Trajectory-based reachability analysis of switched nonlinear systems using matrix measures

John Maidens and Murat Arcak

***Abstract*— Matrix measures, or logarithmic norms, have historically been used to provide bounds on the divergence of trajectories of a system of ordinary differential equations (ODEs). In this paper we use them to compute guaranteed overapproximations of reachable sets for switched nonlinear dynamical systems using numerically simulated trajectories, and to bound the accumulation of numerical errors along simulation traces. To improve the tightness of the computed approximations, we connect these classical tools for ODE analysis with modern techniques for optimization and demonstrate that minimizing the volume of the computed reachable set enclosure can be formulated as a convex problem. Using a benchmark problem for the verification of hybrid systems, we show that this technique enables the efficient computation of reachable sets for systems with over 100 continuous state variables.**

I. Introduction

This paper introduces the classical notion of matrix measures to the reachability problem which is critical for proving safety, liveness or fairness properties of continuous and hybrid systems. Reachability for nonlinear systems is a computationally expensive operation that typically scales poorly with the number of continuous state variables. Existing approaches include level set methods [1], generating linear or piecewise linear models approximating the nonlinear dynamics for which linear reachability techniques can be applied [2], [3], methods based on abstractions [4], [5], interval Taylor series methods [6], [7] and differential inequality methods [8], [9].

Trajectory-based approaches [10], [11], [12], [13] have the advantage that numerical simulation is a relatively inexpensive operation, even for systems with a large number of states. Thus, unlike the more computationally expensive approaches mentioned above, they can scale well with state dimension. In addition, simulation-based approaches are naturally parallelizable because the reachability subproblems corresponding to each trajectory can be solved independently.

We present a new trajectory-based approach where we first sample a number of trajectories of the system. Next, we use matrix measures to establish a bound on the divergence between the samples and neighbouring trajectories. Each sample trajectory provides information about the behaviour of all trajectories initialized from a norm ball centred at the sample's origin. Thus, by sampling enough trajectories such that the norm balls cover the initial set, and using the matrix measure of the system's Jacobian at each point to propagate the norm balls forward along the system's flow, we are able to compute a bound on the set of reachable states (Figure 1).

The papers [11], [12] also take a trajectory-based approach to the reachability problem, but use Lipschitz constants in place of matrix measures to bound the divergence between trajectories. Since Lipschitz constants are always positive, this method leads to reachable set enclosures that expand with time even for systems with negative matrix measures. Unlike [10] which is not able to guarantee that the computed approximation contains the true reachable set, here we provide a guaranteed overapproximation of the set of reachable states. We further account for numerical error, thus our technique can be used to provide formal guarantees of safety. Another related trajectory-based method is [13] which uses finite-time invariant sets computed using sum-of-squares methods to bound the divergence of trajectories. Our approach does not require the symbolic manipulation of polynomials, and thus it scales better as the dimension of the state space increases.

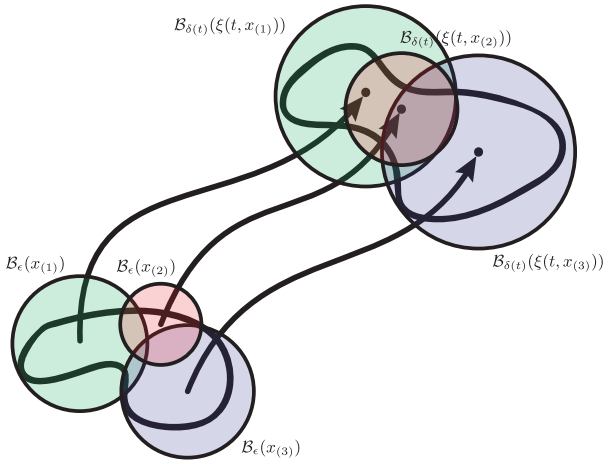

Fig. 1: Trajectory-based overapproximation of the reachable set. The initial set is covered by a set of norm balls and a trajectory from the centre of each ball is computed. The reachable set is then enclosed by a union of balls whose radii are computed using the matrix measure.

Using our method each simulation and the corresponding bound on nearby trajectories can be computed very quickly. Although the bound we provide for each simulation may be conservative, we guarantee that the computed overapproximation approaches the true reachable set asymptotically as the mesh radius of chosen initial states converges to zero. To avoid conservativeness in the approximation, we pursue and optimization approach. We demonstrate that if we cover

J. Maidens and M. Arcak are with the Department of Electrical Engineering & Computer Sciences, University of California, Berkeley, CA, 94720 USA e-mail: {maidens, arcak}@eecs.berkeley.edu. Research supported in part by grants NSF ECCS-1101876, NSF ECCS-1405413, AFOSR FA9550-11-1-0244 and NSERC postgraduate fellowship PGFD3-427610-2012.

the initial set by weighted 1-, 2-, or ∞-norm balls, the problem of minimizing the volume of the resulting reachable set enclosure can be formulated as a convex problem.

We begin with a brief survey of matrix measures in Section II. In Section III we demonstrate that our matrix measure approach to nonlinear reachability analysis provides a guaranteed overapproximation of the set of reachable states of the system. We then provide a method for improving the tightness of the approximation by weighting matrix measures in an optimal manner. In Section IV we use the matrix measure for analysing the accumulation of numerical errors along simulated trajectories of the system. Finally, in Section V we demonstrate our method on the leak test benchmark [14]. This benchmark problem was formulated to compare the scalability of hybrid system verification algorithms as the number of continuous state variables increases. Our algorithm performs well on this benchmark, allowing us to verify a model with over a hundred state variables.

