Guaranteed Reachability for Systems with Unknown Dynamics

Melkior Ornik[1]

***Abstract*— The problem of computing the reachable set for a given system is a quintessential question in nonlinear control theory. Motivated by prior work on safety-critical online planning, this paper considers an environment where the only available information about system dynamics is that of dynamics at a single point. Limited to such knowledge, we study the problem of describing the set of all states that are guaranteed to be reachable regardless of the unknown true dynamics. We show that such a set can be underapproximated by a reachable set of a related known system whose dynamics at every state depend on the velocity vectors that are available in all control systems consistent with the assumed knowledge. Complementing the theory, we discuss a simple model of an aircraft in distress to verify that such an underapproximation is meaningful in practice.**

I. Introduction

Following an adverse inflight event affecting aircraft safety, a pilot operating an aircraft in distress needs to determine whether to divert to an alternate airport, and if so, which airport to choose. While it *may* be possible for the aircraft to continue to its original destination, the pilot is required to determine a diversion airport where the aircraft can *certainly* land [14]. At the time of decision, depending on the nature of the adversity, the pilot might not have a correct model of flight dynamics; in extreme cases such as loss of a wing [2], there may be little prior experience or available knowledge about the system. Thus, it is imperative to immediately provide the pilot, or the flight controller for an autonomous vehicle, with a set of landing points that the aircraft is guaranteed to be able to reach, given the current information about the system.

The above problem describes a variant of a classical question of reachability: given the system's initial state, we wish to describe the set of all states that can be reached by an admissible control input [8], [17]. Previous work on reachability under uncertainty considered computation of reachable sets with dynamics generated by a finite number of uncertain parameters [3] or having bounded disturbances [20]. The framework that we are considering — motivated by the example of an aircraft in sudden distress — contains substantially fewer assumptions on the structure of uncertainty. Namely, drawing from the work of [24] which determines *local controlled dynamics* of a nonlinear system at a given state using solely the information from a single trajectory until the time that the system reached that state, we assume that the only available information at the time of computing the reachable set consists of (i) local dynamics at a single point and (ii) Lipschitz bounds on the rate of change of system dynamics. In light of such limited knowledge, we wish to determine the set of states that are guaranteed to be reachable *regardless* of the true system dynamics, as long as they are consistent with the above knowledge. We call such a set the *guaranteed reachability set (GRS)*.

The primary contribution of this paper is to provide a meaningful underapproximation of the GRS. Our approach for determining such an underapproximation relies on interpreting a control system as a differential inclusion with the *available velocity set* at every state in the state space describing the possible tangents to the system's trajectory when leaving that state. Local dynamics at each state determine the available velocity set at that state. While exact available velocity sets are not known anywhere except for a single point, we can determine the family of all velocity sets that are consistent with our knowledge about the system dynamics. The intersection of the elements of such a family provides the *guaranteed velocity set* at every state in the state space; a set of velocities for which we know there exists a control input generating them at a given state, even if we do not know the exact dynamics at that state. Since such a set may be difficult to compute or express in closed form, we underapproximate it by a ball of maximal possible radius. Moving back from differential inclusions to differential equations with control inputs, such an underapproximation generates a control system with *known* dynamics; the reachable set of such a control system is a subset of the GRS.

Notation. The set of all matrices with m rows and n columns is denoted by $\mathbb{R}^{m\times n}$. For a vector v, $\|v\|$ denotes its Euclidean norm. For a matrix M, M^T denotes its transpose and $\|M\|$ denotes its 2-norm: $\|M\| = \max_{\|v\|=1} \|Mv\|$. Notation $\mathbb{B}^m(x;r)$ denotes a closed ball in $\mathbb{R}^m$ with the center at $x \in \mathbb{R}^m$ and radius $r \geq 0$ under the Euclidean norm; analogously, $\mathbb{B}^{m\times n}(x;r)$ is a closed ball in $\mathbb{R}^{m\times n}$ under the matrix 2-norm. Set $GL(n)$ denotes the general linear group of all matrices $\mathbb{R}^{n\times n}$, i.e., all invertible $n \times n$ matrices. Set $C_L(\mathbb{R}^k, \mathbb{R}^l)$ denotes the set of all functions $f : \mathbb{R}^k \to \mathbb{R}^l$ with a Lipschitz constant L, i.e., the set of all functions f that satisfy $\|f(x) - f(y)\| \leq L\|x - y\|$ for all $x, y \in \mathbb{R}^k$. For a matrix M, vector v, and set $\mathcal{S}$ of vectors of appropriate dimension, notation $v + M\mathcal{S}$ denotes the set $\{v + Ms \mid s \in \mathcal{S}\}$. Notation $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ denotes a diagonal $n \times n$ matrix with elements λ_1, $\ldots$, λ_n on the diagonal, in that order. Vector $e_i \in \mathbb{R}^n$ denotes the i-th coordinate vector, i.e., a vector of all zeros, except a 1 in position i. Matrix I is the identity matrix.

* This work was supported by an Early Stage Innovations grant from NASA's Space Technology Research Grants Program, grant no. 80NSSC19K0209.

[1] M. Ornik is with the Department of Aerospace Engineering and the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. Email: mornik@illinois.edu

II. Prior System Knowledge

The assumptions and objectives of this work are driven by the desire to successfully control a system with entirely unknown dynamics by learning as much as possible about the dynamics solely from the system's behavior during a *single* system run [24]. While previous such work does assume that the system dynamics are known to be control-affine with a particular Lipschitz bound, it significantly differs from classical adaptive and robust control [13], [16], [21] by not assuming almost any knowledge about the magnitude or the structure of the uncertainty in system dynamics, and from classical data-driven learning methods [9], [11] by not allowing collection of data through repeated system runs.

By using a special control input, [24] determines the *local dynamics* $u \mapsto f(x_0)+G(x_0)u$ of a unknown control system $\dot{x} = f(x) + G(x)u$ for any state x_0 lying on a system trajectory, with an arbitrarily small error and using only (i) the knowledge of the trajectory prior to the time at which x_0 is visited and (ii) the bounds on Lipschitz constants and magnitudes of f and G. The latter bounds are not required to be sharp, and may thus come from basic knowledge of the physical laws and the system's environment.

