2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Secondary Controller Design for the Safety of Nonlinear Systems via Sum-of-Squares Programming*

Yankai Lin, Michelle S. Chong, and Carlos Murguia

Abstract— **We consider the problem of ensuring the safety of nonlinear control systems under adversarial signals. Using Lyapunov-based reachability analysis, we first give sufficient conditions to assess safety, i.e., to guarantee that the states of the control system, when starting from a given initial set, always remain in a prescribed safe set. We consider polynomial systems with semi-algebraic safe sets. Using the S-procedure for polynomial functions, safety conditions can be formulated as a Sum-Of-Squares (SOS) programme, which can be solved efficiently. When safety cannot be guaranteed, we provide tools via SOS to synthesize polynomial controllers that enforce safety of the closed-loop system. The theoretical results are illustrated through numerical simulations.**

I. INTRODUCTION

In recent years, cyber-physical systems have gained increasing attention by researchers due to their wide applicability in modern industrial systems. In these systems, computation of control laws and the physical behavior are coupled via networked communications. One of the pertinent vulnerabilities is when the network is compromised and malicious data is injected into the system. In [1], many examples including the well-known StuxNet malware incident were reported. Therefore, investigation on controller design methods that ensure safety is of significant importance. We aim to address one particular instance of the safety-ensuring control problem, which is to keep the states of the control system within a prescribed set, called a *safe set*, for an infinite time horizon. In an earlier work [2], we considered adding an output feedback dynamic secondary controller to a linear system that has already been stabilized by a predesigned primary controller. In this work, we extend the result to polynomial nonlinear systems with semi-algebraic safe sets using static feedback.

There are many approaches to the safe stabilization problem in existing literature. Among them, reachability analysis is one natural way of ensuring that states from a given initial set are steered into the desired location without entering an unsafe region. However, it is in general computationally expensive to solve these problems exactly due to the associated partial differential equations (PDEs) that need to be dealt with [3], [4]. Another approach that bypasses the difficulty of dealing with PDEs is via tools of set invariance [5]. If there exists a subset of the safe set that is forward invariant, then it is guaranteed that the states of the system always remain within the safe set. Sufficient conditions for the forward invariance of autonomous systems can be given by Lyapunov-like sufficient conditions on functions called barrier certificates [6]. These conditions are later extended in [7], where sufficient conditions on the control barrier function (CBF) are given to guarantee robust forward invariance of the safe sets for nonlinear systems driven by control inputs. Based on these conditions, a control law that guarantees safety can be synthesized by solving a quadratic program online. Tools from dissipativity theory can also be used to formulate a similar condition that verifies the safety of interconnected systems [8].

In this work, we consider systems with polynomial dynamics and take the approach of ensuring forward invariance of a given set using tools from Sum-Of-Squares (SOS) programming to address the safe control problem. There has been successful applications of SOS approach to the robustness analysis of control systems [9], [10]. An additional motivation to use SOS programming is that, though some progress has been made recently [11], [12], the synthesis of CBF for general nonlinear systems is still a challenging problem. In the seminal work [13], it is shown that a SOS program can be reformulated as a hierarchy of semidefinite programming by using SOS polynomials of increasing degree. Hence, by restricting the class of Lyapunov-like functions to be polynomial functions, we can efficiently translate the complicated synthesis problem to a convex optimization program. Similar ideas have been successfully applied to various nonlinear control problems such as the search of polynomial Lyapunov functions to check stability [14].

Our contributions are summarized below.

1) We consider the setup of using limited resources (in terms of limited access to system outputs) to design a secondary controller to ensure the safety of a nonlinear controlled system under resource-limited adversaries. Sufficient conditions in terms of SOS programmes are given to synthesise polynomial state feedback controllers. This generalizes our prior work [2] on linear systems to nonlinear systems with polynomial dynamics.

2) We consider the case where control inputs and external signals appear in the system dynamics. Unlike the previous work [15], where it is assumed that the disturbance has finite energy, we consider sensor and actuator attack signals that are constrained by state-dependent upper bounds. This is to capture the fact that intelligent cyber attackers, depending on their available resources [16], are constrained in the class of

*The research leading to these results has received funding from the European Union's Horizon Europe programme under grant agreement No 101069748 – SELFY project.