II. Overview of matrix measures and contraction

Let $|\cdot|$ be a norm on $\mathbb{R}^n$ and $\|\cdot\|$ be its induced norm on the set of real matrices of dimension $n \times n$. The measure $\mu(A)$ of a matrix $A \in \mathbb{R}^{n\times n}$ is the one-sided derivative of $\|\cdot\|$ at $I \in \mathbb{R}^{n\times n}$ in the direction A:

$$\mu(A) = \lim_{t\to 0^+} \frac{\|I + tA\| - \|I\|}{t}. \tag{1}$$

This limit is guaranteed to exist for any norm $|\cdot|$ and $A \in \mathbb{R}^{n\times n}$ (see [15]). The following properties of μ [15], [16] are critical:

1) For all eigenvalues $\lambda_i(A)$ of A we have $-\|A\| \leq -\mu(-A) \leq \Re(\lambda_i(A)) \leq \mu(A) \leq \|A\|$.
2) $\mu(cA) = c\mu(A)$ for all $c \geq 0$.
3) $\mu(A + B) \leq \mu(A) + \mu(B)$.
4) If $P \in \mathbb{R}^{n\times n}$ is nonsingular then the measure μ_P of the norm $|x|_P = |Px|$ is given in terms of μ by $\mu_P(A) = \mu(PAP^{-1})$.

Some familiar vector norms as well as their corresponding induced matrix norms and measures are given in Table I. From these expressions, it is clear that unlike the Lipschitz constant, the matrix measure can take negative values. Further, it follows from property 1 that the matrix measure is bounded above by the Lipschitz constant.

Vector norm	Induced matrix norm	Induced matrix measure
$\|x\|_1 = \sum_j \|x_j\|$	$\|A\|_1 = \max_j \sum_i \|a_{ij}\|$	$\mu_1(A) = \max_j \left(a_{jj} + \sum_{i\neq j} \|a_{ij}\|\right)$
$\|x\|_2 = \sqrt{\sum_j x_j^2}$	$\|A\|_2 = \sqrt{\max_j \lambda_j(A^T A)}$	$\mu_2(A) = \max_j \frac{1}{2}\left(\lambda_j(A + A^T)\right)$
$\|x\|_\infty = \max_j \|x_j\|$	$\|A\|_\infty = \max_i \sum_j \|a_{ij}\|$	$\mu_\infty(A) = \max_i \left(a_{ii} + \sum_{j\neq i} \|a_{ij}\|\right)$

TABLE I: Commonly used vector norms and their corresponding matrix norms and measures

The matrix measure has long been used to provide estimates on solutions of systems of ordinary differential equations [15], [16], [17], [18], [19]. The following proposition, adapted from [19], allows us to bound the distance between trajectories of a system

$$\dot{x}(t) = f(t, x(t)) \tag{2}$$

in terms of their initial distance and the rate of expansion of the system given by the measure of the Jacobian matrix $J(t, x)$ with respect to x.

Proposition 1: Let $\mathcal{D} \subseteq \mathbb{R}^n$ and let the Jacobian $J(t, x) = \frac{\partial f}{\partial x}(t, x)$ satisfy $\mu(J(t, x)) \leq c$ for all $(t, x) \in [0, T] \times \mathcal{D}$. If every trajectory of (4) with initial conditions in the line segment $\{hx_0 + (1 - h)z_0 : h \in [0, 1]\}$ remains in $\mathcal{D}$ until time T then the solutions $\xi(t)$ and $\zeta(t)$ with $\xi(0) = x_0$ and $\zeta(0) = z_0$ satisfy

$$|\xi(t) - \zeta(t)| \leq |\xi(0) - \zeta(0)|e^{ct}. \tag{3}$$

for all $t \in [0, T]$.

This proposition provides global results about the divergence between trajectories of (2) using only information about the system's Jacobian at each point. If there exists $c < 0$ such that for all $(t, x) \in [0, \infty) \times \mathcal{D}$

$$\mu(J(t, x)) \leq c$$

then the system (2) or the vector field $f(t, x)$ is said to be *contracting* with respect to $|\cdot|$ [19], [20], [21]. From (3) it follows that for such systems any two trajectories converge asymptotically. Unlike the literature on contractive or incrementally stable systems which deals primarily with the case where $c < 0$, our results allow the expansion rate c to be positive.

III. Overapproximation of reachable sets

We study a class of switched nonlinear systems

$$\dot{x}(t) = f_{\sigma(t)}(t, x(t)) \tag{4}$$

where $\sigma : \mathbb{R} \to \mathcal{S}$ is a known function describing the switching between modes parametrized by the set $\mathcal{S}$ and $\{f_\sigma : [0, \infty) \times \mathbb{R}^n \to \mathbb{R}^n\}_{\sigma\in\mathcal{S}}$ is a family of functions continuous in t and C^1 in x that describe the continuous-time dynamics within each mode. We assume that on any bounded time interval, only a finite number of switches occur.

Given a compact initial set K and final time T we wish to find a set R such that all trajectories $\xi(\cdot, x_0)$ of (4) with initial condition $\xi(0, x_0) = x_0 \in K$ satisfy $\xi(T, x_0) \in R$. To simplify the presentation, we assume throughout Section III that trajectories of (2) can be computed exactly. In Section IV we will extend the results presented here to ensure that numerical inaccuracies due to floating point arithmetic and discretization of the continuous dynamics are accounted for.

We first introduce the required notation. The ball with radius ϵ and centre x_0 is given by $\{x : |x - x_0| \leq \epsilon\}$ and is denoted $\mathcal{B}_\epsilon(x_0)$. Throughout, we denote the solution to the differential equation (4) with initial condition x_0 as $\xi(t, x_0)$, or when it is clear that we have fixed a particular initial condition x_0, simply as $\xi(t)$. The set reachable at time t from initial set S is denoted $Reach_t(S)$ and is defined as $\{\xi(t, x) : x \in S\}$. The tube reachable from the initial set S over an interval $[0, t]$ is denoted $Reach_{[0,t]}(S) = \{\xi(s, x) : s \in [0, t], x \in S\}$. For a set $X \subseteq Y$, the error with which Y overapproximates X can be quantified via the Hausdorff distance $d(X, Y) = \sup_{y\in Y} \inf_{x\in X} |x - y|$.