Owing to the lack of information about full system dynamics, it may be impossible to determine online whether a particular state can be reached. However, local dynamics coupled with the above bounds yield the following sets:

(i) states that *may* be possible to reach using admissible control signals *(optimistic reachability set)*,

(ii) states that are *guaranteed* to be reachable using admissible control signals *(guaranteed reachability set — GRS)*.

Without discussing the reachable sets, the work in [24] has implicitly focused on optimistic reachability, i.e., attempting to reach the original objective while there appears to be any chance of reaching it. This paper considers guaranteed reachability. Such an investigation is naturally motivated by our running example: a damaged aircraft needs to determine an airport where it can certainly land rather than attempting to continue to an airport which might not be reachable.

We remark the same broad idea of guaranteed reachability is considered in [1], which deems an unknown dynamical system *safe* if it is guaranteed that its trajectory will not enter a particular forbidden set. However, unlike our work, the dynamical system in that work has no control input. In the language of [1], our paper provides a certification of existence of a control input that renders the resulting dynamical system safe.

III. Problem Statement

Throughout the paper, we consider a control system $\mathcal{M}(f, G)$ described by nonlinear control-affine dynamics

$$\dot{x}(t) = f(x(t)) + G(x(t))u(t), \quad x(0) = x_0, \tag{1}$$

where $t \geq 0$, $x(t) \in \mathbb{R}^n$ for all t, admissible inputs $u(t) \in \mathcal{U}$ lie in the set $\mathcal{U} \subseteq \mathbb{R}^m$, and functions $f : \mathbb{R}^n \to \mathbb{R}^n$ and $G : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ are globally Lipschitz-continuous, i.e., there exist $L_f \geq 0$ and $L_G \geq 0$ such that $f \in C_{L_f}(\mathbb{R}^n, \mathbb{R}^n)$ and $G \in C_{L_G}(\mathbb{R}^n, \mathbb{R}^{n\times m})$. Throughout the paper, we only consider control signals u such that all solutions of all relevant differential equations exist and are unique. Without loss of generality, we consider $x_0 = 0$. For technical reasons we make the following assumption.

Assumption 1: System $\mathcal{M}(f, G)$ is fully actuated at $x_0 = 0$, i.e., $m = n$ and $G(0) \in GL(n)$. Set $\mathcal{U}$ equals $\mathbb{B}^n(0; 1)$.

The assumption of full actuation makes estimation of available system velocities at points x around $x_0 = 0$ significantly simpler. We briefly discuss the case of a non-invertible $G(0)$, possibly with $m \neq n$, later in the paper. The assumption of $\mathcal{U} = \mathbb{B}^n(0; 1)$ parallels an assumption in [24] that $\mathcal{U} = [-1, 1]^m$; in our running example, the control inputs available to the aircraft are clearly bounded, and taking $\mathcal{U} = \mathbb{B}^n(0; 1)$ simplifies the technical work.

Building upon [24], we place the following limitations on our information about the system dynamics.

Assumption 2: Bounds L_f and L_G are known, as well as values $f(0)$ and $G(0)$. Nothing else is known about f or G.

We remark that knowing $f(0)$ and $G(0)$ is equivalent to knowing the local system dynamics $u \mapsto f(0) + G(0)u$. Since the uncertainty in the system cannot be boiled down to any finite number of unknown parameters or disturbances, approaches relying on corresponding system structures [3], [20] cannot be used.

Let $\mathcal{D}_{con}$ denote the set of all functions $(\hat{f}, \hat{G})$ consistent with the assumed information: $\mathcal{D}_{con} = \{(\hat{f}, \hat{G}) \mid \hat{f} \in C_{L_f}(\mathbb{R}^n, \mathbb{R}^n), \hat{G} \in C_{L_G}(\mathbb{R}^n, \mathbb{R}^{n\times m}), \hat{f}(0) = f(0), \hat{G}(0) = G(0)\}$. Our goal is to describe a set of states that are guaranteed to be reachable from 0 *regardless* of the functions f and G, but keeping in mind that we know $(f, G) \in \mathcal{D}_{con}$. To formally state this problem, we first define the *(forward) reachable set* $R^{\hat{f},\hat{G}}(T; x_0)$ in the usual manner by $R^{\hat{f},\hat{G}}(T, x_0) = \{\phi_u^{\hat{f},\hat{G}}(T; x_0) \mid u : [0, T] \to \mathcal{U}\}$, where $\phi_u^{\hat{f},\hat{G}}(T; x_0)$ denotes a controlled trajectory of system $\mathcal{M}(\hat{f}, \hat{G})$ with control signal u and $\phi_u^{\hat{f},\hat{G}}(0; x_0) = x_0$.

Problem 1: Let $T \geq 0$. Describe the *guaranteed reachability set (GRS)* defined by

$$R^{\mathcal{G}}(T, 0) = \bigcap_{(\hat{f},\hat{G})\in\mathcal{D}_{con}} R^{\hat{f},\hat{G}}(T, 0). \tag{2}$$

While we pose Problem 1 in terms of finding the GRS at time T, we could analogously ask for guaranteed reachability before time T or guaranteed eventual reachability. As we will show on an example in Section VI, many of the results of this paper can be adapted to the those two sets.

In our running example of an aircraft in distress, Problem 1 is only the first of two steps towards managing the crisis: after determining which airport to land at, the pilot needs to determine in real time *how* to land at that airport. The interest of this paper is primarily in solving Problem 1. However, Section V also provides a method of determining, under technical assumptions, a control input to reach the desired state x_{end} if x_{end} is certified to be reachable. As such a control input will depend on system dynamics, the proposed method exploits the work of [24] on continual learning of local system dynamics.