The authors are with Department of Mechanical Engineering, Eindhoven University of Technology, the Netherlands. {y.lin2,m.s.t.chong,c.g.murguia}@ tue.nl

signals they can inject to remain stealthy, such that they can continue affecting the system without being detected.

II. PRELIMINARIES

A. Notations

Let $\mathbb{R} = (-\infty, \infty)$, $\mathbb{R}_{\geq 0} = [0, \infty)$, $\mathbb{R}_{>0} = (0, \infty)$ and $\mathbb{R}^n$ denotes the n-dimensional Euclidean space. We use $\mathbf{0}$ to denote the zero matrix with appropriate dimensions. For a given square matrix R, $\mathrm{Tr}[R]$ denotes the trace of R. We use $A \succ 0$ ($A \prec 0$) and $A \succeq 0$ ($A \preceq 0$) to denote the matrix A is positive (negative) definite and positive (negative) semidefinite, respectively. Given a polynomial function $p(x) : \mathbb{R}^n \to \mathbb{R}$, p is called SOS if there exist polynomials $p_i : \mathbb{R}^n \to \mathbb{R}$ such that $p(x) = \sum_{i=1}^{k} (p_i(x))^2$. The set of SOS polynomials and set of polynomials with real coefficients in x are denoted by $\Sigma[x]$ and $\mathbb{R}[x]$, respectively. A vector of dimension n composed of SOS (real) polynomial functions of x is denoted by $\Sigma^n[x]$ ($\mathbb{R}^n[x]$).

B. Preliminaries on Polynomial Functions

A standard SOS program is a convex optimization problem of the following form [17]

$$\min_{m} \; b^\top m \text{ such that } \; p_i(x, m) \in \Sigma[x], \; i = 1, 2, \ldots, n, \quad (1)$$

where $p_i(x, m) = c_{i0}(x) + \sum_{j=1}^{k} c_{ij}(x) m_j$, $c_{ij}(x) \in \mathbb{R}[x]$, and b is a given vector. It is shown in [17, p. 74] that (1) is equivalent to a semidefinite program.

A useful tool that will be used extensively in this paper is the generalization of the S-procedure [18] to polynomial functions. This can be done via the Positivstellensatz certificates of set containment [19]. The proof of the following key result can be found in [20, Chapter 2.2].

Lemma 1 *Given p_0, p_1, $\cdots$, $p_m \in \mathbb{R}[x]$, if there exist λ_1, λ_2, $\cdots$, $\lambda_m \in \Sigma[x]$ such that $p_0 - \sum_{i=1}^{m} \lambda_i p_i \in \Sigma[x]$, then we have $\bigcap_{i=1}^{m} \{x | p_i(x) \geq 0\} \subseteq \{x | p_0(x) \geq 0\}$.*

III. PROBLEM FORMULATION

We consider the setup where the plant is modelled as a nonlinear system taking the following form

$$\dot{x}_p = f_p(x_p) + g_p(x_p) u, \qquad y = h_p(x_p), \quad (2)$$

where $x_p \in \mathbb{R}^{n_p}$ is the state vector of the plant, $y \in \mathbb{R}^{n_y}$ is the measured output, $u \in \mathbb{R}^{n_u}$ is the input of the plant, function $f_p : \mathbb{R}^{n_p} \to \mathbb{R}^{n_p}$ is continuous with $f_p(0) = 0$ and function $g_p : \mathbb{R}^{n_p} \to \mathbb{R}^{n_p} \times \mathbb{R}^{n_u}$ is continuous. We assume that system (2) is stabilizable, i.e., there exists a control law u_p generated by the *primary controller* which has already been designed to stabilize the plant (2) and takes the following form