A. Basic Algorithm

We begin by covering the initial set K by a finite number of norm balls $\mathcal{B}_{\epsilon_k}(x_k)$. The set reachable from K is contained in the union of the sets reachable from these norm balls. The number of balls can be chosen so as to achieve the required tightness in the approximation of the reachable set; a larger number of balls provides a more accurate approximation, while covering K by a single ball reduces computation time at the cost of reduced tightness. The computation of the set reachable from each norm ball can be performed in parallel, thus we assume without loss of generality that the initial set is given by a single norm ball $K = \mathcal{B}_\epsilon(x_0)$.

In light of Proposition 1, given a global bound c on $\mu(J_\sigma(t,x))$, we know that all trajectories of (4) with initial conditions in $\mathcal{B}_\epsilon(x_0)$ lie in $\mathcal{B}_{\epsilon e^{cT}}(\xi(T,x_0))$. Since a global bound c on the expansion rate is far too conservative, we provide an iterative method for computing a more accurate approximation based on a local bound on the expansion rate. We begin with the following corollary of Proposition 1, which ensures that our algorithm yields an overapproximation of the set of reachable states.

Corollary 1: Let the system (4) be in mode σ_i for $t \in [t_i, t_{i+1}]$ and let the Jacobian $J_{\sigma_i}(t,x)$ of $f_{\sigma_i}(t,x)$ with respect to x satisfy $\mu(J_{\sigma_i}(t,x)) \le c_i$ for all $(t,x) \in [t_i,t_{i+1}] \times \mathrm{Conv}(Reach_{[t_i,t_{i+1}]}(\mathcal{B}_{\delta_i}(\xi(t_i))))$. Then any solution ζ of (4) with $\zeta(t_i) \in \mathcal{B}_{\delta_i}(\xi(t_i))$ satisfies

$$|\xi(t_{i+1}) - \zeta(t_{i+1})| \le |\xi(t_i) - \zeta(t_i)| e^{c_i(t_{i+1}-t_i)}. \quad (5)$$

Thus given a sequence of local bounds c_i we can compute a guaranteed overapproximation of $Reach_T(\mathcal{B}_\epsilon(x_0))$ as $\mathcal{B}_\delta(\xi(T,x_0))$ where

$$\delta = \left(\prod_{i=0}^{N-1} e^{c_i(t_{i+1}-t_i)} \right) \epsilon.$$

Note that the times t_i are not necessarily switching times, and may be arbitrary so long as all switching times appear in the sequence $\{t_i\}_{i\in\mathbb{N}}$. The t_i could be, for example, times at which a numerical trace of (4) is computed. The set $\mathrm{Conv}(Reach_{[t_i,t_{i+1}]}(\mathcal{B}_{\delta_i}(\xi(t_i))))$ is generally not known, but a crude overapproximation will suffice for the purpose of computing the constant c_i. If we can find some crude bound S on $Reach_{[0,T]}(K)$ (for example an invariant set containing K) such that $|f_{\sigma(t)}(t,x)| \le M$ for all $t \in [0,T]$ and all $x \in S$ then we have the containment

$$\mathrm{Conv}(Reach_{[t_i,t_{i+1}]}(\mathcal{B}_{\delta_i}(\xi(t_i)))) \subseteq \mathcal{B}_{\delta_i + M(t_{i+1}-t_i)}(\xi(t_i)).$$

Note that once an overapproximation of the reach set is computed, this bound can then be used to recompute a smaller M and the method reapplied to generate an even tighter approximation. The proposed method is summarized in Algorithm 1.

The following corollaries of Proposition 1 provide information about the tightness of this approximation. Corollary 2 establishes that the approximation can be made arbitrarily accurate by covering the initial set K by a collection of balls of sufficiently small radius:

Algorithm 1 Basic algorithm for bounding $Reach_T(K)$

Require: Initial ball size $\epsilon > 0$, bound M on magnitude of vector field f, sequence of simulation points $x_i := \xi(t_i)$ for $i = 0, \ldots, N$.
1: Set $\delta_0 = \epsilon$
2: **for** i from 0 to $N-1$ **do**
3: Compute upper bound c_i on expansion rate $\mu(J_{\sigma_i}(t,x))$ within the set with
4: $t_i \le t \le t_{i+1}$ and $|x - x_i| \le \delta_i + M(t_{i+1} - t_i)$.
5: Set $\delta_{i+1} = e^{c_i(t_{i+1}-t_i)}\delta_i$
6: **end for**
7: **return** $\mathcal{B}_{\delta_N}(x_N)$

Corollary 2 (Tightness as a function of mesh size): Let $\mathcal{D} \subseteq \mathbb{R}^n$ be convex and let the Jacobian $J_{\sigma(t)}(t,x)$ of f_σ with respect to x satisfy $\mu(J_{\sigma(t)}(t,x)) \le c$ for all $(t,x) \in [0,\infty) \times \mathcal{D}$. Then the approximation error $d(Reach_t(\mathcal{B}_\epsilon(x_0)), \mathcal{B}_{e^{ct}\epsilon}(\xi(t,x_0)) \to 0$ linearly as $\epsilon \to 0$.

Corollary 3 establishes that for contractive systems, a single ball overapproximating the initial set K is sufficient to generate an approximation that becomes arbitrarily tight as $T \to \infty$:

Corollary 3 (Asymptotic tightness in contractive systems): Let $\mathcal{D} \subseteq \mathbb{R}^n$ be convex and let f_σ satisfy $\mu(J_{\sigma(t)}(t,x)) \le c < 0$ for all $(t,x) \in [0,\infty) \times \mathcal{D}$. Then the approximation error $d(Reach_T(\mathcal{B}_\epsilon(x_0)), \mathcal{B}_{e^{cT}\epsilon}(\xi(T,x_0))) \to 0$ exponentially as $T \to \infty$.

