2018 IEEE Conference on Decision and Control (CDC)
Miami Beach, FL, USA, Dec. 17-19, 2018

Constraining Attackers and Enabling Operators via Actuation Limits

Sahand Hadizadeh Kafash[1], Navid Hashemi[1], Carlos Murguia[2], and Justin Ruths[1]

Abstract— **While all physical systems have bounds on actuator capabilities, imposing artificial limits on the inputs of a system can reduce potential damage caused by system disturbances, including strategic attacks, by avoiding dangerous states. These artificial bounds may, however, affect the performance of the system and may make states required for operation unreachable. To solve this conflict, we construct Linear Matrix Inequalities (LMIs) and exploit some unexpected structure in the solution space to design limits on inputs such that the reachable set of the system state includes desired operation states and avoids dangerous states. We demonstrate the performance of our methods through a numerical example and case study.**

I. INTRODUCTION

Conventional security methods such as encryption and authentication typically guard the outer layer of modern control systems. However, these security systems are complex systems and suffer from the same vulnerabilities as any computing platform, which provides "back doors" into the inner layers of the control system. Unfortunately, the idea that attackers might exploit these vulnerabilities to target control systems is no longer a looming threat, but is a reality and documented in various cases spanning chemical, power, and automobile applications [1], [2], [3]. Our work is guided by the recurring themes in robustness and security: redundancy and orthogonality of defenses. When the outer, cryptographic safeguards fail, we aim to leverage information about the system dynamics and the control loop to identify and mitigate attacks on control systems. The complement of computer science and control theoretic tools will provide an even stronger defense for control applications in cyber-physical systems.

A large part of control theoretic methods for attack detection and mitigation is to use modified tools from fault detection [4], [5], [6]. If an attacker chooses to stay stealthy, or hidden, to these anomaly detectors, this places limits on what he or she is able to do to the system. This current work takes a different perspective. We consider designing limits to the actuation of the system to thereby limit what the attacker can do to the system. Actuators inherently have limits on their actuation ability, e.g., limits on torque and pump rates for motors and pumps, so we study whether we can reduce an attacker's capability to damage the system if we impose more restrictive limits on actuation. This work is a continuation of [7] in which we study this exact question. In past work, we did not consider the obvious complication that by limiting actuation to constrain attackers, we simultaneously constrain the capabilities of operators. Here we begin to address this dual objective (Section III-C) by leveraging some unexpected structure (Section III-B) in the solution of the problems posed in [7]. We improve (simplify) the methods used previously (Section III-A), which are based on Linear Matrix Inequality (LMI) formulations of Lyapunov-type ellipsoidal functions to quantify the reachable set of the system with actuator bounds in effect. In Section II we review the problem formulation, literature on reachable set computation, and summarize past work.

II. REACHABLE SET APPROXIMATIONS

We study discrete-time Linear Time-Invariant (LTI) systems of the form

$$x_{k+1} = Ax_k + Bu_k, \tag{1}$$

with time step $k \in \mathbb{N}$; state vector $x_k \in \mathbb{R}^n$; state matrix $A \in \mathbb{R}^{n\times n}$, input matrix $B \in \mathbb{R}^{n\times m}$ and control input $u_k \in \mathbb{R}^m$.

If the pair (A, B) in (1) is controllable, it is well known that there is an input sequence that drives the state from zero to any state in $\mathbb{R}^n$, i.e., all states are reachable: for all $\xi \in \mathbb{R}^n$ there exits a $N \in \mathbb{N}$ and $\{u_k\}_{k=1}^N$ such that $x_N = \xi$. However, such state transitions may require arbitrarily large input energy and, in any real-world application, actuators have limited capacity (e.g., motors have a maximum torque they can apply). We capture these bounds on actuation ability as symmetric bounds on the control inputs,

$$[u_k]_i^2 \leq \gamma_i, \qquad i = 1, \dots, m, \tag{2}$$

where $\gamma_i \geq 0$ denotes the natural bounds (or - anticipating the next section - artificial bounds) on the i^{th} control input at time k. In bounding the control inputs, the reachable set is in turn bounded, even though the pair (A, B) is controllable. The reachable set is given by

$$\mathcal{R} := \left\{ x_k \in \mathbb{R}^n \;\middle|\; \begin{array}{l} x_{k+1} = Ax_k + Bu_k,\ x_1 = \mathbf{0}, \\ [u_k]_i^2 \leq \gamma_i,\ i = 1, \dots, m,\ \forall k \in \mathbb{N} \end{array} \right\}. \tag{3}$$

There have been many iterative approaches to generating approximately exact reachable sets of LTI systems with bounded input/disturbance (in the sense that the approximation becomes more exact as more iterations are taken) [8], [9], [10]. Depending on the linear transformation described by the matrices A and B, the resulting reachable set can be quite complicated. As we are interested in accomplishing

