Resilience of Linear Systems to Partial Loss of Control Authority ⋆

Jean-Baptiste Bouvier [a], Melkior Ornik [a]

[a] *Department of Aerospace Engineering and Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.*

Abstract

After a loss of control authority over thrusters of the Nauka module, the International Space Station lost attitude control for 45 minutes with potentially disastrous consequences. Motivated by this scenario, we investigate the continued capability of control systems to perform their task despite partial loss of authority over their actuators. We say that a system is resilient to such a malfunction if for any undesirable inputs and any target state there exists an admissible control driving the state to the target. Building on controllability conditions and differential games theory, we establish a necessary and sufficient condition for the resilience of linear systems. As their task might be time-constrained, ensuring completion alone is not sufficient. We also want to estimate how much slower the malfunctioning system is compared to its nominal performance. Relying on Lyapunov theory we derive analytical bounds on the reach times of the nominal and malfunctioning systems in order to quantify their resilience. We illustrate our work on the ADMIRE fighter jet model and on a temperature control system.

Key words: fault-tolerant, resilience, linear systems, reachability, stabilizability, controllability, time-invariant

1 Introduction

After the Nauka module docked to the International Space Station (ISS), a software failure caused a misfire of the module's thrusters, leading to a loss of attitude control of the whole station for 45 minutes [7]. Eventually, other thrusters on the ISS were fired to counteract the uncontrolled and undesirable thrust until the Nauka module ran out of fuel. Motivated by such events, [9] introduced the notion of a *partial loss of control authority over actuators* where some of the actuators of a system start producing uncontrolled and thus possibly undesirable inputs within their full range of actuation. To identify these faulty actuators, we assume sensors monitor each actuator in real time [16]. Our first objective is then one of *resilient reachability*, i.e, verifying whether for all possible outputs of the malfunctioning actuators, the controlled ones can steer the system to its target [9]. Our second objective is to estimate the maximal time penalty caused by such a malfunction.

Classically, changing or unknown dynamics are studied through robust, adaptive, and fault-tolerant control theories. However, robust control needs the undesirable inputs to be significantly smaller than the controls [36]. Since the loss of control authority over actuators may produce large undesirable inputs, robust control performs poorly [10]. In turn, adaptive control tries to estimate unknown parameters before they have time to change significantly [5], which may not be possible for uncontrolled inputs. Such a situation would typically prevent convergence of the estimators and lead to mediocre adaptive control performance [36]. As for fault-tolerant theory, actuator failure investigations are usually limited either to actuators "locking in place" and producing constant inputs [34] or to actuators with reduced effectiveness [4, 37]. Since uncontrolled actuators can still produce a full range of inputs, loss of control authority over actuators is not covered by existing fault-tolerant theory [4].

On the other hand, loss of control authority falls within the framework of differential games because the malfunctioning actuators can be modeled by adversaries as in [23,33]. However, these works do not constitute appropriate starting points for a resilient reachability study due to the unbounded inputs of [23] and the complexity

⋆ Corresponding author J.-B. Bouvier. This work was supported by an Early Stage Innovations grant from NASA's Space Technology Research Grants Program, grant no. 80NSSC19K0209. This material is partially based upon work supported by the United States Air Force AFRL/SBRK under contract no. FA864921P0123.

Email addresses: bouvier3@illinois.edu (Jean-Baptiste Bouvier), mornik@illinois.edu (Melkior Ornik).

of the theory of [33].

Concerning our second objective, *quantitative resilience* was introduced in [13, 14] as the maximal ratio of the minimal reach times for the nominal and malfunctioning systems. However, the exact calculation of quantitative resilience for systems with driftless dynamics [11] does not extend to general linear systems since the minimal reach time in such systems does not have an analytical expression [6].

The main contributions of this work are fourfold. Firstly, relying on the differential games theory of Hájek [21] and the controllability conditions of Brammer [15], we establish simple necessary and sufficient conditions to verify the resilient stabilizability of linear systems, i.e., whether the origin is resiliently reachable from any initial state. Secondly, we extend Hájek's duality theorem in order to study the resilient reachability of affine targets. Thirdly, we use zonotopic underapproximations of reachable sets [2, 19] to determine what states are guaranteed to be resiliently reachable. Finally, we employ Lyapunov theory [25] to establish analytical bounds on the quantitative resilience of linear systems.

The remainder of this work is organized as follows. Section 2 introduces the system dynamics and the problems of interest. Section 3 provides background results. Section 4 establishes necessary and sufficient conditions for resilient stabilizability of linear systems. Section 5 extends these conditions to affine targets and describes zonotopic underapproximations of the resiliently reachable set of linear systems. Section 6 derives analytical bounds on the quantitative resilience of linear systems. Section 7 illustrates our theory on a fighter jet model and a temperature control system.

Notation: We denote the integer interval from a to b, inclusive, with $[\![a, b]\!]$. For a set $\Lambda \subseteq \mathbb{C}$, we say that $Re(\Lambda) \leq 0$ (resp. $Re(\Lambda) = 0$) if the real part of each $\lambda \in \Lambda$ verifies $Re(\lambda) \leq 0$ (resp. $Re(\lambda) = 0$). The norm of a matrix A is $\|A\| := \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \max_{\|x\|=1} \|Ax\|$ and the set of its eigenvalues is $\lambda(A)$. If A is positive definite, denoted $A \succ 0$, then its extremal eigenvalues are λ^A_{min} and λ^A_{max}, and A generates a vector norm $\|x\|_A := \sqrt{x^\top A x}$. The controllability matrix of pair (A, B) is $\mathcal{C}(A, B) = \begin{bmatrix} B & AB & \dots & A^{n-1}B \end{bmatrix}$. The zero matrix of size $n \times m$ is denoted by $0_{n,m}$, the identity matrix of size n is I_n, the vector of ones is $\mathbf{1}_n$, and the vector of zeros except for a 1 in position i is e_i. Set $\mathcal{Z}$ is symmetric if $-\mathcal{Z} = \mathcal{Z}$, its convex hull is denoted by $\text{co}(\mathcal{Z})$, its interior by $\text{int}(\mathcal{Z})$, and its relative interior by $\text{relint}(\mathcal{Z})$. The set of time functions taking value in $\mathcal{Z}$ is denoted $\mathcal{F}(\mathcal{Z}) := \{f : f(t) \in \mathcal{Z} \text{ for all } t \geq 0\}$. The closed ball of dimension b, radius $r \geq 0$, and center c is denoted $\mathbb{B}^b(c, r) := \{x \in \mathbb{R}^b : \|x - c\| \leq r\}$. The Minkowski addition of sets $\mathcal{X}$ and $\mathcal{Y}$ in $\mathbb{R}^n$ is $\mathcal{X} \oplus \mathcal{Y} := \{x + y : x \in \mathcal{X}, y \in \mathcal{Y}\}$, and their Minkowski difference is $\mathcal{X} \ominus \mathcal{Y} := \{z \in \mathbb{R}^n : \{z\} \oplus \mathcal{Y} \subseteq \mathcal{X}\}$. The projection map from $\mathbb{R}^n$ onto $\mathbb{R}^r$ with $r \leq n$ is denoted by $\text{proj}_r(x_1, \dots, x_n) := (x_1, \dots, x_r) \in \mathbb{R}^r$. The operator $\text{span}(\cdot)$ maps a set of vectors to their linear span. The operator $\langle \cdot, \cdot \rangle$ denotes the standard scalar product in $\mathbb{R}^n$.

2 Problem Statement

We consider the linear time-invariant system

$$\dot{x}(t) = Ax(t) + \bar{B}\bar{u}(t),\ x(0) = x_0 \in \mathbb{R}^n,\ \bar{u}(t) \in \bar{\mathcal{U}}, \quad (1)$$

with constant matrices $A \in \mathbb{R}^{n \times n}$ and $\bar{B} \in \mathbb{R}^{n \times (m+p)}$. The admissible controls are assumed to be in $\bar{\mathcal{U}} := [-1, 1]^{m+p}$, in line with previous works [12, 17, 25].

After a loss of control authority over p of the $m+p$ actuators of system (1), the input signal $\bar{u}$ is split between the undesirable input signal $w \in \mathcal{F}(\mathcal{W})$, $\mathcal{W} := [-1, 1]^p$, and the controlled input signal $u \in \mathcal{F}(\mathcal{U}), \mathcal{U} := [-1, 1]^m$. Matrix $\bar{B}$ is accordingly split in $B \in \mathbb{R}^{n \times m}$ and $C \in \mathbb{R}^{n \times p}$ so that the dynamics become

$$\dot{x}(t) = Ax(t) + Bu(t) + Cw(t), \quad x(0) = x_0 \in \mathbb{R}^n. \quad (2)$$

We want to study how the partial loss of control authority affects the *stabilizability* and the *controllability* of the nominal dynamics.

Definition 1 *System* (1) *is* stabilizable *(resp.* controllable*) if there exists an admissible control signal* $\bar{u} \in \mathcal{F}(\bar{\mathcal{U}})$ *driving the state of system* (1) *from any* $x_0 \in \mathbb{R}^n$ *to* $0 \in \mathbb{R}^n$ *(resp. to any* $x_{tg} \in \mathbb{R}^n$*).*

To adapt these two properties to system (2), we first need the notion of *resilient reachability* introduced in [9].

Definition 2 *A target* $x_{tg} \in \mathbb{R}^n$ *is* resiliently reachable *from* $x_0 \in \mathbb{R}^n$ *by system* (2) *if for all* $w \in \mathcal{F}(\mathcal{W})$*, there exists* $T \geq 0$ *and* $u \in \mathcal{F}(\mathcal{U})$ *such that* $u(t)$ *only depends on* $w([0, t])$ *and the solution to* (2) *exists, is unique, and* $x(T) = x_{tg}$*.*

Note that $u(t)$ is allowed to depend on $w(t)$ thanks to real time sensors on all actuators of the system, even on the malfunctioning ones.

Definition 3 *System* (2) *is* resiliently stabilizable *(resp.* resilient*) to the loss of the actuators corresponding to* C *if* $0 \in \mathbb{R}^n$ *(resp. every* $x_{tg} \in \mathbb{R}^n$*) is resiliently reachable from any* $x_0 \in \mathbb{R}^n$ *by system* (2)*.*

We are now led to our first problem.

Problem 1 *Determine whether system* (2) *is resiliently stabilizable and/or resilient.*

Even if system (2) is not resilient, it might still be able to resiliently reach some targets, just not all of $\mathbb{R}^n$.

Problem 2 *Determine the states* $x_{tg} \in \mathbb{R}^n$ *that are resiliently reachable from a given* $x_0 \in \mathbb{R}^n$ *by system* (2)*.*

For time-constrained missions, resilience is not sufficient. We also need to quantify how much slower the malfunctioning system is compared to the nominal one. To do so, we follow [13] and introduce the *nominal reach time*

$$T^*_N(x_0, x_{tg}) := \inf_{\bar{u} \in \mathcal{F}(\bar{\mathcal{U}})} \left\{ \begin{array}{c} T > 0 : x(T) = x_{tg} \\ \text{in system (1)} \end{array} \right\}, \quad (3)$$

the *malfunctioning reach time*

$$T_M^*(x_0, x_{tg}) := \sup_{w \in \mathcal{F}(\mathcal{W})} \left\{ \inf_{u \in \mathcal{F}(\mathcal{U})} \left\{ \begin{matrix} T > 0 : x(T) = x_{tg} \\ \text{in system (2)} \end{matrix} \right\} \right\}, \tag{4}$$

and the *quantitative resilience*

$$r_q(x_{tg}) := \inf_{x_0 \in \mathbb{R}^n} \frac{T_N^*(x_0, x_{tg})}{T_M^*(x_0, x_{tg})}. \tag{5}$$

If $x_0 = x_{tg}$, then $T_N^* = T_M^* = 0$ and we take the convention that their ratio is 1. If x_{tg} is reachable from x_0 by system (1), then Theorem 4.3 of [29] states that the inf in (3) becomes min since $\bar{\mathcal{U}}$ is compact and convex. Similarly, T_M^* in (4) is achieved by optimal signals $w^* \in \mathcal{F}(\mathcal{W})$ and $u^* \in \mathcal{F}(\mathcal{U})$ when system (2) is resilient.