Note that covering a set $S \subseteq \mathbb{R}^n$ by a uniform mesh of radius ϵ requires $\Theta(\epsilon^{-n})$ mesh points and hence the practical applicability of these tightness results is limited. However, methods exist in the literature for choosing non-uniform meshes of trajectories to simulate (*e.g.*, Section 3 of [10]) which help alleviate this issue.

B. Algorithm with norm updating

We now provide a modified scheme that allows us to optimize the norm in which the expansion is measured at a given time and state (t,x). We consider a family of weighted norms $\{|\cdot|_\Gamma\}$ parametrized by weights Γ from some set of real $n \times n$ matrices. Given an initial set B_i described as a norm ball of $|\cdot|_{\Gamma_i}$ we

- overapproximate the initial ball B_i by a ball $\bar{B}_i$ in some new norm $|\cdot|_{\Gamma_{i+1}}$
- compute an expansion rate c_{i+1} at the point (t_i, x_i) satisfying $c_{i+1} \ge \mu_{\Gamma_{i+1}}(J_{\sigma_i}(t_i,x_i))$ where $\mu_{\Gamma_{i+1}}$ is the matrix measure induced by $|\cdot|_{\Gamma_{i+1}}$ (Recall that μ_Γ can be computed in terms of μ via Property 4 from Section II.)
- compute an overapproximation of the set reachable form $\bar{B}_i$ using the expansion rate c_{i+1}. This gives the new set B_{i+1}.

This procedure is illustrated in Figure 2.

It may appear that the next weight Γ_{i+1} should be selected to minimize c_{i+1}. However there is a tradeoff between how small we can make c_{i+1} and how tightly $\bar{B}_i$ approximates B_i. For example, there may be some weight Γ_{i+1} such that the expansion rate c_{i+1} is very small, but if this weight is far

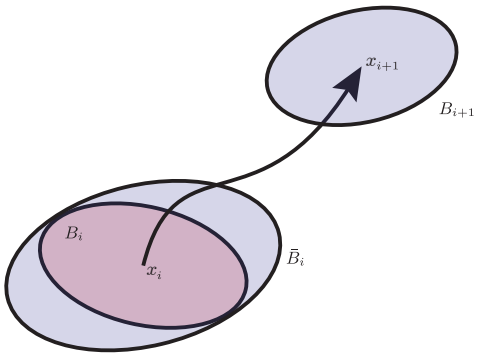

Fig. 2: Given a set B_i which encloses the set reachable at time t_i, we wish to compute a minimum-volume enclosure B_{i+1} of the reachable set at time t_{i+1}. To do this, we overapproximate B_i by a ball $\bar{B}_i$ in a different weighed norm and propagate the resulting ball forward using an expansion rate computed in the new norm to obtain B_{i+1}. By formulating this as an optimization problem, we are able to manage the trade-off between the desire to minimize the expansion rate and the desire to minimize the conservatism induced by overapproximating B_i by $\bar{B}_i$.

from the previous weight Γ_i then B_i and $\bar{B}_i$ will be of very different shapes and thus the overapproximation of B_i by $\bar{B}_i$ will be very conservative. Thus, instead, at each step we choose Γ_{i+1} such that the volume $\mathrm{vol}(B_{i+1})$ is minimized:

$$\begin{array}{ll} \textbf{minimize} & \mathrm{vol}(B_{i+1}) \\ \textbf{subject to} & B_i \subseteq \bar{B}_i \\ & \mu_{\Gamma_{i+1}}(J_{\sigma_i}(t_i, x_i)) \leq c_{i+1}. \end{array} \tag{6}$$

For certain families of norms, this can be formulated as a convex optimization problem in the weighting Γ. We now describe three such cases.

1) Euclidean norms weighted by positive definite matrices: We consider the family of weighted Euclidean norms of the form $x \mapsto |Px|_2$ where P is a positive definite matrix. There is a one-to-one correspondence between such norms and their unit balls, described by the nondegenerate ellipsoid $\{x : x^T \Gamma x \leq 1\}$ and parametrized by $\Gamma = P^2$ from the set of positive definite matrices.

The following proposition relates Γ with the expansion rate of (4) at (t, x).

Proposition 2 (Lemma 2 of [22]): If

$$\Gamma A + A^T \Gamma \leq 2c\Gamma$$

where Γ is a positive definite matrix then $\mu(A) \leq c$ in the norm $x \mapsto |Px|_2$ where $P = \Gamma^{1/2} \succ 0$.

If $B_i = \{x : x^T \Gamma_i x \leq 1\}$ and $\bar{B}_i = \{x : x^T \Gamma_{i+1} x \leq 1\}$ then the constraint $B_i \subseteq \bar{B}_i$ can be expressed in terms of the parameters as $\Gamma_{i+1} \preceq \Gamma_i$.

The set reachable from $\{x_i + x : x^T \Gamma_{i+1} x \leq 1\}$ is approximated by $\{x_{i+1} + x : x^T \Gamma_{i+1} x \leq e^{2c(t_{i+1}-t_i)}\}$ where c satisfies

$$\Gamma_{i+1} J_{\sigma_i}(t_i, x_i) + J_{\sigma_i}(t_i, x_i)^T \Gamma_{i+1} \leq 2c\Gamma_{i+1}.$$

We wish to choose Γ_{i+1} so as to minimize the volume of the computed reach set

$$B_{i+1} = \{x_{i+1} + x : x^T \Gamma_{i+1} x \leq e^{2c(t_{i+1}-t_i)}\}.$$

As the volume of the ellipsoid $\{x : x^T \Gamma x \leq 1\}$ is proportional to $\det(P^{-1})$, the volume $\mathrm{vol}(B_{i+1})$ is proportional to $\frac{1}{\sqrt{\det(e^{-2c(t_{i+1}-t_i)}\Gamma_{i+1})}}$. Problem (6) can now be cast as the following optimization problem in the variables c and Γ_{i+1}:

$$\begin{array}{ll} \textbf{minimize} & -e^{-2c(t_{i+1}-t_i)} \det(\Gamma_{i+1})^{1/n} \\ \textbf{subject to} & \Gamma_{i+1} \preceq \Gamma_i \\ & \Gamma_{i+1} J_{\sigma_i}(t_i, x_i) + J_{\sigma_i}(t_i, x_i)^T \Gamma_{i+1} \preceq 2c\Gamma_{i+1}. \end{array} \tag{7}$$

For fixed $c \in \mathbb{R}$ this is a convex problem in the variable Γ_{i+1}. Thus a solution can be found via a line search over c where each evaluation involves solving this convex problem.

