Resilient Reachability for Linear Systems ⋆

Jean-Baptiste Bouvier * **Melkior Ornik** **

* *Dept. of Aerospace Engineering (e-mail: bouvier3@illinois.edu)*
** *Dept. of Aerospace Engineering & Coordinated Science Laboratory, (e-mail: mornik@illinois.edu)*
University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA

Abstract: A fault-tolerant system is able to reach its goal even when some of its components are malfunctioning. This paper examines tolerance to a specific type of malfunction: the loss of control authority over actuators. Namely, we investigate whether the desired target set for a linear system remains reachable even after such a loss of authority. We frame this problem as the verification of whether a set is reachable from a certain initial state under any undesirable input, without prior knowledge of future undesirable inputs. Building on previous work on reachability with undesirable inputs, this paper develops a reachability condition for linear systems, and obtains a formula that describes reachability of the goal set for driftless linear systems by computing minimum of a concave-convex objective function. From this formulation we establish two novel sufficient conditions for reachability of the goal set.

Keywords: Linear systems, Fault-tolerant systems, Reachability, Zero drift, Non-convex optimisation

1. INTRODUCTION

Fault-tolerant systems are required to be resilient to malfunctioning actuators performing out of their nominal regimes. Among the possible malfunctions, the most widely studied type is actuator failure, which considers an actuator performing with a reduced amplitude or with a fixed unknown magnitude (Tang et al., 2007; Wang and Wen, 2010). Yet, the situation where an actuator becomes unmanageable and produces undesirable, uncontrolled actuator outputs has been less investigated. Such a situation is referred to as *loss of control authority* over an actuator (Bucić et al., 2018). For instance, a damaged rudder flapping in the wind produces undesirable outputs, but cannot be turned off like a defective engine.

We are interested in the case of a system losing control authority over at least one of its actuators. The desire of this paper is to develop simple verification conditions determining whether the system is still able to reach its initial goal. While computation of a reachable set is a classical problem in control theory (Brockett, 1976; Isidori, 1985) and significant computational work has been performed in order to make finding a solution feasible (see, e.g., Kurzhanski and Varaiya (2000); Girard and Guernic (2008)), classical methods often rely on full knowledge of system state and inputs and cannot be directly applied to the case of loss of control authority.

The most widely studied notion of reachability for systems enduring undesirable inputs is *strong reachability* which tackles the problem of how to reach a goal set with a control input that works for all possible undesirable inputs. This approach relates to the field of robust control, and has been studied by, e.g., Bertsekas and Rhodes (1971), Bertsekas (1972) and Raković et al. (2006). Nonetheless, such straightforward robust methods are conservative and often produce meaningful results only when the amount of undesirable "disturbances" in the system is small. Therefore, discussion of reachability in the face of loss of control authority, where the capabilities of the uncontrolled actuators may equal or exceed the capabilities of the remaining control actuators, calls for a different type of reachability. Namely, we say that a goal is *resiliently reachable* from an initial state if for any undesirable inputs, there exists a control law — possibly dependent on current undesirable inputs, but with no knowledge of future ones — able to drive the system to the target set. While not referring to it as resilient reachability, Marzollo and Pascoletti (1973) studied this setting in parallel with strong reachability and described an algorithm to compute resiliently reachable sets. Delfour and Mitter (1969) transformed the problem of resilient reachability into a minimax formula assessing whether a target set is reachable. While our paper heavily draws from the latter work, the resulting reachability conditions of both these approaches are highly abstract, lack intuition, and are difficult to compute.

This paper aims at extending reachability analysis methods to linear systems with loss of control authority. The contributions of this paper are fourfold. First, we consider the reachability condition of Delfour and Mitter (1969) and develop it into a usable equation describing resilient reachability for linear systems. Second, we tackle the specific case of driftless systems, and derive a computable condition for resilient reachability. Third, we analyze the evolution with time of resilient reachability for driftless systems, and show that the resilient reachability problem can be formulated as a minimax optimization of a concave-

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

convex objective function. Fourth, we establish several sufficient conditions that enable us to avoid solving the developed optimization problem.

The remainder of the paper is organized as follows. Section 2 defines the problem of interest and states the related necessary definitions. Section 3 introduces preliminary results obtained by Delfour and Mitter (1969), upon which we build our theory. In Section 4 we develop a resilient reachability condition for linear systems. Section 5 applies this condition to driftless systems, while Section 6 explores how resilient reachability of a target set evolves with time and establishes a sufficient condition for resilient reachability. Two scenarios illustrating a one-dimensional system and an underwater robot illustrate our theory in Section 7.

Notation: The set of real vectors of dimension n is denoted by $\mathbb{R}^n$. The transpose of a matrix M is written M^T, and positive semi-definiteness is denoted by $M \succcurlyeq 0$. We use $\|\cdot\|_X$ to denote the canonical norm on the space X. For $z = (z_1, ...z_n) \in \mathbb{R}^n$, $\|z\|_{\mathbb{R}^n} = \sqrt{\sum z_i^2}$. The unit sphere in $\mathbb{R}^n$ is $\mathbb{U} = \{x \in \mathbb{R}^n : \|x\| = 1\}$, while $\mathbb{B}_X(x, \varepsilon)$ denotes the ball of center x and radius ε in the space X. We use $\langle\cdot,\cdot\rangle_X$ to denote the inner product on X. The empty set is denoted by $\emptyset$, while quantifiers $\exists$ and $\forall$ denote "exists" and "for all", respectively.

The map f from X into Y is described by $f : \left| \begin{matrix} X \longrightarrow Y \\ x \mapsto f(x) \end{matrix} \right.$. The space of continuous linear maps from X into Y is denoted by $\mathcal{L}(X, Y)$, while $\mathcal{L}_2([0,T],\ \mathbb{R}^m)$ denotes the space of the square integrable functions $u : [0,T] \to \mathbb{R}^m$ for $T > 0$. This space may also be simply denoted as $\mathcal{L}_2$.

For a Banach space X, $X^* = \mathcal{L}(X, \mathbb{R})$ denotes its topological dual space. The dual vector of $x \in X$ is $x^* \in X^*$, and denotes the associated linear form from X to $\mathbb{R}$. For $S \in \mathcal{L}(X, Y)$, $S^* \in \mathcal{L}(Y^*, X^*)$ is the adjoint linear map.