The primary contribution of this paper is to provide an underapproximation of $R^{\mathcal{G}}(T,0)$ using the reachable set of the control system $\dot{x} = a + g(\|x\|)u$, where a is a constant and $g : [0,+\infty) \to [0,+\infty)$ is a ramp function. Our attack plan consists of three steps: we will first show, in Proposition 1, that $R^{\mathcal{G}}(T,0)$ can be underapproximated by a reachable set of a related ordinary differential inclusion (ODI). Then, in Theorem 1 and Proposition 2, we will find a maximal underapproximation of the right hand side of this ODI by a ball, yielding a new ODI. Finally, in Theorem 2 we will show that the reachable set of this new ODI is the same as the intersection of the reachable set of an induced control system $\dot{x} = a + g(\|x\|)u$ and a ball.

IV. Guaranteed Velocities

Following the classical approach of interpreting ordinary differential equations with control inputs as inclusions [7], [18], we define an ODI

$$\dot{x} \in \mathcal{V}_x = f(x) + G(x)\mathcal{U}, \quad x(0) = x_0. \tag{3}$$

Clearly, a trajectory $\phi(\cdot; x_0)$ satisfies (3) if and only if it is a solution to the control system (1) for an admissible control signal. Thus, if we define the *reachable set* of (3) as a set of values $\phi(T; x_0)$ of all possible trajectories $\phi(\cdot; x_0)$ that satisfy (3), such a reachable set equals $R^{f,G}(T, x_0)$.

Given Assumption 2, set $\mathcal{V}_{x_0} = \mathcal{V}_0$ is known. The goal of this section is to provide an underapproximation for $\mathcal{V}_x$ for all $x \in \mathbb{R}^n$ using sets $\mathcal{D}_{con}$ and $\mathcal{V}_0$.

Set $\mathcal{V}_x$ describes all the velocities available at state x for a control system (1). Analogously to the GRS, we can define the *guaranteed velocity set*

$$\mathcal{V}_x^{\mathcal{G}} = \bigcap_{(\hat{f},\hat{G})\in\mathcal{D}_{con}} \hat{f}(x) + \hat{G}(x)\mathcal{U} \subseteq \mathcal{V}_x. \tag{4}$$

Let us now consider ODI

$$\dot{x} \in \mathcal{V}_x^{\mathcal{G}}, \quad x(0) = 0. \tag{5}$$

Even though $\mathcal{V}_x^{\mathcal{G}}$ may be empty for certain x, we can still discuss the reachable set of (5); if $\mathcal{V}^{\mathcal{G}}_{\phi(T;x_0)} = \emptyset$, we will consider by convention that trajectory $\phi(\cdot; x_0)$ of (5) ceases to exist at time T. The following proposition then holds directly from (2), (4), and the discussion below (3).

Proposition 1: Let $T \geq 0$. If a function $\phi : [0,+\infty) \to \mathbb{R}^n$ satisfies (5) at all times $t \leq T$, then $\phi(T) \in R^{\mathcal{G}}(T,0)$.

Proposition 1 implies that the reachable set of (5) is a *subset* of $R^{\mathcal{G}}(T,0)$. We neither claim nor assume that these sets are equal: the intersection of reachable sets of $\dot{x} \in F_i(x)$ does not generally equal the reachable set of $\dot{x} \in \cap_i F_i(x)$. Establishing conditions for equality of the reachable set of (5) and $R^{\mathcal{G}}(T,0)$ is an open question for future work.

Motivated by Proposition 1, we can underapproximate $R^{\mathcal{G}}(T,0)$ by considering the reachable set of (5). To describe the latter set, we first examine the geometry of $\mathcal{V}_x^{\mathcal{G}}$.

For a given $x \in \mathbb{R}^n$ it is simple to show that

$$\begin{aligned}&\{(\hat{f}(x), \hat{G}(x)) \mid (\hat{f},\hat{G}) \in \mathcal{D}_{con}\} = \\ &\mathbb{B}^n(f(0); L_f\|x\|) \times \mathbb{B}^n(G(0); L_G\|x\|).\end{aligned} \tag{6}$$

Thus, since $\mathcal{U} = \mathbb{B}^n(0;1)$ and $a + B\mathbb{B}^n(0;1)$ is an ellipsoid [4] for any $a \in \mathbb{R}^n$ and $B \in \mathbb{R}^{n\times n}$, $\mathcal{V}_x^{\mathcal{G}}$ is an intersection of an infinite set of ellipsoids with centers lying in a ball. Since an intersection of any number of ellipsoids is generally not an ellipsoid [19], our next step is to provide a simply computable underapproximation of $\mathcal{V}_x^{\mathcal{G}}$. This underapproximation is given in Theorem 1; the following lemmas build up to it by discussing (i) an underapproximation of an ellipsoid by a ball, (ii) an underapproximation of an intersection of ellipsoids with the same center by a ball, and (iii) the intersection of balls with centers in a ball.

Lemma 1: Let $a \in \mathbb{R}^n$, $B \in GL(n)$, and $\mathcal{U} = \mathbb{B}^n(0;1)$. Then $a + B\mathcal{U} \supseteq \mathbb{B}^n(a; \|B^{-1}\|^{-1})$ and $\mathbb{B}^n(a; \|B^{-1}\|^{-1})$ is a ball of maximal radius contained in $a + B\mathcal{U}$.

Proof: Showing $a + B\mathcal{U} \supseteq \mathbb{B}^n(a; \|B^{-1}\|^{-1})$ is equivalent to showing that for every vector δ, $\|\delta\| \leq \|B^{-1}\|^{-1}$, there exists $u \in \mathbb{B}^n(0;1)$ such that $a + \delta = a + Bu$. Since B is invertible, setting $u = B^{-1}\delta$ yields $a + \delta = a + Bu$. Additionally, $\|u\| = \|B^{-1}\delta\| \leq \|\delta\|\|B^{-1}\|$ by the definition of the matrix 2-norm. Since $\|\delta\| \leq \|B^{-1}\|^{-1}$, we obtain $u \in \mathbb{B}^n(0;1) = \mathcal{U}$ as desired.