The only way to calculate u^* without any future knowledge of w^* is to solve the intractable Isaac's main equation [8], which is the differential games counterpart of the Hamilton-Jacobi-Bellman (HJB) equation. According to [24], Isaac's main equation is even more difficult to solve than the HJB equation, which usually results in intractable partial differential equations [29]. Hence, [8] produces only suboptimal solutions, itself concluding that its practical contribution is minimal.

Instead of the setting of [8], we choose [32], where u^* and w^* are unique, *bang-bang* [31], and make a time-optimal transfer from x_0 to x_{tg}. The controller knows that w^* will be chosen to make T_M^* the longest. Thus, u^* is chosen to react optimally to this worst undesirable input. Then, w^* is chosen, and to make T_M^* the longest, it is the same as the controller had predicted. Hence, from an outside perspective it appears as if the controller built u^* knowing w^* in advance, as reflected by (4). Then, T_M^* is time-optimal and can be meaningfully compared with T_N^*, leading to the following problem.

Problem 3 *Quantify the resilience of system* (2).

We will now provide the background results upon which we build our theory.

3 Background Results

We first introduce Hájek's differential games approach [21] which relies on dynamics

$$\dot{x}(t) = Ax(t) + z(t), \quad x(0) = x_0 \in \mathbb{R}^n, \quad z(t) \in \mathcal{Z}, \tag{6}$$

where $\mathcal{Z} \subseteq \mathbb{R}^n$ is the Minkowski difference between the set of admissible control inputs $B\mathcal{U} := \{Bu : u \in \mathcal{U}\}$ and the opposite of the set of undesirable inputs $C\mathcal{W} := \{Cw : w \in \mathcal{W}\}$, i.e.,

$$\begin{aligned} \mathcal{Z} &:= B\mathcal{U} \ominus (-C\mathcal{W}) \\ &= \{z \in B\mathcal{U} : z - Cw \in B\mathcal{U} \text{ for all } w \in \mathcal{W}\}. \end{aligned}$$

Theorem 1 (Hájek's duality theorem [21]) *The state of system* (2) *can be driven to* $0 \in \mathbb{R}^n$ *at time* T *for all* $w \in \mathcal{F}(\mathcal{W})$ *by control signal* $u \in \mathcal{F}(\mathcal{U})$ *if and only if the state of system* (6) *can be driven to* 0 *at time* T *by a control signal* $z \in \mathcal{F}(\mathcal{Z})$, *and* $Bu(\cdot) = z(\cdot) - Cw(\cdot)$.

Informally, $\mathcal{Z}$ represents the control available after counteracting any undesirable input. Since $\bar{\mathcal{U}}$ is symmetric, compact, and convex, sets $B\mathcal{U}$ and $C\mathcal{W}$ also have these properties by linearity. According to [27], $\mathcal{Z}$ is then also symmetric, compact, and convex.

Theorem 1 transforms the resilient stabilizability of system (2) into the stabilizability of system (6). Because inputs are bounded, Kalman's stabilizability condition [23] do not apply, instead we employ Corollary 3.6 of [15].

Theorem 2 (Stabilizability condition [15]) *If* $\bar{\mathcal{U}} \cap \ker(\bar{B}) \neq \emptyset$ *and* $\text{int}(\text{co}(\bar{\mathcal{U}})) \neq \emptyset$, *then system* (1) *is stabilizable if and only if* $\text{rank}\big(\mathcal{C}(A, \bar{B})\big) = n$, $Re\big(\lambda(A)\big) \leq 0$, *and there is no real eigenvector* v *of* $A^\top$ *satisfying* $v^\top \bar{B}\bar{u} \leq 0$ *for all* $\bar{u} \in \bar{\mathcal{U}}$.

The first condition of Theorem 2 ensures the existence of a control canceling $\bar{B}\bar{u}$ so that the state can be maintained at an equilibrium. The rank condition is Kalman's [15] and the last two conditions guarantee that the drift term Ax does not prevent stabilization. If $\bar{\mathcal{U}} = \mathbb{R}^m$, Theorem 2 reduces to the usual stabilizability condition.

To verify controllability we use Corollary 3.7 of [15], which is very similar to Theorem 2 except that the eigenvalues of A must have a zero real part to avoid creating a drift preventing the reachability of affine targets.

Theorem 3 (Controllability condition [15]) *If* $\bar{\mathcal{U}} \cap \ker(\bar{B}) \neq \emptyset$ *and* $\text{int}(\text{co}(\bar{U})) \neq \emptyset$, *then system* (1) *is controllable if and only if* $\text{rank}\big(\mathcal{C}(A, \bar{B})\big) = n$, $Re\big(\lambda(A)\big) = 0$, *and there is no real eigenvector* v *of* $A^\top$ *satisfying* $v^\top \bar{B}\bar{u} \leq 0$ *for all* $\bar{u} \in \bar{\mathcal{U}}$.

We now have all the background results to start solving Problem 1 by investigating resilient stabilizability.

4 Resilient Stabilizability

In this section, we first establish a simple resilient stabilizability condition before deriving a more complex condition with a wider range of application.

Proposition 1 *If* $\text{int}(\mathcal{Z}) \neq \emptyset$, *then system* (2) *is resiliently stabilizable if and only if* $Re\big(\lambda(A)\big) \leq 0$.

Proof. According to Theorem 1, the resilient stabilizability of system (2) is equivalent to the stabilizability of system (6). We apply Theorem 2 and obtain that if $\mathcal{Z} \cap \ker(I) \neq \emptyset$ and $\text{int}(\text{co}(\mathcal{Z})) \neq \emptyset$ in $\mathbb{R}^n$, then system (6) is stabilizable if and only if $\text{rank}\big(\mathcal{C}(A, I)\big) = n$, $Re\big(\lambda(A)\big) \leq 0$, and there is no real eigenvector v of $A^\top$ satisfying $v^\top Iz \leq 0$ for all $z \in \mathcal{Z}$.

Because $\ker(I) = \{0\}$, the first condition becomes $0 \in \mathcal{Z}$. Since $\mathcal{Z}$ is convex, the second condition becomes $\text{int}(\mathcal{Z}) \neq \emptyset$, which is equivalent to $0 \in \text{int}(\mathcal{Z})$ according to Lemma 1 of Appendix A. This second condition implies the first one, so we only keep $\text{int}(\mathcal{Z}) \neq \emptyset$.

We now assume that $\text{int}(\mathcal{Z}) \neq \emptyset$ and we simplify the last three conditions. Since $\text{rank}(I) = n$, the third condition is always true. Lemma 1 yields $0 \in \text{int}(\mathcal{Z})$. Thus, there exists $\varepsilon > 0$ such that $\mathbb{B}^n(0, \varepsilon) \subseteq \mathcal{Z}$. If $A^\top$ has no real

eigenvector, the last condition is trivially true. Otherwise, for v be a real eigenvector of $A^\top$. Let $z = \varepsilon \frac{v}{\|v\|}$, then $z \in \mathbb{B}^n(0,\varepsilon)$, so $z \in \mathcal{Z}$ and $v^\top I z = \varepsilon \|v\| > 0$. ■

Proposition 1 has a limited range of application because of its requirement $\text{int}(\mathcal{Z}) \neq \emptyset$ in $\mathbb{R}^n$, i.e., $\mathcal{Z}$ must be of dimension n. However, stabilizability does not require $B\mathcal{U}$ to be dimension n, so resilient stabilizability should not require that from $\mathcal{Z}$ either. We then want our condition to rely on the relative interior of $\mathcal{Z}$ instead of its interior.

Definition 4 *The* relative interior $\text{relint}(\mathcal{S})$ *of a set* $\mathcal{S}$ *is the interior of* $\mathcal{S}$ *considered as a subset of its affine hull.*

Definition 5 *The* affine hull *of a set* $\mathcal{S}$ *is the largest subspace included in* $\mathcal{S}$ *with respect to inclusion.*

If we apply Theorem 2 to system (6) as in Proposition 1, then $\text{int}(\mathcal{Z}) \neq \emptyset$ will appear. Instead, we first need to transport system (6) into a basis adapted to $\mathcal{Z}$. Let $r := \dim(\mathcal{Z}) \leq n$. If $\mathcal{Z} = \emptyset$, we take the convention that $r = -\infty$ and $Z := [\,] \in \mathbb{R}^{n\times 0}$, the empty matrix with $\text{Im}([\,]) = \emptyset$. Otherwise, according to Lemma 2 of Appendix A, we have $0 \in \mathcal{Z}$. Then, $\text{span}(\mathcal{Z})$ is a vector space from which we take a basis $\{z_1, \dots, z_r\}$ in $\mathbb{R}^n$. We define the matrix $Z := \big(z_1, \dots, z_r\big) \in \mathbb{R}^{n\times r}$ with the convention that $Z = 0 \in \mathbb{R}^{n\times 1}$ if $r = 0$. Then, $\text{Im}(Z) = \text{span}(\mathcal{Z})$ and we can formulate a resilient stabilizability condition less restrictive than Proposition 1.

Proposition 2 *If* $\text{relint}(\mathcal{Z}) \neq \emptyset$*, then system* (2) *is resiliently stabilizable if and only if* $\text{rank}\big(\mathcal{C}(A,Z)\big) = n$*,* $Re\big(\lambda(A)\big) \leq 0$*, and there is no real eigenvector* v *of* $A^\top$ *satisfying* $v^\top z \leq 0$ *for all* $z \in \mathcal{Z}$*.*

Proof. We apply Theorem 1 and work on system (6). Since $z_1, \dots, z_r$ are linearly independent, we complete this sequence into a basis of $\mathbb{R}^n$ with $V := (v_{r+1}, \dots, v_n)$ and obtain a transition matrix $T_z = (Z, V)$. We change basis in system (6) with $x = T_z^{-1} y$ so that $\dot{x}(t) = T_z^{-1}\dot{y}(t) = T_z^{-1}Ay(t) + T_z^{-1}z(t) = \hat{A}x(t) + s(t)$, with $\hat{A} = T_z^{-1}AT_z$ and $s(t) \in \mathcal{S} := T_z^{-1}\mathcal{Z} = \big\{T_z^{-1}z : z \in \mathcal{Z}\big\}$. By definition, $z_i = T_z e_i$ and thus $\mathcal{S} \subseteq \text{span}(\{e_1, \dots, e_r\})$ in $\mathbb{R}^n$. Let $s \in \mathcal{S}$. Then,

$$s = \begin{pmatrix} s_1 \\ \vdots \\ s_r \\ 0_{n-r,1} \end{pmatrix} = \begin{pmatrix} I_r \\ 0_{n-r,r} \end{pmatrix} \begin{pmatrix} s_1 \\ \vdots \\ s_r \end{pmatrix} := \hat{B}\hat{s},$$

with $\hat{B} = T_z^{-1}Z \in \mathbb{R}^{n\times r}$ and $\hat{s} \in \mathbb{R}^r$, $\hat{s} \in \hat{\mathcal{S}} := \text{proj}_r(\mathcal{S})$, the projection of $\mathcal{S}$ onto $\mathbb{R}^r$. Hence, the stabilizability of system (6) is equivalent to that of system

$$\dot{\hat{x}}(t) = \hat{A}\hat{x}(t) + \hat{B}\hat{s}(t), \quad \hat{x}(0) = T_z^{-1}x_0, \quad \hat{s}(t) \in \hat{\mathcal{S}}. \tag{7}$$

Applying Theorem 2 to system (7) leads to the following stabilizability conditions: $\hat{\mathcal{S}} \cap \ker(\hat{B}) \neq \emptyset$, $\text{int}(\text{co}(\hat{\mathcal{S}})) \neq \emptyset$, $Re(\lambda(\hat{A})) \leq 0$, $\text{rank}\big(\mathcal{C}(\hat{A},\hat{B})\big) = n$, and there is no real eigenvector $\hat{v}$ of $\hat{A}^\top$ satisfying $\hat{v}^\top\hat{B}\hat{s} \leq 0$ for all $\hat{s} \in \hat{\mathcal{S}}$. We now simplify these five conditions.