2. PROBLEM STATEMENT

Consider a system governed by the equation $\dot{x} = Ax + D\bar{u}$, where $A \in \mathbb{R}^{n\times n}$ and $D \in \mathbb{R}^{n\times(m+p)}$ are constant matrices. Let $G \subset \mathbb{R}^n$ be the target set ("goal") to be reached by the system. Assume that, during its mission, the system loses authority over p of its $m+p$ actuators. We can then separate the controlled inputs $u \in \mathbb{R}^m$ from the undesirable inputs $w \in \mathbb{R}^p$ by writing $\bar{u} = [u^T\ w^T]^T$ and $D = [B\ C]$, with $B \in \mathbb{R}^{n\times m}$ and $C \in \mathbb{R}^{n\times p}$. The system's dynamics can thus be rewritten as follows:

$$\dot{x}(t) = Ax(t) + Bu(t) + Cw(t), \qquad x(0) = x_0 \in \mathbb{R}^n. \quad (1)$$

The goal of this paper is to find a simple condition that characterizes whether a target set is reachable in a given time for a system undergoing a loss of control authority, regardless of the inputs imposed by the malfunctioning actuators. We formulate the problems of *resilient reachability* of G from an initial state x_0 below.

Problem 1. Let $T \geq 0$. Determine if, for any undesirable inputs $w \in W$, there exists a control law $u(w) \in U$ that ensures that the system state is in G at time T.

Problem 2. Let $T \geq 0$. Determine if, for any undesirable inputs $w \in W$, there exists a control law $u(w) \in U$ that drives the system from x_0 to G before the time T.

We note the possible dependence of u on the undesirable input w. Unlike the concept of strong reachability in classical robust control (Bertsekas, 1972; Raković et al., 2006), the objective in the Problems 1 and 2 is not to a priori design a control signal that would bring the state to the target set for any undesirable inputs, but instead to guarantee that whatever the undesirable inputs are, one can determine a control signal *dependent on the undesirable inputs* to drive the system to its goal. The intuition behind posing such a problem is that the system inputs, even if not desirable, can often be measured. In turn, counteracting undesirable inputs is simpler when these inputs are known and a subsequent controller can thus handle perturbations of a larger magnitude than a standard robust controller.

Let us formally define the sets of the initial states from which the system can be driven to G at or before time T as follows:

$$\begin{aligned} X_0^T &= \{x_0 \in \mathbb{R}^n : \forall w \in W,\ \exists\ u \in U : x(T) \in G\}, \\ X_0^{\leq T} &= \{x_0 \in \mathbb{R}^n : \exists\ t \in [0,T] : x_0 \in X_0^t\}. \end{aligned} \quad (2)$$

We can now define the notion of resilient reachability associated with our problems:

Definition 1. The target set G is *resiliently reachable from* x_0 *in time* t if $x_0 \in X_0^t$.

Definition 2. The target set G is *resiliently reachable from* x_0 *by time* T if $x_0 \in X_0^{\leq T}$.

We remark that this paper is primarily focused on determining whether a target set is resiliently reachable for a particular initial state, not on computing the control signal to take it to the target set. While the developed work can be adapted to determine such a signal, it is both technically and computationally intensive, and we omit it to preserve the narrative of our manuscript.

The technical work of this paper follows the assumptions of Delfour and Mitter (1969) and considers square integrable control and undesirable signals over their time domain $[0,T]$. Namely, if U is the set of admissible control signals and W is the set of undesirable signals, we assume

$$\begin{aligned} U &= \{u \in \mathcal{L}_2([0,T],\ \mathbb{R}^m) : \|u\|_{\mathcal{L}_2} \leq 1\} &&= \mathbb{B}_{\mathcal{L}_2}(0,1), \\ W &= \{w \in \mathcal{L}_2([0,T],\ \mathbb{R}^p) : \|w\|_{\mathcal{L}_2} \leq 1\} &&= \mathbb{B}_{\mathcal{L}_2}(0,1), \\ G &= \{x \in \mathbb{R}^n : \|x - x_{goal}\|_{\mathbb{R}^n} \leq \varepsilon\} &&= \mathbb{B}_{\mathbb{R}^n}(x_{goal}, \varepsilon), \end{aligned}$$

where $x_{goal} \in \mathbb{R}^n$ and $0 \leq \varepsilon < \infty$.

We now proceed to describe prior results that enable our work.

3. PRELIMINARIES

The main result of this section is a resilient reachability condition derived from Delfour and Mitter (1969). This condition will serve as primary foundation to build our theory.

Delfour and Mitter (1969) worked with the abstract system

$$x = s + Su + Rw, \quad (3)$$

where $x \in X_3$ is the state, $u \in X_1$ is the control and $w \in X_2$ is the disturbance. State $s \in X_3$ represents the system's initial state, while maps $S \in \mathcal{L}(X_1, X_3)$ and $R \in \mathcal{L}(X_2, X_3)$ represent the effects of controlled and

undesirable inputs, respectively. For our case of interest, we consider $X_1 = \mathcal{L}_2\big([0,T],\ \mathbb{R}^m\big)$, $X_2 = \mathcal{L}_2\big([0,T],\ \mathbb{R}^p\big)$, and $X_3 = \mathbb{R}^n$.

We first transform (1) into (3) by applying the process described in Section 7 of Delfour and Mitter (1969). We begin with the notation of (3) and define the following continuous linear operators:

$$S : \left| \begin{array}{c} \mathcal{L}_2\big([0,T],\ \mathbb{R}^m\big) \longrightarrow \mathbb{R}^n \\ u \mapsto \displaystyle\int_0^T e^{A(T-\tau)} B u(\tau) d\tau \end{array} \right. ,$$

$$R : \left| \begin{array}{c} \mathcal{L}_2\big([0,T],\ \mathbb{R}^p\big) \longrightarrow \mathbb{R}^n \\ w \mapsto \displaystyle\int_0^T e^{A(T-\tau)} C w(\tau) d\tau \end{array} \right. .$$

By taking $s = e^{AT} x_0 \in \mathbb{R}^n$, the solution of (1) can then be written as

$$x(T) = s + Su + Rw.$$

Before our first result, we need to define norms on dual spaces, and for that we rely on Conway (1990). Considering a Banach space X and its adjoint X^*, the norm of $f \in X^*$ is defined by

$$\|f^*\|_{X^*} = \sup_{\|x\|_X = 1} \big\{|f^*(x)|\big\} = \sup_{\|x\|_X \leq 1} \big\{|f^*(x)|\big\}. \tag{4}$$

We can now state our first result, which will serve as the basis of the work in the next sections.