To show the latter claim, we use a standard description [5] of the ellipsoid $a + B\mathcal{U}$ using singular values of B. Combining the results in [5], we note that the length of shortest principal semi-axes of the ellipsoid $a + B\mathcal{U}$ equals the smallest singular value of B, which equals exactly $\|B^{-1}\|^{-1}$. Thus, through a change of coordinates, we can assume without loss of generality that $a + B\mathcal{U}$ is given by $\{x \mid \alpha_1 x_1^2 + \alpha_2 x_1^2 + \ldots + \alpha_n x_n^2 \leq 1\}$, where $0 < \alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_n = \|B^{-1}\|^2$.

Consider now a ball of radius $\delta > \|B^{-1}\|^{-1}$ around any state $x' = (x'_1, \ldots, x'_n) \in \mathbb{R}^n$. Then, at least one of the following inequalities holds: $x'_n + \delta > \|B^{-1}\|^{-1}$ or $x'_n - \delta < -\|B^{-1}\|^{-1}$. Hence, $\mathbb{B}^n(x'; \delta) \ni x' + \delta e_n \notin a + B\mathcal{U}$ or $\mathbb{B}^n(x'; \delta) \ni x' - \delta e_n \notin a + B\mathcal{U}$. ∎

Lemma 2: Let $a \in \mathbb{R}^n$, $B \in GL(n)$, $0 \leq r < \|B^{-1}\|^{-1}$, and $\mathcal{U} = \mathbb{B}^n(0;1)$. Define $\varepsilon = \min_{\hat{B}\in\mathbb{B}^{n\times n}(B;r)} \|\hat{B}^{-1}\|^{-1}$. Then $\cap_{\hat{B}\in\mathbb{B}^{n\times n}(B;r)} a + \hat{B}\mathcal{U} \supseteq \mathbb{B}^n(a;\varepsilon)$.

Proof: We note that ε is well-defined: all matrices $\hat{B}$ in $\mathbb{B}^{n\times n}(B;r)$ are invertible by the Eckart-Young-Mirsky theorem [5]. Then, for all $\hat{B} \in \mathbb{B}^{n\times n}(B;r)$, $\|\hat{B}^{-1}\|^{-1} \geq \varepsilon$ obviously holds, so the claim holds by Lemma 1. ∎

Lemma 3: Let $a \in \mathbb{R}^n$, $r \geq 0$, and $R \geq r$. Then $\cap_{\hat{a}\in\mathbb{B}^n(a;r)} \mathbb{B}^n(\hat{a}; R) = \mathbb{B}^n(a; R - r)$.

Proof: Let $v \in \mathbb{B}^n(a; R-r)$. Then, $\|v-a\| \leq R-r$. Thus, for every $\hat{a} \in \mathbb{B}^n(a;r)$, $\|v - \hat{a}\| \leq \|v - a\| + \|a - \hat{a}\| \leq R - r + r \leq R$. Hence, $v \in \mathbb{B}^n(\hat{a}; R)$.

Now, let $v \notin \mathbb{B}^n(a; R-r)$, i.e., $\|v-a\| > R-r$. Choose $\hat{a} = a + r(a-v)/\|a-v\|$. Clearly, $\hat{a} \in \mathbb{B}^n(a;r)$. On the other hand, $\|v - \hat{a}\| = \|v - a - r(a-v)/\|a-v\|\| = \|(a-v)\|(1 + r/\|a-v\|) = r + \|a-v\| > R$. Thus, $v \notin \mathbb{B}^n(\hat{a}; R)$. ∎

We finally combine the above lemmas to obtain an approximation for $\mathcal{V}_x^{\mathcal{G}}$.

Theorem 1: Let $\mathcal{U}$, L_f, L_G, $f(0)$, and $G(0)$ be as defined above. Let $x \in \mathbb{R}^n$ satisfy $(L_f + L_G)\|x\| \leq \|G(0)^{-1}\|^{-1}$.

Define

$$\overline{\mathcal{V}}_x^{\mathcal{G}} = \mathbb{B}^n\left(f(0); \|G(0)^{-1}\|^{-1} - L_f\|x\| - L_G\|x\|\right). \quad (7)$$

Then, $\overline{\mathcal{V}}_x^{\mathcal{G}} \subseteq \mathcal{V}_x^{\mathcal{G}}$.

Proof: By (4) and (6), set $\mathcal{V}_x^{\mathcal{G}}$ satisfies $\mathcal{V}_x^{\mathcal{G}} = \cap_{\hat{a}\in\mathbb{B}^n(f(0);L_f\|x\|)}\left(\cap_{\hat{B}\in\mathbb{B}^{n\times n}(G(0);L_G\|x\|)}\hat{a}+\hat{B}\mathcal{U}\right)$. Combining this result with Lemma 2, $\mathcal{V}_x^{\mathcal{G}}$ is a superset of $\cap_{\hat{a}\in\mathbb{B}^n(f(0);L_f\|x\|)}\mathbb{B}^n\left(\hat{a}; \min_{\hat{B}\in\mathbb{B}^{n\times n}(G(0);L_G\|x\|)}\|\hat{B}^{-1}\|^{-1}\right)$. By the characterization [5] of $\|\hat{B}^{-1}\|^{-1}$ as a singular value of $\hat{B}$ and Weyl's inequality for singular values [26], $\|G(0)^{-1}\|^{-1} - r \leq \min_{\hat{B}\in\mathbb{B}^{n\times n}(G(0);r)}\|\hat{B}^{-1}\|^{-1}$ for any $r \geq 0$. Hence, $\mathcal{V}_x^{\mathcal{G}} \supseteq \cap_{\hat{a}\in\mathbb{B}^n(f(0);L_f\|x\|)}\mathbb{B}^n(\hat{a}; \|G(0)^{-1}\|^{-1} - L_G\|x\|)$. The claim of the theorem now holds by Lemma 3, since we assume $L_f\|x\| \leq \|G(0)^{-1}\|^{-1} - L_G\|x\|$. ■

Theorem 1, illustrated by Fig. 1, is the central result of the current section. Assuming $L_f + L_G > 0$, for each $x \in \mathbb{B}^n(0, 1/((L_f+L_G)\|G(0)^{-1}\|))$ it generates a nonempty set $\overline{\mathcal{V}}_x^{\mathcal{G}}$ of guaranteed velocities, i.e., velocities v for which there certainly exists a control input $u \in \mathcal{U}$ such that $f(x)+G(x)u = v$. The case of $L_f+L_G = 0$ is uninteresting as it results in dynamics (1) being entirely known.

