in [0, T]$ and the initial-value function x_0 is continuously-differentiable on P , with $P \subset \mathbb{R}^{n_p}$ compact. Suppose that nonlinearity bounds Ω_t and Ω_0 satisfying (15) and (16), respectively, are available, with Ω_t locally Lipschitz-continuous in Q on $[0, T]$. Let $\kappa : [0, T] \rightarrow \mathbb{R}_{++}^{n_x}$ be a continuous function, and let $Q : [0, T] \rightarrow \mathbb{S}_+^{n_x}$ be a Lipschitz-continuous function satisfying*

$$\begin{aligned} \text{a.e. } t \in [0, T], \quad \dot{Q}(t) \succeq A(t)Q(t) + Q(t)A(t)^T + \sum_{i=1}^{n_x} \kappa_i(t)Q(t) \\ + \text{diag}(\kappa(t))^{-1} \text{diag rad}(\Omega_t(Q(t)))^2, \end{aligned} \quad (17)$$

with $Q(0) \succeq \text{diag rad}(\Omega_0)^2$. Then, $X(t) \subseteq \{\hat{x}(t)\} \oplus G(t)[P - \hat{p}] \oplus \mathcal{E}(Q(t))$ for all $t \in [0, T]$.

A similar result in the presence of time-varying uncertainties was established in [26]. The state transformation approach used in Theorem 2 exploits the fact that the uncertainty is time invariant, thus avoiding over-conservatism. An alternative proof of this results will be obtained in Sect. 5.1 as a corollary of the generalized differential inequalities introduced in Theorem 3.

Similar to differential inequalities, a practical application of Theorem 2 is in propagating a continuous-time enclosure of the reachable set with equality in (17). We note that a positive semidefinite solution of (17) is guaranteed to exist (at least locally) when the initial value $Q(0)$ is itself positive semidefinite. A clear advantage of this approach over differential inequalities is that it can provide exact bounds for parametric linear ODEs; e.g., the actual reachable set of (11) in Example 1 can be computed. Nonetheless, the number of ODEs in the auxiliary bounding system now scales as $O(n_x^2 + n_x n_p)$, and therefore the size of the bounding system increases quadratically with the number of state variables in the original ODE system. Another aspect that must be taken into account is the differentiability class of the functions f and x_0 in (1), as the ellipsoidal method in Theorem 2 requires continuous differentiability in both x and p , whereas differential inequalities only require local Lipschitz-continuity in x . Nonetheless, we note that the ellipsoidal approach may still

provide valid enclosures in the case that f or x_0 fail to be continuously-differentiable, for instance by choosing $A(t)$ in the Clarke subdifferential $\partial_c f(t, \hat{x}(t), \hat{p})$ [17]—the quadratic convergence property (see Sect. 6.2) would be lost however.

Example 2 Consider the following parametric nonlinear ODE system of the Lotka-Volterra type

$$\dot{x}_1(t, p) = p_1 x_1(t, p) [1 - x_2(t, p)] \quad \text{with } x_1(0, p) = 1.2 + p_2 \quad (18)$$

$$\dot{x}_2(t, p) = p_1 x_2(t, p) [x_1(t, p) - 1] \quad \text{with } x_2(0, p) = 1.1 + p_2^2, \quad (19)$$

where $p \in P := ([2.95, 3.05], [-0.05, 0.05])^T$ and $t \in [0, T]$.

In order to propagate an ellipsoidal enclosure of the reachable set of (18,19), Theorem 2 is applied with equality in (17). Further details about the numerical solution procedure will be given later on in Sect. 7. The results at $t = 1$ are shown in the left plot of Figure 1. The actual reachable set (gray-shaded area) is contained in the reachable set enclosure (thick solid line), which is obtained as the Minkowsky sum of $\{x(t, \hat{p})\} \oplus G(t)[P - \hat{p}]$ (dashed line) and $\mathcal{E}(Q(t))$ (dotted line). The right plot of Figure 1 presents the projections onto x_1 of the time-varying reachable set (shaded area) and of the ellipsoidal enclosure (thick solid line). This enclosure is also to be compared to the best possible bounds obtained with differential inequalities (thick dotted line) as given in Theorem 1, which turn out to be much more conservative in this example. \diamond

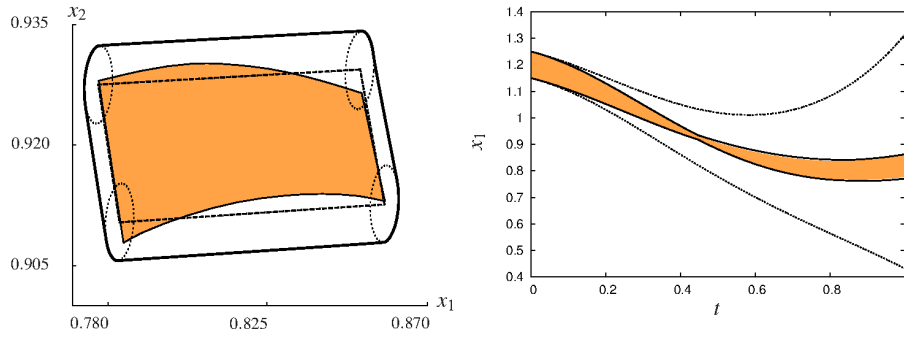


Fig. 1 *Left plot:* Ellipsoidal enclosure $\{x(t, \hat{p})\} \oplus G(t)[P - \hat{p}] \oplus \mathcal{E}(Q(t))$ of the reachable set $X(t)$ at $t = 1$. *Right plot:* Projections onto x_1 of the reachable set $X(t)$, the ellipsoidal enclosure $\{x(t, \hat{p})\} \oplus G(t)[P - \hat{p}] \oplus \mathcal{E}(Q(t))$, and the differential inequalities bounds $y^L(t), y^U(t)$ along $[0, T]$. The reachable set and the ellipsoidal enclosure are represented with a shaded area and a thick solid line, respectively.

4 Generalized Differential Inequalities

This section presents a unified framework for the continuous-time propagation of pointwise-in-time convex enclosures of the reachable set of nonlinear ODEs. The main result is a set of sufficient conditions, in the form of a generalized differential inequality, for a time-varying support function to describe an enclosure of the reachable set. In particular, this framework provides a direct link with standard differential inequalities and other existing time-continuous enclosure techniques that are amenable to numerical implementation.

The analysis that follows is inspired by, and closely related to, Aubin's viability theory [4]. The properties of the propagation operator Γ defined subsequently have been analyzed exhaustively in the context of viability theory, even in the more general case that p is a time-varying input function.

Definition 1 The set propagation operator $\Gamma : [0, T] \times [0, T] \times \Pi(\mathbb{R}^{n_x}) \rightarrow \Pi(\mathbb{R}^{n_x})$ associated to the ODE (1) is defined as

$$\Gamma(t_1, t_2, Y) := \left\{ x(t_2) \left| \begin{array}{l} \exists p \in P : \\ \forall \tau \in [t_1, t_2], \quad \dot{x}(\tau) = f(t, x(\tau), p) \\ x(t_1) \in Y \end{array} \right. \right\}, \quad (20)$$

for all $t_1, t_2 \in [0, T]$ with $t_2 \geq t_1$, and all $Y \subseteq \mathbb{R}^{n_x}$. In particular, for the reachable sets $X(t_1)$ and $X(t_2)$ at $t_1 \leq t_2$, we have

$$X(t_2) \subseteq \Gamma(t_1, t_2, X(t_1)).$$

Our method of proof follows a stepwise procedure, starting by establishing the result under certain regularity and convexity assumptions, before relaxing these assumptions. In order to avoid disrupting the flow of the paper, a number of technical lemmata used in the proof are reported in Appendix A. The following proposition is also instrumental to prove the result.

Proposition 1 Consider the ODE (1), and assume that the right-hand side function f is globally Lipschitz-continuous in x uniformly on $[0, T] \times P$. Let $Y_\varepsilon : [0, T] \rightarrow \mathbb{K}_{\mathbb{C}}^{n_x}$ be a family of set-valued functions, and suppose that there exist two continuous functions $\alpha, \beta : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, with $\alpha(0) = \beta(0) = 0$, such that

$$\Gamma(t, t+h, Y_\varepsilon(t)) \subseteq Y_\varepsilon(t+h) \oplus h[\alpha(h) + \beta(\varepsilon)] \mathcal{B}^{n_x} \quad (21)$$

for all sufficiently small $h \geq 0$ with $t+h \in [0, T]$, all $t \in [0, T]$, and all $\varepsilon \geq 0$. Then,

$$\forall \varepsilon \geq 0, \quad \Gamma(0, T, Y_\varepsilon(0)) \subseteq Y_\varepsilon(T) \oplus \gamma(\varepsilon) \mathcal{B}^{n_x},$$

for some continuous function $\gamma : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ with $\gamma(0) = 0$.

Proof The function f being globally Lipschitz-continuous in x uniformly on $[0, T] \times P$, (unique) solutions $x(t, p)$ of (1) are guaranteed to exist on $[0, T]$ for all $p \in P$. Therefore, the propagation operator Γ is well defined. Moreover, by uniform Lipschitzness of f in x on $[0, T] \times P$, it follows that the set-valued map Γ is itself locally Lipschitz-continuous around every $Y \subset \mathbb{R}^{n_x}$ (see [4], Definition 2.1.3); that is, for all $t, t' \in [0, T]$, with $t \leq t'$, there exists $C < \infty$ such that

$$\Gamma(t, t', Y + \Delta) \subseteq \Gamma(t, t', Y) + C \text{diam}(\Delta) \mathcal{B}^{n_x}, \quad (22)$$

for all $\Delta \subset \mathbb{R}^{n_x}$ with sufficiently small $\text{diam}(\Delta)$.

Consider a subdivision of $[0, T]$ into N equidistant intervals $[\theta_{i-1}, \theta_i]$, $i = 1, \dots, N$, with end-points $\theta_i := ih$ and width $h := \frac{T}{N}$. From (21) and for a large enough N —or, equivalently, for a sufficiently small h —we have

$$\forall i \in \{1, \dots, N\}, \quad \Gamma(\theta_{i-1}, \theta_i, Y_\varepsilon(\theta_{i-1})) \subseteq Y_\varepsilon(\theta_i) \oplus h[\alpha(h) + \beta(\varepsilon)] \mathcal{B}^{n_x}. \quad (23)$$

Consider the sequence of sets X_i , $i = 0, \dots, N$, defined recursively as

$$X_0 := Y_\varepsilon(0), \quad \text{and} \quad X_i := \Gamma(\theta_{i-1}, \theta_i, X_{i-1}) \quad \text{for } i \in \{1, \dots, N\}. \quad (24)$$

If readily follows from Definition 1 and (24) that

$$\forall i \in \{1, \dots, N\}, \quad \Gamma(0, \theta_i, Y_\varepsilon(0)) \subseteq X_i. \quad (25)$$

Next, by finite induction on $i = 0, \dots, N - 1$, assume that

$$X_i \subseteq Y_\varepsilon(\theta_i) \oplus C_i i h [\alpha(h) + \beta(\varepsilon)] \mathcal{B}^{n_x}.$$

for some constant $C_i < \infty$. We have

$$X_{i+1} = \Gamma(\theta_i, \theta_{i+1}, X_i) \subseteq \Gamma(\theta_i, \theta_{i+1}, Y_\varepsilon(\theta_i) \oplus C_i i h [\alpha(h) + \beta(\varepsilon)] \mathcal{B}^{n_x}),$$

and by the Lipschitz-continuity property (22), there are constants $C'_i, C_{i+1} < \infty$ such that

$$\begin{aligned} X_{i+1} &\subseteq \Gamma(\theta_i, \theta_{i+1}, Y_\varepsilon(\theta_i)) \oplus C'_i i h [\alpha(h) + \beta(\varepsilon)] \mathcal{B}^{n_x} \\ &\stackrel{(21)}{\subseteq} Y_\varepsilon(\theta_{i+1}) \oplus C_{i+1} (i+1) h [\alpha(h) + \beta(\varepsilon)] \mathcal{B}^{n_x}, \end{aligned}$$

for sufficiently small h . Noting that $\Gamma(0, 0, Y_\varepsilon(0)) = X_0$, this induction gives

$$\Gamma(0, T, Y_\varepsilon(0)) \stackrel{(25)}{\subseteq} X_N \subseteq Y_\varepsilon(T) \oplus \gamma(\varepsilon) \mathcal{B}^{n_x} \oplus C_N T \alpha(h) \mathcal{B}^{n_x},$$

for sufficiently small h , where we define $\gamma(\varepsilon) := C_N T \beta(\varepsilon)$. Here, The sequence of the constants C_i remains bounded because the function f is globally Lipschitz—this is in analogy with Gronwall's lemma. The result of the proposition follows by a compactness argument, noting that $\lim_{h \rightarrow 0} C_N T \alpha(h) \mathcal{B}^{n_x} = \{0\}$ and $Y_\varepsilon(T) \oplus \gamma(\varepsilon) \mathcal{B}^{n_x}$ is compact by assumption. \square

We are now in a position to state the main result of this section. The following theorem exploits the unique and exact representation of convex sets in terms of their support functions in order to obtain sufficient conditions for a time-varying support function of a convex enclosure of the reachable set. As already noted, this analysis is closely related to Aubin's viability theory [4], where the mathematical properties of convex enclosures are analyzed in detail. In this context, a key contribution of this theorem is a mechanism for constructing the support function of a convex enclosure of the reachable set, which can then be translated into practical algorithms as shown subsequently in Sect. 5.

Theorem 3 Consider the initial value problem (1) for $p \in P$ with $P \subset \mathbb{R}^{n_p}$ compact. Assume that the right-hand side function f is jointly continuous in (t, x, p) and locally Lipschitz-continuous in x uniformly on $[0, T] \times P$, with $P \subset \mathbb{R}^{n_p}$ compact, and that the initial value function x_0 is continuous on P . Let $Y : [0, T] \rightarrow \mathbb{K}_{\mathbb{C}}^{n_x}$ be a set-valued function, with $V[Y(\cdot)](c)$ Lipschitz-continuous for all $c \in \mathbb{R}^{n_x}$, and suppose that

$$a.e. t \in [0, T], \quad \dot{V}[Y(t)](c) \geq \max_{\xi, \rho} \left\{ c^T f(t, \xi, \rho) \mid \begin{array}{l} \xi \in Y(t) \\ c^T \xi = V[Y(t)](c) \\ \rho \in P \end{array} \right\} \quad (26)$$

$$and \quad V[Y(0)](c) \geq \max_{\rho} \{ c^T x_0(\rho) \mid \rho \in P \} \quad (27)$$

for all $c \in \mathbb{R}^{n_x}$. Then, $X(t) \subseteq Y(t)$ for all $t \in [0, T]$.

Proof The proof of the theorem proceeds in three steps. In the first step (S1), we establish the result under the following additional assumptions:

- A1 The set-valued function Y is such that $Y(t) = \{x \mid g(t, x) \leq 0\}$, where the function g is strictly smooth (infinitely-often differentiable) and convex in x for each $t \in [0, T]$;
A2 The function $V[Y(\cdot)](c)$ is differentiable for all $c \in \mathbb{R}^{n_x}$ on $[0, T]$;
A3 The right-hand side function f is globally Lipschitz-continuous in x uniformly on $[0, T] \times P$.

In a second step (S2), we show that the result also holds without imposing Assumption A1, but with Assumptions A2 and A3 still holding. Finally, the last step (S3) shows that the result remains valid without the need for Assumptions A2 or A3 to hold.

At this point, we also note that the conditions (26) and (27) are invariant with respect to scaling of the directions $c \in \mathbb{R}^{n_x}$. Therefore, it is sufficient to prove that the generalized differential inequality (26,27) is satisfied for all directions c with $c^T c = 1$.