*This work was not supported by any organization

[1]These authors are with the Departments of Mechanical and Systems Engineering at the University of Texas at Dallas, Richardson, Texas, USA `Sahand.HadizadehKafash, jruths @utdallas.edu`

[2]This author is with the iTrust Center at Singapore University of Technology and Design, Singapore `murguia_rendon@sutd.edu.sg`

978-1-5386-1395-5/18/$31.00 ©2018 IEEE

design of the actuator bounds using reachable sets, we look for analytic, closed-form tools to approximate the reachable set rather than using iterative tools. In particular, we aim to find an ellipsoid which encapsulates the entire reachable set. This ellipsoid is an upper/outer estimation of the reachable set and defined by

$$\mathcal{E}(P,c) := \left\{x \in \mathbb{R}^n \mid (x-c)^T P (x-c) \leq 1\right\}, \qquad (4)$$

where c is the center of the ellipsoid and P is the positive-definite shape matrix of the ellipsoid. When the center is at origin we omit writing it out, i.e., $\mathcal{E}(P,0) = \mathcal{E}(P)$. This approach for using outer ellipsoidal bounds is also well utilized for approximating the reachable set and our approach follows the methodologies found in, e.g., [11], [12]. Using support functions is another alternative method for analytic reachable set approximations [13].

In past work [7], given natural input bounds we identified the optimal ellipsoidal bound on the reachable set (3). To express this concisely, notice that the individual bounds on control inputs in (2) can be expressed in matrix notation as,

$$\frac{1}{m} u_k^T R u_k \leq 1, \qquad (5)$$

where $R = \operatorname{diag}\left(\frac{1}{\gamma_1}, \ldots, \frac{1}{\gamma_m}\right)$ collects the upper bounds.

Proposition 1 ([7]): For the LTI system (1) with controllable pair (A,B), and upper bounds $\gamma_i \geq 0$, $i = 1, \ldots, m$ collected in R, if there exists an $a \in (0,1)$ for which the positive definite matrix P is a solution of the following convex optimization problem:

$$\begin{cases} \min\limits_{P} \ -\log\det P, \\ \text{s.t. } P > 0, \text{ and} \\ \begin{bmatrix} aP - A^T P A & -A^T P B \\ -B^T P A & \frac{1-a}{m} R - B^T P B \end{bmatrix} \geq 0, \end{cases} \qquad (6)$$

then $\mathcal{R} \subseteq \mathcal{E}(P)$ and $\mathcal{E}(P)$ has minimum volume.

Note that the reachable set, and thus the bounding ellipsoid, are unbounded if the system is open-loop unstable.

III. SYNTHESIS OF ACTUATOR BOUNDS

As discussed in previous section, for a given set of individual actuator bounds we are able to find an outer ellipsoidal approximation of the reachable set of the system. We now turn to design individual actuator bounds to reshape the reachable set of the system. In particular we aim to find a new diagonal matrix $R \geq R_0 = \operatorname{diag}\left(\frac{1}{\gamma_1^0}, \ldots, \frac{1}{\gamma_m^0}\right)$ such that the new bounds $\gamma_i \leq \gamma_i^0$, $i = 1, \ldots, m$, where γ_i^0 are the natural limits of the actuators, lead to a new reachable set that avoids or includes certain portions of state space.

In [7], we developed a method to ensure that the reachable set avoids specific regions of state space, which we call *dangerous states*. At that time we overlooked the inherent complication of imposing new bounds on inputs – that restricting inputs beyond their natural limitations may affect the system's performance, function, and may even compromise stability. We can capture those states that are required for the normal operation of the system in a set of *operation states*. Restricting the bounds such that the reachable set does not include the operation states would, by definition, limit the efficiency, productivity, or stability of the system. Thus the goal of this paper is to synthesize individual actuator bounds to achieve a balance between safety (avoiding dangerous states) and performance (including all necessary operational states) of the system.

The idea of demarcating regions of state space that should be maintained or avoided is not new, nor is their use with reachable sets. However, the work we have found typically identifies *safe states* (or unsafe), which fuses the idea of dangerous and operation states, without allowing for the important gap between the operating states and those states that are detrimental to the system. Moreover, these studies have focused on finite-time reachable sets and the problem of steering the system (e.g., using dynamic programming for a class of discrete time stochastic hybrid systems [14]; or with applications to quadcopter maneuvers [15]) to stay within the set of safe states.

A. Avoiding Dangerous States

We define the set of dangerous states as the union of half-spaces,

$$\mathcal{D} := \left\{ x \in \mathbb{R}^n \;\middle|\; \bigcup_{i=1}^{\kappa} c_i^T x \geq b_i \right\}, \qquad (7)$$

where each pair (c_i, b_i), $c_i \in \mathbb{R}^n$, $b_i \in \mathbb{R}$, $i = 1, \ldots, \kappa$ is used to define a single half-space. We generalize our results in [7] with a simpler formulation of the optimization that determines individual actuator bounds to avoid the dangerous states, $\mathcal{D}$.