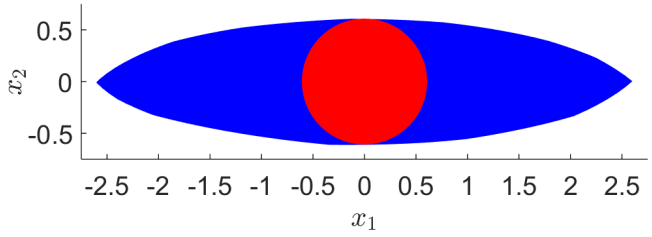

Fig. 1. Guaranteed velocity set $\mathcal{V}_x^{\mathcal{G}}$, with $x = (1, 0)$, $L_f = 0.1$, $L_G = 0.3$, $f(0) = [0\ 0]^T$, and $G(0) = \mathrm{diag}(3, 1)$, is drawn in blue, obtained approximately by Monte Carlo methods. Set $\overline{\mathcal{V}}_x^{\mathcal{G}} = \mathbb{B}^n(0; 0.6)$ is drawn in red.

The following result shows that $\overline{\mathcal{V}}_x^{\mathcal{G}}$ is indeed the *best* underapproximation of $\mathcal{V}_x^{\mathcal{G}}$ by a ball.

Proposition 2: Let $\mathcal{V}_x^{\mathcal{G}}$ and $\overline{\mathcal{V}}_x^{\mathcal{G}}$ be as above. Then, $\overline{\mathcal{V}}_x^{\mathcal{G}}$ is a ball of maximal radius that is contained in $\mathcal{V}_x^{\mathcal{G}}$.

Proof: We follow the same steps as the proof of (ii) in Lemma 1, which proves the proposition in the special case of $L_f = L_G = 0$. As the detailed proof is technically inelegant, if straightforward, we provide its outline.

As in part (ii) of Lemma 1, we may assume without loss of generality that $\mathcal{V}_0$ is an ellipsoid with principal axes parallel to coordinate axes and its shortest principal semi-axes given by $\pm\|G(0)^{-1}\|^{-1}e_n$. By the Eckart-Young-Mirsky theorem and the description of an ellipsoid through an singular value decomposition [5], there exist two ellipsoids $\hat{a} + \hat{B}\mathcal{U}$, where $\hat{a} = \pm L_f\|x\|e_n$ and $\hat{B} \in \mathbb{B}^{n\times n}(G(0); L_G\|x\|)$ such that their principal axes are parallel to the coordinate axes and their shortest principal semi-axes are given by $(\|G(0)^{-1}\|^{-1} - L_G\|x\|)e_n$. There thus exist two control systems $\mathcal{M}(\hat{f}, \hat{G})$ with $(\hat{f}, \hat{G}) \in \mathcal{D}_{con}$, with constructed ellipsoids as their available velocity sets at x.

Now consider any $x' \in \mathbb{R}^n$ and any $\delta > \|G(0)^{-1}\|^{-1} - L_G\|x\| - L_f\|x\|$. Set $\mathbb{B}^n(x'; \delta)$ contains at least one point $y = (y_1, \ldots, y_n)$ with $|y_n| \geq \delta$. On the other hand, any such point is in at most one of the two above ellipsoids $\hat{a} + \hat{B}\mathcal{U}$, and is thus not in $\mathcal{V}_x^{\mathcal{G}}$. ■

While Proposition 2 proves that $\overline{\mathcal{V}}_x^{\mathcal{G}}$ is the maximal approximation of $\mathcal{V}_x^{\mathcal{G}}$ by a ball, such an approximation may still discard a large part of $\mathcal{V}_x^{\mathcal{G}}$. In the case of $x = 0$, where $\mathcal{V}_x^{\mathcal{G}} = \mathcal{V}_x$, the ratio of the volumes of $\overline{\mathcal{V}}_x^{\mathcal{G}}$ and $\mathcal{V}_x^{\mathcal{G}}$ depends on the magnitude of $\|G(0)^{-1}\|^{-1}$ (i.e., the length of the shortest principal semi-axis of $\mathcal{V}_x$) relative to the lengths of other principal semi-axes of $\mathcal{V}_x$. By considering solely the ratio of the lengths of the longest and the shortest principal semi-axes, the fraction of the volume being discarded by our approximation depends on the *condition number* [5] of matrix $G(0)$. In other words, the closer $G(0)$ is to being singular, the worse the provided approximation will be.

Before continuing, let us briefly remark on the situation where system $\mathcal{M}(f, G)$ is underactuated, i.e., Assumption 1 does not hold. If $\mathrm{rank}(G(0)) = k < n$, all the work of this section can be performed by considering the smallest non-zero singular value σ_k of $G(0)$ instead of $\|G(0)^{-1}\|^{-1}$; the set $\mathcal{V}_0$ is then a k-dimensional ellipsoid with shortest principal semi-axes of length σ_k, and $\overline{\mathcal{V}}_x^{\mathcal{G}}$ would consequently be a k-dimensional ball of radius σ_k.

We now proceed to exploit the geometrically simple form of $\overline{\mathcal{V}}_x^{\mathcal{G}}$ to obtain a control system with known dynamics whose reachable set is an underapproximation of the GRS.