Once we have chosen Γ_{i+1}, we proceed as in Section III-A. This leads to Algorithm 2 below for computing an overapproximation of the reachable set using weighted norms, with the weights adjusted at each step.

Algorithm 2 Bounding of reachable set from norm ball based on weighted Euclidean norms

Require: Initial ball shape matrix Γ_0, sequence of simulation points $x_i := \xi(t_i)$ for $i = 0, \ldots, N$.
for i from 0 to $N-1$ **do**
  Find (c, Γ_{i+1}) to solve optimization problem (7)
  Compute bound M on magnitude of vector field f in norm defined by Γ_{i+1}
  Compute upper bound c_i on expansion rate $\mu(J_{\sigma_i}(t, x))$ within the set with
    $t_i \leq t \leq t_{i+1}$ and $\{x_i + x : x^T \Gamma_{i+1} x \leq 1 + M(t_{i+1} - t_i)\}$.
  Set $\Gamma_{i+1} = e^{-2c_i(t_{i+1}-t_i)}\Gamma_{i+1}$

end for
return $\{x_N + x : x^T \Gamma_N x \leq 1\}$

2) 1-norms weighted by positive diagonal matrices: We now show that using norms $|x|_{1,D} = |Dx|_1$ parametrized by positive diagonal matrices $D \succ 0$ also leads to a convex optimization problem. The corresponding induced matrix measure $\mu_{1,D}$ is given by

$$\mu_{1,D}(A) = \mu_1(DAD^{-1}) = \max_j \left(a_{jj} + \sum_{i \neq j} \frac{d_i}{d_j} |a_{ij}| \right)$$

hence the condition $\mu_{1,D}(A) \leq c$ can be expressed as

$$a_{jj} d_j + \sum_{i \neq j} |a_{ij}| d_i \leq c d_j \qquad j = 1, \ldots, n$$

which is linear in the d_i for fixed c. The condition $\{x : |x|_{1,P} \leq 1\} \subseteq \{x : |x|_{1,D} \leq 1\}$ requires that $d_j \leq p_j$ for all $j = 1, \ldots, n$. Finally, the volume of the set $\{x : |x|_{1,D} \leq e^{ct}\}$ is proportional to $e^{nct} \prod_{i=1}^n \frac{1}{d_i}$ which is convex in d. Thus if D is a solution to the problem

$$\begin{array}{ll} \textbf{minimize} & \left(e^{nct} \displaystyle\prod_{i=1}^n \frac{1}{d_i} \right)^{1/n} \\ \textbf{subject to} & D \preceq P \\ & a_{jj} d_j + \displaystyle\sum_{i \neq j} |a_{ij}| d_i \leq c d_j \qquad j = 1, \ldots, n \end{array} \tag{8}$$

where $A = J_{\sigma_i}(t_i, x_i)$ then $|\cdot|_{1,D}$ is a good norm in which to overapproximate the reachable set in a neighbourhood

of the point (t_i, x_i) in order to minimize the accumulation in volume. As with problem (7) in the Euclidean case, the problem (8) is convex in D for fixed c and can be readily solved via a line search over c.

3) ∞-norms weighted by positive diagonal matrices: For norms of the form $|x|_{\infty,D} = |Dx|_\infty$ where $D \succ 0$ is a positive diagonal matrix, a similar procedure yields the problem

$$\begin{aligned} \textbf{minimize} \quad & e^{nct} \prod_{i=1}^{n} \frac{1}{d_i} \\ \textbf{subject to} \quad & \frac{d_i}{p_i} \leq 1 \qquad i = 1, \ldots, n \\ & \frac{1}{c - a_{ii}} \sum_{j \neq i} |a_{ij}| \frac{d_i}{d_j} \leq 1 \qquad i = 1, \ldots, n. \end{aligned} \tag{9}$$

which is a geometric program in posynomial form. This problem is not necessarily convex in the d_i, but can be transformed to an equivalent convex problem (see Section 4.5.3 of [23]).

IV. Bounding error due to numerical integration

Numerical solvers for ordinary differential equations suffer from errors inherent in the discretization of continuous-time systems. Performing simulation-based verification of a continuous-time model thus requires a bound on the accumulated numerical error over subsequent time steps. The matrix measure has been used for the analysis of numerical algorithms for ordinary differential equations [16], [17], [18]. We now extend these results to provide reachability algorithms that are robust against numerical error.

We again consider the system (4) with $f_\sigma : [0, T] \times \mathbb{R}^n \to \mathbb{R}^n$ continuous in t and C^1 in x for all $\sigma \in \mathcal{S}$. We are given a simulation trace $(t_0, x_0), (t_1, x_1), \ldots, (t_l, x_l)$ of this system with initial condition x_0 and an accuracy constant $K_a > 0$ such that $|\xi(t_{i+1} - t_i, x_i) - x_{i+1}| \leq K_a$ for all $i = 0, \ldots, l-1$. We wish to compute a set guaranteed to contain the true system trajectory $\xi(\cdot, x_0)$.