Theorem 1: Consider the LTI system (1) with controllable pair (A,B) and a set of dangerous states $\mathcal{D}$ defined by (7). If there exists an $a \in (0,1)$ for which the positive definite matrix P is a solution of the following convex optimization problem:

$$\begin{cases} \min\limits_{P,R,\lambda} \ \operatorname{trace}(R), \\ \text{s.t. } P > 0, \ R \geq R_0, \text{ and} \\ \begin{bmatrix} aP - A^T P A & -A^T P B \\ -B^T P A & \frac{1}{m}(1-a) R - B^T P B \end{bmatrix} \geq 0, \\ \begin{bmatrix} P & -0.5\lambda c_i \\ -0.5\lambda c_i^T & \lambda b_i - 1 \end{bmatrix} \geq 0, \qquad i = 1, ..., m, \end{cases} \qquad (8)$$

then the new actuator bounds $\gamma_i := (1/[R]_{ii})$, $i = 1, ..., m$, enforce that the resulting reachable set $\mathcal{R}$ does not intersect with the dangerous states $\mathcal{D}$.

Proof: The first LMI in (8) serves to construct P such that it outer bounds the reachable set of the system. This LMI comes directly from Proposition 1 (see [7]).

In order to ensure that the reachable set $\mathcal{R}$ avoids the dangerous states $\mathcal{D}$, a geometrical constraint can be imposed which keeps the ellipsoid $\mathcal{E}(P)$ out of the dangerous states defined by half-spaces. This geometric constraint should guarantee that all states which satisfy $x^T P x \leq 1$ also satisfy

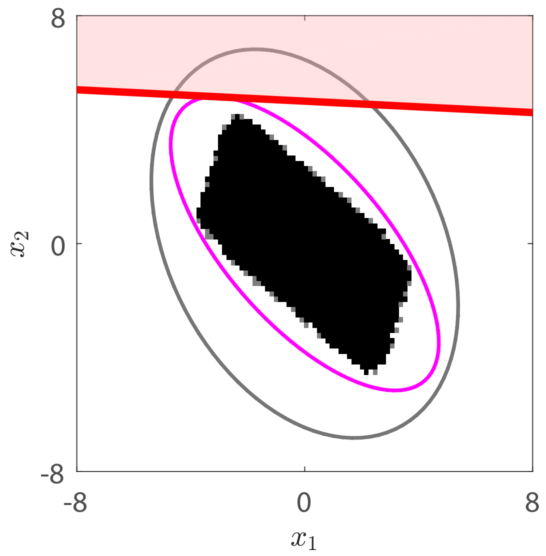

Fig. 1. Proposition 1 provides the gray ellipse that outer bounds the reachable set of the system in Example 1 when inputs are bounded by the natural input bounds $\gamma_1^0 = 75$ and $\gamma_2^0 = 10$. Using Theorem 1 we design new bounds $\gamma_1 = 21.5$ and $\gamma_2 = 8.9$ such that the ellipsoid (magenta) bounding the new reachable set (black) avoids the dangerous states (red). The reachable set (in black) is approximated by Monte Carlo simulations with various bounded inputs.

$c_i^T x \leq b_i$, $i = 1, \dots, m$. The S-procedure provides a way to combine these simultaneous inequalities [16]: these geometrical constraints are satisfied if and only if there exists a non-negative constant λ such that $(x^T P x - 1) - \lambda(c_i^T x - b_i) \geq 0$, which can be written as:

$$\begin{bmatrix} x^T & 1 \end{bmatrix} \underbrace{\begin{bmatrix} P & -0.5\lambda c_i \\ -0.5\lambda c_i^T & \lambda b_i - 1 \end{bmatrix}}_{\mathcal{V}} \begin{bmatrix} x \\ 1 \end{bmatrix} \geq 0. \tag{9}$$

The above inequality is satisfied if $\mathcal{V} \geq 0$. ■

Example 1: Consider the system (1) with matrices

$$A = \begin{bmatrix} -0.33 & -0.1 \\ 0.09 & -0.28 \end{bmatrix}, \quad B = \begin{bmatrix} -0.09 & -0.79 \\ -0.30 & 0.66 \end{bmatrix}, \tag{10}$$

such that the inherent bounds of the system are

$$\gamma_1^0 = 75, \qquad \gamma_2^0 = 10. \tag{11}$$

Moreover consider the dangerous states

$$\mathcal{D} = \{x \mid 0.05x_1 + x_2 \geq 5\} \ \rightarrow \ c_1 = \begin{bmatrix} 0.05 \\ 1 \end{bmatrix}, \ b_1 = 5. \tag{12}$$

First we derive the bounding ellipsoid when the actuator bounds are given by the natural limits, γ^0. Using Proposition 1, we find the gray ellipse in Fig. 1. We observe that some of the reachable states may include some of the dangerous states and then use Theorem 1 to design the bounds $\gamma_1 = 21.5$ and $\gamma_2 = 8.9$ corresponding to the magenta ellipse in Fig. 1, which does not include any dangerous states. This ellipse minimizes the trace of the matrix R.