V. Reachable Set

By Theorem 1, the set of trajectories satisfying ODI (5) is a superset of the set of trajectories of the ODI

$$\dot{x} \in \overline{\mathcal{V}}_x^{\mathcal{G}}, \quad x(0) = x_0. \quad (8)$$

Let $\overline{R}(T, x_0)$ be the reachable set of (8) at time T, i.e., the set of all states achieved by trajectories $\phi : [0, +\infty) \to \mathbb{R}^n$ that satisfy $\phi(0) = x_0$ and (8) for all $t \leq T$. Proposition 1 and Theorem 1 show that $\overline{R}(T, x_0) \subseteq R^{\mathcal{G}}(T, x_0)$.

Set $\overline{\mathcal{V}}_x^{\mathcal{G}}$ is only defined for $x \in \mathbb{B}_{know} = \mathbb{B}^n(0; 1/((L_f + L_G)\|G(0)^{-1}\|))$. In order to exploit previous work on computing reachable sets, we continuously extend (8) to entire $\mathbb{R}^n$ by defining $\overline{\mathcal{V}}_x^{\mathcal{G}} = \{f(0)\}$ outside of $\mathbb{B}_{know}$.

It can be shown that any solution of (8), with extended $\overline{\mathcal{V}}_x^{\mathcal{G}}$, that connects two points in $\mathbb{B}_{know}$ needs to entirely lie within $\mathbb{B}_{know}$; the old $\overline{R}(T, x_0)$ prior to extending $\overline{\mathcal{V}}_x^{\mathcal{G}}$ thus corresponds to $\overline{R}(T, x_0) \cap \mathbb{B}_{know}$ for an extended $\overline{\mathcal{V}}_x^{\mathcal{G}}$. We can thus find an underapproximation $\overline{\overline{R}}(T, x_0) \subseteq R^{\mathcal{G}}(T, x_0)$ by defining $\overline{\overline{R}}(T, x_0) = \overline{R}(T, x_0) \cap \mathbb{B}_{know}$, where $\overline{R}(T, x_0)$ is the reachable set of (8) with extended $\overline{\mathcal{V}}_x^{\mathcal{G}}$. Our central question thus becomes determining the set $\overline{R}(T, x_0)$.

Analogously to our interpretation of dynamics (1) as ODI (3), inclusion (8) can, using the definition of $\overline{\mathcal{V}}_x^{\mathcal{G}}$ in (7), be interpreted as a control system

$$\dot{x} = a + g(\|x\|)u, \quad x(0) = x_0, \quad (9)$$

with $a = f(0)$, $u \in \mathcal{U} = \mathbb{B}^n(0; 1)$, and where $g(s) = \|G(0)^{-1}\|^{-1} - (L_G + L_f)s$ if $s \leq \|G(0)^{-1}\|^{-1}/(L_G +$

L_f) and $g(s) = 0$ otherwise. We thus trivially obtain the following result.

Theorem 2: Set $\overline{\overline{R}}(T, x_0) \subseteq R^{\mathcal{G}}(T, x_0)$ is the intersection of the reachable set of system (9) and $\mathbb{B}_{know}$.

We note that dynamics (9) are *entirely known*. Finding $\overline{\overline{R}}(T, x_0)$ hence becomes a standard — although emphatically nontrivial [25] — problem of determining the reachable set of a nonlinear control system. We follow the approach of [10].

Theorem 3: Let $a \in \mathbb{R}^n$ and $g : [0, +\infty) \to [0, +\infty)$ be as in (9). Let $l : \mathbb{R}^n \to \mathbb{R}$ be any sufficiently smooth function with $\{0\} = \{x \mid l(x) \leq 0\}$. Then, $\overline{\overline{R}}(T, x_0) = \{x \in \mathbb{R}^n \mid V(T, x) \leq 0\} \cap \mathbb{B}_{know}$, where $V : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ is the *viscosity solution* of the Hamilton-Jacobi equation

$$V_t(t,x) - V_x(t,x)^T a - \|V_x(t,x)\| g(\|x\|) = 0 \quad \text{for all } t \in [-T, 0], x \in \mathbb{R}^n, \tag{10}$$
$$V(0, x) = l(x) \text{ for all } x \in \mathbb{R}^n.$$

We direct the reader to [10], [12] for details on viscosity solutions to Hamilton-Jacobi equations. Notation V_t and V_x denotes the Jacobians of V with respect to t and x.

Proof of Theorem 3: The result follows from [10], by "reversing time" and directly adapting the backwards reachability results in [10] to our setting. The adapted results show that $\overline{R}(T, x_0) = \{x \in \mathbb{R}^n \mid V(T, x) \leq 0\}$, where V satisfies $V_t(t, x) + H(x, V_x(t, x)) = 0$ and $V(0, x) = l(x)$ for all $t \in [-T, 0]$ and $x \in \mathbb{R}^n$, and H is given by $H(x, \lambda) = \min_{u \in \mathcal{U}} \lambda^T(-a + g(\|x\|)u)$. The theorem claim now holds from Theorem 2, noticing that $\min_{u \in \mathcal{U}} \lambda^T(-a + g(\|x\|)u)$ equals $-\lambda^T a + g(\|x\|) \min_{u \in \mathcal{U}} \lambda^T u = -\lambda^T a - g(\|x\|)\lambda^T \lambda / \|\lambda\| = -\lambda^T a - g(\|x\|)\|\lambda\|$. ∎

Let us finally briefly discuss the question of determining a control signal that drives the state of system (1) from x_0 to an element of $\overline{\overline{R}}(T, x_0)$. The results of [10] provide an expression for a control signal $\overline{u}$ that drives the state of system (9) from x_0 to the desired objective. We now want to convert that signal into an appropriate signal in (1).

As discussed in Lemma 2, $G(x)$ is invertible for all $x \in \mathbb{B}_{know}$. Thus, the control input u given by $u(t) = (G(x(t)))^{-1}(a + g(\|x(t)\|)\overline{u(t)} - f(x(t)))$ leads to the same trajectory in (1) as $\overline{u}$ does in (9). While f and G are not known in general, the algorithm in [24] enables determination of $f(x(t))$ and $G(x(t))$ at time t with an arbitrarily small error. Thus, the appropriate u that takes the state of (1) from x_0 to any desired state in $\overline{\overline{R}}(T, x_0)$ can be determined online.