$$\dot{x}_c = f_c(x_c, y + a_y), \qquad u_p = h_c(x_c) + a_u, \quad (3)$$

with controller state $x_c \in \mathbb{R}^{n_c}$. The functions f_c and h_c are assumed to satisfy the regularity conditions such that the primary controller (3) exists and the closed-loop is well-posed. Controller (3) is pre-designed to stabilize (2) with the input signal $u_p : \mathbb{R}^{n_c} \to \mathbb{R}^{n_u}$ and is subject to potential adversarial attacks denoted by the attack vector $a = [a_u^\top, a_y^\top]^\top \in \mathbb{R}^{n_u + n_y}$, where $a_u(t) : \mathbb{R}_{\geq 0} \to \mathbb{R}^{n_u}$ and $a_y(t) : \mathbb{R}_{\geq 0} \to \mathbb{R}^{n_y}$ denote actuator and sensor attacks respectively. Since the primary controller (3) is pre-designed without being aware of the attacks, the safety of the closed-loop may be compromised (precise definition is given later in Definition 1). Therefore, we propose introducing a *secondary controller*, that runs in conjunction with the primary controller (3). The secondary controller takes the form of a static output feedback controller that uses measurements of a subset of the plant outputs y which are either available locally or known to be safeguarded against malicious manipulation (e.g., via encryption or watermarking):

$$u_s = h_s(x_s), \quad (4)$$

where $x_s = C_s y = C_s h_p(x_p)$ is a subset of the plant measurements y that are available to the secondary controller (4), and $u_s \in \mathbb{R}^{n_s}$ is the secondary control law. The overall input for our *safe* primary-secondary control scheme is given by

$$u = u_p + E_u u_s, \quad (5)$$

where E_u is a selection matrix we use to denote what entries of the primary controller are affected by the secondary controller. In this work, we assume C_s and E_u are given to model the case where a fixed set of resources are locally available. How to optimally choose C_s and E_u is left for future work. Note that the secondary controller (4) uses attack-free measurements only and it generates an input that will be fed back to the plant reliably. Consequently, no attack signal a appears in (4). The setup is illustrated in Fig. 1.

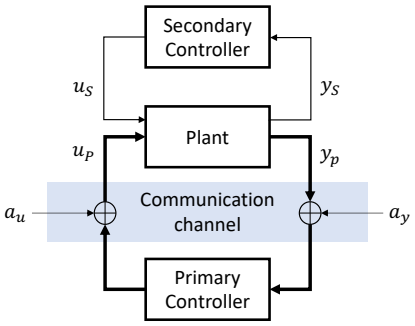

Fig. 1. Ensuring safety with a secondary controller, y_p and y_s denote the measurements used by the primary controller and secondary controller, respectively.

It can be verified that the closed-loop system (2)-(5) can be written in the following form

$$\dot{x} = f(x, a) + g(x) u_s, \quad (6)$$

where $x = [x_p^\top, x_c^\top]^\top$,

$$f(x, a) = \begin{bmatrix} f_p(x_p) + g_p(x_p)(h_c(x_c) + a_u) \\ f_c(x_c, h_p(x_p) + a_y) \end{bmatrix}$$
$$g(x) = \begin{bmatrix} g_p(x_p) E_u \\ \mathbf{0} \end{bmatrix}. \quad (7)$$

It is worth noting that the expression of the closed-loop system (6) is also able to capture the case where static state feedback is used for the primary controller. Suppose we have $u_p = h_c(x_p + a_y) + a_u$, then we have

$$\begin{aligned} f(x,a) &= f_p(x_p) + g_p(x_p)(h_c(x_p + a_y) + a_u) \\ g(x) &= g_p(x_p)E_u. \end{aligned} \tag{8}$$

Throughout the paper, we will make the following assumption about the vector field of the closed-loop system (6).

Assumption 1 *The closed-loop system (2)-(5) written compactly in* (6) *is such that* $f(x,a) \in \mathbb{R}^{n_p+n_c}[x,a]$ *and* $g(x) \in \mathbb{R}^{(n_p+n_c)\times n_s}[x]$.

Remark 1 *Assumption 1 on the closed-loop system (6) makes the analysis more computationally tractable via SOS tools. This assumption can be satisfied if functions* f_p, g_p, h_p, f_c, h_c *are all polynomials of their respective arguments, which may come from least square regression or polynomial approximation of another nonlinear function.*

The goal of the secondary controller (4) is to ensure that, when the overall closed-loop system (2)-(5) is subject to cyber attacks a, the safety of the closed-loop system (6) can be ensured in the following sense.