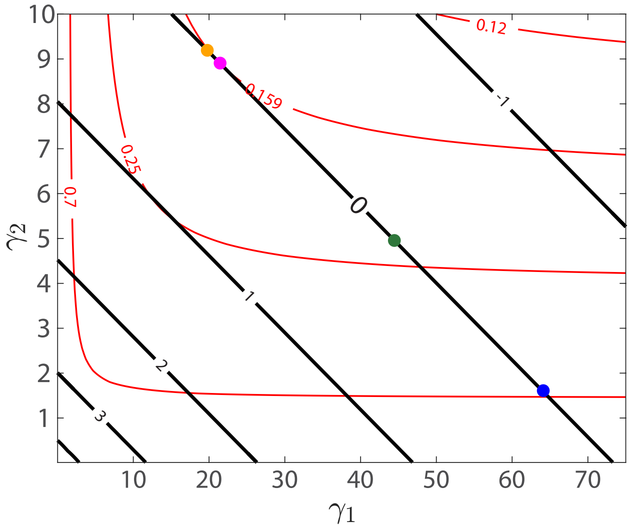

Fig. 2. The level sets (in black) of the minimum distance between bounding ellipsoid and the dangerous states for different choices of input bounds. The surprising linear structure of the level sets provides a way to enumerate all ellipsoids that touch the dangerous states: all choices of bounds along the zero-distance level set. The solution for Example 1 using Theorem 1 (see Fig. 1) is located along this line (magenta dot) and selected to minimize the trace of the matrix R (red curves are level sets of trace(R)). The orange, green, and blue dots denote the bound pairs corresponding to the same color ellipses in Fig. 3 and 4.

B. Insight into Avoiding Dangerous States

In the previous section we defined an optimization with several moving parts. By changing the actuator bounds $\gamma^0 \rightarrow \gamma$, we reshape the set of reachable states $\mathcal{R}$. At the same time, the positive definite matrix P defines an outer ellipsoidal bound $\mathcal{E}(P)$ that contains the new reachable set $\mathcal{R} \subseteq \mathcal{E}(P)$. The glue that connects all these ideas is that P is the solution of a linear matrix inequality that uses a Lyapunov-type function to bound the reachable set. So despite the linearity in the system dynamics, it is rather unexpected that we observe some definitive structure in the solution of the bounds.

To illustrate this structure, we revisit the same system as in (10) with dangerous states given by a single halfspace $\mathcal{D}$ in (12). We use Proposition 1 to generate the bounding ellipsoid corresponding to all possible choices of $\gamma_1 \in [0, \gamma_1^0]$ and $\gamma_2 \in [0, \gamma_2^0]$ by finely discretizing across the values of γ_1 and γ_2. For each we compute the minimum distance between the ellipsoid $\mathcal{E}(P)$ and the dangerous hyperplane defined by c and b

$$d = \frac{|b| - \sqrt{c^T P^{-1} c}}{\sqrt{c^T c}}. \tag{13}$$

The contour plot of this distance as a function of the bounds γ_1 and γ_2 is plotted in Fig. 2 in black. It is easy to see that level-sets of the distance function are lines in the γ_1-γ_2 space. This implies that all choices of γ_1 and γ_2 such that the bounding ellipsoid $\mathcal{E}(P)$ is tangent to the dangerous hyperplane fall along a line. The solution of Theorem 1 provides one point (magenta dot) on this zero-distance line - namely the point for which the value of $\gamma_1^{-1} + \gamma_2^{-1}$

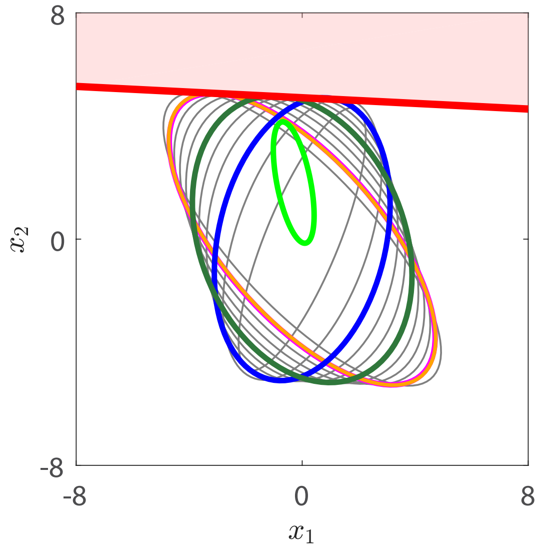

Fig. 3. The pairs of bounds along the zero-distance line in Fig. 2 generate a family of ellipsoids that are tangent to the dangerous states. The point of tangency and the shape matrix of the ellipsoid change with the bounds. For some of the extreme values of the bounds, the ellipsoids do not contain the operation states. The algorithm in Section III-C selects those that do contain the operation states and selects the one that maximizes the containment.