Proposition 1. G is resiliently reachable from x_0 in time T if and only if

$$\sup_{\|x^*\|_{X_3^*}=1} \Big\{ x^*(s - x_{goal}) - \|S^* x^*\|_{X_1^*} + \|R^* x^*\|_{X_2^*} - \varepsilon \Big\} \leq 0.$$

Proof. Let us start from Corollary 5.8 of Delfour and Mitter (1969), which, while not using the same terminology, states that the goal G is resiliently reachable if and only if

$$\begin{aligned} \sup_{\|x^*\|_{X_3^*}=1} \Big\{ & x^*(s) + \inf_{u \in U} \big(S^* x^*(u)\big) \\ & + \sup_{w \in W} \big(R^* x^*(w)\big) - \sup_{y \in G} \big(x^*(y)\big) \Big\} \leq 0. \end{aligned} \tag{5}$$

Since $S^* x^*$ is linear, and for all $u \in U$ it holds that $-u \in U$, we obtain

$$\inf_{u \in U} \big(S^* x^*(u)\big) = -\sup_{u \in U} \big(|S^* x^*(u)|\big) = -\|S^* x^*\|_{X_1^*}, \tag{6}$$

because $U = \big\{u \in X_1 : \|u\|_{X_1} \leq 1\big\}$. Similarly,

$$\sup_{w \in W} \big(R^* x^*(w)\big) = \|R^* x^*\|_{X_2^*}. \tag{7}$$

For $y \in G = \mathbb{B}_{\mathbb{R}^n}(x_{goal}, \varepsilon)$, we can write $y = x_{goal} + \delta y$ with $\delta y \in \mathbb{B}_{\mathbb{R}^n}(0, \varepsilon)$ and, since x^* is linear,

$$x^*(y) = x^*(x_{goal}) + x^*(\delta y).$$

Thus,

$$\sup_{\|\delta y\| \leq \varepsilon} \big(x^*(\delta y)\big) = \varepsilon \sup_{\|\delta y\| \leq 1} \big(x^*(\delta y)\big) = \varepsilon \underbrace{\|x^*\|_{X_3^*}}_{=\,1} = \varepsilon \tag{8}$$

Thus, by (8),

$$\sup_{y \in G} \big(x^*(y)\big) = x^*(x_{goal}) + \sup_{\|\delta y\| \leq \varepsilon} \big(x^*(\delta y)\big) = x^*(x_{goal}) + \varepsilon.$$

We conclude the proof by plugging in identities (6), (7), and (8) into (5). ■

The reachability condition derived in Proposition 1 is highly abstract due to the dual terms and is impractical to use for devising solutions of our two problems of interest. The following two sections aim to develop more workable conditions.

4. INTEGRAL RESILIENT REACHABILITY CONDITION

We will now work on the simplification of Proposition 1. First, we can use the Riesz representation theorem (see Conway (1990)) to simplify x^*. Indeed, $x^* \in \mathcal{L}\big(\mathbb{R}^n, \mathbb{R}\big)$ is bounded in Proposition 1, because $\|x^*\|_{X_3^*} = 1$, so there exists a unique $h \in \mathbb{R}^n$ such that

$$x^*(\cdot) = \langle h, \cdot \rangle \quad \text{and} \quad \|h\|_{\mathbb{R}^n} = \|x^*\|_{X_3^*} = 1.$$

Thus, the supremum in Proposition 1 is over the unit sphere in $\mathbb{R}^n$, for $h \in \mathbb{U} = \big\{x \in \mathbb{R}^n : \|x\| = 1\big\}$. With $s = e^{AT} x_0$, the first term of the reachability condition from Proposition 1 can be rewritten as

$$x^*(s - x_{goal}) = \langle h, e^{AT} x_0 - x_{goal} \rangle_{\mathbb{R}^n}. \tag{9}$$

Now we can simplify the adjoint maps that appear in Proposition 1. By definition of an adjoint map (Conway, 1990), since $S : \mathcal{L}_2 \longrightarrow \mathbb{R}^n$, its adjoint map is given by

$$S^* : \left| \begin{array}{c} \big(\mathbb{R}^n\big)^* \longrightarrow \big(\mathcal{L}_2\big)^* \\ x^* \mapsto x^* \circ S \end{array} \right. . \tag{10}$$

The commutative diagram representing $S^* x^* = x^* \circ S$ is given in Fig. 1.

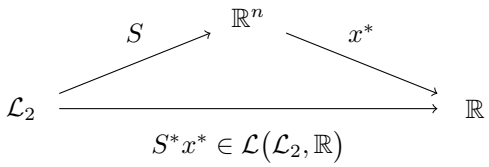

Fig. 1. Commutative diagram of map S and its adjoint S^*.

From (10), for any $u \in \mathcal{L}_2$ we have

$$\begin{aligned} S^* x^*(u) &= \big(x^* \circ S\big)(u) = x^*\big(S(u)\big) = \langle h, S(u) \rangle \\ &= \langle h, \int_0^T e^{A(T-\tau)} B u(\tau) d\tau \rangle. \end{aligned} \tag{11}$$

By (4), the norm on $\mathcal{L}_2^*$ is defined by

$$\|f\|_{\mathcal{L}_2^*} = \sup_{\|u\|_{\mathcal{L}_2}=1} \big\{|f(u)|\big\}, \tag{12}$$

with

$$\|u\|_{\mathcal{L}_2} = \sqrt{\int_0^T \|u(\tau)\|_{\mathbb{R}^m}^2 d\tau}.$$

Putting (11) and (12) together, we obtain

$$\|S^* x^*\|_{\mathcal{L}_2^*} = \sup_{\|u\|_{\mathcal{L}_2}=1} \left\{ \left| \left\langle h, \int_0^T e^{A(T-\tau)} B u(\tau) d\tau \right\rangle \right| \right\}. \tag{13}$$

Similarly,

$$\|R^* x^*\|_{\mathcal{L}_2^*} = \sup_{\|w\|_{\mathcal{L}_2}=1} \left\{ \left| \left\langle h, \int_0^T e^{A(T-\tau)} C w(\tau) d\tau \right\rangle \right| \right\}. \tag{14}$$

We combine (9), (13) and (14) to simplify the resilient reachability condition of Proposition 1.