Definition 1 *The closed-loop system (6) is safe if its state* $x(t)$ *remains within a given safe set* $\mathcal{S}$ *for all* $t \geq 0$.

We describe the safe set by $\mathcal{S} := \{x | s(x) \geq 0\}$, where $s(x) \in \mathbb{R}[x]$. We solve the following problems in this paper.

1) Give sufficient conditions to check if the primary controller (3) alone can render the plant (2) safe in the presence of attacks.
2) Enhance safety of the closed-loop system (6) by synthesizing the secondary controller (4) such that more resources are needed to violate the safety condition.

IV. INVARIANT SET-BASED ANALYSIS

The first step of our analysis is to assess if the primary controller can ensure the safety of the closed-loop system in the presence of the attack signal a. In [2], it is assumed that the attack signal a is norm bounded. This is to take into account that intelligent adversaries often seek to remain stealthy and undetected to be able to continuously send malicious signals to the system. In this work, we impose a more general condition on the attack signal:

$$a \in \mathcal{A} := \{a | A(x,a) \geq 0\}. \tag{9}$$

where $A(x,a) \in \mathbb{R}[x,a]$. The condition (9) covers the situation where adversaries that have access to the states of the system inject attack signals a to the system based on their measurements of x. Although the requirement that A is a polynomial in x and a might be restrictive in some cases, it generalizes the condition used in [2] which is a special case of (9) and can be used as an outer-approximation of the real constraints on the attack signal.

Since we aim to first find the worst case attack signals that ensures the safety of the plant in feedback with the primary controller (3) only, we set $E_u = \mathbf{0}$ and $g(x) = \mathbf{0}$ for all $x \in \mathbb{R}^{n_p+n_c}$. Note that, when $g(x) = \mathbf{0}$, the closed-loop system (6) is a system driven by the attack signal a,

$$\dot{x} = f(x,a). \tag{10}$$

To this end, we define the reachable set of nonlinear systems driven by external signals.

Definition 2 *The (forward) reachable set* $\mathcal{R}_a$ *of the nonlinear system* $\dot{x} = f(x,a)$ *from the initial set* $\mathcal{T}$ *driven by* $a \in \mathcal{A}$ *is defined as the set of all trajectories* $\phi(t, x(0), a)$ *for all* $t \geq 0$ *and* $x(0) \in \mathcal{T}$ *where* $\phi(t, x(0), a)$ *is a solution to* $\dot{x} = f(x,a)$ *at time* t *with the initial condition* $x(0)$.

If we can guarantee that the reachable set of (10) is a subset of the safe set $\mathcal{S}$, then the safety of the system can be ensured since system states can only reach a set that is fully contained in the prescribed safe set. Exact computation of the reachable set of a nonlinear system can be difficult in general. In this work, we construct the set $\mathcal{E}_a$ as an outer approximation of the reachable set such that $\mathcal{R}_a \subseteq \mathcal{E}_a$, where $\mathcal{R}_a$ is the reachable set of (10) from the initial set $\mathcal{T}$. If we manage to find such an $\mathcal{E}_a$ such that $\mathcal{E}_a \subseteq \mathcal{S}$, then it is sufficient to conclude that $\mathcal{R}_a \subseteq \mathcal{S}$. We will make use of the following result which comes from Nagumo's theorem for autonomous systems and its extension to systems with exogenous inputs by Aubin [21], see [5, Section 3.1].

Proposition 1 *Given system (10), if there exists a continuously differentiable function* $V : \mathbb{R}^{n_p+n_c} \to \mathbb{R}$ *such that* $\mathcal{T} \subseteq \mathcal{E}_a := \{x \in \mathbb{R}^{n_p+n_c} | V(x) \leq 1\}$ *and*

$$\frac{\partial V}{\partial x} f(x,a) \leq 0, \ \forall x \in \mathbb{R}^{n_p+n_c}, V(x) = 1, a \in \mathcal{A}; \tag{11}$$

then, we have $\mathcal{R}_a \subseteq \mathcal{E}_a$.