Reference [11] provides a solution to this problem using a Lipschitz constant. We demonstrate here an alternative procedure using the matrix measure. As a consequence of property 1 of the matrix measure: $\mu(A) \leq \|A\|$, our method provides at least as good a bound on the accumulated error than a similar method using the Lipschitz constant. We then extend our results from Section III to develop a verification method that provides guarantees robust against numerical error.

Proposition 3: Define $\epsilon_0 = 0$. For each $i = 0, \ldots, l-1$ suppose that we can find $c_i \in \mathbb{R}$ such that $\mu(J_{\sigma_i}(t, x)) \leq c_i$ for all $(t, x) \in [t_i, t_{i+1}] \times \mathrm{Conv}(Reach_{[t_i . t_{i+1}]}(\mathcal{B}_{\epsilon_i}(x_i)))$. Define $\epsilon_{i+1} = \epsilon_i e^{c_i(t_{i+1} - t_i)} + K_a$. We have the following bounds on the accumulated numerical error

$$|\xi(t_i, x_0) - x_i| \leq \epsilon_i \quad i = 0, \ldots, l \tag{10}$$

In light of Proposition 3 we can account for numerical error in Algorithm 1 by modifying line 5 as

5: $\delta_{i+1} = e^{c_i(t_{i+1} - t_i)} \delta_i + K_a$.

The extension of Algorithm 2 to be robust to numerical error is less straightforward, as we need a bound K'_a on the numerical error in the weighted Euclidean norm $|x|_{\Gamma_i} = x^T \Gamma_i x$ being used at each step. If we have $|\xi(t_{i+1} - t_i, x_i) - x_{i+1}| \leq K_a$ in the initial norm $|x| = x^T \Gamma_0 x$ then the bound K'_a can be computed as $K'_a = \sqrt{s}$ where s is the solution to the optimization problem

$$\begin{aligned} \textbf{minimize} \quad & s \\ \textbf{subject to} \quad & K_a^2 \Gamma_i \preceq s \Gamma_0 \end{aligned} \tag{11}$$

The solution to this problem can then be incorporated to the update of Γ_{i+1} in Algorithm 2 by modifying lines 6 and 7 as

6: Find solution s to optimization problem (11)
7: Set $\Gamma_{i+1} = \frac{1}{e^{2c_i(t_{i+1} - t_i)} + s} \Gamma_i$.

V. Leak Test Benchmark

We now consider a model for the detection of leaks in a pressurized network. This model is presented in [14] as a benchmark problem for comparing the performance of verification algorithms for hybrid systems as the number of continuous states increases. The model is a switched hybrid system with nonlinear dynamics in each mode and time-dependent switching between modes and is therefore well-suited to analysis using our method.

A. Modelling

We model a network of pipe segments designed to transport pressurized gas. The network has the topology of a tree in which gas enters through the root segment and exits to burners situated at the leaf segments. The gas pressure in each segment i of the network is modelled by a state variable x_i. Gas enters segment i through a valve v_i upstream from the segment. At each leaf segment ℓ of the tree is a terminal valve v^*_ℓ regulating the flow from the terminal segment to a gas burner.

We wish to perform a test that determines if any valves in the network leak. To do this, an additional test valve vt_i is attached to each segment i. This valve is connected to a bubbler that empties into a vessel of liquid. When the test valve vt_i is opened, if the pressure in segment i is above a known threshold z_0, gas flows through the valve and can be observed bubbling the vessel. A sample network illustrating this setup is shown in Figure 3.

For a given segment i, we denote the set of neighbouring segments $\mathcal{N}_i$. The system's dynamics are governed by the equations

$$\dot{x}_i = d_{i,\sigma(t)} g(x_i) + \sum_{j \in \mathcal{N}_i} c_{j,\sigma(t)} f(x_i, x_j) \tag{12}$$

where $f(x, y) = \phi(x - y)$ for some odd increasing differentiable function ϕ. The function g is such that $g(x) = 0$ for

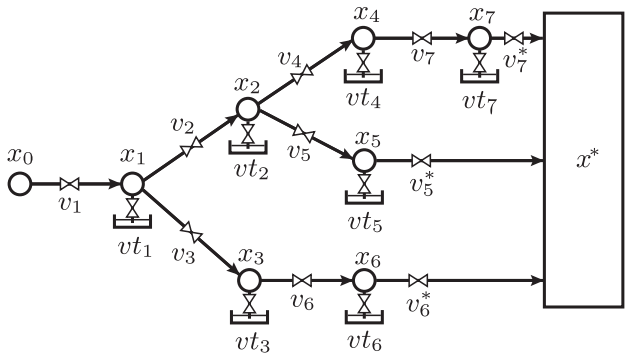

Fig. 3: Sample network for the leak test problem.

x less than some bubbling threshold z_0 and decreasing for $x > z_0$. In [14] the functions

$$f_0(x,y) = \begin{cases} -\sqrt{x-y} & x \geq y \\ \sqrt{y-x} & x < y \end{cases}$$

and

$$g_0(x) = \begin{cases} -\sqrt{x-z_0} & x \geq z_0 \\ 0 & x < z_0. \end{cases}$$

are used. In order to satisfy our differentiability assumption, we generate f and g by taking the convolution of f_0 and g_0 with the bump functions

$$b_\epsilon^2(x,y) = \frac{1}{\epsilon^2} e^{\frac{1}{1-|(x,y)/\epsilon|^2}} 1_{\{|(x,y)|<\epsilon\}}$$

and

$$b_\epsilon^1(x) = \frac{1}{\epsilon} e^{\frac{1}{1-(x/\epsilon)^2}} 1_{\{|x|<\epsilon\}}$$

respectively for $\epsilon = 10^{-3}$. This generates function f and g that approximate f_0 and g_0 but are C^1 everywhere, as is required for our analysis.