VI. Numerical Example

In this section, we will consider a rudimentary model of an aircraft exposed to unexpected actuator deterioration and environmental effects on dynamics, and determine a set of states ("diversion airports") that are guaranteed to be reachable by the aircraft. We direct the reader to an extended version [23] of this paper for an additional example showing that, while our current theory does not establish equality between $\overline{\overline{R}}(T, 0)$, the reachable set of (5), and the GRS $R^{\mathcal{G}}(T, 0)$, those three sets may indeed coincide.

We assume that the aircraft dynamics are known to satisfy

$$\dot{x} = f(t) + G(t)u, \quad x(0) = 0, \tag{11}$$

where $x \in \mathbb{R}^2$ and $u \in \mathcal{U} = \mathbb{B}^2(0; 1)$. Function f models the time-dependent drift caused by the environment, while G models the actuator deterioration. We will use $G(t) = 1 - t/10$ for $t \leq 10$ and $G(t) = 0$ for $t > 10$: the vehicle's actuators deteriorate linearly until they entirely stop functioning at time $t = 10$. We choose $f(t) = [0.1(\cos(t) + 1)\ 0]^T$ to represent the wind flowing in a constant direction with varying strength. Values of functions f and G, except for $f(0)$ and $G(0)$, are considered unknown when determining the GRS. While (11) does not accurately represent aircraft dynamics, versions of the underlying model $\dot{x} = u$ have been extensively used in planning [6], [22].

Our interest is in determining the set of all states that can eventually be reached, i.e., the *guaranteed eventual reachability set (GRES)* $\cup_{T \geq 0} R^{\mathcal{G}}(T, 0)$. We assume that the following information is available: $f(0) = [0.2\ 0]^T$, $G(0) = 1$, $L_f = 0.1$, and $L_G = 1/5$. Such information is pessimistic: the maximal deterioration rate of the actuators is allowed to be twice as large as the true rate.

Dynamics (11) only fall within the class described by (1) after appending a state variable that equals time. By adapting Theorem 1 to the framework of (11), we obtain a subset of the guaranteed velocity set at time t equaling $\overline{\mathcal{V}}_t^{\mathcal{G}} = \mathbb{B}^2((0.2, 0); 1 - 0.3t)$, with $t \leq 10/3$. At times $t > 10/3$ the available knowledge does not provide any information about the effect of actuators on the system. As in Section V, we obtain subsequent dynamics

$$\dot{x} = a + (1 - 0.3t)u, \tag{12}$$

where $a = [0.2\ 0]^T$. If $\overline{\overline{R}}(T, 0)$ denotes the reachable set of (12) at time T from $x(0) = 0$, we are interested in determining the set $\overline{\overline{R}}(0)$ defined by $\overline{\overline{R}}(0) = \cup_{T \in [0, 10/3]} \overline{\overline{R}}(T, 0) \subseteq \cup_{T \in [0, 10/3]} R^{\mathcal{G}}(T, 0)$.

By defining $z(t) = x(t) - at$, equation (12) becomes $\dot{z} = (1 - 0.3t)u$. Its reachable set at time T can be easily computed to equal $\mathbb{B}^2(0; T - 3T^2/20)$. Thus, $\overline{\overline{R}}(T, 0) = \mathbb{B}^2(aT; T - 3T^2/20)$. Fig. 2 compares the approximation $\overline{\overline{R}}(0)$ of the GRES with true reachable sets $\cup_{T \in [0, 10/3]} R(T, 0)$, $\cup_{T \in [0, 10]} R(T, 0)$, and $\cup_{T \geq 0} R(T, 0)$. Naturally, true reachable sets are larger than $\overline{\overline{R}}(0)$. Such a property follows from the lack of available accurate information of true dynamics; only the dynamics at time $t = 0$ are known, and the maximal actuator deterioration rate L_G, is larger than the true rate, yielding smaller GR(E)Ss.

VII. Future Work

This paper provides only a preliminary theoretical discussion of computation of the guaranteed reachable set. While we show that the produced set $\overline{\overline{R}}(T, x_0)$ is indeed a subset of the GRS, we provide no other formal guarantees on the quality of such an approximation. The difference between the GRS and $\overline{\overline{R}}(T, x_0)$ depends on two intermediate approximations: (i) the approximation of the GRS by the reachable

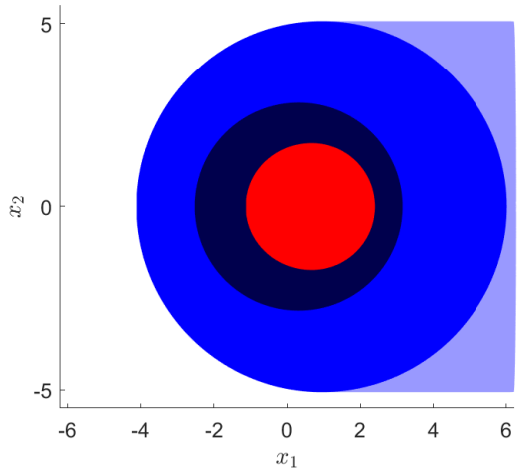

Fig. 2. The underapproximation $\overline{\overline{R}}(0)$ of the GRES is drawn in red. Reachable sets $\cup_{T\in[0,10/3]}R(T,0)$, $\cup_{T\in[0,10]}R(T,0)$, and $\cup_{T\geq 0}R(T,0)$ are drawn in increasingly light shades of blue; set $\cup_{T\geq 0}R(T,0)$ is unbounded as the vehicle can continue "gliding" even after losing all actuation capabilities at time $t = 10$.

set of the ODI $\dot{x} \in \mathcal{V}_x^{\mathcal{G}}$, and (ii) the approximation of sets $\mathcal{V}_x^{\mathcal{G}}$ by balls $\overline{\mathcal{V}}_x^{\mathcal{G}}$. In (i), the example in [23] alludes that the two sets may indeed be equal, potentially under some additional conditions. In (ii), the current paper shows that $\overline{\mathcal{V}}_x^{\mathcal{G}}$ is a maximal ball contained in $\mathcal{V}_x^{\mathcal{G}}$. However, approximating $\mathcal{V}_x^{\mathcal{G}}$ by a geometrical object that more closely fits the complicated shape of $\mathcal{V}_x^{\mathcal{G}}$ would yield a better approximation of the GRS. A possible candidate, potentially simple enough to enable computational work and offering bounds on quality of approximation through John's theorem [15], is a general ellipsoid considered for approximation of related sets in [19]. Another possibility is to consider approximations in norms different from the 2-norm considered by this paper.