We will use the S-procedure for polynomial functions given by Lemma 1 to certify the set containment conditions in Proposition 1. Assuming that the initial set takes the form $\mathcal{T} := \{x \in \mathbb{R}^{n_p+n_c} | T(x) \geq 0\}$ where $T(x) \in \mathbb{R}[x]$. Our first main result is stated below.

Theorem 1 *Consider the closed-loop system (6) with only the primary controller in feedback, i.e.,* $E_u = 0$ *and* $C_s = 0$. *Given* $(s(x), T(x)) \in \mathbb{R}[x] \times \mathbb{R}[x]$ *and* $A(x,a) \in \mathbb{R}[x,a]$, *if there exist* $V(x) \in \mathbb{R}[x]$, $(\lambda_1, \lambda_2) \in \Sigma[x] \times \Sigma[x]$, $\lambda_3 \in \mathbb{R}[x,a]$, *and* $\lambda_4 \in \Sigma[x,a]$ *such that*

$$\begin{aligned} &s(x) - \lambda_1(1 - V(x)) \in \Sigma[x], \\ &1 - V(x) - \lambda_2 T(x) \in \Sigma[x], \\ &-\frac{\partial V}{\partial x} f(x,a) - \lambda_3(V(x) - 1) - \lambda_4 A(x,a) \in \Sigma[x,a], \end{aligned} \tag{12}$$

then we have $\mathcal{R}_a \subseteq \mathcal{S}$.

Proof: Applying the S-procedure for polynomial functions, the first condition in (12) guarantees that, if $1 - V(x) \geq 0$, we have $s(x) \geq 0$ which means that $\mathcal{E}_a \subseteq \mathcal{S}$. Similarly, the second condition in (12) implies $\mathcal{T} \subseteq \mathcal{E}_a$. Lastly, the

third condition guarantees (11). By Proposition 1, we have $\mathcal{R}_a \subseteq \mathcal{E}_a$ which together with $\mathcal{E}_a \subseteq \mathcal{S}$ implies $\mathcal{R}_a \subseteq \mathcal{S}$. ■

Note that the condition (12) is a sufficient condition to guarantee that when the initial state is in $\mathcal{T}$, the state of the closed-loop system never enters the unsafe region of the state space in the presence of attack signals. Any forward invariant set $\mathcal{E}_a$ that verifies (12) will suffice to guarantee the safety of the system. It is also possible to optimize the safety performance of the system by minimizing an appropriately chosen objective function. One example is to minimize the volume of the forward invariant set $\mathcal{E}_a$. However, since V is a polynomial function, finding the exact expression of the volume of $\mathcal{E}_a$ might be intractable. However, we can find an ellipsoid $\mathcal{E}_P := \{x \in \mathbb{R}^{n_p+n_c} | x^\top P x \leq 1\}$, $P \succeq 0$ that fully contains $\mathcal{E}_a$ and minimize the volume of the ellipsoid by minimizing the convex function $-\log\det(P)$, where $\det(P)$ is the determinant of the matrix P. See [18].

Corollary 1 *Consider the closed-loop system (6) with only the primary controller in feedback, i.e., $E_u = 0$ and $C_s = 0$. Given $s(x), T(x) \in \mathbb{R}[x]$ and $A(x,a) \in \mathbb{R}[x,a]$, if there exist $P \succ 0$, $V(x) \in \mathbb{R}[x]$, $\lambda_1, \lambda_2, \lambda_5 \in \Sigma[x]$, $\lambda_3 \in \mathbb{R}[x,a]$, and $\lambda_4 \in \Sigma[x,a]$ such that*

$$\begin{aligned} &\min_{P,V,\lambda_{1-5}} \quad -\log\det(P) \\ &\text{s.t. (12) holds and } x^\top P x - \lambda_5(1-V(x)) \in \Sigma[x], \end{aligned} \tag{13}$$

where λ_{1-5} denotes the set $\{\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5\}$, then we have $\mathcal{R}_a \subseteq \mathcal{S}$. Moreover, $\mathcal{E}_P$ is the ellipsoid with minimal volume such that $\mathcal{E}_a \subseteq \mathcal{E}_P$.