is minimized. In Fig. 2 we overlay the level sets of the cost function trace(R) in red; given the monotonicity of the cost function, there is a unique minimum on the zero-distance line. We could have chosen a different cost function, which would have generated a different solution that falls on this line. The continuum of bounds $\gamma_1 \in [0, \gamma_1^0]$ and $\gamma_2 \in [0, \gamma_2^0]$ on this line generate a family of ellipsoids tangent to the dangerous hyperplane, see Fig. 3. We will exploit this observation to include operation states into our current framework.

This observation trivially generalizes when the dangerous states set is composed of two (or more) half-spaces. In this case the most constraining hyperplane dictates the solution of the bounds. One can imagine overlaying two (or more) distance contour plots, as in Fig. 2, and tracing the most interior zero-distance line (the one closest to $\gamma_1 = \gamma_2 = 0$). Similarly, the extension to generalize this observation to higher-order systems is straightforward in that the linear structure becomes characterized by hyperplanes in the case of higher dimensions.

C. Ensuring Operation States

In this paper we introduce the idea of operation states $\mathcal{O} \subseteq \mathbb{R}^n$ which is a set containing all the states that must be reachable for the proper operation of the system. We define $\mathcal{O}$ by a known ellipsoid:

$$\mathcal{O} := \mathcal{E}(Q, \bar{x}) = \{x \in \mathbb{R}^n \mid (x - \bar{x})^T Q (x - \bar{x}) \le 1\}, \quad (14)$$

where Q is a known positive definite matrix defining the shape of the ellipsoid and $\bar{x}$ is the center of the ellipsoid.

If the problem is posed without dangerous states, then the bounds can be made arbitrarily large and will easily contain the operation states. If the dangerous state constraint is included, the past section observed that all feasible solutions to the problem fall along a line in the space of input bounds. These solutions correspond to the largest ellipsoids possible while avoiding the dangerous states. Thus we leverage the linear structure of these solutions to search through the ellipsoids generated by the bounds along this line.

```
Algorithm III.1: ELLIPSOID($c, b, R_{min}, Q$)

  $\gamma^a \leftarrow$ Theorem 1$(c, b, R_{min})$
  $\gamma^b \leftarrow$ Theorem 1$(c, b, R_{min})$ with different cost
  $\mathcal{L} \leftarrow$ Line$(\gamma^a, \gamma^b)$
  $e \leftarrow 0$;  $P^* \leftarrow \mathbf{0}$
  for $\gamma \in \mathcal{L}$
    do
      $P \leftarrow$ Proposition 1$(\gamma)$
      if $\det H \ge e$                          see (15)
        then
          $e \leftarrow \det H$
          $P^* \leftarrow P$
  return $(P^*)$
```

Algorithm III.1 summarizes the main steps in the method to find the ellipsoid that is tangent to the dangerous states and contains the operation states. Using Theorem 1, we identify a single solution whose input bounds lie on the zero-distance line. Then we can run the same calculation with a slightly modified cost function trace$(R) \rightarrow$ trace(DR), in which the diagonal matrix ($D = I$ is the original definition) specifies a different weighting to be placed between γ_i, $i = 1, \ldots, m$. Using these two points in the space of γ, we can form the line $\mathcal{L}$ that joins these points and extends through to the natural bounds γ^0. We then discretize this line and examine for which choices of γ the optimization in Proposition 1 yields an ellipsoid that contains the operation states. By virtue of the selection of γ along this line, the ellipsoid will be tangent to the dangerous states. We provide the following concise method to determine if one ellipsoid is contained in another.

Lemma 1: Given two ellipsoids $\mathcal{E}(P)$ and $\mathcal{E}(Q, \bar{x})$ characterized by the positive definite matrices P and Q, $\mathcal{E}(Q, \bar{x}) \subset \mathcal{E}(P)$ if and only if

$$H := \begin{bmatrix} \tau Q - P & -\tau Q \bar{x} \\ -\tau \bar{x}^T Q & 1 + \tau(\bar{x}^T Q \bar{x} - 1) \end{bmatrix} > 0 \quad (15)$$

Proof: Given the definition of both ellipsoids, i.e., all values of x for which $x^T P x \le 1$ and $(x - \bar{x})^T Q (x - \bar{x}) \le 1$, respectively, we combine these inequalities together using the S-procedure; the procedure is nearly identical to the proof of Theorem 1. ■