- S1 Suppose that Assumptions A1, A2, and A3 are satisfied. Since (27) holds for all $c \in \mathbb{R}^{n_x}$, we have $X(0) \subseteq Y(0)$. If we can prove that, for all $c \in \mathbb{R}^{n_x}$ with $c^T c = 1$, there exist two continuous functions $\alpha, \beta : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, with $\alpha(0) = \beta(0) = 0$, such that

$$\forall t \in [0, T], \quad V[\Gamma(t, t+h, Y(t))](c) \leq V[Y(t+h)](c) + h[\alpha(h) + \beta(\varepsilon)] \quad (28)$$

for all sufficiently small $h \geq 0$, and all sufficiently small $\varepsilon > 0$, then it will directly follow from Definition 1 and Proposition 1—after passing to the limit as $\varepsilon \rightarrow 0$ —that

$$\forall t \in [0, T], \quad X(t) \subseteq \Gamma(0, t, Y(0)) \subseteq Y(t).$$

Let $\chi(\cdot, t, x_t, p)$ denote the solution of the initial value problem

$$\forall \tau \in [t, T], \quad \dot{\chi}(\tau, t, x_t, p) := f(\tau, \chi(\tau, t, x_t, p), p) \quad \text{with} \quad \chi(t, t, x_t, p) = x_t.$$

Solutions to this problem are guaranteed to exist for all $p \in P$ and all $x_t \in \mathbb{R}^{n_x}$ by Assumption A3, and we have

$$\begin{aligned} V[\Gamma(t, t+h, Y(t))](c) &\stackrel{(3)}{=} \max_{\xi} \{c^T \xi \mid \xi \in \Gamma(t, t+h, Y(t))\} \\ &\stackrel{(20)}{=} \max_{\xi, \rho} \left\{ c^T \chi(t+h, t, \xi, \rho) \mid \begin{array}{l} \xi \in Y(t) \\ \rho \in P \end{array} \right\}. \\ &\leq \max_{\xi, \rho} \left\{ c^T [\xi + hf(t, \xi, \rho)] \mid \begin{array}{l} \xi \in Y(t) \\ \rho \in P \end{array} \right\} + h\alpha_1(h), \end{aligned}$$

for all $c \in \mathbb{R}^{n_x}$ with $c^T c = 1$, and for some continuous function $\alpha_1 : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, with $\alpha_1(0) = 0$. Here, the last inequality follows from the fact that $\chi(\cdot, t, \xi, \rho)$ exists and is jointly continuous in all the variables.

Now, by Lemma 1, approximation functions $f_\varepsilon : [0, T] \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_p} \rightarrow \mathbb{R}^{n_x}$ of f can be constructed that are jointly continuous in (t, x, p) , smooth in x on $[0, T] \times P$, and such that

$$\forall x \in \mathbb{R}^{n_x}, \forall t \in [0, T], \forall \rho \in P, \quad \|f_\varepsilon(t, x, \rho) - f(t, x, \rho)\| \leq \beta(\varepsilon), \quad (29)$$

for some continuous function $\beta : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, with $\beta(0) = 0$. Then, by Assumption A1, we have

$$\begin{aligned} V[\Gamma(t, t+h, Y(t))](c) &\leq \max_{\rho \in P} \max_{\xi} \{c^T \xi + hc^T f_\varepsilon(t, \xi, \rho) \mid g(t, \xi) \leq 0\} \\ &\quad + h[\alpha_1(h) + \beta(\varepsilon)], \end{aligned} \quad (30)$$

for all $c \in \mathbb{R}^{n_x}$ with $c^T c = 1$. We now have a closer look at the inner maximization problem

$$\max_{\xi} \{ c^T \xi + hc^T f_{\varepsilon}(t, \xi, \rho) \mid g(t, \xi) \leq 0 \}. \quad (31)$$

Since $g(t, \cdot)$ is strictly convex (Assumption A1), the maximizer $\xi_t^*(c) := \arg \max_{\xi} \{ c^T \xi \mid g(t, \xi) \leq 0 \}$ is unique and well-defined for all $c \in \mathbb{R}^{n_x} \setminus \{0\}$. Moreover, the following properties hold at $\xi_t^*(c)$ [22]: (i) the constraint $g(t, \xi) \leq 0$ is strongly active, i.e., its KKT multiplier v_t^* is nonzero; (ii) the linear independence constraint qualification (LICQ) holds since $\frac{\partial g}{\partial \xi}(t, \xi_t^*(c)) = \frac{1}{v_t^*} c$; and, (iii) the strong second-order sufficiency conditions (SSOSC) hold. It follows from NLP sensitivity theory [22] that the inner-maximization problem in (31) itself has a unique maximizer $\hat{\xi}_t^*(h, \rho, c)$ for all $(c, \rho) \in \mathbb{R}^{n_x} \times P$ and all sufficiently small $h \geq 0$. Moreover, $\hat{\xi}_t^*(\cdot, \rho, c)$ is differentiable and g is strongly active in that neighborhood of $h = 0$. Therefore, we have

$$0 = g(t, \hat{\xi}_t^*(h, \rho, c)) = g(t, \xi_t^*(c)) + h \frac{\partial g}{\partial \xi}(t, \xi_t^*(c)) \frac{\partial \hat{\xi}_t^*}{\partial h}(0, \rho, c) + O(h^2),$$

so that

$$\frac{\partial g}{\partial \xi}(t, \xi_t^*(c)) \frac{\partial \hat{\xi}_t^*}{\partial h}(0, \rho, c) = O(h). \quad (32)$$

From the stationarity condition at $\xi_t^*(c)$, we also get

$$c - v_t^* \frac{\partial g}{\partial \xi}(t, \xi_t^*(c)) = 0,$$

which, when multiplied with $\frac{\partial \hat{\xi}_t^*}{\partial h}(0, \rho, c)$ from the right and using (32), gives

$$c^T \frac{\partial \hat{\xi}_t^*}{\partial h}(0, \rho, c) = O(h), \quad (33)$$

for all $c \in \mathbb{R}^{n_x}$ with $c^T c = 1$. It follows that

$$\begin{aligned} & \max_{\xi} \{ c^T \xi + hc^T f_{\varepsilon}(t, \xi, \rho) \mid g(t, \xi) \leq 0 \} \\ &= c^T \hat{\xi}_t^*(h, \rho, c) + hc^T f_{\varepsilon}(t, \hat{\xi}_t^*(h, \rho, c), \rho) \\ &= c^T \left[\xi_t^*(c) + h \frac{\partial \hat{\xi}_t^*}{\partial h}(0, \rho, c) \right] + hc^T f_{\varepsilon}(t, \xi_t^*(c), \rho) + O(h^2) \\ &\stackrel{(33)}{=} c^T \xi_t^*(c) + hc^T f_{\varepsilon}(t, \xi_t^*(c), \rho) + h \alpha_2(h), \end{aligned}$$

for all sufficiently small $h \geq 0$, all ρ in P , all $c \in \mathbb{R}^{n_x}$ with $c^T c = 1$, and for some continuous function $\alpha_2 : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ with $\alpha_2(0) = 0$. Observing that $c^T \xi_t^*(c) = V[Y(t)](c)$, we

then obtain

$$\begin{aligned}
& \max_{\rho \in P} \max_{\xi} \{ c^T \xi + h c^T f_\varepsilon(t, \xi, \rho) \mid g(t, \xi) \leq 0 \} \\
&= V[Y(t)](c) + h \max_{\rho \in P} c^T f_\varepsilon(t, \xi^*(c), \rho) + h \alpha_2(h) \\
&\leq V[Y(t)](c) + h \max_{\xi, \rho} \left\{ c^T f_\varepsilon(t, \xi, \rho) \left| \begin{array}{l} \xi \in Y(t) \\ c^T \xi = V[Y(t)](c) \\ \rho \in P \end{array} \right. \right\} + h \alpha_2(h) \\
&\stackrel{(29)}{\leq} V[Y(t)](c) + h \max_{\xi, \rho} \left\{ c^T f(t, \xi, \rho) \left| \begin{array}{l} \xi \in Y(t) \\ c^T \xi = V[Y(t)](c) \\ \rho \in P \end{array} \right. \right\} + h [\alpha_2(h) + \beta(\varepsilon)] \\
&\stackrel{(26)}{\leq} V[Y(t)](c) + h \dot{V}[Y(t)](c) + h [O(h) + \alpha_2(\varepsilon)] \\
&\stackrel{(A2)}{\leq} V[Y(t+h)](c) + h [\alpha_2(h) + \alpha_3(h) + \beta(\varepsilon)] ,
\end{aligned}$$

for all sufficiently small $h \geq 0$, and all $c \in \mathbb{R}^{n_x}$ with $c^T c = 1$. In the last equality, we have used that $V[Y(\cdot)](c)$ is differentiable on $[0, T]$ for all c by Assumption A2. The condition (28) thus follows from (30).

S2 Suppose now that Assumptions A2 and A3 still hold, but not Assumption A1. Following Lemma 2 (Appendix A), we construct a family of set-valued functions $Y_\varepsilon : [0, T] \rightarrow \mathbb{K}_{\mathbb{C}}^{n_x}$, parameterized by $\varepsilon \geq 0$, such that $Y_\varepsilon(t) := \{x \in \mathbb{R}^{n_x} \mid g_\varepsilon(t, x) \leq 0\} \supseteq Y(t)$, with $g_\varepsilon(t, \cdot)$ strictly convex and smooth for all $\varepsilon > 0$ and all $t \in [0, T]$, and

$$\dot{V}[Y_\varepsilon(t)](c) \geq \dot{V}[Y(t)](c) + L\alpha(\varepsilon)$$

for all $c \in \mathbb{R}^{n_x}$ with $c^T c = 1$ and some $L < \frac{1}{T}$. Suppose for a moment that T is sufficiently small in order for the chosen L to remain larger than the uniform Lipschitz constant of the right-hand side function f on the compact sets $\bigcup_{t \in [0, T]} Y_\varepsilon(t)$. This way, we have

$$\begin{aligned}
\dot{V}[Y_\varepsilon(t)](c) &\stackrel{(26)}{\geq} \max_{\xi, \rho} \left\{ c^T f(t, \xi, \rho) \left| \begin{array}{l} \xi \in Y(t) \\ c^T \xi = V[Y(t)](c) \\ \rho \in P \end{array} \right. \right\} + L\alpha(\varepsilon) \\
&\geq \max_{\xi, \rho} \left\{ c^T f(t, \xi, \rho) \left| \begin{array}{l} \xi \in Y_\varepsilon(t) \\ c^T \xi = V[Y_\varepsilon(t)](c) \\ \rho \in P \end{array} \right. \right\}
\end{aligned}$$

for all $t \in [0, T]$ and all $c \in \mathbb{R}^{n_x}$ with $c^T c = 1$. Then, we can apply the result from part S1 above to show that $X(t) \subseteq Y_\varepsilon(t)$ for all $t \in [0, T]$ and all $\varepsilon > 0$, and the result of the theorem follows by noting that the sets $Y_\varepsilon(t)$ converge to the compact sets $Y(t)$ in the Hausdorff metric as $\varepsilon \rightarrow 0$. Finally, the auxiliary assumption that T is sufficiently small—but strictly larger than 0—can be made without loss of generality, since we can otherwise divide $[0, T]$ into a finite number of sufficiently small subintervals onto which the foregoing procedure can be applied.

S3 In order to complete the proof of the theorem, we finally relax Assumptions A2 and A3. Assume first that the right-hand side function f is locally Lipschitz at each $x \in D$ uniformly on $[0, T] \times P$, where $D \in \mathcal{H}^{n_x}$ is such that $D \supseteq \cup_{t \in [0, T]} Y(t)$ —Note that such a set D always exists since the sets $Y(t)$ are bounded by assumption. A modified function \tilde{f} that is globally Lipschitz-continuous in x uniformly on $[0, T] \times P$ can then be constructed, which coincides with f on D . In particular, using \tilde{f} instead of f does not modify the generalized differential inequality (26), showing that the assumption that f is globally Lipschitz-continuous in x can be replaced by a local Lipschitzness condition without loss of generality. Finally, because any (locally) Lipschitz-continuous function is differentiable almost everywhere and since perturbing a differential inequality on a set with Lebesgue measure equal to zero does not affect the result, one can drop Assumption A2 as well. \square

Remark 2 Theorem 3 remains true if the time-invariant parameter $p \in P \subset \mathbb{R}^{n_p}$ is replaced with time-varying measurable function $p : [0, T] \rightarrow \mathbb{R}^{n_p}$ such that $p(t) \in P$. This extension can for example be established by dividing the integration horizon $[0, T]$ into small subintervals of width $h := \frac{T}{N}$ for $N \in \mathbb{N}$, where the uncertain parameter is kept constant and Theorem 3 can thus be applied. Since the reachable set enclosures $Y(t)$ are compact, one can then consider the limit as $h \rightarrow 0$ to prove that the result is valid for general bounded functions, not merely piecewise constant functions. \diamond

5 Applications of Generalized Differential Inequalities

This section describes how the generalized differential inequality introduced in Theorem 3 can be specialized to yield alternative proofs for the differential inequalities and ellipsoidal propagation results recalled in Sect. 3. An approach based on Taylor models with convex remainder terms is also described in order to propagate nonconvex enclosures of the reachable set.

5.1 Link with Standard Differential Inequalities and Ellipsoidal Bounding Approach

The case of interval enclosures is addressed first, by specializing the convex enclosure as $Y(t) := [y^L(t), y^U(t)]$. The following proposition establishes that the result of the standard differential inequalities in Theorem 1 is implied by the generalized differential inequalities in Theorem 3.

Proposition 2 Any pair of functions $y^L(t), y^U(t) : [0, T] \rightarrow \mathbb{R}^{n_x}$ satisfying Theorem 1 also satisfies Theorem 3 with $Y(t) := [y^L(t), y^U(t)]$.

Proof Noting that

$$\sum_{i=1}^{n_x} \left\{ \begin{array}{l} c_i y_i^L(t) \text{ if } c_i \leq 0 \\ c_i y_i^U(t) \text{ otherwise} \end{array} \right\} = \frac{1}{2} c^T (y^L(t) + y^U(t)) + \frac{1}{2} \text{abs}(c)^T (y^U(t) - y^L(t)), \quad (34)$$

with $\text{abs}(c) := (|c_1|, \dots, |c_{n_x}|)^T$, the support function of the interval vector $Y(t)$ is given by

$$\forall c \in \mathbb{R}^{n_x}, \quad V[Y(t)](c) = \frac{1}{2} c^T (y^L(t) + y^U(t)) + \frac{1}{2} \text{abs}(c)^T (y^U(t) - y^L(t)). \quad (35)$$

The initial-value condition (27) is trivially satisfied with this definition. Moreover, rewriting the (right-hand) derivative of the support function in terms of the (right-hand) derivatives of the bounding trajectories $y^L(t)$, $y^U(t)$ and using the standard differential inequalities result in Theorem 1 gives

$$\begin{aligned}
\dot{V}[Y(t)](c) &\stackrel{(35)}{=} \frac{1}{2}(c - \text{abs}(c))^T y^L(t) + \frac{1}{2}(c + \text{abs}(c))^T y^U(t) \\
&\stackrel{(7),(8)}{\geq} \sum_{i=1}^{n_x} \left[\frac{1}{2} \min_{\xi, \rho} \left\{ (c_i - |c_i|) f_i(t, \xi, \rho) \left| \begin{array}{l} \xi \in [y^L(t), y^U(t)] \\ \xi_i = y_i^L(t), \rho \in P \end{array} \right. \right\} \right. \\
&\quad \left. + \frac{1}{2} \max_{\xi, \rho} \left\{ (c_i + |c_i|) f_i(t, \xi, \rho) \left| \begin{array}{l} \xi \in [y^L(t), y^U(t)] \\ \xi_i = y_i^U(t), \rho \in P \end{array} \right. \right\} \right] \\
&= \sum_{i=1}^{n_x} \max_{\xi, \rho} \left\{ c_i f_i(t, \xi, \rho) \left| \begin{array}{l} \xi \in Y(t), \rho \in P, \\ \xi_i = \begin{cases} y_i^L(t) & \text{if } c_i \leq 0 \\ y_i^U(t) & \text{otherwise} \end{cases} \end{array} \right. \right\} \\
&\geq \max_{\xi, \rho} \left\{ c^T f(t, \xi, \rho) \left| \begin{array}{l} \xi \in Y(t), \rho \in P, \\ \sum_{i=1}^{n_x} c_i \xi_i = \sum_{i=1}^{n_x} \begin{cases} c_i y_i^L(t) & \text{if } c_i \leq 0 \\ c_i y_i^U(t) & \text{otherwise} \end{cases} \end{array} \right. \right\} \\
&\stackrel{(34)}{=} \max_{\xi, \rho} \left\{ c^T f(t, \xi, \rho) \left| \begin{array}{l} \xi \in Y(t), \rho \in P, \\ c^T \xi = \frac{1}{2} c^T (y^L(t) + y^U(t)) + \frac{1}{2} \text{abs}(c)^T (y^U(t) - y^L(t)) \end{array} \right. \right\} \\
&\stackrel{(35)}{=} \max_{\xi, \rho} \left\{ c^T f(t, \xi, \rho) \mid \xi \in Y(t), c^T \xi = V[Y(t)](c), \rho \in P \right\}.
\end{aligned}$$

□

The case of ellipsoidal enclosures is addressed by specializing the convex enclosure as $Y(t) := \mathcal{E}(Q(t))$. Like previously with standard differential inequalities, the result of Theorem 2 is implied by the generalized differential inequalities in Theorem 3.