Proof: By the S-procedure for polynomial functions, the last condition in (13) guarantees that $\mathcal{E}_a \subseteq \mathcal{E}_P$. The rest of the proof follows from the proof of Theorem 1. ■

It can be seen that the condition (12) contains bilinear SOS constraints involving decision variables (λ_1, V) and (λ_3, V) rendering the optimization problem non-convex. However, the constraints are linear in λ_1 and λ_3 when V is fixed and linear in V if λ_1 and λ_3 are fixed. As a result, in practice we can solve (12) and similarly (13) in an alternating fashion between variables (λ_1, λ_3) and V as done in many existing results, see [14], [22], [23] for example.

In this work, we adopt an approach similar to [23, Algorithm 2] by introducing a positive slack variable ϵ to the last condition in (12). The modified conditions takes the form:

$$\begin{aligned} &s(x) - \lambda_1(1-V(x)) \in \Sigma[x], \\ &1 - V(x) - \lambda_2 T(x) \in \Sigma[x], \\ &-\frac{\partial V}{\partial x} f(x,a) + \epsilon - \lambda_3(V(x)-1) - \lambda_4 A(x,a) \in \Sigma[x,a]. \end{aligned} \tag{14}$$

The role of ϵ is to relax the decreasing condition on V and allow $\dot{V}$ to be positive by the margin characterized by ϵ. Then we alternately minimize ϵ over two bilinear groups of decision variables and repeat until $\epsilon \leq 0$ is satisfied, which can be done by the following steps.

1) Specify the orders of polynomials V and λ_{1-4} to be found.
2) Start with an initial guess $\bar{V}$ and minimize ϵ over λ_1 and λ_3 subject to (14).
3) Set λ_1 and λ_3 to the values found in the previous step and minimize ϵ over V subject to (14).
4) Repeat the previous two steps until an $\epsilon \leq 0$ is found.

We will refer to Steps $1) - 4)$ as the alternating search algorithm.

Remark 2 *Given the specified orders of polynomials, the alternating search algorithm guarantees that after each step, ϵ is non-increasing. However, there is no guarantee that ϵ will decrease to be non-positive. It is worth noting that finding a V and λ_{1-4} that satisfy (12) is only a sufficient condition to guarantee that the reachable set $\mathcal{R}_a$ is contained within the safe set $\mathcal{S}$. If the alternating algorithm fails to give a feasible solution to (12), then one can try to increase the order of the polynomials and start the algorithm again until a maximum allowable order is reached, in which case (12) is claimed to be infeasible (though it can be possibly feasible).*

Remark 3 *Once a valid V is found that verifies the safety of (10), this V can be used as the initial value when solving (13). A similar alternating algorithm can be constructed to minimize the volume of $\mathcal{E}_P$. First, given V, $-\log\det(P)$ is minimized with respect to λ_1, λ_3, and λ_5. Then, the obtained λ_1, λ_3, and λ_5 are used to minimize $-\log\det(P)$ over V. The process is repeated until the decrease in $-\log\det(P)$ is within a specified tolerance.*

V. SECONDARY CONTROLLER SYNTHESIS

When the primary controller alone is insufficient to guarantee the safety of the overall closed-loop system (6), introducing a secondary controller may achieve safety. To this end, we aim to systematically design the secondary controller in this section. To be able to employ the SOS programming tools to computationally solve the synthesis problem, we restrict our class of secondary controller (4) to *polynomial* static feedback, i.e., $h_s(x_s) \in \mathbb{R}[x_s]$. With the secondary controller (4) included, the closed-loop system takes the form

$$\dot{x} = f(x,a) + g(x)h_s(x_s) := \tilde{f}(x,a), \tag{15}$$

where the expressions of $f(x,a)$ and $g(x)$ are given in (7). Note that the new closed-loop system (15) with the secondary controller included takes a form similar to (10). Therefore, we can employ Proposition 1 again to conclude the following result, whose proof follows from the proof of Theorem 1 by replacing f by $\tilde{f}$.