In addition, we append states with dynamics $\dot{x}_0 = 0$ and $\dot{x}^* = 0$ that represent the constant pressures in the chamber used to pressurize the network and the ambient pressure respectively. This gives us a system with full state $x = [x_0, x_1, \ldots, x_n, x^*]^T$. The switching is determined by whether each valve is turned on or off. If valve v_j between segment i and a neighbouring segment j is closed at time t then $c_{j,\sigma(t)} = c_{closed}$, otherwise $c_{j,\sigma(t)} = c_{open}$. Similarly, $d_{i,\sigma(t)} = d_{open}$ whenever the bubbler valve for segment i is open, and $d_{i,\sigma(t)} = d_{closed}$ otherwise.

A leak test proceeds as follows. The network is pressurized for an initialization period of t_{init} time units, during which all internal valves are opened and gas enters the network through the root from a chamber with a fixed pressure. After the initialization period is complete, the internal valves are closed and segments are tested, beginning from the leaf segments. To test a given segment, we first wait for a period of t_{wait} time units. Afterwards, the bubble valve and the upstream valves are opened. Bubbling is expected initially, and if bubbling does not occur a leak in an upstream valve is suspected. We now wait for a period of t_{test}. We expect the bubbling to stop within this period, otherwise a leak in the downstream valve is suspected.

B. Computing contraction rate

The interaction terms between states satisfy

$$\frac{\partial f}{\partial x}(x,y) = \phi'(x-y) = -\frac{\partial f}{\partial y}(x,y)$$

and the leak term satisfies $\frac{\partial g}{\partial x}(x) \leq 0$. Now, if $J_{\sigma(t)}(t,x)$ is the system Jacobian then the rows corresponding to states x_0 and x^* all have $J_{ij}(t,x) = 0$ for all j. The remaining rows then satisfy

$$\begin{aligned} J_{ii}(t,x) \quad + \quad & \sum_{j \neq i} |J_{ij}(t,x)| \\ = \quad & \left(d_{i,\sigma(t)} \frac{\partial g}{\partial x_i}(x_i) + \sum_{j \in \mathcal{N}_i} c_{j,\sigma(t)} \frac{\partial f}{\partial x_i}(x_i, x_j) \right) \\ & + \sum_{j \in \mathcal{N}_i} \left| c_{j,\sigma(t)} \frac{\partial f}{\partial x_j}(x_i, x_j) \right| \\ = \quad & d_{i,\sigma(t)} \frac{\partial g}{\partial x_i}(x_i) \leq 0. \end{aligned}$$

Therefore

$$\begin{aligned} \mu_\infty(J_{\sigma(t)}(t,x)) \quad &= \quad \max_i \left(J_{ii}(t,x) + \sum_{j \neq i} |J_{ij}(t,x)| \right) \\ &= \begin{cases} \max_i d_{i,\sigma(t)} \frac{\partial g}{\partial x_i}(x_i) & \text{if } d_{j,\sigma(t)} > 0 \text{ for all } j \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

C. Verification of the example network

We wish to verify a leak test procedure for an example network with the topology presented in Figure 3. The valve and test parameters are given in Table II.

c_{open}	c_{closed}	d_{open}	d_{closed}	z_0	t_{init}	t_{wait}	t_{test}
1.0	0.01	0.1	0	1.1	10	3	3

TABLE II: Valve and test parameters for verification of sample network.

We assume that for each segment i, the initial states satisfy $0.95 \leq x_i(0) \leq 1.05$, $x_0 = 2$, $x^* = 1$. We wish to verify that for any set of initial conditions within this range, this leak test procedure correctly determines that no valves leak. Using the value of $\mu_\infty(J_{\sigma(t)}(t,x))$ computed in Section V-B, we can perform the verification using a single simulation trace. This trace, and the corresponding over-approximations of the reachable set are shown in Figure 4.

D. Demonstration of scalability

We use the leak test benchmark to demonstrate that our reachability algorithm scales well, allowing us to verify the correctness of a leak test procedure for systems with a large number of state variables. We consider a sequence of binary tree networks of increasing depth. The valve and test parameters are given in Table II but with t_{init} changed to a value of 40. Note that the initialization time t_{init} must be greater than for the sample network in Section V-C to generate a successful test because it takes a longer amount of time to pressurize a large network.

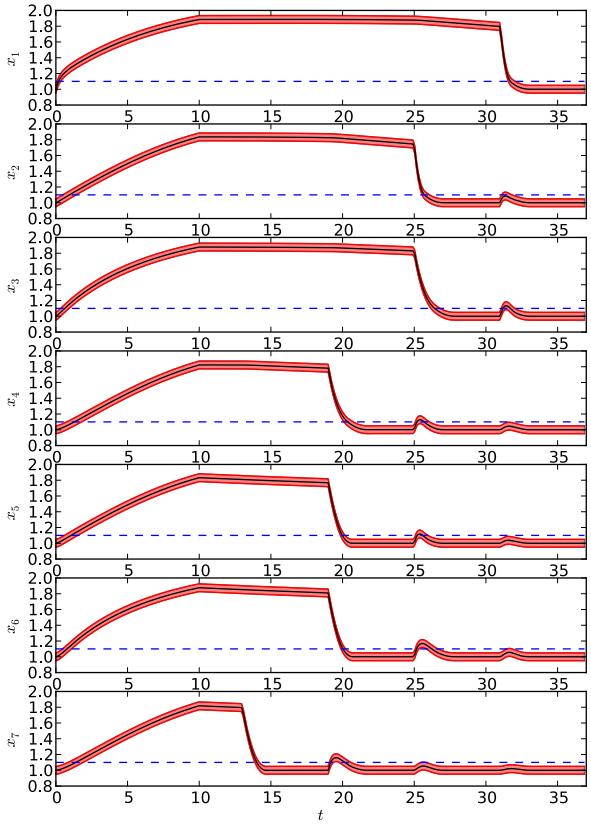

Fig. 4: Verification of a leak test scheme using a single simulation trace. The single computed trace is shown in black. The computed overapproximation of set reachable from the set of initial conditions is shown red. The dashed blue line depicts the bubbling threshold.