Theorem 2. G is resiliently reachable from x_0 in time T if and only if the following inequality holds:

$$\max_{h\in\mathbb{U}} \Bigg\{ \langle h, e^{AT}x_0 - x_{goal}\rangle - \sup_{\|u\|_{\mathcal{L}_2}=1} \left\{ \left| \left\langle h, \int_0^T e^{A(T-\tau)}Bu(\tau)d\tau \right\rangle \right| \right\} + \sup_{\|w\|_{\mathcal{L}_2}=1} \left\{ \left| \left\langle h, \int_0^T e^{A(T-\tau)}Cw(\tau)d\tau \right\rangle \right| \right\} - \varepsilon \Bigg\} \leq 0. \tag{15}$$

Proof. After using (9), (13) and (14) in Proposition 1, the only work left is to prove that the supremum from Proposition 1 turns into $\max_{h\in\mathbb{U}}$, which follows from the discussion preceding (9), $\mathbb{U}$ being closed, and the function to maximize being continuous in h. ■

Because it directly uses matrices A, B and C instead of the adjoints of maps derived from those matrices, the condition from Theorem 2 is more direct than the equation (5) we started from. Yet, computing the two supremums on the unit sphere of $\mathcal{L}_2$ is a difficult task because of the infinite dimension of $\mathcal{L}_2$. We now focus on driftless systems where the integrals in (15) can be simplified.

5. DRIFTLESS SYSTEMS

Driftless systems are widely studied in robotics, as their dynamics represent the kinematics constraints of the system: numerous examples are described in Siciliano and Khatlib (2016). For these systems matrix A equals 0, so that (1) becomes

$$\dot{x}(t) = Bu(t) + Cw(t). \tag{16}$$

The removal of A enables us to distill Theorem 2 into a simpler resilient reachability condition.

Theorem 3. $G = \mathbb{B}_{\mathbb{R}^n}(x_{goal}, \varepsilon)$ is resiliently reachable at T from x_0 if and only if

$$\max_{h\in\mathbb{U}} \Big\{ \langle h, x_0 - x_{goal}\rangle - \sqrt{T}\left\|B^T h\right\|_{\mathbb{R}^m} + \sqrt{T}\left\|C^T h\right\|_{\mathbb{R}^p} \Big\} \leq \varepsilon.$$

Proof. When $A = 0$, the leftmost term in (15) clearly equals $\langle h, x_0 - x_{goal}\rangle$. We now simplify the next term by using the Cauchy-Schwarz inequality:

$$\left| \langle h, B\int_0^T u(\tau)d\tau\rangle_{\mathbb{R}^n} \right| = \left| \langle B^T h, \int_0^T u(\tau)d\tau\rangle_{\mathbb{R}^m} \right| \leq \left\|B^T h\right\|_{\mathbb{R}^m} \left\| \int_0^T u(\tau)d\tau \right\|_{\mathbb{R}^m}. \tag{17}$$

The equality case happens when $B^T h$ and $\int_0^T u(\tau)d\tau$ are *positively collinear*, i.e., $\int_0^T u(\tau)d\tau$ is a nonnegative scalar multiple of $B^T h$ (Conway, 1990).

If $(e_1, ..., e_m)$ is the canonical basis of $\mathbb{R}^m$, there exist $u_1, ..., u_m \in \mathcal{L}_2([0,T], \mathbb{R})$ such that $u = \sum_{i=1}^m u_i e_i$. The norm of the integral of u can then be simplified:

$$\begin{aligned} \left\| \int_0^T u(\tau)d\tau \right\|_{\mathbb{R}^m} &= \sqrt{\sum_{i=1}^m \left(\int_0^T u_i(\tau) \times 1 \, d\tau \right)^2} \\ &\leq \sqrt{\sum_{i=1}^m \left(\int_0^T u_i^2(\tau)d\tau \right)\left(\int_0^T 1^2 d\tau \right)} \\ &= \sqrt{T}\sqrt{\int_0^T \sum_{i=1}^m u_i^2(\tau)d\tau} = \sqrt{T}\|u\|_{\mathcal{L}_2}. \end{aligned} \tag{18}$$

In (18), we again use the Cauchy-Schwarz inequality and the equality case happens when each u_i is almost everywhere (in the measure-theoretical sense) collinear with the function $\tau \mapsto 1$, i.e., when u is almost everywhere constant.

By combining (17) and (18), we proved that

$$\sup_{\|u\|_{\mathcal{L}_2}=1} \left\{ \left| \left\langle h, \int_0^T Bu(\tau)d\tau \right\rangle \right| \right\} \leq \|B^T h\|_{\mathbb{R}^m}\sqrt{T}. \tag{19}$$

If we can find a function u_h of unit norm in $\mathcal{L}_2$ for which the inequality in (19) is an equality, then the supremum in (19) would be a maximum. The function u_h must realize both equality cases of the Cauchy-Schwarz inequality used previously. Hence, for $h \in \mathbb{U}$ we define the following constant function:

$$u_h : \left| \begin{array}{c} [0,T] \longrightarrow \mathbb{R}^m \\ t \mapsto \dfrac{B^T h}{\sqrt{T}\|B^T h\|_{\mathbb{R}^m}} \end{array} \right. .$$

We note that $\|u_h(t)\|_{\mathbb{R}^m} = \frac{1}{\sqrt{T}}$ for all t. Thus, u_h is of unit norm on $\mathcal{L}_2$:

$$\|u_h\|_{\mathcal{L}_2} = \sqrt{\int_0^T \|u_h(t)\|_{\mathbb{R}^m}^2 dt} = \sqrt{\int_0^T \frac{1}{T}dt} = 1.$$