Theorem 2 *Consider the closed-loop system (15). Given $(s(x), T(x)) \in \mathbb{R}[x] \times \mathbb{R}[x]$ and $A(x,a) \in \mathbb{R}[x,a]$, if there exist $h_s(x_s) \in \mathbb{R}[x_s]$, $V(x) \in \mathbb{R}[x]$, $(\lambda_1, \lambda_2) \in \Sigma[x] \times \Sigma[x]$, $\lambda_3 \in \mathbb{R}[x,a]$, and $\lambda_4 \in \Sigma[x,a]$ such that*

$$\begin{aligned} &s(x) - \lambda_1(1-V(x)) \in \Sigma[x], \\ &1 - V(x) - \lambda_2 T(x) \in \Sigma[x], \\ &-\frac{\partial V}{\partial x} \tilde{f}(x,a) - \lambda_3(V(x)-1) - \lambda_4 A(x,a) \in \Sigma[x,a], \end{aligned} \tag{16}$$

where $\tilde{f}(x,a) = f(x,a) + g(x)h_s(x_s)$ depends on $h_s(x_s)$, then we have $\tilde{\mathcal{R}}_a \subseteq \mathcal{S}$, where $\tilde{\mathcal{R}}_a$ is the reachable set of (15) from the initial set $\mathcal{T}$.

If condition (16) is satisfied by a set of decision variables $\{h_s, V, \lambda_1, \lambda_2, \lambda_3, \lambda_4\}$, then it is sufficient to conclude that the state of the closed-loop system (15) never leaves the safe set $\mathcal{S}$ when initialized in $\mathcal{T}$. In the synthesis of the secondary controller, we can also find an ellipsoidal outer-approximation of $\mathcal{E}_a$, $\mathcal{E}_P := \{x \in \mathbb{R}^{n_p+n_c} | x^\top P x \leq 1\}$ for some $P \succeq 0$. Then we minimize the volume of the ellipsoid by minimizing the convex function $-\log\det(P)$.

Corollary 2 *Consider the closed-loop system (15). Given $(s(x), T(x)) \in \mathbb{R}[x] \times \mathbb{R}[x]$ and $A(x,a) \in \mathbb{R}[x,a]$, if there exist $P \succ 0$, $h_s(x_s) \in \mathbb{R}[x_s]$, $V(x) \in \mathbb{R}[x]$, $(\lambda_1, \lambda_2, \lambda_5) \in \Sigma[x] \times \Sigma[x] \times \Sigma[x]$, $\lambda_3 \in \mathbb{R}[x,a]$, and $\lambda_4 \in \Sigma[x,a]$ such that*

$$\begin{aligned} &\min_{P, h_s, V, \lambda_{1-5}} \quad -\log\det(P) \\ &\text{s.t. (16) holds and } x^\top P x - \lambda_5(1 - V(x)) \in \Sigma[x], \end{aligned} \tag{17}$$

where $\tilde{f}(x,a) = f(x,a) + g(x)h_s(x_s)$ depends on $h_s(x_s)$, then we have $\tilde{\mathcal{R}}_a \subseteq \mathcal{S}$, where $\tilde{\mathcal{R}}_a$ is the reachable set of (15) from the initial set $\mathcal{T}$. Moreover, $\mathcal{E}_P$ is the ellipsoid with minimal volume such that $\tilde{\mathcal{R}}_a \subseteq \mathcal{E}_P$.

By introducing the secondary control term $h_s(x_s)$, a new bilinear term $\frac{\partial V}{\partial x} h_s(x_s)$ appears in (16) and (17). This, together with other bilinear terms in (12) and (13), makes (16) and (17) non-convex. Nevertheless, the constraints are linear in λ_{1-5} and h_s when V is fixed and linear in V if λ_{1-5} and h_s are fixed. There are no bilinear terms in (16) and (17) involving products of λ_{1-5} and h_s. Thus, there is no need to perform an additional round of alternation. The variables λ_{1-5} and h_s can be solved simultaneously given V. We again introduce the slack variable ϵ to (16) such that the modified condition takes the form

$$\begin{aligned} &s(x) - \lambda_1(1 - V(x)) \in \Sigma[x], \\ &1 - V(x) - \lambda_2 T(x) \in \Sigma[x], \\ &-\frac{\partial V}{\partial x}\tilde{f}(x,a) + \epsilon - \lambda_3(V(x) - 1) - \lambda_4 A(x,a) \in \Sigma[x,a]. \end{aligned} \tag{18}$$

We alternately minimize ϵ over $(\lambda_{1,3}, h_s)$ given V and over V given $(\lambda_{1,3}, h_s)$ and repeat until $\epsilon \leq 0$ is satisfied:

1) Specify the orders of polynomials V, h_s, λ_{1-4}.
2) Start with an initial guess of $\bar{V}$ and minimize ϵ over h_s, λ_1 and λ_3 subject to (18).
3) Set h_s, λ_1 and λ_3 to the values found in the last step and minimize ϵ over V subject to (18).
4) Repeat the previous two steps until an $\epsilon \leq 0$ is found.