The reachability computations are performed in Python 2.7.1 running on a machine with a 2.3 GHz Intel Core i7 processor and 8.0 GB RAM. The time required for computation is given in Table III for networks of varying dimension.

State dimension	3	7	15	31	63	127
Time for verification (s)	1.520	3.275	7.603	17.994	51.407	107.525

TABLE III: Computation time to verify binary tree networks with varying dimension.

We see that using Algorithm 1 we can verify the leak test procedure for as many as 127 states relatively quickly.

VI. Conclusions

We provided a method of using matrix measures to compute reachable set enclosures for switched nonlinear dynamical systems. We showed that this method provides guaranteed overapproximations of the reachable sets and provided a method of improving the tightness of the approximation by choosing optimally weighted norms. Our example illustrates that this method can be applied to a benchmark problem in the verification of high-dimensional nonlinear hybrid systems, allowing us to compute reachable sets for systems with as many as 127 continuous state variables. This degree of scalability is remarkable in the context of nonlinear reachability, but comes at the cost of a guarantee on the tightness of the approximation; we can only guarantee convergence to the true reachable set asymptotically as the size of the mesh from which initial conditions are chosen tends to zero. However, we show that our method computes reachable set enclosures of sufficient tightness to verify safety for this benchmark problem.

References

[1] I. Mitchell, A. Bayen, and C. Tomlin, "A time-dependent Hamilton-Jacobi formulation of reachable sets for continuous dynamic games," *IEEE Trans. Autom. Control*, vol. 50, no. 7, pp. 947–957, 2005.

[2] M. Althoff, O. Stursberg, and M. Buss, "Reachability analysis of nonlinear systems with uncertain parameters using conservative linearization," in *IEEE Conf. Decision Control*, 2008, pp. 4042–4048.

[3] A. Chutinan and B. Krogh, "Computational techniques for hybrid system verification," *IEEE Trans. Autom. Control*, vol. 48, no. 1, pp. 64–75, 2003.

[4] A. Girard and G. J. Pappas, "Verification using simulation," in *Hybrid Sys.: Comp. Control*, 2006, pp. 272–286.

[5] J. Kapinski, A. Donzé, F. Lerda, H. Maka, S. Wagner, and B. H. Krogh, "Control software model checking using bisimulation functions for nonlinear systems." in *IEEE Conf. Decis. Control*, 2008, pp. 4024–4029.

[6] Y. Lin and M. A. Stadtherr, "Validated solutions of initial value problems for parametric ODEs," *Appl. Numer. Math.*, vol. 57, no. 10, pp. 1145–1162, 2007.

[7] M. Neher, K. R. Jackson, and N. S. Nedialkov, "On Taylor model based integration of ODEs," *SIAM J. Numer. Anal*, vol. 45, 2007.

[8] V. Lakshmikantham and S. Leela, *Differential and Integral Inequalities, volume 1*. Academic Press, New York, 1969.

[9] J. K. Scott and P. I. Barton, "Bounds on the reachable sets of nonlinear control systems," *Automatica*, vol. 49, no. 1, pp. 93–100, 2013.

[10] A. Donzé and O. Maler, "Systematic simulation using sensitivity analysis," in *Hybrid Sys.: Comp. Control*. Springer-Verlag, 2007, pp. 174–189.

[11] Z. Huang and S. Mitra, "Computing bounded reach sets from sampled simulation traces," in *Hybrid Sys.: Comp. Control*, 2012, pp. 291–294.

[12] Z. Huang, "On simulation based verification of nonlinear nondeterministic hybrid systems," Master's Thesis, University of Illinois at Urbana-Champaign, 2013.

[13] A. Julius and G. Pappas, "Trajectory based verification using local finite-time invariance," in *Hybrid Sys.: Comp. Control*. Springer, 2009, pp. 223–236.

[14] A. Fehnker and F. Ivančić, "Benchmarks for hybrid systems verification," in *Hybrid Systems: Computation and Control*. Springer, 2004, pp. 326–341.

[15] C. Desoer and M. Vidyasagar, *Feedback Systems: Input-Output Properties*, ser. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, 2009.

[16] C. Desoer and H. Haneda, "The measure of a matrix as a tool to analyze computer algorithms for circuit analysis," *IEEE Trans. Circuit Theory*, vol. 19, no. 5, pp. 480–486, 1972.

[17] G. Dahlquist, *Stability and Error Bounds in the Numerical Integration of Ordinary Differential Equations*. Almqvist & Wiksells, 1959.

[18] S. M. Lozinskiĭ, "Error estimates for numerical integration of ordinary differential equations (Russian)," *Izv. Vysš. Učebn. Zaved. Matematika*, vol. 5, no. 5, pp. 52–90, 1958.

[19] E. D. Sontag, "Contractive systems with inputs," in *Perspectives in Mathematical System Theory, Control, and Signal Processing*. Springer-Verlag, 2010, pp. 217–228.

[20] W. Lohmiller and J.-J. Slotine, "On contraction analysis for non-linear systems," *Automatica*, vol. 34, no. 6, pp. 683–696, 1998.

[21] D. Angeli, "A Lyapunov approach to incremental stability properties," *IEEE Trans. Autom. Control*, vol. 47, no. 3, pp. 410–421, 2002.

[22] Z. Aminzare, Y. Shafi, M. Arcak, and E. Sontag, "Guaranteeing spatial uniformity in reaction-diffusion systems using weighted L_2-norm contractions," in *A Systems Theoretic Approach to Systems and Synthetic Biology: Models and System Characterizations*, V. Kulkarni, G.-B. Stan, and K. Raman, Eds. Springer-Verlag, 2014, pp. 73–102.

[23] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.