A second natural area of future work is in considering a different class of available prior knowledge of system dynamics. While we currently exploit only local dynamics at a single point, the method in [24] yields local dynamics at arbitrarily many points on a single trajectory. Possibly combined with additional prior knowledge about system dynamics (e.g., bounds on higher partial derivatives of f and G), exploiting such information will result in larger sets of guaranteed velocities compared to current work, and thus in GRSs that are closer underapproximations of the true reachable sets. The primary obstacle to successfully improving the estimates using such additional information is geometric. Namely, with heterogeneous constraints describing all the information available about the system dynamics, set $\{(\hat{f}(x), \hat{G}(x)) \mid (\hat{f}, \hat{G}) \in \mathcal{D}_{con}\}$, crucial in approximating $\mathcal{V}_x^{\mathcal{G}}$, may be difficult to simply describe or underapproximate.

Acknowledgment

The author thanks Ufuk Topcu and Franck Djeumou for discussions on related topics.

Correction

By the author's error, the final published version contains an incorrect illustration of the guaranteed velocity set in Fig. 1. The issue is fixed in the present version. We apologize for this mistake.

References

[1] M. Ahmadi, A. Israel, and U. Topcu, "Safety assessment based on physically-viable data-driven models," in *56th IEEE Conference on Decision and Control*, 2017, pp. 6409–6414.
[2] S. Aloni, *Israeli F-15 Eagle Units in Combat*. Osprey Publishing, 2006.
[3] M. Althoff, O. Stursberg, and M. Buss, "Reachability analysis of nonlinear systems with uncertain parameters using conservative linearization," in *47th IEEE Conference on Decision and Control*, 2008, pp. 4042–4048.
[4] A. Barvinok, *A Course in Convexity*. American Mathematical Society, 2002.
[5] A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications*. Springer, 2003.
[6] D. J. Bertsimas and G. van Ryzin, "A stochastic and dynamic vehicle routing problem in the Euclidean plane," *Operations Research*, vol. 39, no. 4, pp. 601–615, 1991.
[7] A. Bressan and B. Piccoli, *Introduction to the Mathematical Theory of Control*. Springer, 2007.
[8] R. W. Brockett, "Nonlinear systems and differential geometry," *Proceedings of the IEEE*, vol. 64, no. 1, pp. 61–72, 1976.
[9] S. L. Brunton, J. L. Proctor, and J. N. Kutz, "Discovering governing equations from data by sparse identification of nonlinear dynamical systems," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 113, no. 15, pp. 3932–3937, 2016.
[10] M. Chen, "High dimensional reachability analysis: Addressing the curse of dimensionality in formal verification," Ph.D. dissertation, University of California, Berkeley, 2017.
[11] Y. Chen, H. Peng, J. Grizzle, and N. Ozay, "Data-driven computation of minimal robust control invariant set," in *57th IEEE Conference on Decision and Control*, 2018, pp. 4052–4058.
[12] M. G. Crandall and P.-L. Lions, "Viscosity solutions of Hamilton-Jacobi equations," *Transactions of the American Mathematical Society*, vol. 277, no. 1, pp. 1–42, 1983.
[13] G. E. Dullerud and F. Paganini, *A Course in Robust Control Theory*. Springer, 2000.
[14] C. Ekstrand and M. Pandey, "New ETOPS regulations," *Aero*, no. 22, pp. 3–11, 2003.
[15] S. Har-Peled, *Geometric Approximation Algorithms*. American Mathematical Society, 2011.
[16] P. A. Ioannou and J. Sun, *Robust Adaptive Control*. Prentice Hall, 1996.
[17] A. Isidori, *Nonlinear Control Systems*. Springer, 1985.
[18] A. B. Kurzhanski and P. Varaiya, "Ellipsoidal techniques for reachability analysis," in *Hybrid Systems: Computation and Control*, 2000, pp. 202–214.
[19] A. A. Kurzhanskiy and P. Varaiya, "Ellipsoidal toolbox," University of California at Berkeley, Tech. Rep. UCB/EECS-2006-46, 2006.
[20] I. M. Mitchell, A. M. Bayen, and C. J. Tomlin, "A time-dependent Hamilton-Jacobi formulation of reachable sets for continuous dynamic games," *IEEE Transactions on Automatic Control*, vol. 50, no. 7, pp. 947–957, 2005.
[21] I. M. Mitchell and C. J. Tomlin, "Overapproximating reachable sets by Hamilton-Jacobi projections," *Journal of Scientific Computing*, vol. 19, no. 1–3, 2003.
[22] K.-K. Oh and H.-S. Ahn, "Distance-based undirected formations of single-integrator and double-integrator modeled agents in n-dimensional space," *International Journal of Robust and Nonlinear Control*, vol. 24, no. 12, pp. 1809–1820, 2014.
[23] M. Ornik, "Guaranteed reachability for systems with unknown dynamics," *arXiv:1910.00803 [math.OC]*, 2019.
[24] M. Ornik, S. Carr, A. Israel, and U. Topcu, "Control-oriented learning on the fly," *IEEE Transactions on Automatic Control*, 2019.
[25] E. D. Sontag, "Controllability is harder to decide than accessibility," *SIAM Journal on Control and Optimization*, vol. 26, no. 5, pp. 1106–1118, 1988.

[26] G. W. Stewart, “Perturbation theory for the singular value decomposition,” University of Maryland Institute for Advanced Computer Studies, Tech. Rep. UMIACS-TR-90-124, 1990.