Proposition 3 Any functions $\kappa : [0, T] \rightarrow \mathbb{R}_{++}$ and $Q : [0, T] \rightarrow \mathbb{S}_+^{n_x}$ satisfying Theorem 2, with associated nonlinearity bounds Ω_0 and Ω_t , also satisfy Theorem 3 with $Y(t) := \{\hat{x}(t)\} \oplus G(t)[P - \hat{p}] \oplus \mathcal{E}(Q(t))$.

Proof Noting that, for all $c \in \mathbb{R}^{n_x}$, $V[\mathcal{E}(Q(t))](c) = \sqrt{c^T Q(t) c}$, we have

$$\dot{V}[\mathcal{E}(Q(t))](c) = \frac{1}{2\sqrt{c^T Q(t) c}} c^T \dot{Q}(t) c. \quad (36)$$

Assume, for a moment, that $Q(t)$ is positive definite and $c \neq 0$, so that $\dot{V}[\mathcal{E}(Q(t))](c)$ exists almost everywhere. Substituting (17) in (36), and using the fact that $\sum_{i=1}^{n_x} \kappa_i(t) Q(t) + \text{diag}(\kappa(t))^{-1} \text{diag rad}(\Omega_t(Q(t)))^2$ is minimized by choosing

$$\kappa(t) = \frac{c^T \text{diag rad}(\Omega_t(Q(t)))}{\sqrt{c^T Q(t) c}},$$

we have that

$$\dot{V}[\mathcal{E}(Q(t))](c) \geq \frac{c^T A(t) Q(t) c}{\sqrt{c^T Q(t) c}} + \|c^T \text{diag rad}(\Omega_t(Q(t)))\|_1, \quad (37)$$

for all $\kappa(t) > 0$.

The shifted state trajectories $z(t, p) := x(t, p) - \hat{x}(t) - G(t)[p - \hat{p}]$ satisfy the ODE

$$\begin{aligned} \dot{z}(t, p) &= f(t, z(t, p) + \hat{x}(t) + G(t)(p - \hat{p}), p) - f(t, \hat{x}(t), \hat{p}) - [A(t)G(t) + B(t)](p - \hat{p}) \\ &=: \phi(t, z(t, p), p) \\ \text{with } z(0, p) &= x_0(p) - \hat{x}(0) - G(0)(p - \hat{p}) =: z_0(p). \end{aligned} \quad (38)$$

By construction of the non-linearity bounder Ω_t in (15) at a given $t \in [0, T]$, we thus have

$$\phi(t, \xi, \rho) - A(t)\xi \in \Omega_t(Q(t)),$$

and

$$\|c^T \text{diag rad}(\Omega_t(Q(t)))\|_1 = \max_{\omega \in \Omega_t(Q(t))} c^T \omega \geq c^T [\phi(t, \xi, \rho) - A(t)\xi],$$

for all $(\xi, \rho) \in \mathcal{E}(Q(t)) \times P$ and all $c \in \mathbb{R}^{n_x}$. In particular, choosing $\xi := \frac{Q(t)c}{\sqrt{c^T Q(t)c}} \in \mathcal{E}(Q(t))$ gives

$$\|c^T \text{diag rad}(\Omega_t(Q(t)))\|_1 \geq \max_{\rho \in P} c^T \left[\phi \left(t, \frac{Q(t)c}{\sqrt{c^T Q(t)c}}, \rho \right) - A(t) \frac{Q(t)c}{\sqrt{c^T Q(t)c}} \right],$$

which after substitution into (37) yields

$$\begin{aligned} \forall c \in \mathbb{R}^{n_x}, \quad \dot{V}[\mathcal{E}(Q(t))](c) &\geq \max_{\rho \in P} c^T \phi \left(t, \frac{Q(t)c}{\sqrt{c^T Q(t)c}}, \rho \right) \\ &= \max_{\rho, \xi} \left\{ c^T \phi(t, \xi, \rho) \mid \begin{array}{l} \xi \in \mathcal{E}(Q(t)) \\ c^T \xi = V[\mathcal{E}(Q(t))](c) \\ \rho \in P \end{array} \right\}. \end{aligned}$$

Likewise, by construction of the initial value non-linearity bounder Ω_0 in (16), we have

$$V[\mathcal{E}(Q(0))](c) = \sqrt{c^T Q(0)c} \geq \|c^T \text{diag rad}(\Omega_0)\|_1 = \max_{\omega \in \Omega_0} c^T \omega \geq c^T z_0(\rho),$$

for all $\rho \in P$ and all $c \in \mathbb{R}^{n_x}$. Therefore, $V[\mathcal{E}(Q(t))](c)$ satisfies the generalized differential inequalities (26) and corresponding initial condition.

The case that $Q(t)$ is only positive semi-definite can be treated by approximating $Q(t)$ with a positive definite matrix $Q_\varepsilon \succ 0$ with $\|Q_\varepsilon(t) - Q(t)\| < \varepsilon$ for any sufficiently small $\varepsilon > 0$, then repeating the above construction with Q_ε instead of Q , and finally taking the limit $\varepsilon \rightarrow 0$. Notice that this argumentation is analogous to the strongly convex relaxation of the set $Y(t)$ with the set $Y_\varepsilon(t)$ that was used in the proof of Theorem 3. \square

5.2 Propagation of Nonconvex Enclosures using Taylor Models

Convex enclosures of reachable sets can be conveniently computed using the unified framework of generalized differential inequalities, for instance as interval vectors or ellipsoids. Nonetheless, approximating nonconvex enclosures with convex sets can result in large overestimation, due to both the wrapping effect and the dependency problem. Clearly, the ability to propagate nonconvex enclosures can help mitigate this overestimation.

This section describes a novel way of propagating nonconvex enclosures of the reachable set of the parametric ODE (1), whereby high-order polynomial approximations of the parametric ODE solutions are used in combination with the framework of generalized differential inequalities for bounding the approximation error. Specifically, let $\mathcal{P}_x^q : [0, T] \times P \rightarrow \mathbb{R}^{n_x}$ denote a q th-order polynomial approximant, and let the remainder function $r_x^q : [0, T] \times P \rightarrow \mathbb{R}^{n_x}$ be defined such that

$$\forall (t, p) \in [0, T] \times P, \quad r_x^q(t, p) := x(t, p) - \mathcal{P}_x^q(t, p). \quad (39)$$

It follows from differentiating (39) with respect to time that the remainder function r_x^q satisfies the parametric ODE

$$\dot{r}_x^q(t, p) = f(t, \mathcal{P}_x^q(t, p) + r_x^q(t, p), p) - \dot{\mathcal{P}}_x^q(t, p) =: r_f^q(t, r_x^q(t, p), p), \quad (40)$$

with initial condition $r_x^q(0, p) = x_0(p) - \mathcal{P}_{x_0}^q(p)$, where $\mathcal{P}_{x_0}^q : P \rightarrow \mathbb{R}^{n_x}$ denotes a q th-order polynomial approximation of x_0 on P . The idea is to apply Theorem 3 for characterizing pointwise-in-time convex enclosures $\mathcal{R}_x^q(t, P) \supseteq \{r(t, p) | p \in P\}$, and thus obtain a nonconvex enclosure of the reachable set in the form

$$X(t, P) \subseteq \{\mathcal{P}_x^q(t, p) \mid p \in P\} \oplus \mathcal{R}_x^q(t, P).$$

The focus hereafter is on Taylor models [12, 39, 48], although alternative types of polynomial approximation can be used in principle as long as these constructions can be automated for general factorable functions. In the Taylor model approach, the polynomial approximant $\mathcal{P}_x^q(t, \cdot)$ matches the q th-order Taylor expansion of $x(t, \cdot)$ on P at a given reference point $\hat{p} \in P$ [14]:

$$\forall i \in \{1, \dots, n_x\}, \quad \forall t \in [0, T], \quad \forall p \in P, \quad \mathcal{P}_{x_i}^q(t, p) := \sum_{\substack{\gamma \in \mathbb{N}^{n_p} \\ |\gamma| \leq q}} \frac{\partial^\gamma x_i(t, \hat{p})}{\gamma!} (p - \hat{p})^\gamma. \quad (41)$$

That is, the time-varying coefficients of \mathcal{P}_x^q are expressed as functions of the state-sensitivities $\partial^\kappa x_i(\cdot, \hat{p})$ of (1) up to order q at the reference points \hat{p} . In particular, this construction requires that the right-hand side function f and initial-value function x_0 be at least $(q+1)$ -times continuously-differentiable in both x and p . Moreover, it requires that a system of state-sensitivity equations of size $O(n_x n_p^q)$ be integrated on $[0, T]$.

The application of differential inequalities to propagate interval bounds on the remainder function r_x^q as $\mathcal{R}_x^q(t, P) := [r_{x_i}^L(t), r_{x_i}^U(t)]$ was first investigated in [14]. The following proposition follows directly on application of Theorem 1 to bound the solutions of (40).

Proposition 4 *Consider the initial value problem (1), and assume that the right-hand side function f is jointly continuous in (t, x, p) and $(q+1)$ -times continuously-differentiable in (x, p) on $[0, T]$, with $q \geq 1$, and the initial-value function x_0 is $(q+1)$ -times continuously-differentiable on P , with $P \subset \mathbb{R}^{n_p}$ compact. Let $\mathcal{P}_x^q : [0, T] \times P \rightarrow \mathbb{R}^{n_x}$ be defined as in (41),*

for a given reference point $\hat{p} \in P$. Let the functions $r_x^{q,L}, r_x^{q,U} : [0, T] \rightarrow \mathbb{R}^{n_x}$ be Lipschitz-continuous, with $r_x^{q,L} \leq r_x^{q,U}$, and satisfy

$$\text{a.e. } t \in [0, T], \quad \dot{r}_{x_i}^{q,L}(t) \leq \min_{\xi, \rho} \left\{ f_i(t, \mathcal{P}_x^q(t, \rho) + \xi, \rho) - \dot{\mathcal{P}}_x^q(t, \rho) \left| \begin{array}{l} \xi_i = r_{x_i}^{q,L}(t) \\ \xi \in \mathcal{R}_x^q(t, P) \\ \rho \in P \end{array} \right. \right\} \quad (42)$$

$$\dot{r}_{x_i}^{q,U}(t) \geq \max_{\xi, \rho} \left\{ f_i(t, \mathcal{P}_x^q(t, \rho) + \xi, \rho) - \dot{\mathcal{P}}_x^q(t, \rho) \left| \begin{array}{l} \xi_i = r_{x_i}^{q,U}(t) \\ \xi \in \mathcal{R}_x^q(t, P) \\ \rho \in P \end{array} \right. \right\} \quad (43)$$

for each $i = 1, \dots, n_x$, and $[r_{x_i}^L(0), r_{x_i}^U(0)] \supseteq \{x_0(p) - \mathcal{P}_{x_0}^q \mid p \in P\}$, where $\mathcal{P}_{x_0}^q : \mathbb{R}^{n_p} \rightarrow \mathbb{R}^{n_x}$ is a truncated Taylor expansion of x_0 on P at \hat{p} . Then, $X(t) \subseteq \{\mathcal{P}_x^q(t, p) \mid p \in P\} \oplus [r_x^{q,L}(t), r_x^{q,U}(t)]$, for all $t \in [0, T]$.

In practice, the resulting enclosures are often found to be tighter than upon application of the standard differential inequalities given by Theorem 1, especially as the expansion order q increases. This trend is confirmed by the convergence analysis results in Sect. 6, where it is established that the overestimation is of order (no less than) $O(\text{diam}(P)^{q+1})$ for Taylor models combined with differential inequalities, while it is normally of order $O(\text{diam}(P))$ with standard differential inequalities. Nonetheless, it is noteworthy that this trend is only asymptotic as $\text{diam}(P) \rightarrow 0$, and so standard differential inequalities can outperform their Taylor model counterparts for large parameter host sets. It is also worth noting that the combination with Taylor models entails a scaling in the size of the bounding system not only with n_x but also with n_p and q as $O(n_x n_p^q)$. A trade-off can therefore be expected between the enclosure tightness and the increase in computational time for larger expansion orders.

The propagation of ellipsoidal enclosures for the remainder function r_x^q as $\mathcal{R}_x^q(t, P) := \mathcal{E}(\mathcal{Q}_x^q(t))$ was recently investigated in [66] as well. One way of applying the ellipsoidal bounding technique of Theorem 2 involves rewriting the ODE (40) in the form

$$\dot{r}_x^q(t, p) = A_f(t) r_x^q(t, p) + N_f(t, r_x^q(t, p), p, \hat{p}), \quad (44)$$

$$\text{with } A_f(t) := \frac{\partial f}{\partial x}(t, \mathcal{P}_x^q(t, \hat{p}), \hat{p}) \quad \text{and} \quad N_f(t, r, p) := r_f^q(t, r, p) - A_f(t) r.$$

Then, assuming that nonlinearity bounders $\Omega_t^q(Q), \Omega_0^q \in \mathbb{I}\mathbb{R}^{n_x}$ can be constructed at each $t \in [0, T]$ such that

$$\forall (r, p) \in \mathcal{E}(Q) \times P, \quad N_f(t, r, p) \in \Omega_t^q(Q) \quad (45)$$

$$\text{and } \forall p \in P, \quad x_0(p) - \mathcal{P}_{x_0}^q(p) \in \Omega_0^q, \quad (46)$$

the following proposition follows directly on application of Theorem 2 to (44).

Proposition 5 Consider the initial value problem (1), and assume that the right-hand side function f is jointly continuous in (t, x, p) and $(q+1)$ -times continuously-differentiable in (x, p) on $[0, T]$, with $q \geq 1$, and the initial-value function x_0 is $(q+1)$ -times continuously-differentiable on P , with $P \subset \mathbb{R}^{n_p}$ compact. Let $\mathcal{P}_x^q : [0, T] \times P \rightarrow \mathbb{R}^{n_x}$ be defined as in (41), for a given reference point $\hat{p} \in P$. Suppose that nonlinearity bounders $\Omega_t^q(Q)$ and Ω_0^q satisfying (45) and (46), respectively, are available, with Ω_t^q locally Lipschitz-continuous in

Q on $[0, T]$. Let $\kappa : [0, T] \rightarrow \mathbb{R}_{++}^{n_x}$ be a continuous function, and let $Q_x^q : [0, T] \rightarrow \mathbb{S}_+^{n_x}$ be a Lipschitz-continuous function satisfying

$$\begin{aligned} \text{a.e. } t \in [0, T], \quad \dot{Q}_x^q(t) \succeq & A_f(t)Q_x^q(t) + Q_x^q(t)A_f(t)^T + \sum_{i=1}^{n_x} \kappa_i(t)Q_x^q(t) \\ & + \text{diag}(\kappa(t))^{-1} \text{diag rad}(\Omega_t^q(Q_x^q(t)))^2, \end{aligned} \quad (47)$$

with $Q_x^q(0) \succeq \text{diag rad}(\Omega_0^q)$. Then, $X(t) \subseteq \{\mathcal{P}_x^q(t, p) | p \in P\} \oplus \mathcal{E}(Q_x^q(t))$ for all $t \in [0, T]$.

Like with differential inequalities, the propagation of an ellipsoidal enclosure of the remainder term in a higher-order Taylor model is often found to improve upon the standard ellipsoidal bounds given by Theorem 2. In particular, it can be shown that the enclosures obtained by considering a first-order Taylor model with ellipsoidal remainder or a standard ellipsoidal approach are identical when the nonlinearity bounders Ω_0, Ω_t in the latter are derived from first-order Taylor models; that is, $\{x(t, \hat{p})\} \oplus G(t)(p - \hat{p})$ and $\mathcal{P}_x^1(t, p)$ match and so do $Q(t)$ and $Q_x^1(t)$. Here again, a main limitation of Taylor models with ellipsoidal remainders is the size of the bounding system, which scales as $O(n_x n_p^q + n_x^2)$.