Moreover, u_h is positively collinear with $B^T h$ and is constant over time, therefore it satisfies both of the Cauchy-Schwarz equality cases in (17) and (18), which leads to

$$\begin{aligned} \left| \left\langle h, \int_0^T Bu_h(\tau)d\tau \right\rangle \right| &= \left\|B^T h\right\|_{\mathbb{R}^m} \left\| \int_0^T u_h(\tau)d\tau \right\|_{\mathbb{R}^m} \\ &= \left\|B^T h\right\|_{\mathbb{R}^m} \sqrt{T}. \end{aligned} \tag{20}$$

From (19) and (20), we clearly obtain

$$\max_{\|u\|_{\mathcal{L}_2}=1} \left\{ \left| \left\langle h, \int_0^T Bu(\tau)d\tau \right\rangle \right| \right\} = \sqrt{T}\left\|B^T h\right\|_{\mathbb{R}^m}, \tag{21}$$

The same process can be applied to the final term in (15), yielding the theorem claim. ■

The reader desiring intuition on Theorem 3 should recall that $\mathbb{U}$ is the unit sphere in $\mathbb{R}^n$, so the maximum over $\mathbb{U}$ explores every direction for h. The scalar product $\langle h, x_0 - x_{goal}\rangle$ gives the intuition that h represents a direction of the system's travel in the state space. The h maximizing this scalar product is positively collinear with $x_0 - x_{goal}$, so it is driving the system away from

x_{goal}. On the other hand, the terms $B^T h$ and $C^T h$ represent how the controls and the undesirable inputs drive the system when they are both along the direction h. Hence, the h that maximizes $\|C^T h\| - \|B^T h\|$ is the direction giving the most strength to the undesirable inputs over the controls. Therefore, the h that realizes the overall maximum represents the worst direction for resilient reachability.

We can strengthen our faith in Theorem 3 by looking at a few special cases. If we assume that $x_0 = x_{goal}$, then G is reachable at time $T = 0$ as, for all $h \in \mathbb{U}$, $\langle h, x_0 - x_{goal} \rangle = 0$. Another simple case is when $B = 0$ and $C = 0$, so $f(T) = \|x_0 - x_{goal}\|$, which represents the distance the system has to travel to reach the center of the target set. In this case, $\dot{x} = 0$, so for all t, $x(t) = x_0$ and the reachability condition becomes as expected $\|x_0 - x_{goal}\| \leq \varepsilon$, i.e., equivalent to $x_0 \in G$.

Theorem 3 gives a condition on resilient reachability at time T. We now have all the tools to start working on Problem 2, and study how the resilient reachability of G evolves with time.

6. EVOLUTION OF REACHABILITY WITH TIME

To simplify the notation of Theorem 3, let us first write $d = x_0 - x_{goal}$ and define functions

$$J : \left| \begin{array}{ll} \mathbb{U} \times \mathbb{R}^+ & \longrightarrow \mathbb{R} \\ (h,\ t) & \mapsto \langle h, d \rangle + \sqrt{t}\big(\|C^T h\| - \|B^T h\| \big) \end{array} \right.$$

and $f : \left| \begin{array}{ll} \mathbb{R}^+ & \longrightarrow \mathbb{R} \\ t & \mapsto \max\limits_{h \in \mathbb{U}} \big\{ J(h,t) \big\} \end{array} \right.$. Thus, the condition of Theorem 3 is equivalent to $f(T) \leq \varepsilon$.

For a given d, the inner product $\langle h, d \rangle$ in J is bounded for $h \in \mathbb{U}$, so, as time T becomes sufficiently large, resilient reachability of G largely depends on the sign of the coefficient of $\sqrt{T}$. If there are controls ($B \neq 0$), but no undesirable inputs ($C = 0$), the time coefficient is $-\|B^T h\| < 0$, so J decreases with time, which intuitively means that the reachable set grows. The opposite also happens as expected when $B = 0$ and $C \neq 0$. We now wish to analyze resilient reachability of G over time for general driftless systems.

Note first that for $t > 0$, $J(\cdot, t)$ is not a concave function, and thus its maximization over $\mathbb{U}$ is not an easy task. Indeed, both functions $h \mapsto \|C^T h\|$ and $h \mapsto \|B^T h\|$ are convex, so $J(\cdot, t)$ is the difference between two convex functions. This type of maximization is referred to as a *difference of convex* (DC) problem, and analytical solutions are only available for a few special cases. Numerous algorithms have been developed by, e.g. Tuy (1987) and Tao and An (1997). In particular, the simple algorithm devised by Yuille and Rangarajan (2003) to minimize a function composed of a concave and a convex part has been of great interest and is called the *concave-convex procedure*. While these numerical results, combined with Theorem 3, enable us to determine whether set G is resiliently reachable *at* every given time, they do not enable us to directly gain insight regarding reachability *by* a certain time.

In order to discuss reachability by a certain time, we apply Theorem 3 to note that G is resiliently reachable from x_0 by time T if and only if

$$\min_{t \in [0,T]} \Big\{ \max_{h \in \mathbb{U}} \big\{ J(h,t) \big\} \Big\} \leq \varepsilon.$$

Hence the reachability by time T can be described as a minimax problem with a DC cost function. We will omit the discussion of possible numerical solutions to such a problem and instead focus on analytical results.

Let us define the function

$$g : \left| \begin{array}{c} \mathbb{U} \longrightarrow \mathbb{R} \\ h \mapsto \|C^T h\| - \|B^T h\| \end{array} \right. ,$$

so that $J(h,t) = h^T d + g(h)\sqrt{t}$. For a given goal and initial state, $\|h^T d\|$ is bounded. So, as time grows, $\sqrt{t}$ becomes the leading term in J, with its sign determined by $g(h)$. We therefore study the sign of $\max\limits_{h \in \mathbb{U}} \big\{ g(h) \big\}$. We will show the following:

- if $\max\limits_{h \in \mathbb{U}} \big\{ g(h) \big\} > 0$, G is only resiliently reachable up to a certain time,
- if $\max\limits_{h \in \mathbb{U}} \big\{ g(h) \big\} = 0$, G can be either always resiliently reachable, never resiliently reachable, or its resilient reachability depends on time,
- if $\max\limits_{h \in \mathbb{U}} \big\{ g(h) \big\} < 0$, G is resiliently reachable from some time onwards.

We prove these claims in the following three subsections.