Beyond an easy check for containment, Lemma 1 provides a quantitative measure of how contained $\mathcal{E}(Q, \bar{x})$ is in $\mathcal{E}(P)$. The larger the value of $\det H$, the more contained the inner ellipsoid is (with the inner ellipsoid tangent from the inside when $\det H = 0$). We chose to select the ellipsoid that maximizes this value so that the operation states are well

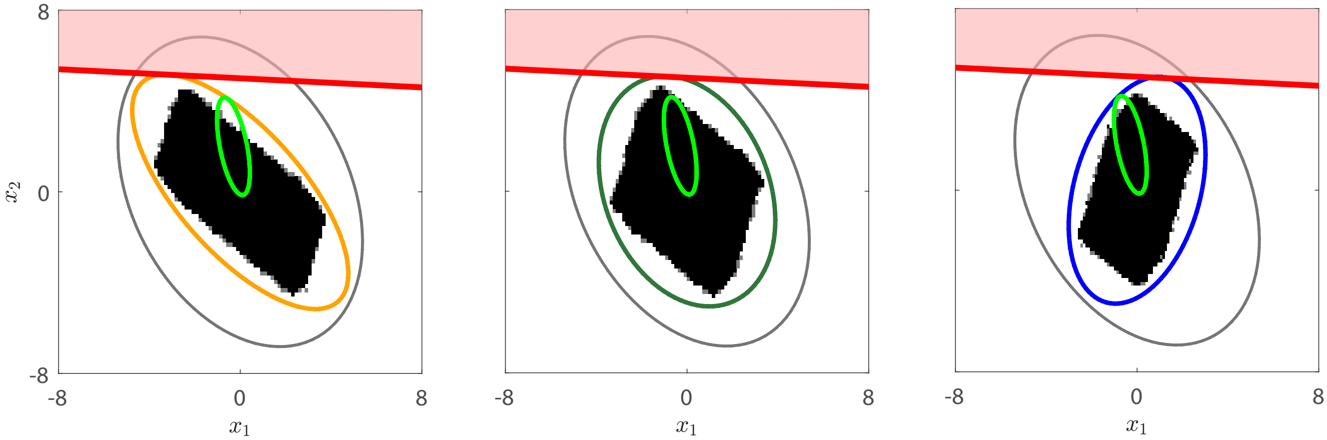

Fig. 4. Three different choices of actuator bounds that cause the bounding ellipsoid to touch the hyperplane that defines the dangerous halfspace. The first ($\gamma_1 = 20$, $\gamma_2 = 9.2$) and third ($\gamma_1 = 65$, $\gamma_2 = 1.7$) show different extremes of bound choices that do include the operation states; the middle case ($\gamma_1 = 45$, $\gamma_2 = 5$) is the proposed solution that aims to locate the operation states as deep as possible into the bounding ellipsoid by maximizing $\det H$, with H defined in (15).

within the reachable set.

The astute reader will now have realized that the bounding ellipsoid containing the operation states does not necessarily imply that the reachable set contains the operation states. Mathematically if $\mathcal{E}(Q,\bar{x}) \subseteq \mathcal{E}(P)$ and $\mathcal{R} \subseteq \mathcal{E}(P)$, then it is not necessarily true that $\mathcal{E}(Q,\bar{x}) \subseteq \mathcal{R}$. The combined "looseness" and "inexact fit" of the ellipsoid to the reachable set means there are portions of the ellipsoid that are not in fact reachable by the system. Our selection of maximizing $\det H$ is a proxy to drive the operation states as deeply within the bounding ellipsoid as possible. This work is a stepping stone to a more comprehensive analytic framework in which we incorporate the operation states into the convex optimization problem.

The results and working of the proposed algorithm are depicted in Fig. 4, building off of Example 1. Here we use operation states $\mathcal{E}(Q,\bar{x})$ characterized by

$$Q = \begin{bmatrix} 2.72 & 0.50 \\ 0.50 & 0.31 \end{bmatrix}, \qquad \bar{x} = \begin{bmatrix} -0.3 \\ 2 \end{bmatrix}. \tag{16}$$

The middle panel shows the result of the algorithm, which designs the input bounds to maximize the containment of the operational states inside the outerbound ellipsoid. The operation states (bright green) fit within both the bounding ellipsoid $\mathcal{E}(P)$ (dark green) and the reachable set $\mathcal{R}$ (black; Monte Carlo simulation). The first and third panels show other feasible solutions, in which the operation states are contained in the bounding ellipsoid and the bounding ellipsoid touches the dangerous states. In the first panel γ_1 is set to its maximum possible value and in the third panel γ_2 is set to its maximum possible value. It is easy to see that in the first and third case, although it is a feasible solution, there are parts of the operation states that are not actually reachable by the system.