(1) Since $\hat{B} = \left(\begin{smallmatrix} I_r \\ 0 \end{smallmatrix}\right)$, $\text{rank}(\hat{B}) = r$, and hence $\ker(\hat{B}) = \{0\}$ in $\mathbb{R}^r$. Then, $\hat{\mathcal{S}} \cap \ker(\hat{B}) \neq \emptyset$ is equivalent to $0 \in \hat{\mathcal{S}} = \text{proj}_r(T_z^{-1}\mathcal{Z})$. In turn, this is equivalent to the existence of $v \in \mathbb{R}^{n-r}$ such that $T_z \left(\begin{smallmatrix} 0 \\ v \end{smallmatrix}\right) \in \mathcal{Z}$, i.e., $Vv \in \mathcal{Z}$. By definition of V, $\text{Im}(V) \cap \text{span}(\mathcal{Z}) = \{0\}$. Thus, $\hat{\mathcal{S}} \cap \ker(\hat{B}) \neq \emptyset$ is equivalent to $0 \in \mathcal{Z}$, i.e., $\text{relint}(\mathcal{Z}) \neq \emptyset$ according to Lemma 2 of Appendix A.

(2) By definition of $\mathcal{S}$, $\text{int}(\hat{\mathcal{S}}) \neq \emptyset$ in $\mathbb{R}^r$ is equivalent to $\text{relint}(\mathcal{Z}) \neq \emptyset$ since T_z is invertible.

(3) Because $\hat{A} = T_z^{-1}AT_z$, $\lambda(A) = \lambda(\hat{A})$, and thus the third condition becomes $Re(\lambda(A)) \leq 0$.

(4) For $i \in [\![0, n-1]\!]$, $T_z\hat{A}^i\hat{B} = T_z\big(T_z^{-1}AT_z\big)^i\hat{B} = A^iT_z\hat{B} = A^iZ$ because $T_z\hat{B} = Z$. Hence, $\text{Im}\big(T_z\mathcal{C}(\hat{A},\hat{B})\big) = \text{Im}\big(\mathcal{C}(A,Z)\big)$. The invertibility of T_z leads to $\text{rank}\big(\mathcal{C}(\hat{A},\hat{B})\big) = \text{rank}\big(\mathcal{C}(A,Z)\big)$ [20].

(5) Assume that $\hat{v}$ is a real eigenvector of $\hat{A}^\top$ associated to the eigenvalue $\hat{\lambda}$. Then, $v := T_z^{-\top}\hat{v}$ is an eigenvector of $A^\top$ associated to the same eigenvalue $\hat{\lambda}$ [20]. For $\hat{s} \in \hat{\mathcal{S}}$, we have $\hat{B}\hat{s} \in \mathcal{S}$ by definition. Hence, if we define $z := T_z\hat{B}\hat{s}$, we have $z \in \mathcal{Z}$. Then, $\hat{v}^\top\hat{B}\hat{s} = v^\top T_z\hat{B}\hat{s} = v^\top z$. ■

To further expand the applicability of our resilient stabilizability condition, we now remove the requirement $\text{relint}(\mathcal{Z}) \neq \emptyset$ from Proposition 2 and obtain a necessary and sufficient condition.

Theorem 4 (Resilient stabilizability condition) *System* (2) *is resiliently stabilizable if and only if* $\text{rank}\big(\mathcal{C}(A,Z)\big) = n$*,* $Re\big(\lambda(A)\big) \leq 0$*, and there is no real eigenvector* v *of* $A^\top$ *satisfying* $v^\top z \leq 0$ *for all* $z \in \mathcal{Z}$*.*

Proof. Let us define the three properties stated in Proposition 2 as $\mathcal{P}_1$:="$\text{relint}(\mathcal{Z}) \neq \emptyset$", $\mathcal{P}_2$:="System (2) is resiliently stabilizable", and $\mathcal{P}_3$:="$\text{rank}\big(\mathcal{C}(A,Z)\big) = n$, $Re\big(\lambda(A)\big) \leq 0$, and there is no real eigenvector v of $A^\top$ satisfying $v^\top z \leq 0$ for all $z \in \mathcal{Z}$". Proposition 2 states that if $\mathcal{P}_1$ holds, then $\mathcal{P}_2$ is equivalent to $\mathcal{P}_3$. We will now show that when $\mathcal{P}_1$ is false, so are $\mathcal{P}_2$ and $\mathcal{P}_3$, which leads to $\mathcal{P}_2$ equivalent to $\mathcal{P}_3$ no matter the status of $\mathcal{P}_1$, which is exactly the statement of this theorem.

Assume that $\mathcal{P}_1$ is false. Then, according to Lemmas 2, 5, and 6 of Appendix A, system (2) is not resiliently stabilizable, i.e., $\mathcal{P}_2$ is false. We took the convention that $Z = [\,]$ with $\text{rank}([\,]) = -\infty$, so $\mathcal{P}_3$ is false too. ■

Note that the rank condition in Theorem 4 concerns the pair (A, Z) and not (A, B) as one might have wanted. For the stabilizability of these pairs to be equivalent, we need $\mathcal{Z}$ and $B\mathcal{U}$ to have the same dimension.

Corollary 1 *If* $\dim(\mathcal{Z}) = \text{rank}(B)$*, then system* (2) *is resiliently stabilizable if and only if* $\text{rank}\big(\mathcal{C}(A,B)\big) = n$*,*

$Re\big(\lambda(A)\big) \leq 0$, and there is no real eigenvector v of $A^\top$ satisfying $v^\top z \leq 0$ for all $z \in \mathcal{Z}$.

Proof. If $\mathcal{Z} = \emptyset$, then $\text{rank}(B) = -\infty$, i.e., $B = []$, so (2) is not resiliently stabilizable and $\text{rank}\big(\mathcal{C}(A,B)\big) \neq n$.

Now assume that $\mathcal{Z} \neq \emptyset$. From Lemma 4 of Appendix A we get $\text{Im}(B) = \text{Im}(Z)$. Then, $\text{Im}\big(\mathcal{C}(A,B)\big) = \text{Im}\big(\mathcal{C}(A,Z)\big)$. In the proof of Proposition 2 we had $\text{Im}\big(\mathcal{C}(A,Z)\big) = \text{Im}\big(T\mathcal{C}(\hat{A},\hat{B})\big)$. Since T is invertible, we obtain $\text{rank}\big(\mathcal{C}(A,B)\big) = \text{rank}\big(\mathcal{C}(\hat{A},\hat{B})\big)$, and we conclude with the rest of the proof of Proposition 2. ■

Notice how the three conditions listed in Corollary 1 are similar to the stabilizability conditions from Theorem 2. We are then led to the following result.

Corollary 2 *If $\dim(\mathcal{Z}) = \text{rank}(B)$, then system (2) is resiliently stabilizable if and only if system (1) is stabilizable.*

Proof. Let v be a real eigenvector of $A^\top$. Assume first that there exists $z \in \mathcal{Z}$ such that $v^\top z > 0$. By construction of B, $\mathcal{U}$, and $\mathcal{Z}$, we have $\mathcal{Z} \subseteq B\mathcal{U} \subseteq \bar{B}\bar{\mathcal{U}}$. Hence, there exists $\bar{u} \in \bar{\mathcal{U}}$ such that $z = \bar{B}\bar{u}$ and $v^\top \bar{B}\bar{u} > 0$.

On the other hand, assume that there exists $\bar{u} \in \bar{\mathcal{U}}$ such that $v^\top \bar{B}\bar{u} > 0$. According to Lemma 4, $\text{span}(\mathcal{Z}) = \text{Im}(\bar{B})$. Then, the convexity of $\mathcal{Z}$ yields the existence of $\alpha \in \mathbb{R}$ and $z \in \mathcal{Z}$ such that $\bar{B}\bar{u} = \alpha z$. Note that $\alpha \neq 0$ by definition of $\bar{u}$. If $\alpha > 0$, we have $v^\top z > 0$. Otherwise, $\alpha < 0$ but we use the symmetry of $\mathcal{Z}$ to obtain $-z \in \mathcal{Z}$ and $v^\top(-z) > 0$.

Thus, the condition "there is no real eigenvector v of $A^\top$ satisfying $v^\top z \leq 0$ for all $z \in \mathcal{Z}$" is equivalent to "there is no real eigenvector v of $A^\top$ satisfying $v^\top \bar{B}\bar{u} \leq 0$ for all $\bar{u} \in \bar{\mathcal{U}}$" when $\dim(\mathcal{Z}) = \text{rank}(B)$. According to Lemma 4 of Appendix A, $\text{Im}(B) = \text{Im}(\bar{B})$. Hence, $\text{rank}\big(\mathcal{C}(A,B)\big) = \text{rank}\big(\mathcal{C}(A,\bar{B})\big)$. Then, applying Corollary 1 to system (2) and Theorem 2 to system (1) concludes the proof. ■

We have established several resilient stabilizability conditions, hence solving the first half of Problem 1. We will now tackle its second part concerning affine targets.

5 Resilient Reachability

In this section we extend Hájek's duality theorem [21] to affine targets and study the resilience of linear systems.

Theorem 5 (Extended duality theorem)
The state of system (2) can be driven to $x_{tg} \in \mathbb{R}^n$ at time T for all $w \in \mathcal{F}(\mathcal{W})$ by control signal $u \in \mathcal{F}(\mathcal{U})$ if and only if the state of system (6) can be driven to x_{tg} at time T by a control signal $z \in \mathcal{F}(\mathcal{Z})$, and $Bu(\cdot) = z(\cdot) - Cw(\cdot)$.

Proof. Consider system (2) with a target state $x_{tg} \in \mathbb{R}^n$, $x_{tg} \neq 0$. Let $X(t) := \begin{pmatrix} x(t) - x_{tg} \\ Ax_{tg} \end{pmatrix} \in \mathbb{R}^{2n}$. Then,

$$\begin{aligned} \dot{X}(t) &= A_2 X(t) + B_2 u(t) + C_2 w(t), \\ X(0) = X_0 \in \mathbb{R}^{2n}, &\quad u(t) \in \mathcal{U}, \quad w(t) \in \mathcal{W}, \end{aligned} \tag{8}$$

$A_2 = \begin{pmatrix} A & I_n \\ 0_{n,n} & 0_{n,n} \end{pmatrix}$, $B_2 = \begin{pmatrix} B \\ 0_{n,m} \end{pmatrix}$, $C_2 = \begin{pmatrix} C \\ 0_{n,p} \end{pmatrix}$ and $X_0 = \begin{pmatrix} x_0 - x_{tg} \\ Ax_{tg} \end{pmatrix}$. Let the target set be $\mathcal{G} = \left\{ \begin{pmatrix} 0 \\ a \end{pmatrix} \in \mathbb{R}^{2n} \right\} = \{0\}^n \times \mathbb{R}^n$. Since $0 \in C_2\mathcal{W}$, we can apply Hájek's second duality theorem of [21] stating that $\mathcal{G}$ is resiliently reachable in time T from X_0 by system (8) if and only if $\mathcal{G}$ is reachable in time T from X_0 by the following system

$$\dot{X}(t) = A_2 X(t) + v_2(t), \quad X(0) = X_0, \tag{9}$$

$v_2(t) \in \mathcal{V}_2 := B_2\mathcal{U} \cap \big[(B_2\mathcal{U} \oplus \mathcal{G}_{A_2}) \ominus (-C_2\mathcal{W})\big] \subseteq \mathbb{R}^{2n}$, where $\mathcal{G}_{A_2}$ is the largest subspace of $\mathcal{G}$ invariant by A_2. Take $g = \begin{pmatrix} 0 \\ a \end{pmatrix} \in \mathcal{G}$, then $A_2 g = \begin{pmatrix} A & I_n \\ 0_{n,n} & 0_{n,n} \end{pmatrix} \begin{pmatrix} 0 \\ a \end{pmatrix} = \begin{pmatrix} a \\ 0 \end{pmatrix}$. Hence, $A_2 g \in \mathcal{G} \iff a = 0$, i.e., $\mathcal{G}_{A_2} = \{0\}^{2n}$. Thus, $\mathcal{V}_2 = \big\{v \in B_2\mathcal{U} : v - C_2 w \in B_2\mathcal{U}, \text{for all } w \in \mathcal{W}\big\} = \mathcal{Z} \times \{0\}^n$, because of the architecture of B_2 and C_2. Then, system (9) is related to system (6) the same way that system (8) is related to system (2). Therefore, the following statements are equivalent:

- x_{tg} is resiliently reachable by system (2),
- $\mathcal{G}$ is resiliently reachable by system (8),
- $\mathcal{G}$ is reachable by system (9),
- x_{tg} is reachable by system (6). ■

Theorem 5 transforms resilience of system (2) into bounded controllability of system (6), which we verify with Theorem 3.