Example 2 (continued) We revisit the Lotka-Volterra system (18,19) in the context of non-convex enclosures based on Taylor models. Second-order Taylor models ($q = 2$) are considered, as expanded at the midpoint $\hat{p} = (3, 0)^T$ and with centered remainder bounds. Both differential inequalities-based bounds and ellipsoidal bounds are considered for the remainder function r_x^q and are computed via, respectively, (42,43) and (47), both with equalities. Further details about the numerical solution procedure will be given later on in Sect. 7.

The results at $t = 1$ are shown in Figure 2. The actual reachable set (shaded area) is contained in the reachable set enclosure (thick solid line), for Taylor models with either interval (left plot) or ellipsoidal (right plot) remainder bounds. The former is obtained as the Minkowsky sum of $\{\mathcal{P}_x^2(t, p) | p \in P\}$ (dashed line) and the remainder interval $[r_x^{2,L}(t), r_x^{2,U}(t)]$ (dotted lines); the latter, as the Minkowsky sum of $\{\mathcal{P}_x^2(t, p) | p \in P\}$ (dashed line) and the remainder enclosure $\mathcal{E}(Q_x^2(t))$ (dotted lines). In comparing Taylor models with different remainder bounds, it is evident that the wrapping effect and the dependency problem transfers to the enclosures of the remainder function, although the over-estimation is greatly reduced in this example by the use of Taylor models. Also note that Taylor models with ellipsoidal remainders outperform their interval remainder counterparts.

◇

6 Convergence Analysis

We already argued in the introduction section that convergence properties of the computed enclosures to the actual reachable set are important in application areas such as global or robust optimization. This section investigates conditions under which a family of enclosures $Y(\cdot, P)$ that satisfy the generalized differential inequality in Theorem 3 exhibits a certain convergence order $k \geq 1$ with respect to the parameter host set P . This generic result is then specialized to interval enclosures, ellipsoidal enclosures, and Taylor models with convex remainders as a means for deriving sharp bounds on their convergence order. The results are illustrated with a numerical example.

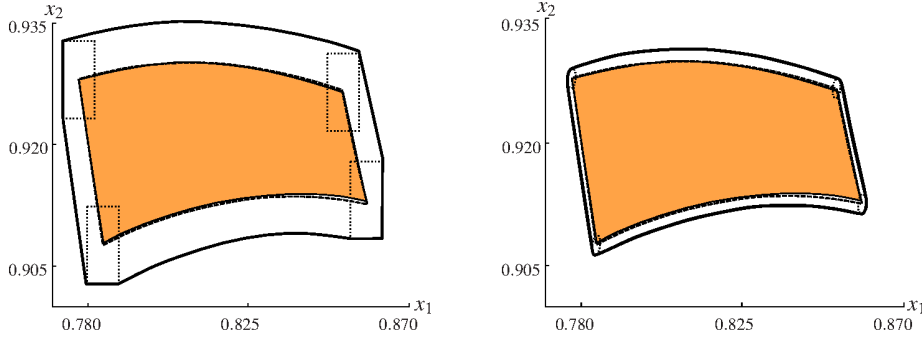


Fig. 2 *Left plot:* Second-order Taylor model with interval remainder $\{\mathcal{P}_x^2(t, p) | p \in P\} \oplus [r_x^{2,L}(t), r_x^{2,U}(t)]$ of the reachable set $X(t, P)$ at $t = 1$. *Right plot:* Second-order Taylor model with ellipsoidal remainder $\{\mathcal{P}_x^2(t, p) | p \in P\} \oplus \mathcal{E}(Q_x^2(t))$ of the reachable set $X(t, P)$ at $t = 1$. The reachable set and the ellipsoidal enclosure are represented with a shaded area and a thick solid line, respectively.

6.1 General Convergence Theorem

This subsection investigates the convergence properties of enclosure functions that satisfy the generalized differential inequality (26,27). A formal definition of the convergence order of (the image of) a set-valued function is given below. We formulate this definition in terms of *representable sets*, which can be described exactly by a finite number of parameters. The class of representable sets of interest hereafter are interval vectors and ellipsoids.

Definition 2 Let $\mathcal{S}(\mathbb{K}^n)$ be a class of representable sets in \mathbb{K}^n . A set-valued function $\Psi : \mathcal{S}(\mathbb{K}^n) \rightarrow \mathbb{K}^m$ is said to have convergence order $k \geq 1$ on a set $Z \in \mathcal{S}(\mathbb{K}^n)$, if there exists a constant $C < \infty$ such that

$$\text{diam}(\Psi(W)) \leq C \text{diam}(W)^k,$$

for all $W \in \mathcal{S}(\mathbb{K}^n)$ such that $W \subseteq Z$ with sufficiently small $\text{diam}(W)$.

In practice, the maximization problem defining the right-hand side of (26) may not be solved exactly and the application of Theorem 3 typically relies on the availability of an upper-bounding function $\mathcal{B}_t : \mathbb{R}^{n_x} \times \mathbb{K}^{n_x} \times \mathbb{K}^{n_p} \rightarrow \mathbb{R}$ such that

$$\forall (c, Y, Z) \in \mathbb{R}^{n_x} \times \mathbb{K}^{n_x} \times \mathbb{K}^{n_p}, \quad \mathcal{B}_t(c, Y, Z) \geq \max_{\xi, \rho} \left\{ c^T f(t, \xi, \rho) \left| \begin{array}{l} \xi \in Y \\ \rho \in Z \\ c^T \xi = V[Y](c) \end{array} \right. \right\},$$

as well as another upper-bounding function $\mathcal{B}_0 : \mathbb{R}^{n_x} \times \mathbb{K}^{n_p} \rightarrow \mathbb{R}$ such that

$$\forall (c, Z) \in \mathbb{R}^{n_x} \times \mathbb{K}^{n_p}, \quad \mathcal{B}_0(c, Z) \geq \max_{\rho} \{ c^T x_0(\rho) \mid \rho \in Z \}.$$

With this notation, the generalized differential inequality (26) can be rewritten in the form

$$\text{a.e. } t \in [0, T], \quad \dot{V}[Y(t, Z)](c) = \mathcal{B}_t(c, Y(t, Z), Z) \quad \text{with} \quad V[Y(0, Z)](c) = \mathcal{B}_0(c, Z). \quad (48)$$

The following technical assumptions on the upper-bounding functions \mathcal{B}_t and \mathcal{B}_0 are made in order to analyze the convergence rate of the enclosure family $Y(\cdot, Z)$ with $Z \subseteq P$ on $[0, T]$.

Assumption 1 Let $\mathcal{S}(\mathbb{K}_C^{n_x})$ a class of representable sets in $\mathbb{K}_C^{n_x}$. There exist integers $m, n \geq 1$ and a set of constants $0 \leq L_0, \dots, L_n < \infty$ such that the functions \mathcal{B}_t and \mathcal{B}_0 satisfy

$$\begin{aligned} [V[Y](c) + V[Y](-c)]^{n-1} [\mathcal{B}_t(c, Y, Z) + \mathcal{B}_t(-c, Y, Z)] \\ \leq \sum_{i=0}^n L_i \text{diam}(Y)^{n-i} \text{diam}(Z)^{im} \end{aligned} \quad (49)$$

$$\text{and } \mathcal{B}_0(c, Z) + \mathcal{B}_0(-c, Z) = O(\text{diam}(Z)^m) \quad (50)$$

for all $c \in \mathbb{R}^{n_x}$ with $\|c\| \leq 1$, all $Y \in \mathcal{S}(\mathbb{K}_C^{n_x})$, and all $Z \subseteq P$.

Theorem 4 Let Assumption 1 be satisfied with $m \geq 1$ for a class $\mathcal{S}(\mathbb{K}_C^{n_x})$ of representable sets, and let the right-hand side function $f(t, \cdot, \cdot)$ be locally Lipschitz-continuous. If $V[Y(\cdot, Z)](c) : [0, T] \rightarrow \mathbb{R}$ denotes the solution of the initial value problem (48) for all $c \in \mathbb{R}^{n_x}$ and all $Z \subseteq P$, then $Y(t, \cdot)$ has convergence order m on P , for all $t \in [0, T]$.

Proof From conditions (49,50) of Assumption 1 and by property (4), there exist an integer $n \geq 1$ and a set of constants $0 \leq L_0, \dots, L_n < \infty$, $C_0, \geq 1$ and $0 \leq C_1 < \infty$ such that the solutions $V[Y(\cdot, Z)](c)$ to the initial value problem (48) also satisfy

$$\begin{aligned} \text{a.e. } t \in [0, T], \quad \frac{d}{dt} [V[Y(t, Z)](c) + V[Y(t, Z)](-c)]^n \\ \leq n \sum_{i=0}^n \frac{L_i}{C_1} \left[\max_{c' \in \mathbb{R}^n, \|c'\| \leq 1} V[Y(t, Z)](c') + V[Y(t, Z)](-c') \right]^{n-i} \text{diam}(Z)^{im} \\ \text{and } V[Y(0, Z)](c) \leq C_0 \text{diam}(Z)^m, \end{aligned}$$

for all $c \in \mathbb{R}^{n_x}$ with $\|c\| \leq 1$, and all $Z \subseteq P$ with sufficiently small $\text{diam}(Z)$. Then, we have

$$\begin{aligned} \text{a.e. } t \in [0, T], \quad \frac{d}{dt} \left[\max_{c \in \mathbb{R}^n, \|c\| \leq 1} V[Y(t, Z)](c) + V[Y(t, Z)](-c) \right]^n \\ \leq \max_{c \in \mathbb{R}^n, \|c\| \leq 1} \frac{d}{dt} [V[Y(t, Z)](c) + V[Y(t, Z)](-c)]^n \\ \leq n \sum_{i=0}^n \frac{L_i}{C_1} \left[\max_{c \in \mathbb{R}^n, \|c\| \leq 1} V[Y(t, Z)](c) + V[Y(t, Z)](-c) \right]^{n-i} \text{diam}(Z)^{im}. \end{aligned}$$

The result of the theorem follows from the application of Lemma 3, with the choice of $u(t) := \max_{c \in \mathbb{R}^n, \|c\| \leq 1} V[Y(t, Z)](c) + V[Y(t, Z)](-c)$, and property (4). \square

6.2 Convergence Rate of Standard Differential Inequalities and Ellipsoidal Bounds

Starting with the case of standard differential inequalities as given in Theorem 1, we follow the same approach as in the proof of Proposition 2 by defining

$$\mathcal{B}_t(c, Y, Z) := c^T \text{mid}(F_t(Y, Z)) + \text{abs}(c)^T \text{rad}(F_t(Y, Z)),$$

with the range bounder $F_t(Y, Z) := [f_t^L(Y, Z), f_t^U(Y, Z)] \in \mathbb{IR}^{n_x}$ such that

$$[f_t^L(Y, Z), f_t^U(Y, Z)] \supseteq \left[\min_{\xi, \rho} \left\{ f_i(t, \xi, \rho) \left| \begin{array}{l} \xi_i = y_i^L \\ \xi \in Y \\ \rho \in Z \end{array} \right. \right\}, \max_{\xi, \rho} \left\{ f_i(t, \xi, \rho) \left| \begin{array}{l} \xi_i = y_i^U \\ \xi \in Y \\ \rho \in Z \end{array} \right. \right\} \right].$$

for each $i \in \{1, \dots, n_x\}$. Provided that F_t is such that

$$\forall Y \in \mathbb{IR}^{n_x}, \forall Z \subseteq P, \quad \text{diam}(F_t^i(Y, Z)) \leq O(\text{diam}(Y)) + O(\text{diam}(Z)), \quad (51)$$

it is easy to see that condition (49) in Assumption (1) can be satisfied with $m = n = 1$ for the class of representable sets $\mathcal{S}(\mathbb{K}_{\mathbb{C}}^{n_x}) := \mathbb{IR}^{n_x}$. In particular, we note that if the right-hand side function $f(t, \cdot, \cdot)$ is locally Lipschitz-continuous on $Y \times Z$, then its range converges with order at least, but in general no better than, $m = 1$; see, e.g., [11, Lemma 1]. Moreover, F_t can be evaluated using natural interval extensions or other convergent forms of interval extensions.

Likewise, provided that a range bounder $X_0(Z) := [x_0^L(Z), x_0^U(Z)] \in \mathbb{IR}^{n_x}$ is available such that

$$\begin{aligned} X_0(Z) &\supseteq \left[\min_{\rho} \{x_0(\rho) \mid \rho \in Z\}, \max_{\rho} \{x_0(\rho) \mid \rho \in Z\} \right], \\ \text{and } \forall Z \subseteq P, \quad \text{diam}(X_0(Y, Z)) &\leq O(\text{diam}(Y)) + O(\text{diam}(Z)), \end{aligned} \quad (52)$$

condition (50) can be satisfied with $m = 1$ as well. This is the case, in general, when the initial value function x_0 is (locally) Lipschitz-continuous and natural interval extensions are used.

With this in mind, the following convergence result follows readily from Theorem 4.

Corollary 1 *Consider the initial value problem (1), and let the right-hand side function f and the initial value function x_0 be locally Lipschitz-continuous in all their arguments. For any subset $Z \subseteq P$ with sufficiently small $\text{diam}(Z)$, let the trajectories $y^L(\cdot, Z), y^U(\cdot, Z) : [0, T] \rightarrow \mathbb{R}^{n_x}$, with $-\infty < y^L(\cdot, Z) \leq y^U(\cdot, Z) < \infty$, be Lipschitz-continuous and satisfy*

$$\begin{aligned} \text{a.e. } t \in [0, T], \quad \dot{y}_i^L(t, Z) &= f_{t,i}^L([y^L(t, Z), y^U(t, Z)], Z) \\ \dot{y}_i^U(t, Z) &= f_{t,i}^U([y^L(t, Z), y^U(t, Z)], Z), \\ \text{and } y_i^L(0, Z) &= x_0^L(Z), \quad y_i^U(0, Z) = x_0^U(Z). \end{aligned}$$

If conditions (51) and (52) hold, then the reachable set enclosure $Y(t, \cdot) := [y^L(t, \cdot), y^U(t, \cdot)]$ has convergence order $m \geq 1$ on P for all $t \in [0, T]$.

A direct consequence of Corollary 1 is that the convergence rate of $Y(t, P)$ towards the actual reachable set $X(t, P)$ as $\text{diam}(P) \rightarrow 0$, in the sense of the Hausdorff metric, is itself (at least) linear; that is,

$$d([y^L(t, P), y^U(t, P)], X(t, P)) \leq O(\text{diam}(P)).$$

We show in the following example that this bound is sharp.