6.1 Maximum of g is positive

When $\max \big\{ g(h) \big\} > 0$, there exists a h such that $\|C^T h\| > \|B^T h\|$, i.e., in line with our intuition, there is an input direction where the matrix C produces a stronger undesirable input than what the control matrix B is capable of counteracting. Since we want to guarantee reaching the goal for *any* undesirable input, a single direction where the undesirable inputs are stronger than the controlled ones is sufficient to prevent reachability. We formalize this intuition as follows.

Theorem 4. Let $x_0 \in \mathbb{R}^n$. If $\max\limits_{h \in \mathbb{U}} \big\{ g(h) \big\} > 0$, then there exists $t_{lim} > 0$ such that for all $t \geq t_{lim}$, $x_0 \notin X_0^t$.

Proof. We use the notation as given above. Because $\max\limits_{h \in \mathbb{U}} \big\{ g(h) \big\} > 0$, there is a $h_+ \in \mathbb{U}$ such that $g(h_+) > 0$. Then $f(t) \geq \langle h_+, d \rangle + g(h_+)\sqrt{t} \xrightarrow[t \to \infty]{} +\infty$.

So, $\lim\limits_{t \to \infty} f(t) = +\infty$. Then, there exists $t_{lim} > 0$ such that for $t \geq T$, $f(t) > \varepsilon$, so G is not reachable at time t from x_0. Thus $x_0 \notin X_0^t$. ■

Theorem 4 states that, for a fixed initial state x_0 and a goal G, there exists a time T after which the target set is not resiliently reachable anymore. Thus, all resilient reachability can only happen in finite time.

6.2 Maximum of g equals zero

In this subsection, we assume that $\max \big\{ g(h) \big\} = 0$. Thus, there is at least one $h \in \mathbb{U}$ such that $g(h) = 0$.

Intuitively, in this direction h the undesirable inputs match the controls. But, in directions where g is negative, the controls have a greater magnitude than the undesirable inputs. Thus, overall the controls can at least compensate the effects of the undesirable inputs.

Since undesirable inputs can match the controls in certain directions, the resiliently reachable region does not expand in every direction with time. Thus, the resilient reachability of G depends on its location.

Let us define $H_0 = \{h \in \mathbb{U} : g(h) = 0\}$. The set H_0 is closed, bounded, and nonempty by the assumption of $\max\{g(h)\} = 0$. So with $d = x_0 - x_{goal}$, we can define

$$h_0 = \arg\max_{h \in H_0} \{h^T d\}.$$

We note that vector h_0 need not be uniquely defined. The theorem below holds for every h_0.

Theorem 5. Assume $\max_{h\in\mathbb{U}}\{g(h)\} = 0$. If $\varepsilon \geq \|d\|$, then $x_0 \in X_0^t$ for all $t \geq 0$, and if $\varepsilon < h_0^T d$, then $x_0 \notin X_0^t$ for all $t \geq 0$.

Proof. We note that $\max_{h\in\mathbb{U}}\{h^T d\} = f(0) = \|d\|$. Thus,

$$f(t) \leq \max_{h\in\mathbb{U}}\{h^T d\} + \max_{h\in\mathbb{U}}\{g(h)\sqrt{t}\} = \|d\| + 0 = \|d\|,$$

so $\max_{t\geq 0}\{f(t)\} = \|d\|$.

Additionally, $h_0 \in \mathbb{U}$, so $f(t) \geq h_0^T d + g(h_0)\sqrt{t} = h_0^T d$. Thus, $h_0^T d \leq f(t) \leq \|d\|$ for all $t \geq 0$.

If $\varepsilon \geq \|d\|$, then for $t \geq 0$, $f(t) \leq \varepsilon$, i.e., by Theorem 3, $x_0 \in X_0^t$. On the other hand, if $\varepsilon < h_0^T d$, then for $t \geq 0$, $f(t) > \varepsilon$, i.e., $x_0 \notin X_0^t$. ■

So, if $\varepsilon \geq \|d\|$, G is resiliently reachable from the start and remains always resiliently reachable, while if $\varepsilon < h_0^T d$, G is never resiliently reachable. There is obviously an intermediate case for $\varepsilon \in [h_0^T d, \|d\|]$ where the resilient reachability of G depends on time.

6.3 Maximum of g is negative

We can now tackle the third case, where $\max\{g(h)\} < 0$. In this situation, our intuition stipulates that controls are stronger than the undesirable inputs in every direction, so the reachable set grows unbounded with time. The theorem below confirms this intuition.

Theorem 6. If $\max_{h\in\mathbb{U}}\{g(h)\} < 0$, then there exists $t_{lim} \geq 0$ such that $x_0 \in X_0^t$ for all $t \geq t_{lim}$.

Proof.

Let $\max_{h\in\mathbb{U}}\{g(h)\} = -\gamma < 0$. Then f can be bounded by above:

$$\begin{aligned} f(t) &= \max_{h\in\mathbb{U}}\{h^T d + g(h)\sqrt{t}\} \\ &\leq \max_{h\in\mathbb{U}}\{h^T d\} + \max_{h\in\mathbb{U}}\{g(h)\}\sqrt{t} \\ &= \|d\| - \gamma\sqrt{t}. \end{aligned}$$

We compare this upper bound with ε to obtain a reachability condition

$$\|d\| - \gamma\sqrt{t} \leq \varepsilon \iff \left(\frac{\|d\| - \varepsilon}{\gamma}\right)^2 \leq t.$$

We can define a time

$$t_{lim} = \left(\frac{\|d\| - \varepsilon}{\gamma}\right)^2 = \left(\frac{\|d\| - \varepsilon}{\max_{h\in\mathbb{U}}\{g(h)\}}\right)^2 \qquad (22)$$

such that, for all $t \geq t_{lim}$, $f(t) \leq \varepsilon$, i.e., $x_0 \in X_0^t$. ■

The t_{lim} defined in (22) might not be the minimal time for resilient reachability since a first inequality has been used in order to decouple $h^T d$ and $g(h)\sqrt{t}$. Nonetheless, Theorem 6 proves that, after some time, any target set becomes resiliently reachable.

Theorems 4, 5 and 6 show that the sign of the maximum of g leads to interesting conclusions. It is thus natural to attempt to analytically determine an upper bound for g.