We can easily adapt the results of Section 4 to the resilience case by reusing the same proofs, except that we use Theorems 5 and 3 instead of Theorems 1 and 2.

Proposition 3 *If $\text{int}(\mathcal{Z}) \neq \emptyset$, then system (2) is resilient if and only if $Re(\lambda(A)) = 0$.*

Corollary 3 *If $\dim(\mathcal{Z}) = \text{rank}(B)$, then system (2) is resilient if and only if system (1) is controllable.*

Theorem 6 (Resilience condition)
System (2) is resilient if and only if $Re\big(\lambda(A)\big) = 0$, $\text{rank}\big(\mathcal{C}(A,Z)\big) = n$, and there is no real eigenvector v of $A^\top$ satisfying $v^\top z \leq 0$ for all $z \in \mathcal{Z}$.

We now have all the results necessary to solve Problem 1. However, the condition $Re\big(\lambda(A)\big) = 0$ in Theorem 6 is not satisfied by most systems, that are hence not resilient. This reasoning led us to Problem 2, i.e., the determination of the resiliently reachable set of system (2). Following Theorem 5, we will now study the reachable set of system (6) given by

$$R(T, x_0) := \left\{ \begin{array}{c} e^{AT}\left(x_0 + \int_0^T e^{-At} z(t)\, dt\right) \\ \text{with } z(t) \in \mathcal{Z} \text{ for all } t \in [0,T] \end{array} \right\}.$$

Because analytical study of $R(T, x_0)$ is difficult, most of the research tries to approximate it (see [19] and references therein). We want inner approximations of $R(T, x_0)$ in order to determine the states that are guaranteed to be resiliently reachable. We will then present a method of *zonotopic* underapproximation of $R(T, x_0)$ combining the approaches of [19] and [2].

Definition 6 *A* zonotope $\mathcal{S} \subseteq \mathbb{R}^n$ *is a set parametrized by a center* $c \in \mathbb{R}^n$ *and generators* $g_1, \dots, g_q \in \mathbb{R}^n$ *expressed as* $\mathcal{S} := \{c + \sum_{i=1}^{q} \alpha_i g_i : \alpha_i \in [-1, 1]\}$ *and is denoted* $\mathcal{S} = (c, g_1, \dots g_q)$.

Note that $B\mathcal{U}$ is a zonotope of center 0 and generators B_i, the columns of B. Similarly, $C\mathcal{W} = (0, C_1, \dots, C_p)$. However, $\mathcal{Z}$ is not a zonotope in general since these sets are not closed under Minkowski difference except for some specific scenarios, as detailed in [2].

Following [2], we build an underapproximation of $\mathcal{Z}$ with a symmetric zonotope $\big(0, g_1, \dots, g_r\big) \subseteq \mathcal{Z}$ by removing or contracting the generators of $B\mathcal{U}$. We apply the method described in [19] to compute efficiently an inner approximation of $R(T, x_0)$. For $N \in \mathbb{N}$, $N \geq 1$, we define $\delta t := \frac{T}{N}$, $\Omega_0 := \{x_0\}$, $V := \left\{\int_0^{\delta t} e^{A(\delta t - t)} z(t)\, dt : z(t) \in \mathcal{Z} \text{ for } t \in [0, \delta t]\right\}$, and the recursion $\Omega_{i+1} := e^{A\delta t}\Omega_i \oplus V$. Note that Ω_i is the exact reachable set $R(i\,\delta t, x_0)$.

However, V is not a zonotope and cannot be computed exactly. Thus, we define the zonotope $\tilde{V} := \Big(0, \int_0^{\delta t} e^{A(\delta t - t)} g_1\, dt, \dots, \int_0^{\delta t} e^{A(\delta t - t)} g_r\, dt\Big)$, and $\tilde{V} \subseteq V$ since $\tilde{V}$ corresponds to piecewise constant components of $z(t)$ in $\big(0, g_1, \dots, g_r\big)$.

Then, we build $\tilde{\Omega}_0 = \Omega_0 = \{x_0\}$ and $\tilde{\Omega}_{i+1} := e^{A\delta t}\tilde{\Omega}_i \oplus \tilde{V}$, which yields $\tilde{\Omega}_i \subseteq \Omega_i$ for all $i \geq 0$. Since linear maps and Minkowski sums are straightforward on zonotopes [2,19], $\tilde{\Omega}_i$ is an easily computable inner approximation of the reachable set $R(i\,\delta t, x_0)$. Note that the precision of the approximation increases with N.

Before implementing this solution to Problem 2 in Section 7.1, we need to answer Problem 3 by quantifying the resilience of linear systems.

6 Quantitative Resilience

Let us now investigate more complex missions where the target needs to be reached by a certain time. In such scenarios it is crucial to evaluate the maximal time penalty incurred by the malfunctioning system.

Unlike in the driftless case [13], the optimal reach times T_N^* (3) and T_M^* (4) cannot be reduced to a linear optimization and elude analytical expressions [6]. Following [17] and [32] we could numerically compute these reach times, but not the quantitative resilience r_q (5) since it would require computing $T_N^*(x_0)$ and $T_M^*(x_0)$ for all $x_0 \in \mathbb{R}^n$. Instead, using Lyapunov theory [25], we establish analytical bound on these two reach times for the target $x_{tg} = 0$ and analytically approximate r_q.

6.1 Nominal reach time

Assume that A is Hurwitz. Then, for any $Q \succ 0$ there exists $P \succ 0$ such that $PA + A^\top P = -Q$ [25]. Let us consider any such pair (P, Q). We define the Lyapunov function $V(x) := x^\top P x = \|x\|_P^2$ [26]. Then, for x following (1) we have

$$\begin{aligned}\dot{V}(x) &= \dot{x}^\top P x + x^\top P \dot{x} = x^\top (A^\top P + PA)x + 2x^\top P\bar{B}\bar{u} \\ &= -x^\top Q x + 2x^\top P\bar{B}\bar{u}. \end{aligned} \tag{10}$$

We will now bound $T_N^*(x_0)$.

Proposition 4 *If system* (1) *is stabilizable and* A *is Hurwitz, then*

$$T_N^*(x_0) \;\geq\; 2\frac{\lambda_{min}^P}{\lambda_{max}^Q} \ln\left(1 + \frac{\lambda_{max}^Q \|x_0\|_P}{2\lambda_{min}^P b_{max}^P}\right), \tag{11}$$

with $b_{max}^P := \max\big\{\|\bar{B}\bar{u}\|_P : \bar{u} \in \bar{\mathcal{U}}\big\}$.

Proof. Because $\bar{\mathcal{U}}$ is compact and convex, and system (1) is stabilizable, there exists a time-optimal control signal $\bar{u}^* \in \mathcal{F}(\bar{\mathcal{U}})$ driving the state from x_0 to the origin in a finite time $T_N^*(x_0)$ [29].

We now bound $\dot{V}$ using (10). Since $P \succ 0$, there exists $M \in \mathbb{R}^{n\times n}$ such that $P = M^\top M$ [20]. Then, $x^\top P\bar{B}\bar{u} = (Mx)^\top M\bar{B}\bar{u} \;\geq\; -\|Mx\|_2\|M\bar{B}\bar{u}\|_2$, by the Cauchy-Schwarz inequality [20]. Notice $\|Mx\|_2^2 = x^\top M^\top Mx = x^\top Px = \|x\|_P^2$. Similarly, $\|M\bar{B}\bar{u}\|_2 = \|\bar{B}\bar{u}\|_P$.

The maximum b_{max}^P exists since $\bar{\mathcal{U}}$ is compact and the map $\bar{u} \mapsto \|\bar{B}\bar{u}\|_P$ is continuous. Since $Q \succ 0$, we have $x^\top Qx \leq \lambda_{max}^Q \|x\|_2^2$ and $\|x\|_2^2 \leq \|x\|_P^2/\lambda_{min}^P$ because $P \succ 0$. For $x \neq 0$, we have now lower bounded (10)

$$\dot{V}(x) = \frac{d}{dt}\|x\|_P^2 \geq -\frac{\lambda_{max}^Q}{\lambda_{min}^P}\|x\|_P^2 - 2b_{max}^P\|x\|_P. \tag{12}$$

Let $y(t) := \|x(t)\|_P$, $\alpha := \frac{\lambda_{max}^Q}{2\lambda_{min}^P} > 0$, and $\beta := b_{max}^P > 0$. For $x \neq 0$ we divide (12) by $2y > 0$ so that $\dot{y} \geq f(y) := -\alpha y - \beta$. The solution of the differential equation $\dot{s}(t) = f\big(s(t)\big)$ with $s(0) = y(0)$ is given by $s(t) = e^{-\alpha t}\left(y(0) + \frac{\beta}{\alpha}\right) - \frac{\beta}{\alpha}$.

Since f is Lipschitz, we can apply the comparison lemma of [26] and we obtain $y(t) \geq s(t)$ for all $t \geq 0$. At time $T = \frac{1}{\alpha}\ln\Big(1 + \frac{\alpha}{\beta}y(0)\Big)$, we have $s(T) = 0$. Because $\|x(t)\|_P \geq s(t) > 0$ for all $t \in [0, T]$, we have $T_N^*(x_0) \geq T$. Substituting α and β yields (13). ■

The proof of Propositions 4, as well as subsequent Propositions 5, 6, and 7, is shorter than presented in the conference paper [12] due to our use of the comparison lemma [26]. We now upper bound $T_N^*(x_0)$.

Proposition 5 *If* $\operatorname{rank}(\bar{B}) = n$ *and* A *is Hurwitz, then*

$$T_N^*(x_0) \;\le\; 2\frac{\lambda_{max}^P}{\lambda_{min}^Q} \ln\left(1 + \frac{\lambda_{min}^Q \|x_0\|_P}{2\lambda_{max}^P b_{min}^P}\right), \qquad (13)$$

with $b_{min}^P := \min\left\{\|\bar{B}\bar{u}\|_P : \bar{u} \in \partial\bar{\mathcal{U}}\right\}$.

Proof. The minimum b_{min}^P exists since map $\bar{u} \mapsto \|\bar{B}\bar{u}\|_P$ is continuous and $\partial\bar{\mathcal{U}}$ is compact.

Because $\operatorname{rank}(\bar{B}) = n$, we can choose $\bar{u} \in \mathcal{F}(\bar{\mathcal{U}})$ such that $\bar{B}\bar{u}(t) = -\frac{x(t)}{\|x(t)\|_P} b_{min}^P$ for $x(t) \neq 0$. Indeed, assume for contradiction purposes that for some $\tau \geq 0$, $\bar{u}(\tau) \notin \bar{\mathcal{U}}$, i.e., $\|\bar{u}(\tau)\|_\infty > 1$. Let $\hat{u} := \frac{\bar{u}(\tau)}{\|\bar{u}(\tau)\|_\infty}$. Then, $\|\hat{u}\|_\infty = 1$, so $\hat{u} \in \partial\bar{\mathcal{U}}$, but $\|\bar{B}\hat{u}\|_P = \frac{\|\bar{B}\bar{u}(\tau)\|_P}{\|\bar{u}(\tau)\|_\infty} = \frac{b_{min}^P}{\|\bar{u}\|_\infty} < b_{min}^P$, which is a contradiction. Hence, the proposed control signal is admissible and we implement it in (10).