Example 1 (continued) We revisit the linear ODE system (9) in the context of convergence analysis. For any interval vector $Z := \rho[-1, 1]^2$, with $0 \leq \rho \leq 1$, we have

$$X(t, Z) = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix} Z \quad \text{and} \quad Y(t, Z) = \exp(t)Z, \quad (53)$$

from which it can be established that $d_H(X(t, Z), Y(t, Z)) \geq (\exp(t) - 1) \text{diam}(Z)$. \diamond

Next, we analyze the convergence rate of the ellipsoidal bounding method as formulated in Theorem 2. The class of representable sets of interest in this case is $\mathcal{S}(\mathbb{K}_C^{n_x}) := \mathcal{E}(Q)$, with $Q \in \mathbb{S}_+^{n_x}$. Using (36), the tight version of the sufficient condition (17)—that is, with equality—can be rewritten in the form of (48) with

$$\mathcal{B}_t(c, \mathcal{E}(Q), Z) := \frac{c^T \left[A(t)Q + QA(t)^T + \sum_{i=1}^{n_x} \kappa_i(t)Q + (\text{diag}(\kappa(t))^{-1} \text{diag rad}(\Omega_t(Q)))^2 \right] c}{2\sqrt{c^T Q c}},$$

as parameterized by some continuous function $\kappa : [0, T] \rightarrow \mathbb{R}_{++}^{n_x}$. It follows that

$$\begin{aligned} & [V[\mathcal{E}(Q)](c) + V[\mathcal{E}(Q)](-c)] [\mathcal{B}_t(c, \mathcal{E}(Q), Z) + \mathcal{B}_t(-c, \mathcal{E}(Q), Z)] \quad (54) \\ &= 2c^T \left[A(t)Q + QA(t)^T + \sum_{i=1}^{n_x} \kappa_i(t)Q + (\text{diag}(\kappa(t))^{-1} \text{diag rad}(\Omega_t(Q)))^2 \right] c \\ &\leq \|c\|^2 (C_1 \|Q\| + C_2 \|\text{diag rad}(\Omega_t(Q))\|), \end{aligned}$$

for some constants $0 \leq C_1, C_2 < \infty$. By definition, a nonlinearity bounder Ω_t satisfying (15) can always be constructed such that

$$\forall Q \in \mathbb{S}^{n_x}, \forall Z \subseteq P, \quad \|\text{diag rad}(\Omega_t(Q))\| \leq O(\text{diam}(\mathcal{E}(Q))^2) + O(\text{diam}(Z)^2). \quad (55)$$

Combining (54) and (55) gives

$$\begin{aligned} & [V[\mathcal{E}(Q)](c) + V[\mathcal{E}(Q)](-c)] [\mathcal{B}_t(c, \mathcal{E}(Q), Z) + \mathcal{B}_t(-c, \mathcal{E}(Q), Z)] \\ &\leq O(\text{diam}(\mathcal{E}(Q))^2) + O(\text{diam}(Z)^4), \end{aligned}$$

for all $c \in \mathbb{R}^{n_x}$ with $\|c\| \leq 1$, all $Q \in \mathbb{S}^{n_x}$ with sufficiently small $\text{diam}(\mathcal{E}(Q))$, and all $Z \subseteq P$ with sufficiently small $\text{diam}(Z)$. Therefore, condition (49) is satisfied with $m = n = 2$ when (55) holds. In practice, Ω_t can be computed as the remainder of a first-order Taylor model of $f(t, \cdot, \cdot)$ at $(\hat{x}(t), \hat{p})$ on $\{\hat{x}(t)\} \oplus G(t)[P - \hat{p}] \oplus \mathcal{E}(Q) \times P$ in the case that $f(t, \cdot, \cdot)$ is twice continuously-differentiable [12, 48].

Likewise, a nonlinearity bounder Ω_0 satisfying (16) can be constructed such that

$$\forall Z \subseteq P, \quad \|\text{diag rad}(\Omega_0)\| \leq O(\text{diam}(Z)^2), \quad (56)$$

and condition (50) can thus be satisfied with $m = 2$ as well. For instance, Ω_0 can be computed as the remainder of a first-order Taylor model of x_0 at \hat{p} on P provided that x_0 is twice continuously-differentiable.

The following convergence result follows readily from Theorem 4 based on the foregoing considerations.

Corollary 2 *Consider the initial value problem (1), and let the right-hand side function f and the initial value function x_0 be continuously-differentiable in all their arguments. For any subset $Z \subseteq P$ with sufficiently small $\text{diam}(Z)$, let the matrix-valued functions $Q(\cdot, Z) : [0, T] \rightarrow \mathbb{S}^{n_x}$ be Lipschitz-continuous and such that*

$$\begin{aligned} \text{a.e. } t \in [0, T], \quad \dot{Q}(t, Z) &= A(t)Q(t, Z) + Q(t, Z)A(t)^T + \sum_{i=1}^{n_x} \kappa_i(t)Q(t, Z) \\ &\quad + \text{diag}(\kappa(t))^{-1} \text{diag rad}(\Omega_t(Q(t, Z)))^2 \\ \text{and } Q(0, Z) &= \text{diag rad}(\Omega_0)^2, \end{aligned}$$

where Ω_t and Ω_0 satisfy the conditions (55) and (56), and $\kappa : [0, T] \rightarrow \mathbb{R}^{n_x}$ is such that $\kappa(t) \in [\kappa^L, \kappa^U]$ with $0 < \kappa^L \leq \kappa^U < \infty$. Then, the ellipsoidal bound $\mathcal{E}(Q(t, \cdot))$ has convergence order $m \geq 2$ on P for all $t \in [0, T]$.

A direct consequence of Corollary 2 is that the convergence rate of $Y(t, P) := \{\hat{x}(t)\} \oplus G(t)[P - \hat{p}] \oplus \mathcal{E}(Q(t, P))$ towards the actual reachable set $X(t, P)$ when $\text{diam}(P) \rightarrow 0$ is itself (at least) quadratic; that is,

$$d(\{\hat{x}(t)\} \oplus G(t)[P - \hat{p}] \oplus \mathcal{E}(Q(t, P)), X(t, P)) \leq O(\text{diam}(P)^2).$$

The sharpness of this bound will be established indirectly in Example 3, by checking the convergence rate of first-order Taylor models with ellipsoidal bounds.

6.3 Convergence of Taylor Models with Convex Remainder Bound

The following analysis of the convergence properties of the remainder bounds in Taylor models assumes that the right-hand side function $f(t, \cdot, \cdot)$ and the initial value function x_0 are both $(q + 1)$ -times continuously-differentiable. Here, the bounds \mathcal{B}_t are defined such that

$$\mathcal{B}_t(c, Y, Z) \geq \max_{\xi, \rho} \left\{ c^T r_f^q(t, \xi, \rho) \left| \begin{array}{l} \xi \in Y \\ \rho \in Z \\ c^T \xi = V[Y](c) \end{array} \right. \right\},$$

with the remainder function r_f^q given by (40). By construction, r_f^q satisfies

$$\begin{aligned} \left\| r_f^q(t, \xi, \rho) \right\| &\leq \left\| f(t, \mathcal{P}_x^q(t, \rho), \rho) - \dot{\mathcal{P}}_x^q(t, \rho) + \frac{\partial f}{\partial x}(t, \eta \xi, \rho) \xi \right\| \quad \text{for some } \eta \in [0, 1] \\ &\leq O(\|\rho - \hat{p}\|^{q+1}) + O(\|\xi\|), \end{aligned}$$

for all $(\xi, \rho) \in \mathcal{E}(Q_x^q(t)) \times P$, where the last inequality follows by the convergence properties of q th-order Taylor models [12, 48]. An enclosure $\mathcal{R}_{f,t}(Y, Z) \supseteq \{r_f^q(t, \xi, \rho) \mid (\xi, \rho) \in Y \times Z\}$ in $\mathbb{K}_{\mathbb{C}}^{n_x}$ satisfying

$$\forall Y \in \mathbb{K}_{\mathbb{C}}^{n_x}, \forall Z \subseteq P, \quad \text{diam}(\mathcal{R}_{f,t}(Y, Z)) \leq O(\text{diam}(Y)) + O(\text{diam}(Z)^{q+1}) \quad (57)$$

can thus be obtained, e.g., based on a q th-order Taylor model of $f(t, \mathcal{P}_x^q(t, \cdot), \cdot)$ on Z as well as on an interval extension of $\frac{\partial f}{\partial x}(t, \cdot, \cdot)$ on $Y \times Z$. It follows that condition (49) can be satisfied with $m = q + 1$ and $n = 1$ if condition (57) holds.

Likewise, the bound \mathcal{B}_0 is defined such that

$$\mathcal{B}_0(c, Z) \geq \max_{\rho} \{ c^T [x_0(\rho) - \mathcal{P}_{x_0}^q(\rho)] \mid \rho \in Z \},$$

and an enclosure $\mathcal{R}_{x_0}(Z) \supseteq \{x_0(\rho) - \mathcal{P}_{x_0}^q(\rho) \mid \rho \in Z\}$ in $\mathbb{K}_{\mathbb{C}}^{n_x}$ satisfying

$$\forall Z \subseteq P, \quad \text{diam}(\mathcal{R}_{x_0}(Z)) \leq O(\text{diam}(Z)^{q+1}) \quad (58)$$

can be obtained, e.g., as the remainder of a q th-order Taylor model of x_0 . Therefore, condition (50) can be satisfied with $m = q + 1$ if (58) holds.

At this stage, we note that various parameterization of the remainder enclosure $Y(\cdot, Z)$ can be used, given that conditions (49) and (50) are satisfied. For instance, both interval remainders and ellipsoidal remainders can be propagated, as described in Sect. 5.2. These variants differ in the way the bounders \mathcal{B}_t and \mathcal{B}_0 are constructed, similar in essence to the constructions in Sect. 6.2.

The following convergence result follows readily from Theorem 4 and the above considerations.

Corollary 3 *Consider the initial value problem (1), and let the right-hand side function f and the initial value function x_0 be $(q+1)$ -times continuously-differentiable in all their arguments. For any subset $Z \subseteq P$ with sufficiently small $\text{diam}(Z)$, let the set-valued functions $Y(\cdot, Z) : [0, T] \rightarrow \mathbb{K}_{\mathbb{C}}^{n_x}$ be such that $V[Y(\cdot, Z)](c)$ is locally Lipschitz-continuous and satisfy (48) for all $c \in \mathbb{R}^{n_x}$. If conditions (57) and (58) hold, then the remainder bound $\mathcal{R}_x^q(Y(t, \cdot))$ has convergence order $m \geq q+1$ on P for all $t \in [0, T]$.*

In turn, the convergence rate of the q th-order Taylor models $\mathcal{P}_x^q(t, P) \oplus \mathcal{R}_x^q(t, P)$ to the actual reachable set $X(t, P)$ as $\text{diam}(P) \rightarrow 0$ is of order (at least) $q+1$. For applications in global optimization however, bounding the multivariate polynomial part too presents some challenges. A variety of range-bounding strategies for the Taylor polynomial part have been proposed, some of which enjoy quadratic or higher convergence rate to the actual polynomial range; see, e.g., [33, 35, 48].

In order to keep our considerations general, we assume here that a set-valued function $\mathcal{T}_x^q(\cdot, P) : [0, T] \rightarrow \mathbb{R}^{n_x}$ is available such that

$$\forall t \in [0, T], \quad \mathcal{T}_x^q(t, P) \supseteq \{\mathcal{P}_x^q(t, p) \mid p \in P\} \text{ and } d(\mathcal{P}_x^q(t, P), \mathcal{T}_x^q(t, P)) \leq O(\text{diam}(P)^{\phi(q)}),$$

for a known order function $\phi : \mathbb{N} \rightarrow \mathbb{N}$. It follows from the triangular inequality (for the Hausdorff metric d) that

$$\forall t \in [0, T], \quad d(\mathcal{T}_x^q(t, P) \oplus \mathcal{R}_x^q(t, P), X(t, P)) \leq O(\|P\|^{\min\{\phi(q), q+1\}}).$$

The convergence properties of Taylor models with interval and ellipsoidal remainder bounds are illustrated in the following example, which assumes exact range bounding of the multivariate polynomial part.

Example 3 The dynamic model of an artificial genetic circuit with three states [69], known as *repressilator*, is considered:

$$\dot{x}_1 = \frac{p_1}{1+x_3^3} - p_2x_1 + p_3, \quad \text{with } x_1(0) = 5.5 \quad (59)$$

$$\dot{x}_2 = \frac{p_1}{1+x_1^3} - p_2x_2 + p_3, \quad \text{with } x_2(0) = 3.5 \quad (60)$$

$$\dot{x}_3 = \frac{p_1}{1+x_2^3} - p_2x_3 + p_3, \quad \text{with } x_3(0) = 4.5, \quad (61)$$

with $p \in P := ([215, 216] [0.995, 1.005] [1.495, 1.505])^T$. Taylor models are propagated continuously in time to enclose the reachable set $X(t, P)$ of (59-61). These estimators are constructed in the same way as earlier in Example 2 and further details about the numerical solution procedure will be given in Sect. 7 below.

The left plot of Figure 3 shows the projection of $X(t, P)$ (shaded area) onto x_1 , as well as the projections of the corresponding Taylor model enclosures $\mathcal{P}_x^q(t, P) \oplus \mathcal{R}_x^q(t, P)$ for both

interval remainder bounds (dashed lines) and ellipsoidal remainder bounds (solid lines). Notice how increasing the Taylor model order q , here from $q = 1$ to 5, delays the time when the enclosure size blows up. Moreover, enclosing the remainder within ellipsoidal bounds instead of interval bounds provides tighter bounds in this case. The right plot of Figure 3 illustrates the corresponding convergence rate of $d_{\text{H}}(\mathcal{P}_x^q(t, P) \oplus \mathcal{R}_x^q(t, P), X(t, P))$ at time $t = 10$ as $\text{diam}(P) \rightarrow 0$. It is found that all q th-order Taylor enclosure converge with order $q + 1$, in agreement with the theory, for both types of remainder enclosures. This also establishes that the convergence bounds derived previously are indeed sharp. \diamond

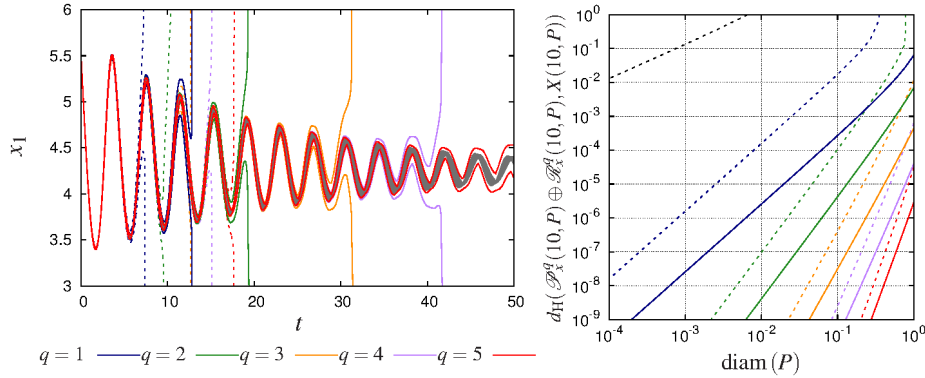


Fig. 3 *Left plot:* Projections onto x_1 of the reachable set $X(t)$ and the enclosures $\mathcal{P}_x^q(t, P) \oplus \mathcal{R}_x^q(t, P)$ for both interval and ellipsoidal remainder bounds. *Right plot:* Convergence of the Taylor model enclosure to the actual reachable set at $t = 10$ in the Hausdorff metric $d_{\text{H}}(\mathcal{P}_x^q(t, P) \oplus \mathcal{R}_x^q(t, P), X(t, P))$. The reachable set is represented with a shaded area and the Taylor model enclosures with interval and ellipsoidal remainders with, respectively, dashed lines and solid lines.

7 Numerical Implementation and Case Study

This section describes the main implementation details for the bounding approaches presented in Sect. 5, before investigating a more challenging case study.

The main difficulty in using the sufficient conditions in Theorems 1 and 2 and Propositions 4 and 5 for computing bounds on the parametric solutions of (1) is constructing and evaluating the right-hand sides of the auxiliary bounding systems (7,8), (17), (42,43) and (47), respectively. Our implementation assumes that the right-hand side function f in (1) and the corresponding initial value function x_0 are both factorable. We use verified libraries such as PROFIL (<http://www.ti3.tu-harburg.de/>) for interval analysis as well as the library MC++ [43] for Taylor model arithmetic. For numerical ODE integration, we use explicit Runge-Kutta schemes as this avoids computing the Jacobian of the right-hand side of the auxiliary ODEs. More specifically, our implementation relies on the explicit solvers with adaptive step-size control available as part of the GNU Scientific Library (GSL)—function `gsl_odeiv2` in GSL ver. 1.5. The developed continuous-time set-propagation code comes in the form of a C++ class called `ODEBND_GSL`, which is made freely available at: <http://www3.imperial.ac.uk/environmentenergyoptimisation/software>.

Standard Differential Inequalities Technique Theorem 1 is applied with equalities in (7,8), and the resulting $2n_x$ lower- and upper-bounding ODEs are integrated all together, forward in time, as they are coupled. At a given time $t \in [0, T]$, the right-hand sides of these ODEs are evaluated component-wise. For instance, the right-hand side of the i th lower-bounding ODE (7) is evaluated by taking the lower bound from the natural interval extension of $f_i(t, \cdot, \cdot)$ on $L^i \times P$, with $L_j^i := \{\min Y_j(t)\}$ and $L_j^i := Y_j(t)$ if $j \neq i$.