IV. CASE STUDY: PLATOONING

To demonstrate our methods further, we consider a platoon of two vehicles whose relative position error $e(t)$ (deviation from the desired distance between two vehicles) and two velocities $v_1(t)$ and $v_2(t)$ evolve according to,

$$\begin{cases} \dot{e}(t) = v_2(t) - v_1(t), \\ \dot{v_1}(t) = w_1(t) + u_1(t) + \beta v_1(t), \\ \dot{v_2}(t) = w_2(t) + u_2(t) + \beta v_2(t), \end{cases} \tag{17}$$

where $w_i(t)$ and $u_i(t)$, $i = 1, 2$, are the feedback and feedforward control signals, respectively; $\beta < 0$ is a friction coefficient. Feedback inputs are typically locally-generated controls using on-vehicle sensor measurements. Here we consider a PD feedback form with k_p and k_d the proportional and derivative gains,

$$\begin{cases} w_1(t) = k_p e(t) + k_d(v_2(t) - v_1(t)), \\ w_2(t) = -k_p e(t) - k_d(v_2(t) - v_1(t)). \end{cases} \tag{18}$$

To improve the string-stability of platoons vehicle-to-vehicle communication is often used, which we model here as the open-loop control of the system. It is this communicated signal that an adversary can potentially intercept and modify with the intent to make the vehicles crash. The bounds on this communicated feedforward "acceleration" signal can be interpreted as either physical constraints of the engine/brakes or as a crude validation check on the magnitude of signal (i.e., if it is too large or too small, we interpret it as corrupted). Our goal here is to adjust those virtual constraints on the feedforward signal in order to avoid the dangerous states.

The continuous time dynamic in (17) can be discretized using a zero-order hold to form a discrete-time LTI system as in (1), with state $x(t) = \begin{bmatrix} e(t), & v_1(t), & v_2(t) \end{bmatrix}^T$ and matrices

$$A = \begin{bmatrix} 0.95 & -0.07 & 0.07 \\ 0.4 & 0.68 & 0.28 \\ -0.4 & 0.28 & 0.68 \end{bmatrix}, \; B = \begin{bmatrix} -0.006 & 0.006 \\ 0.069 & 0.027 \\ 0.027 & 0.069 \end{bmatrix}, \tag{19}$$

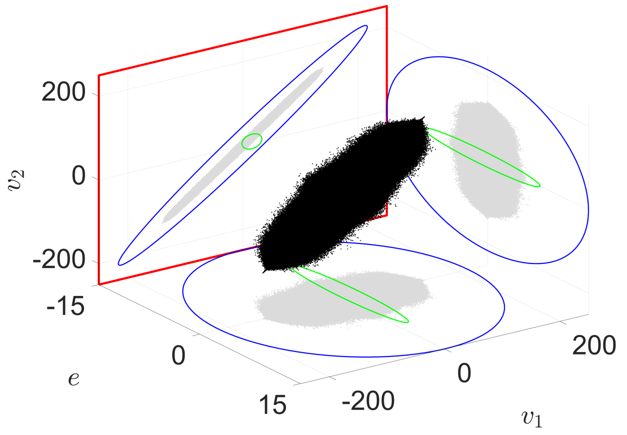

Fig. 5. The operation states (green) are contained within the bounding ellipsoid (blue) for the platoon of two vehicles, while simultaneously avoiding the dangerous state set ($e = -15$ plane) that corresponds to the vehicles crashing. The empirical reachable set (black; Monte Carlo simulation) contains most of the operation states.

when $k_p = 6$, $k_d = 4$, and $\beta = -0.4$. We select the desired distance to be 15 meters, therefore, the vehicles crash when the error $e(t) = -15$. Since we want to avoid crashes, we define the dangerous states

$$\mathcal{D} := \left\{ x \in \mathbb{R}^3 \mid e \leq -15 \right\}. \tag{20}$$

We select operation states such that reasonable vehicle speeds and distance error are included,

$$\begin{aligned} 0 \text{ (m/s)} \leq v_1, v_2 &\leq 20 \text{ (m/s)}, && (21) \\ -5 \text{ (m)} \leq \quad e \quad &\leq 5 \text{ (m)}. && (22) \end{aligned}$$

The minimal ellipsoid that contains the rectangular operation region above is $\mathcal{E}(Q, \bar{x})$ where

$$Q = \begin{bmatrix} \frac{1}{75} & 0 & 0 \\ 0 & \frac{1}{300} & 0 \\ 0 & 0 & \frac{1}{300} \end{bmatrix}, \qquad \bar{x} = \begin{bmatrix} 0 \\ 10 \\ 10 \end{bmatrix}. \tag{23}$$