We obtain $2x^\top P\bar{B}\bar{u} = -2b_{min}^P\|x\|_P$, so that

$$\frac{d}{dt}\|x\|_P^2 = \dot{V}(x) \leq -\frac{\lambda_{min}^Q}{\lambda_{max}^P}\|x\|_P^2 - 2b_{min}^P\|x\|_P. \qquad (14)$$

Let $y(t) := \|x(t)\|_P$, $\gamma := \frac{\lambda_{min}^Q}{2\lambda_{max}^P} > 0$, and $\kappa := b_{min}^P > 0$. For $x \neq 0$, dividing (14) by $2y > 0$, yields $\dot{y} \leq f(y) := -\gamma y - \kappa$. As in Proposition 4, the comparison lemma of [26] yields $y(t) \leq s(t) = e^{-\gamma t}\left(y(0) + \frac{\kappa}{\gamma}\right) - \frac{\kappa}{\gamma}$ for all $t \geq 0$ as long as $y(t) > 0$. At time $T = \frac{1}{\gamma}\ln\left(1 + \frac{\gamma}{\kappa}y(0)\right)$, $s(T) = 0$. Since $y\big(T_N^*(x_0)\big) = 0$, $T_N^*(x_0) \leq T$. ■

We now bound the malfunctioning reach time T_M^* following the same method applied to T_N^*.

6.2 Malfunctioning reach time

We use the same Lyapunov function as above, but with x following (2), so $\dot{V}(x) = -x^\top Qx + 2x^\top P(Bu + Cw)$. We can now lower bound T_M^* as we have done for T_N^*.

Proposition 6 *If system* (2) *is resiliently stabilizable and* A *is Hurwitz, then*

$$T_M^*(x_0) \;\ge\; 2\frac{\lambda_{min}^P}{\lambda_{max}^Q} \ln\left(1 + \frac{\lambda_{max}^Q \|x_0\|_P}{2\lambda_{min}^P z_{max}^P}\right), \qquad (15)$$

with $z_{max}^P := \max\left\{\|z\|_P : z \in \mathcal{Z}\right\}$.

Proof. Since $B\mathcal{U}$ and $C\mathcal{W}$ are compact, $\mathcal{Z}$ is compact [27], so z_{max}^P exists. Since system (2) is resiliently stabilizable, $T_M^*(x_0)$ exists. Let $w^* \in \mathcal{F}(\mathcal{W})$ and $u^* \in \mathcal{F}(\mathcal{W})$ be the arguments of the optimizations in (4). By definition of $\mathcal{Z}$, $z = Cw^* + Bu^* \in \mathcal{F}(\mathcal{Z})$. Then, $\|Cw^*(t) + Bu^*(t)\|_P \leq z_{max}^P$, which yields

$$\dot{V}(x) \geq -\frac{\lambda_{max}^Q}{\lambda_{min}^P}\|x\|_P^2 - 2z_{max}^P\|x\|_P.$$

We now proceed as in the second half of the proof of Proposition 4 to obtain (15). ■

Similarly, we upper bound the malfunctioning reach time.

Proposition 7 *If* $\operatorname{int}(\mathcal{Z}) \neq \emptyset$ *and* A *is Hurwitz, then*

$$T_M^*(x_0) \;\le\; 2\frac{\lambda_{max}^P}{\lambda_{min}^Q} \ln\left(1 + \frac{\lambda_{min}^Q \|x_0\|_P}{2\lambda_{max}^P z_{min}^P}\right), \qquad (16)$$

with $z_{min}^P := \min\left\{\|z\|_P : z \in \partial\mathcal{Z}\right\}$.

Proof. According to Proposition 1, system (2) is resiliently stabilizable, hence a finite T_M^* exists.

Since $\mathcal{Z}$ is compact, so is $\partial\mathcal{Z}$, and thus z_{min}^P exists. Because $\operatorname{int}(\mathcal{Z}) \neq \emptyset$, according to Lemma 1, $0 \in \operatorname{int}(\mathcal{Z})$. Then, the convexity of $\|\cdot\|_P$ yields $\left\{z \in \mathbb{R}^n : \|x\|_P \leq z_{min}^P\right\} \subseteq \mathcal{Z}$, so $z(t) := \frac{-x(t)}{\|x(t)\|_P} z_{min}^P \in \mathcal{Z}$.

Let $w^* \in \mathcal{F}(\mathcal{W})$ be the argument of the maximum in (4). Since $z(t) \in \mathcal{Z}$, there exists $u \in \mathcal{F}(\mathcal{U})$ such that $z(t) = Cw^*(t) + Bu(t)$. Then, applying w^* and u leads to an upper bound of T_M^* since u is not necessarily optimal, while w^* is optimal. Hence

$$\dot{V}(x) \leq -\frac{\lambda_{min}^Q}{\lambda_{max}^P}\|x\|_P^2 - 2z_{min}^P\|x\|_P.$$

We now proceed as in the second half of the proof of Proposition 5 to obtain (16). ■

We can now bound $T_N^*(x_0)/T_M^*(x_0)$ for all $x_0 \in \mathbb{R}^n$ and hence obtain an approximate of quantitative resilience r_q which cannot be done with prior algorithms [17,32] that only compute a single instance of $T_N^*(x_0)$ or $T_M^*(x_0)$.

6.3 Bounding quantitative resilience

If the system's quantitative resilience r_q is bounded by $\gamma \leq r_q$, then in the worst case, the malfunctioning system will take less than $1/\gamma$ times longer than the nominal system to reach the origin from the same initial state.

Theorem 7 *If* $\operatorname{int}(\mathcal{Z}) \neq \emptyset$ *and* A *is Hurwitz, then*

$$r_q \geq \max\left(\frac{\lambda_{min}^P \lambda_{min}^Q}{\lambda_{max}^P \lambda_{max}^Q},\; \frac{z_{min}^P}{b_{max}^P}\right), \qquad (17)$$

for any $P \succ 0$ *and* $Q \succ 0$ *such that* $A^\top P + PA = -Q$.

Proof. According to Proposition 1, system (2) is resiliently stabilizable. Since $\operatorname{int}(\mathcal{Z}) \neq \emptyset$, we have $\dim(\mathcal{Z}) = n$, and $\mathcal{Z} \subseteq B\mathcal{U} \subseteq \mathbb{R}^n$ yields $\operatorname{rank}(B) = n$. According to Corollary 2, system (1) is stabilizable, so we can use (11) and (16). We define the positive constants $a := \frac{\lambda_{min}^P \lambda_{min}^Q}{\lambda_{max}^P \lambda_{max}^Q}$, $b := \frac{\lambda_{max}^Q}{2\lambda_{min}^P b_{max}^P}$, and $c := \frac{\lambda_{min}^Q}{2\lambda_{max}^P z_{min}^P}$, so that for $x_0 \in \mathbb{R}^n$, $x_0 \neq 0$, (11) and (16) yield

$$\frac{T_N^*(x_0)}{T_M^*(x_0)} \geq a\frac{\ln(1 + b\|x_0\|_P)}{\ln(1 + c\|x_0\|_P)} := f(\|x_0\|_P).$$

Then, according to (5), $r_q \geq \inf_{x_0 \in \mathbb{R}^n} f(\|x_0\|_P)$.

If $b = c$, then $f(s) = a$ for all $s \geq 0$, so $r_q \geq a$. If $b > c$, then f is increasing, so $\inf\{f(s) : s > 0\} = \lim_{s\to 0} f(s)$. L'Hôpital's Rule [28] yields

$$\lim_{s\to 0} f(s) = \lim_{s\to 0} a\frac{\ln(1+bs)}{\ln(1+cs)} = \lim_{s\to 0} a\frac{\frac{b}{1+bs}}{\frac{c}{1+cs}} = \frac{ab}{c}.$$

Then, $f(0) = \frac{ab}{c} = \frac{z^P_{min}}{b^P_{max}} > a$. If $c > b$, then f is decreasing, so $\inf\{f(s) : s \geq 0\} = \lim_{s\to+\infty} f(s) = a$ by L'Hôpital's Rule [28]. To sum up, $\inf_{s\geq 0} f(s) = \max\left(a, \frac{ab}{c}\right) \leq r_q$. ■

We can upper bound r_q using a similar approach.

Theorem 8 *If* $\text{rank}(\bar{B}) = n$, *A is Hurwitz, and system (2) is resiliently stabilizable, then*

$$r_q \leq \max\left(\frac{\lambda^P_{max}\lambda^Q_{max}}{\lambda^P_{min}\lambda^Q_{min}}, \frac{z^P_{max}}{b^P_{min}}\right), \tag{18}$$

for any $P \succ 0$ and $Q \succ 0$ such that $A^\top P + PA = -Q$.

Proof. With our assumptions we are allowed to use Propositions 5 and 6. We define the positive constants $a := \frac{\lambda^P_{max}\lambda^Q_{max}}{\lambda^P_{min}\lambda^Q_{min}}$, $b := \frac{\lambda^Q_{min}}{2\lambda^P_{max}b^P_{min}}$, and $c := \frac{\lambda^Q_{max}}{2\lambda^P_{min}z^P_{max}}$, so that for $x_0 \in \mathbb{R}^n$, $x_0 \neq 0$, (13) and (15) yield

$$\frac{T^*_N(x_0)}{T^*_M(x_0)} \leq a\frac{\ln(1+b\|x_0\|_P)}{\ln(1+c\|x_0\|_P)} := g(\|x_0\|_P).$$

Then, according to (5), $r_q \leq \inf_{x_0\in\mathbb{R}^n} g(\|x_0\|_P)$. This function g is similar to f in the proof of Theorem 7, and thus $r_q \leq \inf_{x_0\in\mathbb{R}^n} g(\|x_0\|_P) = \max\left(a, a\frac{b}{c}\right)$, yielding (18). ■

Theorems 7 and 8 bound r_q and hence solve Problem 3. We will now apply the developed theory to two examples.

7 Numerical Results

We will first study the resilient reachability of the ADMIRE fighter jet model [18], before quantifying the resilience of a temperature control system.

7.1 Resilient reachability of the ADMIRE fighter jet model

The ADMIRE model has already served as an application case in several control frameworks [10, 22] and is illustrated on Fig. 1.

Relying on the simulation package *Admirer4p1* [1] we run the ADMIRE simulation in MATLAB and obtain the linearized dynamics at Mach 0.3 and altitude 2000 m. We scale $\bar{B}$ so that the input set of each actuator from [18] is scaled to $[-1, 1]$. The states and matrices of the system $\dot{X}(t) = AX(t) + \bar{B}\bar{u}(t)$ are given below.

[1] `https://app.box.com/s/r9wfyjd9o4pq2if9xhd17yxeqc36j7ei`

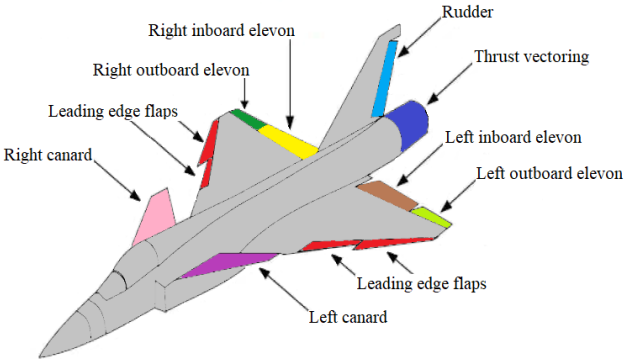
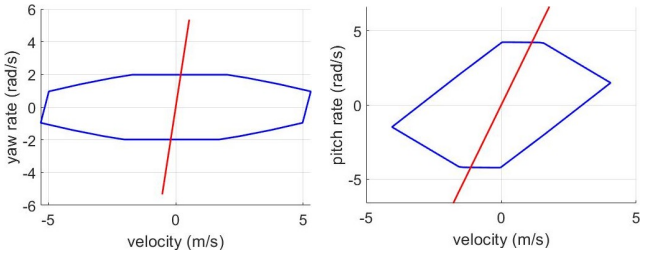


Fig. 1. The ADMIRE fighter jet model. Image modified from [18] with a different color for each independent actuator.