Ellipsoidal Bounding Technique Theorem 2 is applied with equality in (17), and the ODEs describing $\hat{x}(t)$, $G(t)$ and $Q(t)$ are integrated all together, forward in time. Of the alternatives to construct the right-hand sides of these auxiliary ODEs at a given time $t \in [0, T]$, our implementation considers a q' -th-order Taylor model of $f(t, \cdot, \cdot)$ at $(\hat{x}(t), \hat{p})$ on $\Xi_t \times P$, with $q' \geq 1$ and where Ξ_t denotes the interval hull of $\hat{x}(t) \oplus G(t)[P - \hat{p}] \oplus \mathcal{E}(Q(t))$ and the reference point \hat{p} is the midpoint of P . Writing such a Taylor model in the form

$$\forall (\xi, \rho) \in \Xi_t \times P, \quad f(t, \xi, \rho) \in c_{f_i} + L_{f_i}^x (\xi - \hat{x}(t)) + L_{f_i}^p (\rho - \hat{p}) + B_{f_i},$$

with $c_{f_i} \in \mathbb{R}^{n_x}$, $L_{f_i}^x \in \mathbb{R}^{n_x \times n_x}$, $L_{f_i}^p \in \mathbb{R}^{n_x \times n_p}$ and $B_{f_i} \in \mathbb{I}\mathbb{R}^{n_x}$, it follows that B_{f_i} can be used as the nonlinearity bounder Ω_t in (15) and the right-hand side of the auxiliary bounding ODE system can be evaluated as

$$\begin{aligned} \dot{\hat{x}}(t) &= c_{f_i} \\ \dot{G}(t) &= L_{f_i}^p \\ \dot{Q}(t) &= (L_{f_i}^x) Q(t) + Q(t) (L_{f_i}^x)^T + \sum_{i=1}^{n_x} \kappa_i(t) Q(t) + \text{diag}(\kappa(t))^{-1} \text{diag rad}(B_{f_i})^2. \end{aligned}$$

The ellipsoidal parameterization $\kappa(t)$ in the latter ODE is so chosen as to minimize $\text{tr } \dot{Q}(t)$. A direct evaluation of $\kappa(t)$ as

$$\forall t \in [0, T], \quad \kappa(t) = \frac{\text{rad}(B_{f_i})}{\sqrt{\text{tr } Q(t)}},$$

would nonetheless result in $\kappa(t) \rightarrow 0$ as $\text{diam}(\mathcal{E}(Q(t))) \rightarrow 0$, thereby impairing the quadratic convergence property established in Corollary 2. The following modification

$$\forall t \in [0, T], \quad \forall i \in \{1, \dots, N\}, \quad \kappa_i(t) = \frac{\sqrt{a_i(t) + \varepsilon}}{\text{TOL} \sqrt{a_i(t) + \varepsilon} + \sqrt{b(t) + \varepsilon}} \quad (62)$$

$$\text{with } a_i(t) = \frac{1}{s_i(t)} \text{rad}(B_{f_i}), \quad b(t) = \text{tr} \left(\text{diag}(s(t))^{-1} Q(t) \right),$$

and with scaling factors $s_i(t) = 1$ or $s_i(t) = x_i(t)^2$, ensures that $\kappa(t) \geq \frac{1}{1+\text{TOL}} > 0$ for all $t \in [0, T]$ —and hence quadratic convergence. Our implementation uses $\varepsilon = 2^{-52}$ and $\text{TOL} = 10^{-6}$ as default values. Finally, initial conditions for the auxiliary bounding ODEs are evaluated as

$$\hat{x}(0) = c_{x_0}, \quad G(0) = L_{x_0} \quad \text{and} \quad Q(0) = \text{diag rad}(B_{x_0})^2,$$

based on a q' -th-order Taylor model of x_0 at \hat{p} on P given in the form

$$\forall \rho \in P, \quad x_0(\rho) \in c_{x_0} + L_{x_0}(\rho - \hat{p}) + B_{x_0},$$

with $c_{x_0} \in \mathbb{R}^{n_x}$, $L_{x_0} \in \mathbb{R}^{n_x \times n_p}$ and $B_{x_0} \in \mathbb{I}\mathbb{R}^{n_x}$. Our implementation considers second-order Taylor models of $f(t, \cdot, \cdot)$ and x_0 by default, i.e. $q' = 2$. The use of higher-order Taylor models can provide tighter enclosures by capturing more dependencies in some cases, but this also causes a significant computational overhead.

Taylor Model with Differential Inequalities Bounding Technique A q th-order Taylor model with interval remainder, $(\mathcal{P}_x^q(t, \cdot), [r_x^{q,L}(t), r_x^{q,U}(t)])$, is computed based on Proposition 4 with equalities in (42,43). The $2n_x$ remainder bounds are propagated together with the coefficients $\partial^{\kappa}x(\cdot, \hat{p})$ of the Taylor polynomial (41), forward in time. At a given time $t \in [0, T]$ and for each component $i = 1, \dots, n_x$, our implementation computes a q th-order Taylor model on P at \hat{p} of the composite function

$$f_i(t, (\mathcal{P}_x^q(t, \cdot), L_x^i), \cdot) =: (f_{f_i}^q, L_{f_i}^i), \quad \text{with} \quad L_{x_j}^i := \begin{cases} [r_{x_j}^{q,L}(t), r_{x_j}^{q,U}(t)] & \text{if } i \neq j, \\ \{r_{x_i}^{q,L}(t)\} & \text{otherwise.} \end{cases}$$

By construction, the coefficients of the multivariate polynomial $\mathcal{P}_{f_i}^q$ match the time derivatives $\partial^{\kappa}\dot{x}(t, \hat{p})$ of the polynomial coefficients in (41), and the lower bound $(\min L_{f_i}^i)$ provides the right-hand side of (42). The right-hand side of (43) is constructed in a likewise manner.

Taylor Model with Ellipsoidal Remainder Bounding Technique A q th-order Taylor model with ellipsoidal remainder, $(\mathcal{P}_x^q(t, \cdot), Q_x^q(t))$, is computed based on Proposition 5 with equality in (47). Here, the shape matrix Q_x^q of the ellipsoidal remainder is propagated with the coefficients $\partial^{\kappa}x(\cdot, \hat{p})$ of the Taylor polynomial (41), forward in time. Of the alternatives to construct the right-hand sides of these auxiliary ODEs at a given time $t \in [0, T]$, one approach—referred to as the *full approach* subsequently—involves computing a q th-order Taylor model (with interval remainder) of the composite function $f(t, \mathcal{P}_x^q(t, p) + r, p)$ at $(0, \hat{p})$ for $(r, p) \in R_t \times P$, where R_t denotes the interval hull of $\mathcal{E}(Q_x^q(t))$. Writing such a Taylor model in the form

$$\forall (r, \rho) \in R_t \times P, \quad f(t, \mathcal{P}_x^q(t, \rho) + r, \rho) \in \sum_{\substack{\gamma \in \mathbb{N}^{n_p} \\ |\gamma| \leq q}} \alpha_{f_i, \gamma} (\rho - \hat{p})^\gamma + L_{f_i}^r r + B_{f_i},$$

with $\alpha_{f_i, \gamma} \in \mathbb{R}^{n_x}$, $L_{f_i}^r \in \mathbb{R}^{n_x \times n_x}$ and $B_{f_i} \in \mathbb{IR}^{n_x}$, we have that $L_{f_i}^r$ and B_{f_i} can be used, respectively, as the matrix $A_{f_i}(t)$ in (44) and as the nonlinearity bound $\Omega_{f_i}^q$ in (45); that is, the right-hand side of the auxiliary bounding ODE system can be evaluated as

$$\begin{aligned} \partial^\gamma \dot{x}(t, \hat{p}) &= \alpha_{f_i, \gamma}, \quad \text{for all } \gamma \in \mathbb{N}^{n_p}, \text{ with } |\gamma| \leq q \\ \dot{Q}_x^q(t) &= (L_{f_i}^r) Q_x^q(t) + Q_x^q(t) (L_{f_i}^r)^\top + \sum_{i=1}^{n_x} \kappa_i(t) Q_x^q(t) + \text{diag}(\kappa(t))^{-1} \text{diag rad}(B_{f_i})^2. \end{aligned}$$

Moreover, the ellipsoidal parameterization $\kappa(t)$ in (47) is chosen in a similar way as (62), here for minimizing $\text{tr} Q_x^q(t)$, possibly after scaling.

The full approach involves Taylor expanding the function $f(t, \mathcal{P}_x^q(t, p) + r, p)$ jointly in (r, p) up to order q , which can prove computationally demanding when either the dynamic system size n_x or the expansion order q is large. An alternative approach—referred to as the *mean-value approach* hereafter—proceeds in two steps:

- i. Compute a q th-order Taylor model (with interval remainder) of the composite function $f(t, \mathcal{P}_x^q(t, \cdot), \cdot)$ at \hat{p} on P :

$$\forall \rho \in P, \quad f(t, \mathcal{P}_x^q(t, \rho), \rho) \in \sum_{\substack{\gamma \in \mathbb{N}^{n_p} \\ |\gamma| \leq q}} \alpha_{f_i, \gamma} (\rho - \hat{p})^\gamma + B_{f_i}^0,$$

with $\alpha_{f_i, \gamma} \in \mathbb{R}$ and $B_{f_i}^0 \in \mathbb{IR}^{n_x}$;

- ii. Compute a q' -th-order Taylor model (with interval remainder) of the Jacobian matrix $\frac{\partial f}{\partial x}(t, \mathcal{P}_{x_i}^{q'}(\rho) + r, \rho)$ at $(0, \hat{p})$ for $(r, \rho) \in \mathbb{R}_{x_i}^{q'} \times P$, with $0 \leq q' \leq q$ and where $(\mathcal{P}_{x_i}^{q'}, \mathbb{R}_{x_i}^{q'})$ is a q' -th-order Taylor model of $x(t, \cdot)$ on P at \hat{p} :

$$\forall (r, \rho) \in \mathbb{R}_{x_i}^{q'} \times P, \quad \frac{\partial f}{\partial x}(t, \mathcal{P}_{x_i}^{q'}(\rho) + r, \rho) \in C_{\partial_{x_i} f} + B_{\partial_{x_i} f},$$

with $C_{\partial_{x_i} f} \in \mathbb{R}^{n_x \times n_x}$ and $B_{\partial_{x_i} f} \in \mathbb{I}\mathbb{R}^{n_x \times n_x}$.

By construction, $C_{\partial_{x_i} f}$ can be used as the matrix $A_f(t)$ in (44), whereas $B_{f_i}^0 + (B_{\partial_{x_i} f} - C_{\partial_{x_i} f})\mathbb{R}_{x_i}^{q'}$ can be used as the nonlinearity bounder Ω_i^q in (45). This way, the right-hand side of the auxiliary bounding ODE system can be evaluated as

$$\begin{aligned} \partial^\gamma \dot{x}(t, \hat{p}) &= \alpha_{f_i, \gamma}, \quad \text{for all } \gamma \in \mathbb{N}^{n_p}, \text{ with } |\gamma| \leq q \\ \dot{Q}_x^q(t) &= (C_{\partial_{x_i} f}) Q_x^q(t) + Q_x^q(t) (C_{\partial_{x_i} f})^\top + \sum_{i=1}^{n_x} \kappa_i(t) Q_x^q(t) \\ &\quad + \text{diag}(\kappa(t))^{-1} \text{diag rad} \left(B_{f_i}^0 + (B_{\partial_{x_i} f} - C_{\partial_{x_i} f})\mathbb{R}_{x_i}^{q'} \right)^2. \end{aligned}$$

Our implementation of the simplified approach uses the automatic differentiation package FADBAD++ (<http://www.fadbad.com/fadbad.html>) in order to compute the Jacobian matrix $\frac{\partial f}{\partial x}$ and first-order Taylor models are used to bound the entries of this matrix by default ($q' = 1$). The use of second- or higher-order Taylor models can provide tighter enclosures by capturing more dependencies, but this also causes a significant computational overhead.

7.1 Case Study: Anaerobic Digestion

Consider the following six-state model representing the dynamics of an anaerobic digester, as originally proposed by [7]. Enclosing the solutions of this model is challenging as it features complex dynamics due to pH self-regulation and liquid-gas transfer. Moreover, the system presents both fast dynamics on a time-scale of hours and slow dynamics on a time-scale of days.

$$\dot{X}_1 = (\mu_1(S_1) - \alpha D)X_1 \quad (63)$$

$$\dot{X}_2 = (\mu_2(S_2) - \alpha D)X_2 \quad (64)$$

$$\dot{S}_1 = D(S_1^{\text{in}} - S_1) - k_1 \mu_1(S_1)X_1 \quad (65)$$

$$\dot{S}_2 = D(S_2^{\text{in}} - S_2) + k_2 \mu_1(S_1)X_1 - k_3 \mu_2(S_2)X_2 \quad (66)$$

$$\dot{Z} = D(Z^{\text{in}} - Z) \quad (67)$$

$$\dot{C} = D(C^{\text{in}} - C) - q_{\text{CO}_2} + k_4 \mu_1(S_1)X_1 + k_5 \mu_2(S_2)X_2, \quad (68)$$

with

$$q_{\text{CO}_2} := k_L a(C + S_2 - Z - K_H P_{\text{CO}_2}) \quad (69)$$

$$P_{\text{CO}_2} := \frac{\phi_{\text{CO}_2} - \sqrt{\phi_{\text{CO}_2}^2 - 4K_H P_1(C + S_2 - Z)}}{2K_H} \quad (70)$$

$$\phi_{\text{CO}_2} := C + S_2 - Z + K_H P_1 + \frac{k_6}{k_L a} \mu_2(S_2) X_2 \quad (71)$$

$$\mu_1(S_1) := \bar{\mu}_1 \frac{S_1}{S_1 + K_{S_1}} \quad (72)$$

$$\mu_2(S_2) := \bar{\mu}_2 \frac{S_2}{S_2 + K_{S_2} + S_2^2/K_{I_2}}. \quad (73)$$

In this case study, uncertainty is in the initial conditions only, while all of the other model parameters are considered constant; these values are taken from [7] and summarized in Table 1 for the sake of reproducibility.

Table 1 Parameters in the anaerobic digestion model (63,73).

Parameter	Value	Parameter	Value
$\bar{\mu}_1$	1.2 day ⁻¹	k_1	42.14 g(COD)g(cell) ⁻¹
K_{S_1}	7.1 g(COD)L ⁻¹	k_2	116.5 mmolg(cell) ⁻¹
$\bar{\mu}_2$	0.74 day ⁻¹	k_3	268.0 mmolg(cell) ⁻¹
K_{S_2}	9.28 mmolL ⁻¹	k_4	50.6 mmolg(cell) ⁻¹
K_{I_2}	256 mmolL ⁻¹	k_5	343.6 mmolg(cell) ⁻¹
$k_L a$	19.8 day ⁻¹	k_6	453.0 mmolg(cell) ⁻¹
K_H	16 mmolL ⁻¹ atm ⁻¹	S_1^{in}	5 g(COD)L ⁻¹
P_1	1 atm	S_2^{in}	80 mmolL ⁻¹
α	0.5 –	Z^{in}	50 mmolL ⁻¹
D	0.4 day ⁻¹	C^{in}	0 mmolL ⁻¹

First, we consider initial conditions as $X_1(0) \in 0.5 \times [0.94, 1.06]$ g(COD)L⁻¹, $X_2(0) \in [0.94, 1.06]$ g(COD)L⁻¹, $S_1(0) = 1$ g(COD)L⁻¹, $S_2(0) = 5$ mmolL⁻¹, $Z(0) = 50$ mmolL⁻¹, and $C(0) \in 40 \times [0.94, 1.06]$ mmolL⁻¹; that is, the model contains three uncertain quantities. Projections onto the variables X_1 and S_2 of the actual reachable set and of the enclosures obtained with standard differential inequalities and the ellipsoidal bounding technique as well as their Taylor model counterparts for various expansion orders are shown in Fig. 4. Standard differential inequalities are found to produce rather weak bounds here, blowing up after about $t \approx 0.09$ day. Combining Taylor models with differential inequalities delays this blow up time significantly, up to about $t \approx 4.01$ day with 4th-order Taylor models. In comparison, bounds computed with the ellipsoidal technique blow up around $t \approx 1.61$ day, thus outperforming 2nd-order Taylor models with differential inequalities remainder bounds. When used in combination with ellipsoidal remainder bounds, Taylor models delay the blow up time significantly as well. It is even observed that 4th- or higher-order Taylor models with ellipsoidal remainder bounds can stabilize the reachable set enclosure for this level of uncertainty; that is, the bounds converge to the actual steady-state values as $t \rightarrow \infty$. These remarkable stability properties of the auxiliary bounding system will be the topic of a follow-up paper.