6.4 Bounding g

Let $\sigma_{max}^{C^T}$ be the maximal singular value of C^T, and $\sigma_{min}^{B^T}$ be the minimal singular value of B^T. We claim that the relationship between these two values impacts the maximal value of g.

Theorem 7. If $\sigma_{max}^{C^T} < \sigma_{min}^{B^T}$, then $\max_{h\in\mathbb{U}}\{g(h)\} < 0$.

Proof. Let us define $M = CC^T \succcurlyeq 0$. Matrix M is symmetric, and we can use the following classical inequality (Horn and Johnson, 2012):

$$\lambda_{min}^M \|x\|^2 \leq x^T M x \leq \lambda_{max}^M \|x\|^2, \qquad \forall x \in \mathbb{R}^n,$$

with λ_{min}^M and λ_{max}^M respectively, the minimum and maximum eigenvalues of M. Since M is trivially positive semidefinite, $\lambda_{min}^M \geq 0$. Note that $\|C^T h\| = \sqrt{h^T C C^T h} = \sqrt{h^T M h}$. Thus we obtain

$$\sqrt{\lambda_{min}^M} \leq \|C^T h\| \leq \sqrt{\lambda_{max}^M} = \sigma_{max}^{C^T}, \qquad \forall h \in \mathbb{U}.$$

By doing the same for B^T, g can be bounded as follows:

$$\sqrt{\lambda_{min}^{C^T}} - \sqrt{\lambda_{max}^{B^T}} \leq g(h) \leq \sigma_{max}^{C^T} - \sigma_{min}^{B^T}, \qquad \forall h \in \mathbb{U}.$$

So if $\sigma_{max}^{C^T} < \sigma_{min}^{B^T}$, then $\max_{h\in\mathbb{U}}\{g(h)\} < 0$. ■

Theorems 6 and 7 trivially imply the following corollary.

Corollary 8. If all singular values of C^T are strictly smaller than those of B^T, then the target set is resiliently reachable in finite time.

We now proceed to computationally confirm the above theoretical results.

7. NUMERICAL EXAMPLES

7.1 1D system

To illustrate Theorem 2, we consider the following one-dimensional system:

$$\begin{aligned} &\dot{x} = x + bu + cw, \qquad b \geq 0, \qquad c \geq 0, \\ &x_0 = 0, \qquad x_{goal} = 1, \\ &U = W = \{v \in \mathcal{L}_2([0,T],\ \mathbb{R}) : \|v\|_{\mathcal{L}_2} \leq 1\}. \end{aligned}$$

For the above system we will calculate the resilient reachability condition given by Theorem 2, and verify whether the goal can actually be reached in that case.

First, supremums from (13) and (14) can be simplified, since the inner product on $\mathbb{R}$ is just multiplication:

$$\begin{aligned}\left|\left\langle h, \int_0^T e^{T-t} b u(t) dt \right\rangle\right| &= \underbrace{|h|}_{=1} \cdot b e^T \left| \int_0^T e^{-t} u(t) dt \right| \\ &\leq b e^T \sqrt{\int_0^T e^{-2t} dt} \sqrt{\int_0^T u^2 dt} \\ &= b e^T \sqrt{\frac{1 - e^{-2T}}{2}} \times \|u\|_{\mathcal{L}_2}.\end{aligned}$$

We use the Cauchy-Schwarz inequality above, and note that the equality case is realized for $u(t) = e^{-t}$. We proceed the same way as we did to transform (19) into (21) and obtain

$$\sup_{\|u\|_{\mathcal{L}_2}=1} \left\{ \left| \left\langle h, \int_0^T e^{T-t} b u(t) dt \right\rangle \right| \right\} = b \sqrt{\frac{e^{2T} - 1}{2}}.$$

The same process is applied to (14). The last term left to calculate in Theorem 2 is the scalar product from (9):

$$\langle h, e^t x_0 - x_{goal} \rangle = -h.$$

Theorem 2 can then be simplified to

$$1 + \sqrt{\frac{e^{2T} - 1}{2}} (c - b) \leq \varepsilon. \tag{23}$$

Defining $m(\varepsilon, T) = (1 - \varepsilon)\sqrt{2}/(\sqrt{e^{2T} - 1})$, (23) is equivalent to

$$c + m(\varepsilon, T) \leq b. \tag{24}$$

Intuitively, if (24) holds, then the control magnitude b is larger than the magnitude c of undesirable inputs, plus the minimum margin $m(\varepsilon, T)$ required to reach the goal at time T. The more time is allowed, the less margin is necessary: $m(\varepsilon, T)$ decreases with T.

We can now test condition (24), i.e., Theorem 2, by taking $w = \frac{-1}{\sqrt{T}}$ and $u = \frac{1}{\sqrt{T}}$. These inputs were chosen to be of unit norm in $\mathcal{L}_2$, with w pulling the state away from its goal and u counteracting w. We can then calculate the final state:

$$x(T) = \int_0^T e^{T-t} \left(b \frac{1}{\sqrt{T}} + c \frac{-1}{\sqrt{T}} \right) dt = (b - c) \frac{(e^T - 1)}{\sqrt{T}}.$$

The closest point of the target set from the initial state is $x_{goal} - \varepsilon = 1 - \varepsilon$. Thus, goal G is reached at T *for the particular control input* u *and the particular uncontrolled signal* w, if $x(T) \geq 1 - \varepsilon$, i.e., $b - c \geq (1 - \varepsilon)\sqrt{T}/(e^T - 1)$.

Since

$$\frac{(1 - \varepsilon)\sqrt{T}}{e^T - 1} \geq m(\varepsilon, T)$$

for all $T \geq 0$, the above result validates Theorem 2: if G is reached above (in which case G is certainly resiliently reachable), (24) will hold.

We now proceed to illustrate the developed theory for driftless systems.

7.2 A driftless underwater vehicle

We consider an underwater robot propelled by a main engine and two side engines for operations in a 2D plane, as shown in Fig. 7.2. We consider the scenario where the main engine u_1 has a small bias in the y direction. The system dynamics are thus given by

$$\dot{X} = \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} 10 & 1 & 1 \\ 0.2 & -1 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}.$$

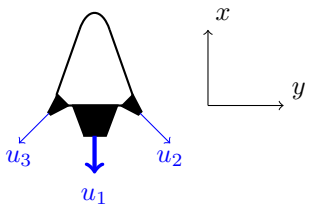

Fig. 2. A model of an underwater robot with three actuators.