Consider a scenario in which, after sustaining damage, an actuator of the fighter jet starts producing uncontrolled and possibly undesirable inputs. By studying $\bar{B}$, we gain intuition on the resilience of the jet. The effect of the yaw (resp. pitch) thrust vectoring on the yaw (resp. pitch) rate is larger than that of all the other actuators combined, which gives the intuition that the jet is not resilient to the loss control over thrust vectoring. None of the other actuators produce such a dominant effect, hence giving the intuition that the jet is resilient to the loss of control over any one of the first eight actuators.

Following Lemma 6, we test our intuition by verifying whether $C\mathcal{W} \subseteq B\mathcal{U}$. These sets are zonotopes of dimension 9, represented in MATLAB using function *zonotope*$(\cdot)$ from the CORA package [3]. The associated function *in*$(\cdot)$ is employed to verify their inclusion. As expected, $C\mathcal{W} \subseteq B\mathcal{U}$ for the loss of control over any one actuator except for the thrust vectoring ones, as shown on Fig. 2.



(a) Yaw thrust vectoring. (b) Pitch thrust vectoring.

Fig. 2. 2D projection of sets $B\mathcal{U}$ (blue) and $C\mathcal{W}$ (red) for the loss of control over the two thrust vectoring actuators.

The eigenvalues of A do not verify either $Re(\lambda(A)) = 0$ or $Re(\lambda(A)) \leq 0$. Thus, the system is neither resilient nor resiliently stabilizable. However, as anticipated with Problem 2, the linearized model is only valid locally and hence we should only study the resilient reachability of targets close to the linearization equilibrium.

We follow the method detailed in Section 5 to approximate the resiliently reachable set of the malfunctioning system. Assume the pilot lost control over the right

$$X = \begin{pmatrix} v \\ \alpha \\ \beta \\ p \\ q \\ r \\ \psi \\ \theta \\ \varphi \end{pmatrix} \begin{matrix} \text{velocity } (m/s), \\ \text{angle of attack } (rad), \\ \text{sideslip angle } (rad), \\ \text{roll rate } (rad/s), \\ \text{pitch rate } (rad/s), \\ \text{yaw rate } (rad/s), \\ \text{heading angle } (rad), \\ \text{pitch angle } (rad), \\ \text{roll angle } (rad), \end{matrix} \quad A = \begin{pmatrix} -0.02 & -4.65 & 0.37 & 0 & -0.3 & 0 & 0 & -9.81 & 0 \\ 0 & -0.78 & 0.01 & 0 & 0.97 & 0 & 0 & 0 & 0 \\ 0 & 0 & -0.19 & 0.12 & 0 & -0.98 & 0 & 0 & 0.1 \\ 0 & 0 & -15.47 & -1.5 & 0 & 0.54 & 0 & 0 & 0 \\ 0 & 4.18 & -0.01 & 0 & -0.78 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.95 & -0.09 & 0 & -0.34 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1.01 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0.12 & 0 & 0 & 0 \end{pmatrix}$$

$$\bar{B}^\top = \begin{pmatrix} -0.62 & 0 & 0 & 0.37 & 0.67 & -0.19 & 0 & 0 & 0 \\ -0.62 & 0 & 0 & -0.37 & 0.67 & 0.19 & 0 & 0 & 0 \\ -0.4 & -0.02 & 0 & -2.27 & -0.55 & -0.1 & 0 & 0 & 0 \\ -0.62 & -0.04 & 0.01 & -1.96 & -0.88 & -0.22 & 0 & 0 & 0 \\ -0.62 & -0.04 & -0.01 & 1.96 & -0.88 & 0.22 & 0 & 0 & 0 \\ -0.4 & -0.02 & 0 & 2.27 & -0.55 & 0.1 & 0 & 0 & 0 \\ -0.16 & 0 & 0.02 & 1.59 & 0 & -0.96 & 0 & 0 & 0 \\ 0.08 & 0 & 0 & 0 & -0.02 & 0 & 0 & 0 & 0 \\ -0.53 & 0 & 0.11 & -0.64 & 0.01 & -5.34 & 0 & 0 & 0 \\ -1.78 & -0.11 & 0 & 0 & -6.63 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{matrix} \text{right canard}, \\ \text{left canard}, \\ \text{right outboard elevon}, \\ \text{right inboard elevon}, \\ \text{left inboard elevon}, \\ \text{left outboard elevon}, \\ \text{rudder}, \\ \text{leading edge flaps}, \\ \text{yaw thrust vectoring}, \\ \text{pitch thrust vectoring}. \end{matrix}$$

outboard elevon $\bar{u}_3$. We use the CORA [3] function $minus(\cdot,\cdot)$ to underapproximate the Minkowski difference $\mathcal{Z} = B\mathcal{U} \ominus C\mathcal{W}$ as a zonotope $(0, g_1, \ldots, g_9)$, following the method of [2]. We take $T = 0.2\,s$ and $N = 5$. Then, we underapproximate $R(T, x_0)$ with $\tilde{\Omega}_N$ using the recursion $\tilde{\Omega}_{i+1} = e^{A\delta t}\tilde{\Omega}_i \oplus \tilde{V}$ of Section 5.

Since the malfunctioning actuator $\bar{u}_3$ has a strong impact on the roll rate p of the jet, we want to see what range of roll rates is reachable. We compute $\tilde{\Omega}_1, \ldots, \tilde{\Omega}_N$ and project them in 2D as shown on Fig. 3. Then, in time T the jet can change its roll rate up to $\pm 1.2\,rad/s$, despite the loss of control over the right outboard elevon.

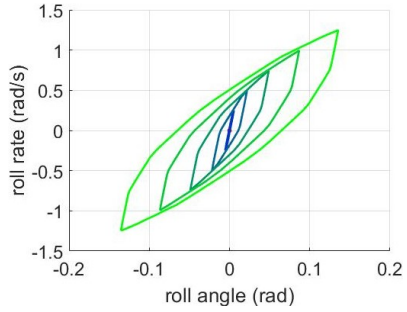
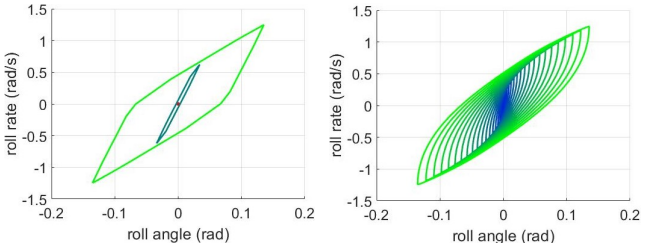


Fig. 3. Projection of $\tilde{\Omega}_1, \ldots, \tilde{\Omega}_5$ on the (ϕ, p) plane.

We now study the impact of N, i.e., of δt on the precision of $\tilde{\Omega}_N$ to approximate the real reachable set $R(T, x_0)$ when keeping T constant. Since $\dim\big(R(T, x_0)\big) = 9$, we will only study the impact on the range of roll rates reachable at roll angle $\phi = 0\,rad$. For $N = 2$ the reachable range of roll rates is $\pm 0.37\,rad/s$, while for $N = 5$ it is $\pm 0.42\,rad/s$, and $\pm 0.43\,rad/s$ for $N = 20$, as illustrated on Fig. 3 and 4. Hence, as explained in Section 5, increasing N raises nonlinearly the precision of $\tilde{\Omega}_N$ and increases linearly the computational cost since $\tilde{\Omega}_N$ is a zonotope with $9N$ generators.



(a) $N = 2$. (b) $N = 20$.

Fig. 4. Projection of $\tilde{\Omega}_1, \ldots, \tilde{\Omega}_N$ on the (ϕ, p) plane for different values of N.

Now assume that the in-flight damage responsible for the loss of control over the elevon $\bar{u}_3$ also initially caused it to jerk resulting in a sudden jump in roll rate. Then, instead of $X(0) = 0$ we have $p(0) = 0.44\,rad/s$ and the goal is to stabilize the jet at the origin X_{tg}.

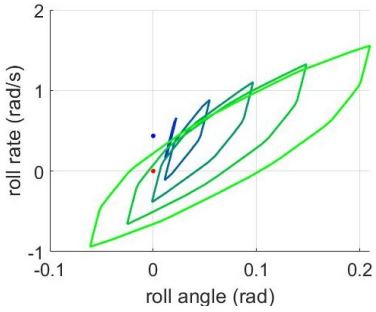


Fig. 5. Projection of $\tilde{\Omega}_1, \ldots, \tilde{\Omega}_5$ on the (ϕ, p) plane. Initial state X_0 is the blue dot, target X_{tg} is the red dot, and $N = 5$.

We can see on Fig. 5 that the target only enters the projection of the reachable set after 4 iterations of $\delta t = 0.04\,s$, i.e., for $t \geq 0.16\,s$. By choosing a smaller δt we can refine the precision on the minimal entering time. However, to calculate the reachable time $T_M^*(X_0, X_{tg})$ we need to use the CORA function $in(\cdot)$ to verify whether $X_{tg} \in \tilde{\Omega}_N$ since Fig. 5 is only a 2D projection of the 9D reachable set and could be deceiving. Indeed, for $p(0) = 0.5\,rad/s$, the 2D projection is similar to Fig. 5 with the red dot inside the projection of $\tilde{\Omega}_N$, but $X_{tg} \notin \tilde{\Omega}_N$.

We successfully demonstrated the developed resilience theory and the zonotopic method to underapproximate the resiliently reachable set of the ADMIRE jet model.

7.2 Temperature control system

We now illustrate our quantitative resilience bounds on a temperature control system motivated by [35] and illustrated on Fig. 6.

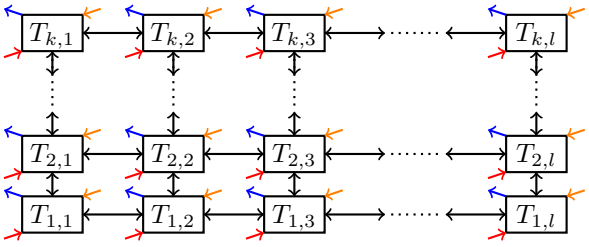


Fig. 6. Heat exchange graph of an office building with k floors of l rooms, each at a temperature $T_{i,j}$.

We study a scenario where a worker remains in their office after hours and manually opens or closes their door and window, thus overriding the building heat controller which aims at maintaining a target temperature T_{tg}. After this loss of control, we will compare our analytical bounds on the nominal and malfunctioning reach times with the numerical results of [17,32]. We will also bound the quantitative resilience of the system which could not be done with prior work and motivated the analytical bounds of Section 6.

The controller uses a central heater q_h, central AC q_{AC}, and incrementally opens doors q_d and windows q_w for room specific adjustments. The controller also takes advantage of solar heating q_S, heat losses through the outside wall q_l, and heat transfers between adjoining rooms q_{adj}. The temperature dynamics are then

$$mC_p\dot{T}_{i,j} = q_h - q_{AC} + q_{d_{i,j}} - q_{w_{i,j}} + q_{S_{i,j}} - q_{l_{i,j}} + \sum q_{adj}$$

with m the mass of air in each room, C_p its specific heat capacity, $q_{adj} = aU(T_{adj} - T_{i,j})$, with a the area of the wall between rooms, and U the overall heat transfer coefficient between adjoining rooms, which depends on the wall materials. To have symmetric inputs, we combine the heat transfers in pairs: $q_h - q_{AC} =: Q_{hAC}u_{hAC}$, $q_{d_{i,j}} - q_{w_{i,j}} =: Q_{dw}u_{dw}^{i,j}$, and $q_{S_{i,j}} - q_{l_{i,j}} =: Q_{Sl}u_{Sl}^{i,j}$ with u_{hAC}, $u_{dw}^{i,j}$, and $u_{Sl}^{i,j} \in [-1, 1]$.