The results reported in Table 2 are for reduced uncertainty in the initial conditions as $X_1(0) \in 0.5 \times [0.98, 1.02]$ g(COD)L⁻¹, $X_2(0) \in [0.98, 1.02]$ mmolL⁻¹, $S_1(0) = 1$ mmolL⁻¹, $S_2(0) = 5$ mmolL⁻¹, $Z(0) = 50$ mmolL⁻¹, and $C(0) \in 40 \times [0.98, 1.02]$ mmolL⁻¹. Here again, standard differential inequalities produce the weakest bounds, with a blow up time of $t \approx 0.40$ day. The use of ellipsoidal calculus improves the bounds significantly, by delaying the blow up time to $t \approx 1.99$ day and $t \approx 13.52$ day when the nonlinearity bounder

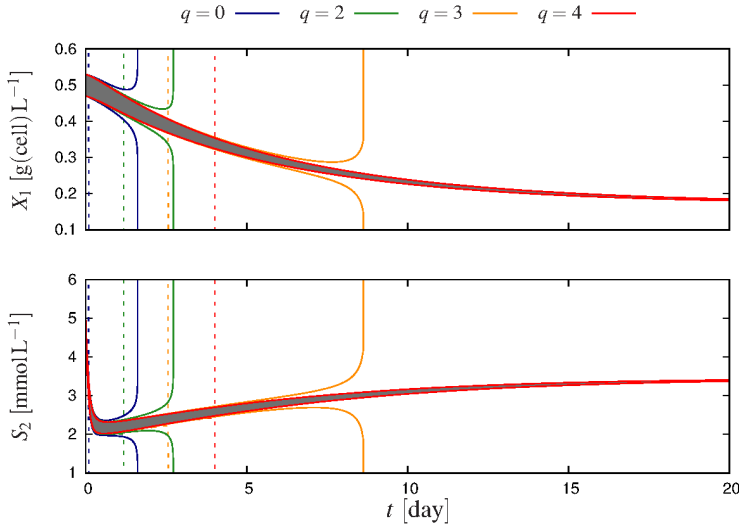


Fig. 4 Projections onto X_1 (top plot) and C (bottom plot) of the reachable set $X(t)$ and the enclosures $Y(t)$ computed with various approaches. The reachable set is represented with a shaded area. The enclosures labelled with $q = 0$ correspond to standard differential inequalities (dashed lines) and the ellipsoidal bounding approach (solid lines). The enclosures labelled $q = 2, 3$ and 4 are Taylor model enclosures with interval remainders (dashed lines) and ellipsoidal remainders (with full approach; solid lines).

Ω_t is constructed by using, respectively, 1st-order and 2nd-order Taylor models. The use of Taylor models as a means of propagating nonconvex enclosures provides tighter bounds as the Taylor model order is increased, which illustrates well the advantage of bounding techniques enjoying higher-order Hausdorff convergence. Ellipsoidal remainder bounds are found to outperform differential inequalities remainder bounds in this case as well. In particular, all three variants of Taylor models with ellipsoidal remainders appear to stabilize the reachable set enclosures for small Taylor expansion orders—as small as $q = 2$ in the mean-value approach with $q' = 1$.

Comparisons of computational performance are made for a fixed integration horizon of $t \in [0, 4]$. As far as Taylor model-based techniques are concerned, the number of right-hand side evaluations—and therefore the number of integration steps—increases slowly with the Taylor expansion order between $q = 2, \dots, 5$, which is due to a larger size of the auxiliary bounding system while the tolerances remain unchanged. For a particular Taylor expansion order q , the use of various bounding schemes for the Taylor remainder does not affect the number of right-hand side evaluations though; this suggests that the step-size is mainly determined by the propagation of the Taylor polynomial here. Regarding CPU times, a significant increase is observed in the full approach of ellipsoidal remainder bounds with the Taylor expansion order between $q = 2, \dots, 5$ compared to differential inequalities remainder bounds. This computational burden is greatly reduced by the use of the mean-value approach with either $q' = 0$ or 1 . In particular, the ellipsoidal remainder bound variant with $q' = 0$ appears to be both faster and to provide tighter bounds than the differential inequalities remainder bound technique as well as the standard ellipsoidal bounding technique.

Table 2 Comparison of computational performance for various continuous-time enclosure methods. The numerical integration algorithm used is the explicit embedded Runge-Kutta-Fehlberg (4, 5) method in GSL, with relative tolerance $\text{RTOL} = 10^{-6}$ and absolute tolerance $\text{ATOL} = 10^{-8}$. The reported CPU times are for an IntelTM CORETMi7 vProTM (4×2.1 GHz) computer with 8Gb RAM, running Ubuntu 12.04.3 LTS (GNU/Linux 3.5.0-45-generic x86_64) and gcc version 4.7.3 (with `-O2` compilation option).

Differential inequalities (DI) and ellipsoidal bounding (EB) techniques				
Technique q	DI	EB ($q' = 1$)	EB ($q' = 2$)	
Blow up time [day]	0.39	1.99	13.55	
RHS evaluations for $t \in [0, 4]$	N/A	N/A	812	
CPU time in [s] for $t \in [0, 4]$	N/A	N/A	0.54	
Taylor model (TM) w/ differential inequalities remainder bounds				
TM Order q	2	3	4	5
Blow up time [day]	3.07	5.31	7.58	9.89
RHS evaluations for $t \in [0, 4]$	N/A	1,077	1,191	1,272
CPU time in [s] for $t \in [0, 4]$	N/A	0.39	0.72	1.33
Taylor model (TM) w/ ellipsoidal remainder bounds – Mean-value approach ($q' = 0$)				
TM Order q	2	3	4	5
Blow up time [day]	4.23	11.12	∞	∞
RHS evaluations for $t \in [0, 4]$	972	1,077	1,191	1,272
CPU time in [s] for $t \in [0, 4]$	0.14	0.23	0.41	0.71
Taylor model (TM) w/ ellipsoidal remainder bounds – Mean-value approach ($q' = 1$)				
TM Order q	2	3	4	5
Blow up time [day]	∞	∞	∞	∞
RHS evaluations for $t \in [0, 4]$	954	1,077	1,191	1,272
CPU time in [s] for $t \in [0, 4]$	0.98	1.20	1.47	1.85
Taylor model (TM) w/ ellipsoidal remainder bounds – Full approach				
TM Order q	2	3	4	5
Blow up time [day]	39.84	∞	∞	∞
RHS evaluations for $t \in [0, 4]$	954	1,077	1,191	1,272
CPU time in [s] for $t \in [0, 4]$	0.35	1.49	6.35	26.37

8 Conclusions

This paper has presented a unified framework for continuous-time propagation of enclosures for the reachable set of parametric nonlinear ODEs. The main contribution is the formulation of a generalized differential inequality, which provides sufficient conditions for the construction of convex enclosures. This framework has been shown to encompass the classical theory of differential inequalities as well as the ellipsoidal bounding technique. Besides being directly applicable for the construction of other types of convex enclosures, it can also be used for constructing nonconvex enclosures via its combination with Taylor models. Another principal contribution has been using this generalized differential inequality to analyze the convergence properties of various kinds of enclosures, namely interval, ellipsoidal, and Taylor model bounds. Sharp bounds on the convergence order of these methods have been obtained and illustrated with numerical examples. Finally, implementation details for the various bounding techniques have been outlined, and we recall that the developed C++ code can be obtained freely at: <http://www3.imperial.ac.uk/environmentenergyoptimisation/software>. The numerical case study of a six-state dynamic model of anaerobic digestion has shown the superior stability properties of nonconvex set-propagation techniques based on Taylor models with ellipsoidal remainder bounds compared to either Taylor models with differential inequalities remainder bounds or convex

set-propagation techniques. On the whole, this approach appears to be well suited for use in global and robust dynamic optimization.

Acknowledgements This paper is based upon work supported by the Engineering and Physical Sciences Research Council (EPSRC) under Grant EP/J006572/1. Financial support from Marie Curie Career Integration Grant PCIG09-GA-2011-293953 and from the Centre of Process Systems Engineering (CPSE) of Imperial College is gratefully acknowledged. MEV thanks CONACYT for doctoral scholarship. The authors are grateful to the anonymous reviewers and the associate editor for their thoughtful comments that led to substantial improvement of the article.

References

1. Adjiman CS, Dallwig S, Floudas CA, Neumaier A (1998) A global optimization method, α BB, for general twice-differentiable constrained NLPs - I. Theoretical advances. *Computers & Chemical Engineering* 22(9):1137–1158
2. Alamo T, Bravo JM, Camacho EF (2005) Guaranteed state estimation by zonotopes. *Automatica* 41(6):1035–1043
3. Althoff M, Stursberg O, Buss M (2008) Reachability analysis of nonlinear systems with uncertain parameters using conservative linearization. In: 47th IEEE Conference on Decision and Control, 2008. CDC 2008, pp 4042–4048
4. Aubin JP (1991) *Viability Theory*. Birkhauser
5. Aubin JP, Cellina A (1984) *Differential Inclusions: Set-Valued Maps and Viability Theory*. Grundlehren der mathematischen Wissenschaften, #264, Springer-Verlag, Berlin
6. Azagra D, Ferrera J (2002) Every closed convex set is the set of minimizers of some C^∞ -smooth function. *Proceedings of the American Mathematical Society* 130(12):3687–3692
7. Bernard O, Hadj-Sadok Z, Dochain D, Genovesi A, Steyer JP (2001) Dynamical model development and parameter identification for an anaerobic wastewater treatment process. *Biotechnology & Bioengineering* 75(4):424–438
8. Berz M, Makino K (1998) Verified integration of ODEs and flows using differential algebraic methods on high-order Taylor series. *Reliable Computing* 4:361–369
9. Berz M, Makino K (2006) Performance of Taylor model methods for validated integration of ODEs. *Lecture Notes in Computer Science* 3732:65–74
10. Blanchini F, Miani S (2008) *Set-Theoretic Methods in Control*. Birkhäuser, Basel
11. Bompadre A, Mitsos A (2012) Convergence rate of McCormick relaxations. *Journal of Global Optimization* 52(1):1–28
12. Bompadre A, Mitsos A, Chachuat B (2013) Convergence analysis of Taylor and McCormick-Taylor models. *Journal of Global Optimization* 57(1):75–114
13. Chachuat B, Latifi MA (2003) A new approach in deterministic global optimization of problems with ordinary differential equations. In: Floudas CA, Pardalos PM (eds) *Frontiers in Global Optimization*, Kluwer Academic Publishers, *Nonconvex Optimization and Its Applications*, vol 74, pp 83–108
14. Chachuat B, Villanueva ME (2012) Bounding the solutions of parametric ODEs: When Taylor models meet differential inequalities. In: Bogle IDL, Fairweather M (eds) *22nd European Symposium on Computer Aided Process Engineering*, vol 30, Elsevier, pp 1307–1311
15. Chachuat B, Singer AB, Barton PI (2006) Global methods for dynamic optimization and mixed-integer dynamic optimization. *Industrial & Engineering Chemistry Research* 45(25):8373–8392
16. Chutinan A, Krogh BH (1999) Verification of polyhedral-invariant hybrid automata using polygonal flow pipe approximations. In: Vaandrager FW, van Schuppen JH (eds) *Hybrid Systems: Computation and Control*, no. 1569 in *Lecture Notes in Computer Science*, Springer, Berlin, pp 76–90
17. Clarke FH (1975) Generalized gradients and applications. *Transactions of the American Mathematical Society* 205:247–262
18. Coddington EA, Levinson N (1955) *Theory of Ordinary Differential Equations*. McGraw-Hill, New York
19. Corliss GF, Rihm R (1996) Validating an a priori enclosure using high-order Taylor series. In: Alefeld G, Frommer A, Lang B (eds) *Proceedings of the International Symposium on Scientific Computing, Computer Arithmetic and Validated Numerics (SCAN'95)*, Akademie Verlag, Berlin, Germany, pp 228–238
20. Du K, Kearfott RB (1994) The cluster problem in multivariate global optimization. *Journal of Global Optimization* 5(3):253–265

21. Eijgenraam P (1981) The solution of initial value problems using interval arithmetic: formulation and analysis of an algorithm, vol Volume 144 of *Mathematical Centre tracts*. Mathematisch Centrum
22. Fiacco AV (1983) *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming*, Mathematics in Science and Engineering, vol 165. Academic Press, New York
23. Friedrichs K (1944) The identity of weak and strong extensions of differential operators. *Transactions of the American Mathematical Society* 55(1):132–151
24. Gronwall TH (1919) Note on the derivatives with respect to a parameter of the solutions of a system of differential equations. *Annals of Mathematics* 20(2):292–296
25. Hairer E, Nørsett SP, Wanner G (1993) *Solving Ordinary Differential Equations I: Nonstiff Problems*, 2nd edn. Springer
26. Houska B, Logist F, Van Impe J, Diehl M (2012) Robust optimization of nonlinear dynamic systems with application to a jacketed tubular reactor. *Journal of Process Control* 22:1152–1160
27. Jaulin L (2002) Nonlinear bounded-error state estimation of continuous-time systems. *Automatica* 38(6):1079 – 1082
28. Kieffer M, Walter E, Simeonov I (2006) Guaranteed nonlinear parameter estimation for continuous-time dynamical models. In: Ninness B, Hjalmarsson H (eds) *Proceeding of the 14th IFAC Symposium on System Identification (SYSID)*, pp 843–848
29. Kurzhanski AB (2006) Comparison principle for equations of the Hamilton-Jacobi type in control theory. *Proceedings of the Steklov Institute of Mathematics* 253(1):S185–S195
30. Kurzhanski AB, Varaiya P (2002) Reachability analysis for uncertain systems - The ellipsoidal technique. *Dynamics of Continuous, Discrete and Impulsive System, Series B* 9(3):347,368
31. Lakshmikantham V, Leela S (1969) *Differential and Integral Inequalities, Theory and Applications: Volume I, Ordinary Differential Equations*. Academic Press
32. Limon D, Bravo JM, Alamo T, Camacho EF (2005) Robust MPC of constrained nonlinear systems based on interval arithmetic. *IEE Proceedings—Control Theory and Applications* 152(3):325–332
33. Lin Q, Rokne JG (1995) Methods for bounding the range of a polynomial. *Journal of Computational & Applied Mathematics* 58:193–199
34. Lin Y, Stadtherr MA (2007) Deterministic global optimization of nonlinear dynamic systems. *AIChE Journal* 53(4):866–875
35. Lin Y, Stadtherr MA (2007) Validated solutions of initial value problems for parametric ODEs. *Applied Numerical Mathematics* 57(10):1145–1162
36. Lin Y, Stadtherr MA (2009) Rigorous model-based safety analysis for nonlinear continuous-time systems. *Computers & Chemical Engineering* 33(2):493–502
37. Lohner RJ (1992) Computation of guaranteed enclosures for the solutions of ordinary initial and boundary value problems. In: Cash JR, Gladwell I (eds) *Computational Ordinary Differential Equations*, Clarendon Press, vol 1, pp 425–436
38. Lygeros J (2004) On reachability and minimum cost optimal control. *Automatica* 40(6):917–927
39. Makino K, Berz M (1999) Efficient control of the dependency problem based on Taylor model methods. *Reliable Computing* 5(1):3–12
40. Maranas CD, Floudas CA (1994) Global minimum potential energy conformations of small molecules. *Journal of Global Optimization* 4:135–170
41. McCormick GP (1976) Computability of global solutions to factorable nonconvex programs: Part I – Convex underestimating problems. *Mathematical Programming* 10:147–175
42. Mitchell IM, Bayen AM, Tomlin CJ (2005) A time-dependent Hamilton-Jacobi formulation of reachable sets for continuous dynamic games. *IEEE Transactions on Automatic Control* 50(7):947–957
43. Mitsos A, Chachuat B, Barton PL (2009) McCormick-based relaxations of algorithms. *SIAM Journal on Optimization* 20(2):573–601
44. Nedialkov NS, Jackson KR (1999) An interval hermite-obreschkoff method for computing rigorous bounds on the solution of an initial value problem for an ordinary differential equation. *Reliable Computing* 5(3):289–310
45. Nedialkov NS, Jackson KR, Corliss G (1998) Validated solutions of initial value problems for ordinary differential equations. *Applied Mathematics and Computation* 105:21–68
46. Nedialkov NS, Jackson KR, Pryce JD (2001) An effective high-order interval method for validating existence and uniqueness of the solution of an IVP for an ODE. *Reliable Computing* 7:449–465
47. Neher M, Jackson KR, Nedialkov NS (2007) On Taylor model based integration of ODEs. *SIAM Journal on Numerical Analysis* 45:236–262
48. Neumaier A (2002) Taylor forms – Use and limits. *Reliable Computing* 9(1):43–79
49. Papamichail I, Adjiman CS (2002) A rigorous global optimization algorithm for problems with ordinary differential equations. *Journal of Global Optimization* 24(1):1–33