We employ the proposed algorithm to identify the bounds on the feedforward input that will avoid the dangerous states (vehicles crashing) and ensure the proper operation of the system (by including the operation states). Fig. 5 shows the reachable sets and bounding ellipsoid corresponding to the bounds $\gamma_1 = \gamma_2 = 5183$ m/s^2. The bounds are identical due to the structure of the system (a positive action on one vehicle is equivalent to a negative action of the same magnitude on the other vehicle). The bounds are large in this example, which means that it should be relatively easy to secure this system from attacks against the communication channel. In Fig. 5, we show the projections of the operation state ellipsoid (green) and the bounding ellipsoid (blue) onto the e-v_1, e-v_2, and v_1-v_2 planes (this is visually more clear than plotting the three-dimensional ellipsoids). An empirical approximation of the reachable set (in black and projections in gray) suggests that the reachable set is, indeed, included within the bounding ellipsoid. It also suggests that all operation states might not be included in the reachable set. If we consider the original rectangular bounds in (21)-(22), the containment is much closer, however, potentially not perfect.

V. Conclusions

In this work we have developed tools to design actuator limits to simultaneously avoid dangerous states and ensure the states required for operation of the system. These tools are a mixture of convex optimization techniques expressed as linear matrix inequalities and grid search methods to utilize an unexpected structure in the solution of the partial solution. Future work will aim to incorporate the operation states as part of the main LMI. To do this, we need to develop similar LMI techniques to provide for internal ellipsoidal approximations of the reachable set.

References

[1] R. Langner, "Stuxnet: Dissecting a cyberwarfare weapon," *IEEE Security & Privacy*, vol. 9, no. 3, pp. 49–51, 2011.

[2] K. Zetter, "Inside the cunning, unprecedented hack of ukraines power grid," *Wired Magazine*, 2016.

[3] S. Checkoway, D. McCoy, B. Kantor, D. Anderson, H. Shacham, S. Savage, K. Koscher, A. Czeskis, F. Roesner, T. Kohno *et al.*, "Comprehensive experimental analyses of automotive attack surfaces." in *USENIX Security Symposium*. San Francisco, 2011.

[4] N. Hashemi, C. Murguia, and J. Ruths, "A comparison of stealthy sensor attacks on control systems," in *American Control Conference*, 2018.

[5] A. Cardenas, S. Amin, B. Sinopoli, A. Giani, A. Perrig, and S. Sastry, "Challenges for securing cyber physical systems," in *Workshop on Future Directions in Cyber-physical Systems Security*, 2009.

[6] A. Teixeira, S. Amin, H. Sandberg, K. H. Johansson, and S. S. Sastry, "Cyber security analysis of state estimators in electric power systems," in *Decision and Control (CDC), 2010 49th IEEE Conference on*, 2010, pp. 5991–5998.

[7] S. H. Kafash, J. Giraldo, C. Murguia, A. Cardenas, and J. Ruths, "Constraining attacker capabilities through actuator saturation," in *American Control Conference*, 2018. [Online]. Available: https://arxiv.org/abs/1710.02576

[8] Y. Mo and B. Sinopoli, "On the performance degradation of cyber-physical systems under stealthy integrity attacks," *IEEE Transactions on Automatic Control*, vol. 61, no. 9, pp. 2618–2624, 2016.

[9] A. A. Kurzhanskiy and P. Varaiya, "Ellipsoidal techniques for reachability analysis of discrete-time linear systems," *IEEE Transactions on Automatic Control*, vol. 52, no. 1, pp. 26–38, 2007.

[10] O. Matviychuk, "Ellipsoidal estimates of reachable sets of impulsive control systems with bilinear uncertainty," *Cybernetics and Physics Journal*, vol. 5, no. 3, pp. 96–104, 2016.

[11] N. D. That, P. T. Nam, and Q. P. Ha, "Reachable set bounding for linear discrete-time systems with delays and bounded disturbances," *Journal of Optimization Theory and Applications*, vol. 157, pp. 96–107, 2013.

[12] F. Chernousko, "Ellipsoidal state estimation for dynamical systems," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 63, no. 5-7, pp. 872–879, 2005.

[13] C. Le Guernic and A. Girard, "Reachability analysis of linear systems using support functions," *Nonlinear Analysis: Hybrid Systems*, vol. 4, no. 2, pp. 250–262, 2010.

[14] A. Abate, M. Prandini, J. Lygeros, and S. Sastry, "Probabilistic reachability and safety for controlled discrete time stochastic hybrid systems," *Automatica*, vol. 44, no. 11, pp. 2724–2734, 2008.

[15] J. H. Gillula, H. Huang, M. P. Vitus, and C. J. Tomlin, "Design of guaranteed safe maneuvers using reachable sets: Autonomous quadrotor aerobatics in theory and practice," in *Robotics and Automation (ICRA), 2010 IEEE international conference on*. IEEE, 2010, pp. 1649–1654.

[16] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, ser. Studies in Applied Mathematics. Philadelphia, PA: SIAM, 1994, vol. 15.