We write the dynamics as $\dot{T} = AT + \bar{B}\bar{u}$, with

$$A = \frac{a}{mC_p}\begin{pmatrix} -2U & U & 0 & 0 & \dots & 0 & U & 0 & 0 & \dots \\ U & -3U & U & 0 & \dots & 0 & 0 & U & 0 & \dots \\ 0 & \ddots & \ddots & \ddots & & \ddots & \ddots & \ddots & & \end{pmatrix},$$

$$\bar{B} = \frac{1}{mC_p}\begin{pmatrix} Q_{Sl}I_{kl,kl} & Q_{dw}I_{kl,kl} & Q_{hAC}\mathbf{1}_{kl} \end{pmatrix},$$

$\bar{u}^\top = \left(u_{Sl}^{1,1}, \dots, u_{Sl}^{k,l}, u_{dw}^{1,1}, \dots, u_{dw}^{k,l}, u_{hAC}\right) \in \mathbb{R}^{2kl+1}$ and $T^\top = \left(T_{1,1}, \dots, T_{k,l}\right) \in \mathbb{R}^{kl}$. To perform numerical calculations, we restrict our building to $k = 1$ and $l = 3$, as schematized in Fig. 7.

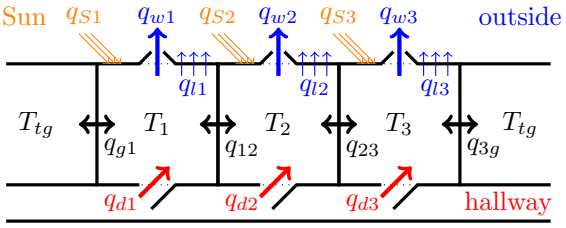


Fig. 7. Scheme of the rooms and of the heat transfers. The heater q_h and AC transfers q_{AC} are not shown for clarity.

Taking $x := T - T_{tg}$, the heat dynamics of the system illustrated on Fig. 7 are $\dot{x} = Ax + \bar{B}\bar{u}$ with $x_{tg} = 0$ and

$$A = \frac{a}{mC_p}\begin{pmatrix} -U_{g1} - U_{12} & U_{12} & 0 \\ U_{12} & -U_{12} - U_{23} & U_{23} \\ 0 & U_{23} & -U_{23} - U_{3g} \end{pmatrix}.$$

Based on [35], we use the following values: $a = 12\,m^2$, $mC_p = 42186\,J/K$, $U_{g1} = 6.27\,W/K$, $U_{12} = 5.08\,W/K$, $U_{23} = 5.41\,W/K$, $U_{3g} = 6.27\,W/K$, $Q_{hAC} = 350\,W$, $Q_{dw} = 300\,W$, $Q_{Sl} = 200\,W$, and $T_{tg} = 293\,K$.

Since $\lambda(A) = \{-0.052, -0.033, -0.010\} \subseteq \mathbb{R}^-$, A is Hurwitz. Then, according to Theorem 6, the system is not resilient, but it might be resiliently stabilizable. For the loss of any one column C, $\text{rank}(B) = 3$ and we numerically verify that $-C\mathcal{W} \subseteq \text{int}(B\mathcal{U})$. Then, following Lemma 3, $\dim(\mathcal{Z}) = 3$, so $\text{int}(\mathcal{Z}) \neq \emptyset$. According to Proposition 1, the system is resiliently stabilizable.

The controller wants to cool the building overnight from an initial state $x_0^\top = \left(0.8°C,\ 0.7°C,\ 0.9°C\right)$. However, a worker is overriding u_{dw}^1 by manually opening the door and window in room 1. We now compare the analytical bounds on the nominal and malfunctioning reach times of Section 6 with the numerical results of [17, 32]. Our bounds require pairs $P \succ 0$ and $Q \succ 0$ solutions of $A^\top P + PA = -Q$. We generate randomly a thousand of such pairs (P, Q) and compute bounds on T_N^* with (11) and (13), and on T_M^* with (15) and (16). Another way of choosing P relies on the linearization of (15), which yields $T_M^* \geq \frac{\|x_0\|_P}{z_{max}^P}$. This bound is maximized when $P \succ 0$ is the tightest ellipsoidal approximation of

$\mathcal{Z}$, which results in much tighter bound than stochastic P, as shown on Fig. 8.

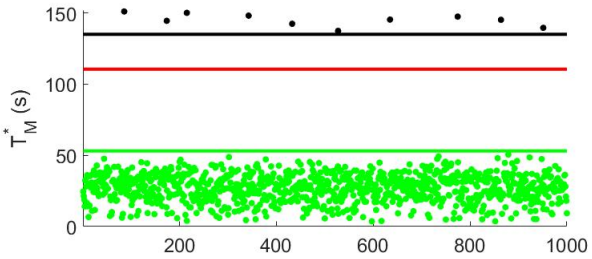


Fig. 8. Bounds on the malfunctioning reach time $T^*_M(x_0)$ in red. The dots are the upper (16) and lower bounds (15) for 1000 stochastic pairs (P, Q). The tightest bounds in green and black result from the ellipsoidal approximations of $\mathcal{Z}$.

For the given x_0 the best bounds on the reach times are $35.5\,s \leq T^*_N(x_0) = 42.5\,s \leq 54.1\,s$ and $53\,s \leq T^*_M(x_0) = 110.5\,s \leq 135\,s$. Then, the rooms can take up to $T^*_M(x_0)/T^*_N(x_0) = 2.6$ times longer to all reach T_{tg} from the initial state $T_{tg} + x_0$ after the loss of control authority over u^1_{dw}, while our bounds predict a worst-case factor of 3.8.

We were able to compute numerically $T^*_N(x_0)$ [17] and $T^*_M(x_0)$ [32], but accessing r_q can only be done analytically with Theorems 7 and 8. Over all $x_0 \in \mathbb{R}^3$, they predict $r_q \in [0.166, 0.979]$. Hence, the loss of control over u^1_{dw} can render the damaged system up to $1/0.166 = 6$ times slower to reach the target temperature from any initial state. This information could not be obtained with prior work and is the motivation for our analytical bounds in Section 6.

If instead of losing control over u^1_{dw} a disgruntled worker takes over the central heating/AC unit u_{hAC}, the rooms can take as much as $T^*_M(x_0)/T^*_N(x_0) = 4.7$ times longer to reach T_{tg} from the same initial temperature, while our bound predicts a max ratio of 9.3. These values are larger than for the loss of u^1_{dw} because $Q_{hAC} > Q_{dw}$ and the central heating/AC affects directly all 3 rooms. Additionally, Theorem 7 yields $r_q \in [0.1, 0.37]$, so the malfunctioning controller can take between 2.7 and 10 times longer than nominally to enforce the target temperature from any initial condition.

8 Conclusion and Future Work

This paper establishes novel necessary and sufficient conditions for the resilient stabilizability and reachability of affine targets by linear systems. Additionally, we quantified the resilience of control systems to the loss of authority over some of their actuators.

There are several avenues of future work. Building on our resilient stabilizability conditions, we have started to work on the resilience of networks to a partial loss of control authority over actuators of a subsystem. Another interesting problem is to ensure the safety of critical systems by preventing them from visiting dangerous locations while completing their mission even after enduring a loss of control. Future work should also aim at extending resilience theory to nonlinear systems. The main hurdle to this last project is to establish a new proof of Hájek's duality theorem. Indeed, this result is essential for resilience theory and its current proof relies on the linearity of the dynamics, hence preventing a straightforward extension to nonlinear systems.

A Supporting Lemmata

In this appendix we provide supporting results concerning sets $B\mathcal{U}$, $C\mathcal{W}$, and $\mathcal{Z}$ defined in Section 3.

Lemma 1 *The interior of $\mathcal{Z}$ is non-empty if and only if* $0 \in \text{int}(\mathcal{Z})$.

Proof. Since $\mathcal{Z}$ is convex and symmetric, so is its interior [30]. If $\text{int}(\mathcal{Z}) \neq \emptyset$, there exists $z \in \text{int}(\mathcal{Z})$, by symmetry $-z \in \text{int}(\mathcal{Z})$, and $0 \in \text{int}(\mathcal{Z})$ by convexity. The reverse implication is trivial. ■

Lemma 2 *The following statements are equivalent: (a)* $0 \in \text{relint}(\mathcal{Z})$, *(b)* $0 \in \mathcal{Z}$, *(c)* $\mathcal{Z} \neq \emptyset$, *(d)* $\text{relint}(\mathcal{Z}) \neq \emptyset$.

Proof. Since $\text{relint}(\mathcal{Z}) \subseteq \mathcal{Z}$, we have $(a) \implies (b)$ and trivially, $(b) \implies (c)$. Since $\mathcal{Z}$ is a convex subset of $\mathbb{R}^n$, $(c) \implies (d)$ according to Lemma 7.33 of [1]. Because $\mathcal{Z}$ is convex and symmetric, so is its relative interior according to [30]. Then, the same proof as for Lemma 1 yields $(d) \implies (a)$ which completes the proof. ■

Definition 7 *The dimension of a compact set $\mathcal{S}$ is the dimension of the smallest affine subspace (with respect to inclusion) containing $\mathcal{S}$ [1].*

Lemma 3 *The relative interior of $B\mathcal{U}$ contains $-C\mathcal{W}$ if and only if* $\dim(\mathcal{Z}) = \text{rank}(B)$.

Proof. Let $q := \dim(B\mathcal{U}) \leq n$. Since $\mathcal{U} = [-1, 1]^{m-p}$, its interior is not empty in $\mathbb{R}^{m-p}$ and thus $q = \text{rank}(B)$. Take q linearly independent vectors of $B\mathcal{U}$ denoted by $B_q := (b_1, \ldots, b_q)$ and pick $V := (v_{q+1}, \ldots, v_n) \in \mathbb{R}^{n\times(n-q)}$ such that $T_b := (B_q, V)$ is invertible. Then, T_b is a transition matrix with $T_b e_i = b_i$ for $i \in [\![1, q]\!]$.

Assume first that $-C\mathcal{W} \subseteq \text{relint}(B\mathcal{U})$. Then, there exists $\varepsilon > 0$ such that $T_b\big(\mathbb{B}^q(0, \varepsilon) \times \{0\}^{n-q}\big) \oplus -C\mathcal{W} \subseteq B\mathcal{U}$. Informally, $-C\mathcal{W}$ remains in $B\mathcal{U}$ when it is 'extended' by ε in all q dimensions of $B\mathcal{U}$. Because $\mathcal{Z} = \big\{z \in \mathbb{R}^n : \{z\} \oplus -C\mathcal{W} \subseteq B\mathcal{U}\big\}$, we have $T_b\big(\mathbb{B}^q(0, \varepsilon) \times \{0\}^{n-q}\big) \subseteq \mathcal{Z}$. Then, $q \leq \dim(\mathcal{Z})$. Since $0 \in -C\mathcal{W}$, $\mathcal{Z} \subseteq B\mathcal{U}$, and hence $\dim(\mathcal{Z}) \leq q$. Thus, $\dim(\mathcal{Z}) = q = \text{rank}(B)$.

On the other hand, assume that $\dim(\mathcal{Z}) = q$. Since $0 \in -C\mathcal{W}$, $\mathcal{Z} \subseteq B\mathcal{U}$. Then, $\mathcal{Z}$ being of same dimension and included in $B\mathcal{U}$ yields that $(b_1, \ldots, b_q)$ is also a basis of $\text{span}(\mathcal{Z}) = \text{Im}(B)$. Hence, T_b is a transition matrix from $\mathbb{R}^n$ to $\text{span}(\mathcal{Z})$. According to Lemma 2, $0 \in \text{relint}(\mathcal{Z})$, i.e, there exists $\delta > 0$ such that $T_b\big(\mathbb{B}^q(0, \delta) \times \{0\}^{n-q}\big) \subseteq$

$\mathcal{Z}$. As above, the definition of $\mathcal{Z}$ yields $T_b\big(\mathbb{B}^q(0,\delta)\times\{0\}^{n-q}\big)\oplus(-C\mathcal{W})\subseteq B\mathcal{U}$. Because $\dim(\mathbb{B}^q(0,\varepsilon))=q=\dim(B\mathcal{U})$, we have $-C\mathcal{W}\subseteq\operatorname{relint}(B\mathcal{U})$. ■

Lemma 4 *If* $\dim(\mathcal{Z})=\operatorname{rank}(B)$, *then* $\operatorname{span}(\mathcal{Z})=\operatorname{Im}(B)=\operatorname{Im}(\bar{B})$.