50. Papamichail I, Adjiman CS (2004) Global optimization of dynamic systems. *Computers & Chemical Engineering* 28(3):403–415
51. Ramdani N, Meslem N, Candau Y (2009) A hybrid bounding method for computing an over-approximation for the reachable set of uncertain nonlinear systems. *IEEE Transactions on Automatic Control* 54(10):2352–2364
52. Rauh A, Hofer EP, Auer E (2006) VALENCIA-IVP: A comparison with other initial value problem solvers. In: *Proceedings of the 12th GAMM-IMACS International Symposium on Scientific Computing, Computer Arithmetic and Validated Numerics (SCAN'2006)*, Duisburg, Germany
53. Rauh A, Westphal R, Aschemann H (2013) Verified simulation of control systems with interval parameters using an exponential state enclosure technique. In: *2013 18th International Conference on Methods and Models in Automation and Robotics (MMAR)*, pp 241–246
54. Sahlodin AM (2012) Global optimization of dynamic process systems using complete search methods. PhD thesis, McMaster University, Ontario, Canada
55. Sahlodin AM, Chachuat B (2011) Convex/concave relaxations of parametric ODEs using Taylor models. *Computers & Chemical Engineering* 35(5):844–857
56. Sahlodin AM, Chachuat B (2011) Discretize-then-relax approach for convex/concave relaxations of the solutions of parametric ODEs. *Applied Numerical Mathematics* 61:803–820
57. Scott JK, Barton PI (2013) Bounds on the reachable sets of nonlinear control systems. *Automatica* 49(1):93–100
58. Scott JK, Barton PI (2013) Improved relaxations for the parametric solutions of ODEs using differential inequalities. *Journal of Global Optimization* 57(1):143–176
59. Scott JK, Stuber M, Barton PI (2011) Generalized McCormick relaxations. *Journal of Global Optimization* 51(4):569–606
60. Scott JK, Chachuat B, Barton PI (2013) Nonlinear convex and concave relaxations for the solutions of parametric ODEs. *Optimal Control Applications & Methods* 34(2):145–163
61. Singer AB, Barton PI (2004) Global solution of optimization problems with parameter-embedded linear dynamic systems. *Journal of Optimization Theory and Applications* 121(3):613–646
62. Singer AB, Barton PI (2006) Bounding the solutions of parameter dependent nonlinear ordinary differential equations. *SIAM Journal on Scientific Computing* 27(6):2167–2182
63. Tomlin CJ (2011) Verification and control of hybrid systems using reachability analysis. In: *19th Mediterranean Conference on Control and Automation (MED)*, Corfu, Greece
64. Tomlin CJ, Mitchell I, Bayen AM, Oishi M (2003) Computational techniques for the verification of hybrid systems. *Proceedings of the IEEE* 91(7):986–1001
65. Varaiya P, Kurzhanski AB (2006) Ellipsoidal methods for dynamics and control. Part I. *Journal of Mathematical Sciences* 139(5):6863–6901
66. Villanueva ME, Paulen R, Houska B, Chachuat B (2013) Enclosing the reachable set of parametric ODEs using Taylor models and ellipsoidal calculus. In: *Kraslawski A, Turunen I (eds) 23rd European Symposium on Computer Aided Process Engineering (ESCAPE)*, Elsevier, vol Volume 31
67. Walter W (1970) *Differential and Integral Inequalities*. Springer-Verlag, Berlin, Germany
68. Wechsung A, Schaber SD, Barton PI (in press) The cluster problem revisited. *Journal of Global Optimization* pp doi: 10.1007/s10,898–013–0059–9
69. Zhou TS, Zhang JJ, Yuan ZJ, Chen LN (2008) Synchronization of genetic oscillators. *Chaos* 18(3):037,126

A Technical Lemmata

The following two lemmata are used in the proof of Theorem 3. Although variants of these results can be found in the literature [23], we provide short proofs for the sake of completeness.

Lemma 1 *Let $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a continuous function. For any compact set $D \subset \mathbb{R}^n$ and any finite tolerance $\varepsilon > 0$, there exists a smooth function $\varphi_\varepsilon : \mathbb{R}^n \rightarrow \mathbb{R}^m$ such that*

$$\forall x \in D, \quad \|\varphi_\varepsilon(x) - \varphi(x)\| \leq \alpha(\varepsilon), \quad (74)$$

for some continuous function $\alpha : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ with $\alpha(0) = 0$.

Proof The proof follows by applying well-known standard analysis techniques [23], and we only summarize the main idea here. Let $\sigma_\varepsilon : \mathbb{R}^n \rightarrow \mathbb{R}$, $\varepsilon > 0$ be a family of smooth functions parameterized in $\varepsilon > 0$, such that

$\sigma_\varepsilon(x) = 0$ for all x with $\|x\| \geq \varepsilon$ and $\int_{\mathbb{R}^n} \sigma_\varepsilon(x) dx = 1$. Of the alternatives for constructing such a family of ‘mollifier’ functions, we consider the function

$$\sigma_\varepsilon(x) := \begin{cases} C(\varepsilon) \exp\left(\frac{1}{\|x\|^2 - \varepsilon^2}\right) & \text{if } \|x\| < \varepsilon \\ 0 & \text{otherwise} \end{cases} \quad \text{with } C(\varepsilon) := \int_{\|x\| \leq \varepsilon} \exp\left(\frac{1}{\|x\|^2 - \varepsilon^2}\right) dx.$$

In turn, the function φ_ε can be defined as the convolution

$$\forall x \in \mathbb{R}^n, \quad \varphi_\varepsilon(x) := \int_{\mathbb{R}^n} \sigma_\varepsilon(x-y) \varphi(y) dy,$$

which is smooth by construction for any $\varepsilon > 0$. Observe that the function $\alpha : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ defined by

$$\forall \varepsilon \geq 0, \quad \alpha(\varepsilon) := \max_{x \in D} \max_y \{ \|\varphi(x) - \varphi(y)\| \mid \|y - x\| \leq \varepsilon \},$$

is continuous since φ is itself continuous and D is compact, and such that $\alpha(0) = 0$. In particular, this choice of α satisfies the condition (74). \square

Lemma 2 *Let $Y : [0, T] \rightarrow \mathbb{K}_C^{n \times 1}$ be a set-valued function such that $V[Y(\cdot)](c)$ is differentiable and $\dot{V}[Y(\cdot)](c)$ is bounded for all $c \in \mathbb{R}^n$ with $c^T c = 1$. Then, there exists a family of functions $g_\varepsilon : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$ parameterized by $\varepsilon \geq 0$, such that $g_\varepsilon(t, \cdot)$ is strictly convex and smooth for all $\varepsilon > 0$ and all $t \in [0, T]$, and the associated sets $Y_\varepsilon(t) := \{x \in \mathbb{R}^n \mid g_\varepsilon(t, x) \leq 0\}$ satisfy*

$$\begin{aligned} \forall \varepsilon \geq 0, \forall t \in [0, T], \quad Y(t) \subseteq Y_\varepsilon(t) \quad \text{and} \quad d_H(Y(t), Y_\varepsilon(t)) \leq \alpha(\varepsilon), \\ \forall \varepsilon \geq 0, \forall t \in [0, T], \forall c \in \mathbb{R}^n \text{ with } c^T c = 1, \quad \dot{V}[Y_\varepsilon(t)](c) \geq \dot{V}[Y(t)](c) + L\alpha(\varepsilon) \end{aligned} \quad (75)$$

for some continuous function $\alpha : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ with $\alpha(0) = 0$, and any constant $0 \leq L < \frac{1}{T}$.

Proof A proof can be obtained by passing through two steps.

S1 We start with any smooth function $v_\varepsilon(t, \cdot) : \mathbb{R}^n \rightarrow \mathbb{R}$ such that

$$\forall \varepsilon > 0, \forall c \in \mathbb{R}^n \text{ with } c^T c = 1, \forall t \in [0, T], \quad v_\varepsilon(t, c) \geq \dot{V}[Y(t)](c) \quad \text{and} \quad \|v_\varepsilon(t, c) - \dot{V}(t, c)\| \leq \alpha_1(\varepsilon),$$

for some continuous function $\alpha_1 : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ with $\alpha_1(0) = 0$. Such a function is guaranteed to exist by Lemma 1. Then, we define the set-valued function $Z_\varepsilon : [0, T] \rightarrow \mathbb{K}_C^{n \times 1}$ such that

$$\forall c \in \mathbb{R}^n \text{ with } c^T c = 1, \quad V[Z_\varepsilon(t)](c) := V[Y(0)](c) + \int_0^t v_\varepsilon(\tau, c) d\tau.$$

The following properties hold by construction of Z_ε , for every $\varepsilon > 0$:

a) For all $c \in \mathbb{R}^n$ with $c^T c = 1$, the function $V[Z_\varepsilon(\cdot)](c)$ is differentiable on $[0, T]$, and we have

$$\forall c \in \mathbb{R}^n \text{ with } c^T c = 1, \forall t \in [0, T], \quad \dot{V}[Z_\varepsilon(t)](c) \geq \dot{V}[Y(t)](c).$$

b) $d_H(Z_\varepsilon(t), Y(t)) \leq T\alpha_1(\varepsilon)$, by property (6).

S2 We construct the set-valued function

$$Y_\varepsilon(t) := Z_\varepsilon(t) \oplus [T\alpha_1(\varepsilon) + tL\alpha(\varepsilon)] \mathcal{B}^n \quad \text{with} \quad \alpha(\varepsilon) := \frac{2T}{1-TL} \alpha_1(\varepsilon),$$

with $0 \leq L < \frac{1}{T}$. Note that the function α is continuous and non-negative and it satisfies $\alpha(0) = 0$ by definition. Therefore, we have $Y_\varepsilon(t) \supseteq Y(t)$ since $Y_\varepsilon(t) \supseteq Z_\varepsilon(t)$ and $d_H(Y_\varepsilon(t), Z_\varepsilon(t)) \geq d_H(Z_\varepsilon(t), Y(t))$. It follows from Property a) that

$$\begin{aligned} \dot{V}[Y_\varepsilon(t)](c) &= \dot{V}[Z_\varepsilon(t)](c) + L\alpha(\varepsilon) \\ &\geq \dot{V}[Y(t)](c) + L\alpha(\varepsilon), \end{aligned}$$

for all $t \in [0, T]$, all $c \in \mathbb{R}^{n_x}$ with $c^T c = 1$, and all $\varepsilon \geq 0$. Moreover, by Property b), we have

$$\begin{aligned} d_H(Y_\varepsilon(t), Y(t)) &\leq d_H(Z_\varepsilon(t), Y(t)) + \alpha_0(\varepsilon) + T\alpha_1(\varepsilon) + TL\alpha(\varepsilon) \\ &\leq 2\alpha_0(\varepsilon) + 2T\alpha_1(\varepsilon) + TL\alpha(\varepsilon) \\ &= \alpha(\varepsilon), \end{aligned}$$

for all $t \in [0, T]$, and all $\varepsilon \geq 0$. Finally, by Theorem 1 in [6], there exists a functions $g_\varepsilon : [0, T] \times \mathbb{R}^{n_x} \rightarrow \mathbb{R}$ such that $Y_\varepsilon(t) =: \{x \in \mathbb{R}^{n_x} \mid g_\varepsilon(t, x) \leq 0\}$ and $g_\varepsilon(t, \cdot)$ is convex and smooth for all $\varepsilon \geq 0$ and all $t \in [0, T]$. In order for $g_\varepsilon(t, \cdot)$ to be strictly convex for all $\varepsilon > 0$, one can always add a strictly convex and smooth term of order $O(\varepsilon)$ that is negative on the compact sets $\bigcup_{t \in [0, T]} Y_\varepsilon(t)$. \square

The following lemma is used in the proof of Theorem 4. The result allows one to bound the solution of a particular parametric differential inequality and can be regarded as a generalization of Gronwall's lemma [24].

Lemma 3 *Let $v \in \mathbb{R}_+$ and let $u : [0, T] \rightarrow \mathbb{R}$ be a Lipschitz-continuous function satisfying the parametric differential inequality*

$$\text{a.e. } t \in [0, T], \quad \frac{d}{dt}[u(t)]^n \leq n \sum_{i=0}^n L_i u(t)^{n-i} v^{im} \quad \text{with } u(0) \leq C_0 v^m, \quad (76)$$

for some integers $m, n \geq 1$ and a set of constants $0 \leq L_0, \dots, L_n < \infty$ and $C_0 \geq 1$. Then, $u(t) \leq C_0 \exp(\sum_{i=0}^n L_i t) v^m$, for all $t \in [0, T]$.

Proof The proof proceeds in two steps. It is assumed first that the function u is differentiable on $[0, T]$. Then, it is argued that the result still holds in extending this class of functions to Lipschitz-continuous.

Assuming that u is differentiable on $[0, T]$ and discretizing the differential inequality (76) with a step-size $h := \frac{T}{N}$ for a large enough $N \in \mathbb{N}$ gives

$$\forall k \in \{0, N-1\}, \quad [u((k+1)h)]^n \leq [u(kh)]^n + hn \sum_{i=0}^n L_i (u(kh))^{n-i} v^{im} + h\alpha(h),$$

for some continuous function $\alpha : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ with $\alpha(0) = 0$. Now, supposing that $u(kh) \leq C_k v^m$ with $C_k \geq 1$, we have

$$[u((k+1)h)]^n \leq (C_{k+1})^n v^{nm} + h\alpha(h) \quad \text{with } (C_{k+1})^n := \left(1 + hn \sum_{i=0}^n L_i\right) (C_k)^n \geq 1.$$

In particular, the definition of $(C_{k+1})^n$ uses the result that

$$(C_k)^n + hn \sum_{i=0}^n L_i (C_k)^{n-i} \leq \left(1 + hn \sum_{i=0}^n L_i\right) (C_k)^n,$$

for all $C_k \geq 1$. It follows by induction that $u(kh) \leq C_k v^m + h\alpha(h)$ for each $k = 0, \dots, N$, with

$$C_k = \left(1 + hn \sum_{i=0}^n L_i\right)^{k/n} C_0.$$

Let $\bar{t} \in [0, T]$ be such that $\bar{t} := \frac{k_0}{N_0} T$ for given $0 \leq k_0 \leq N_0$, and consider the sequence $\{\bar{C}_j\}$ given by

$$\bar{C}_j := \left(1 + \frac{nT}{jN_0} \sum_{i=0}^n L_i\right)^{j k_0 / n} C_0,$$

so that $u(\bar{t}) \leq \bar{C}_j v^m + \frac{T}{jN_0} \alpha(\frac{T}{jN_0})$ for all $j \geq 1$. It follows from the definition of the exponential function as $\exp(x) := \lim_{j \rightarrow \infty} (1 + \frac{x}{j})^j$ that this sequence is convergent, and we have

$$\lim_{j \rightarrow \infty} \bar{C}_j = \exp\left(\sum_{i=0}^n L_i \bar{t}\right) C_0.$$

As u is continuous on $[0, T]$, and since the rationals are a dense subset of the real numbers, it follows that

$$\forall t \in [0, T], \quad u(t) \leq C(t)v^m \quad \text{with} \quad C(t) := \exp\left(\sum_{i=0}^n L_i t\right) C_0.$$

In a second step, the assumption of differentiability for u can be relaxed to Lipschitz-continuity, by a similar argument as in part S3 of the proof of Theorem 3, namely that any (locally) Lipschitz-continuous function is differentiable almost everywhere. \square