Our example is motivated by the work of Vela et al. (2002) and Yu et al. (2016), which also have also considered driftless dynamics. The assumption of driftlessness can intuitively be justified by the viscosity of the water combined with a small speed of the robot.

We assume that during its mission the controller loses authority over the third actuator. The terms in (16) can thus be written as follows:

$$u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \quad w = u_3, \quad B = \begin{bmatrix} 10 & 1 \\ 0.2 & -1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Intuitively, the robot should still be able to reach any goal set, since the second actuator u_2 can counteract the undesirable inputs of u_3, and the small bias of u_1 on y provides a net motion on y, while the desired displacement along x is also realized by the main engine. While the conditions of Theorem 7 are not satisfied, we can compute $\max_{h \in \mathbb{U}} \{g(h)\} = -0.02$, and use Theorem 6 to show that any goal is eventually resiliently reachable, as suggested by our intuition.

In the situation where the controller loses authority over both the second and third actuators, our intuition suggests that a controlled motion along x is still possible, but the displacements along y cannot be controlled. Therefore, we cannot guarantee to reach any target position. We numerically compute g and find that $\max_{h \in \mathbb{U}} \{g(h)\} = 1.4 > 0$. The conclusion of Theorem 4 validates our intuition.

If the controller only loses authority over the first actuator, then $\max_{h \in \mathbb{U}} \{g(h)\} = 8.6 > 0$. Of course none of the side engines can make up for the loss of the main one, as predicted by Theorem 4.

Another interesting case to note is when u_1 thrusts only along x without bias on y, i.e.,

$$\dot{X} = \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} 10 & 1 & 1 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}.$$

Then, a loss of control authority over one of the side engines results in $\max_{h \in \mathbb{U}} \{g(h)\} = 0.02 > 0$. Indeed, we cannot guarantee to reach a goal that is not on the x axis, because no net motion on y is guaranteed, since both side engines can cancel each other out.

8. CONCLUSION

This paper described the problem of resilient reachability: deciding whether a system can always be driven to a desired goal, given that some of its actuators act in an undesirable manner and without prior knowledge of these undesirable inputs. To solve this problem, we derived a resilient reachability condition for linear systems and a more specific condition for driftless systems. We investigated the evolution of resilient reachability with time and rewrote the problem as a minimax optimization with a concave-convex objective function. We then derived results that do not require directly solving the optimization problem, at the price of providing sufficient or necessary conditions.

We foresee that immediate future work will build a more advanced analysis of the time-evolution of the reachability condition using the envelope theorems by Milgrom and Segal (2002). This manuscript, however, presents only the first step in our long-term goal of *resilient system synthesis*, i.e., design of actuator functionality (in the context of this paper represented by system matrices) for which the system retains resilient reachability to loss of one or more actuators.

REFERENCES

Bertsekas, D. (1972). Infinite-time reachability of state-space regions by using feedback control. *IEEE Transactions on Automatic Control*, 17(5), 604 – 612.

Bertsekas, D. and Rhodes, I. (1971). On the minimax reachability of target sets and target tubes. *Automatica*, 7, 233 – 247.

Brockett, R.W. (1976). Nonlinear systems and differential geometry. *Proceedings of the IEEE*, 64(1), 61–72.

Bucić, M., Ornik, M., and Topcu, U. (2018). Graph-based controller synthesis for safety-constrained, resilient systems. In *56th Annual Allerton Conference on Communication, Control, and Computing*, 297 – 304.

Conway, J.B. (1990). *A Course in Functional Analysis*. Springer.

Delfour, M.C. and Mitter, S.K. (1969). Reachability of perturbed systems and min sup problems. *SIAM Journal on Control and Optimization*, 7(4), 521 – 533.

Girard, A. and Guernic, C.L. (2008). Efficient reachability analysis for linear systems using support functions. In *17th IFAC World Congress*, 8966 – 8971.

Horn, R.A. and Johnson, C.R. (2012). *Matrix Analysis*. Cambridge University Press.

Isidori, A. (1985). *Nonlinear Control Systems*. Springer.

Kurzhanski, A.B. and Varaiya, P. (2000). Ellipsoidal tehniques for reachability analysis. In *Hybrid Systems: Computation and Control*, 202 – 214.

Marzollo, A. and Pascoletti, A. (1973). On the reachability of a given set under disturbances. *Control and Cybernetics*, 2(3), 99 – 106.

Milgrom, P. and Segal, I. (2002). Envelope theorems for arbitrary choice sets. *Econometrica*, 70(2), 583–601.

Raković, S., Kerrigan, E., Mayne, D., and Lygeros, J. (2006). Reachability analysis of discrete-time systems with disturbances. *IEEE Transactions on Automatic Control*, 51(4), 546 – 561.

Siciliano, B. and Khatlib, O. (2016). *Springer Handbook of Robotics*. Springer.

Tang, X., Tao, G., and Joshi, S.M. (2007). Adaptive actuator failure compensation for nonlinear MIMO systems with an aircraft control application. *Automatica*, 43, 1869 – 1883.

Tao, P.D. and An, L.T.H. (1997). Convex analysis approach to d.c. programming: Theory, algorithms and applications. *Acta Mathematica Vietnamica*, 22(1), 289 – 355.

Tuy, H. (1987). Global minimization of a difference of two convex functions. *Mathematical Programming Study*, 30, 150 – 182.

Vela, P.A., Morgansent, K.A., and Burdick, J.W. (2002). Underwater locomotion from oscillatory shape deformations. In *41st IEEE Conference on Decision and Control*, volume 2, 2074 – 2080.

Wang, W. and Wen, C. (2010). Adaptive actuator failure compensation control of uncertain nonlinear systems with guaranteed transient performance. *Automatica*, 46, 2082 – 2091.

Yu, J., Wang, C., and Xie, G. (2016). Coordination of multiple robotic fish with applications to underwater robot competition. *IEEE Transactions on Industrial Electronics*, 63(2), 1280 – 1288.

Yuille, A.L. and Rangarajan, A. (2003). The concave-convex procedure. *Neural Computation*, 15(4), 915 – 936.