Proof. In the proof of Lemma 3 we showed that $\operatorname{span}(\mathcal{Z})=\operatorname{Im}(B)$. The inclusion $-C\mathcal{W}\subseteq\operatorname{relint}(B\mathcal{U})$ holds according to Lemma 3 and yields $\operatorname{Im}(C)\subseteq\operatorname{Im}(B)$, and since $\bar{B}=[B\ C]$ after adequate column permutations, we have $\operatorname{Im}(\bar{B})=\operatorname{Im}([B\ C])=\operatorname{Im}(B)$. ■

Lemma 5 *Set* $\mathcal{Z}$ *is empty if and only if set* $C\mathcal{W}$ *is not entirely included in* $B\mathcal{U}$, *i.e.,* $\mathcal{Z}=\emptyset\iff C\mathcal{W}\nsubseteq B\mathcal{U}$.

Proof. If $\mathcal{Z}=\emptyset$, then by definition, for all $z\in B\mathcal{U}$, there exists $w\in\mathcal{W}$ such that $z-Cw\notin B\mathcal{U}$. Taking $z=0$ yields $C\mathcal{W}\nsubseteq B\mathcal{U}$.

On the other hand, assume that there exists $w\in\mathcal{W}$ such that $Cw\notin B\mathcal{U}$. Assume for contradiction purposes that $\mathcal{Z}\neq\emptyset$. Then, we can take $z\in\mathcal{Z}$ and $z-Cw\in B\mathcal{U}$. Since $B\mathcal{U}$ is symmetric, we thus have $-z+Cw\in B\mathcal{U}$. Because $z\in\mathcal{Z}$ and $-w\in\mathcal{W}$, we also have $z+Cw\in B\mathcal{U}$. The convexity of $B\mathcal{U}$ yields $\frac{1}{2}(-z+Cw)+\frac{1}{2}(z+Cw)\in B\mathcal{U}$, i.e., $Cw\in B\mathcal{U}$ which contradicts our first assumption. Hence, $\mathcal{Z}=\emptyset$. ■

Lemma 6 *If* $C\mathcal{W}\nsubseteq B\mathcal{U}$, *then system* (2) *is not resiliently stabilizable.*

Proof. Since $C\mathcal{W}\nsubseteq B\mathcal{U}$, there exists $w\in\mathcal{W}$ such that $Cw\notin B\mathcal{U}$. The sets $\{Cw\}$ and $B\mathcal{U}$ are nonempty, disjoint, convex, and compact, hence they are strongly separated according to Theorem 5.79 of [1]. Then, there exists $v\in\mathbb{R}^n$, $v\neq0$, $c>0$, and $\varepsilon>0$ such that $\langle Cw,v\rangle\geq c+\varepsilon$, and for all $u\in\mathcal{U}$, $\langle Bu,v\rangle\leq c-\varepsilon$. Because $B\mathcal{U}$ and $C\mathcal{W}$ are symmetric, $\{-Cw\}$ and $B\mathcal{U}$ are also strongly separated by the symmetric hyperplane: $\langle -Cw,v\rangle\leq -c-\varepsilon$ and for all $u\in\mathcal{U}$, $\langle Bu,v\rangle\geq -c+\varepsilon$.

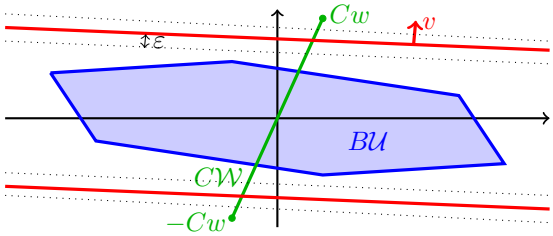


Fig. A.1. Illustration of the strong separation of sets $B\mathcal{U}$ (blue) and $\{\pm Cw\}$ (green) by symmetric hyperplanes.

If $A\neq0$, then $\|A\|>0$. Since $v\neq0$, we can define $r:=\frac{\varepsilon}{\|v\|\,\|A\|}>0$. We will show that if $x\in\mathbb{B}^n(0,r)$, then no controls $u\in\mathcal{U}$ can bring the state x closer to the origin. Let $x\in\mathbb{B}^n(0,r)$ and first assume that $\langle x,v\rangle\geq0$. Then, we apply the undesirable input w and any control $u\in\mathcal{U}$ to system (2)

$$\begin{aligned}\langle\dot{x},v\rangle&=\langle Ax,v\rangle+\langle Bu,v\rangle+\langle Cw,v\rangle\\&\geq-\|Ax\|\,\|v\|-c+\varepsilon+c+\varepsilon\\&\geq-\|A\|\,\|x\|\,\|v\|+2\varepsilon\geq\varepsilon,\end{aligned}$$

where we used the Cauchy-Schwarz inequality [20], the definition of $\|A\|$ and $\|x\|\leq r$. Similarly, if $\langle x,v\rangle<0$, we apply the undesirable input $-w$ and any control $u\in\mathcal{U}$ to system (2)

$$\begin{aligned}\langle\dot{x},v\rangle&=\langle Ax,v\rangle+\langle Bu,v\rangle+\langle -Cw,v\rangle\\&\leq\|A\|\,\|x\|\,\|v\|+c-\varepsilon-c-\varepsilon\\&\leq r\|A\|\,\|v\|-2\varepsilon=-\varepsilon.\end{aligned}$$

Thus, the state $x\in\mathbb{B}^n(0,r)$ can be pushed away from the origin along v. Hence, system (2) is not stabilizable.

If $A=0$, we can take any $x\in\mathbb{R}^n$ such that $\langle x,v\rangle\geq0$ (resp. ≤0) and obtain $\langle\dot{x},v\rangle\geq2\varepsilon$ (resp. $\leq-2\varepsilon$) so the same conclusion holds. ■

Acknowledgment

The authors thank Dr. Bordignon and Dr. Durham for providing us with the ADMIRE model and Dr. Althoff for his help concerning zonotopes.

References

[1] C. Aliprantis and K. Border. *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer, New York, 2006.

[2] M. Althoff. On computing the Minkowski difference of zonotopes. *arXiv preprint arXiv:1512.02794*, 2015.

[3] M. Althoff, N. Kochdumper, and M. Wetzlinger. CORA 2020 manual. *TU Munich*, 2016.

[4] A. A. Amin and K. M. Hasan. A review of fault tolerant control systems: advancements and applications. *Measurement*, 143:58 – 68, 2019.

[5] B. Anderson and A. Dehghani. Challenges of adaptive control–past, permanent and future. *Annual Reviews in Control*, 32:123 — 135, 2008.

[6] M. Athans. The status of optimal control theory and applications for deterministic systems. *IEEE Transactions on Automatic Control*, 11(3):580 – 596, 1966.

[7] M. Bartels. Russia says 'software failure' caused thruster misfire at space station. *https://www.space.com/space-station-nauka-arrival-thruster-fire-update*, 2021.

[8] W. Borgest and P. Varaiya. Target function approach to linear pursuit problems. *IEEE Transactions on Automatic Control*, 16(5):449 – 459, 1971.

[9] J.-B. Bouvier and M. Ornik. Resilient reachability for linear systems. In *21st IFAC World Congress*, pages 4409 – 4414, 2020.

[10] J.-B. Bouvier and M. Ornik. Designing resilient linear systems. *IEEE Transactions on Automatic Control*, 67(9):4832 – 4837, 2022.

[11] J.-B. Bouvier and M. Ornik. The maximax minimax quotient theorem. *Journal of Optimization Theory and Applications*, 192:1084 – 1101, 2022.

[12] J.-B. Bouvier and M. Ornik. Quantitative resilience of linear systems. In *20th European Control Conference*, pages 485 – 490, 2022.

[13] J.-B. Bouvier, K. Xu, and M. Ornik. Quantitative resilience of linear driftless systems. In *SIAM Conference on Control and its Applications*, pages 32 – 39, 2021.

[14] J.-B. Bouvier, K. Xu, and M. Ornik. Quantitative resilience of generalized integrators. *in review*, https://arxiv.org/abs/2111.04163.

[15] R. F. Brammer. Controllability in linear autonomous systems with positive controllers. *SIAM Journal on Control*, 10(2):339 – 353, 1972.

[16] J. Davidson, F. Lallman, and T. Bundick. Real-time adaptive control allocation applied to a high performance aircraft. In *5th SIAM Conference on Control and Its Applications*, 2001.

[17] J. H. Eaton. An iterative solution to time-optimal control. *Journal of Mathematical Analysis and Applications*, 5(2):329 – 344, 1962.

[18] L. Forssell and U. Nilsson. ADMIRE: The aero-data model in a research environment version 4.0, model description. Technical report, FOI - Swedish Defence Research Agency, December 2005.

[19] A. Girard, C. Le Guernic, and O. Maler. Efficient computation of reachable sets of linear time-invariant systems with inputs. In *International Workshop on Hybrid Systems: Computation and Control*, pages 257 – 271. Springer, 2006.

[20] G. Golub and C. Van Loan. *Matrix Computations*. John Hopkins University Press, 2013.

[21] O. Hájek. Duality for differential games and optimal control. *Mathematical Systems Theory*, 8(1):1 – 7, 1974.

[22] O. Härkegård and S. T. Glad. Resolving actuator redundancy - optimal control vs. control allocation. *Automatica*, 41:137 – 144, 2005.

[23] M. Heymann, M. Pachter, and R. Stern. Max-min control problems: A system theoretic approach. *IEEE Transactions on Automatic Control*, 21(4):455 – 463, 1976.

[24] Y.-C. Ho. Review of the book Differential Games by R. Isaacs. *IEEE Transactions on Automatic Control*, 10:501 – 503, 1965.

[25] R. E. Kalman and J. E. Bertram. Control system analysis and design via the "second method" of Lyapunov: continuous-time systems. *Journal of Basic Engineering*, 82(2):371 – 393, 1960.

[26] H. K. Khalil. *Nonlinear Systems*. Prentice Hall, 2002.

[27] I. Kolmanovsky and E. G. Gilbert. Theory and computation of disturbance invariant sets for discrete-time linear systems. *Mathematical Problems in Engineering*, 4(4):317 – 367, 1998.

[28] S. G. Krantz. *A handbook of real variables: with applications to differential equations and Fourier analysis*. Springer Science & Business Media, 2011.

[29] D. Liberzon. *Calculus of Variations and Optimal Control Theory: a Concise Introduction*. Princeton University Press, 2011.

[30] M. Moszynska. *Selected Topics in Convex Geometry*. Springer, 2006.

[31] E. Rechtschaffen. Equivalences between differential games and optimal controls. *Journal of Optimization Theory and Applications*, 18(1):73 – 79, 1976.

[32] Y. Sakawa. Solution of linear pursuit-evasion games. *SIAM Journal on Control*, 8(1):100 – 112, 1970.

[33] W. Schmitendorf and B. Elenbogen. Constrained max-min controllability. *IEEE Transactions on Automatic Control*, 27(3):731 – 733, 1982.

[34] G. Tao, S. Chen, and S. M. Joshi. An adaptive actuator failure compensation controller using output feedback. *IEEE Transactions on Automatic Control*, 47(3):506 – 511, 2002.

[35] S. H. Trapnes. Optimal Temperature Control of Rooms. Master's thesis, Norwegian University of Science and Technology, 2012.

[36] L. Y. Wang and J.-F. Zhang. Fundamental limitations and differences of robust and adaptive control. In *2001 American Control Conference*, pages 4802 – 4807, 2001.

[37] B. Xiao, Q. Hu, and P. Shi. Attitude stabilization of spacecrafts under actuator saturation and partial loss of control effectiveness. *IEEE Transactions on Control Systems Technology*, 21(6):2251 – 2263, 2013.