Remark 4 *The initial guess of V can be taken from the result of checking conditions (13) and (14). Specifically, if there does not exist a V that satisfies (14) for a non-positive ϵ, then the initial value of V can be the one that minimizes ϵ subject to (14). Hopefully, with the additional term h_s, ϵ can be made negative after several iterations. If there does exist a V that verifies the safety of the closed-loop system (10), then the solution to (13) can be used. In such cases, (14) will always be feasible since $h_s = 0$ is a trivial secondary controller that ensures the safety of the closed-loop system.*

Remark 5 *When the alternating algorithm does not find a non-positive ϵ that satisfies (18), one can increase the order of the variables including the new term h_s. Moreover, changing the values of C_s and E_u might also be helpful in synthesizing a secondary controller that ensures the safety of the closed-loop system.*

Remark 6 *Once valid V and h_s are found to satisfy (16), this V and h_s can be used as the initial guess to solve (17) following the discussion in Remark 3. However, it should be noted that the conditions (16) in Theorem 2, though being sufficient to guarantee the safety of the closed-loop system, are not sufficient to guarantee that the origin is asymptotically stable in the absence of the attack signal a. This is in contrast to the linear counterpart shown in [2]. The design of a secondary controller aiming to recover the performance of the primary controlled system while ensuring safety will be left for future work.*

VI. NUMERICAL SIMULATION

In this section, we illustrate our main results via numerical simulations of a second order nonlinear system. Suppose a primary controller has been designed such that the closed-loop system (15) takes the following form

$$\begin{aligned} \dot{x}_1 &= -x_1 + x_2 + a_1 \\ \dot{x}_2 &= -x_2 - x_1^2 x_2 + a_2 + h_s(x_2), \end{aligned} \tag{19}$$

where $a = [a_1, a_2]^\top$ is the attack vector and measurement $x_s = x_2$ is used to design the secondary controller. For simplicity, the attack signals are assumed to satisfy $A(x,a) = A(a) = 1 - a_1^2 + a_2^2 \geq 0$. Moreover, we assume that the initial set $\mathcal{T}$ is the singleton containing the origin, i.e., $T(x) = -x_1^2 - x_2^2 \geq 0$. This captures the steady state for a globally asymptotically stable system when there are no attack signals. Under these conditions, we aim to keep the state $x = [x_1, x_2]^\top$ within the safe set $\mathcal{S}$ characterized by $s(x) = 1.3 - x_1^2 - x_2^2 \geq 0$.

First, we set $h_s(x_s) = 0$ and test if the primary controller alone can keep the states x inside the safe set $\mathcal{S}$. We alternately solve the condition (12) in Theorem 1 using SOSTOOLS [24] with SeDuMi being the solver [25]. In this example, we restrict our search of all polynomial variables to polynomials of orders no higher than 4. It turns out that, under these conditions, the condition (12) is infeasible. To explore the limitations of the primary controller, instead of insisting that the state stays in the safe set, we impose the condition that the state x remains within the set $\{x \in \mathbb{R}^2 | s(x) + \gamma \geq 0\}$ for some $\gamma > 0$. We then minimize γ subject to condition (12), alternately over V and λ_1, λ_3. After 8 alternating iterations, γ reaches the value of 0.19 with $V = V_1(x) = 0.67315x_1^2 + 0.70356x_2^2$. Thus, we have failed to find a V which certifies the safety of the closed-loop system via Theorem 1. However, as discussed before, this does not mean the primarily controlled system is unsafe

since there might exist polynomials of higher orders that satisfy (12). In the simulation, we have attempted to increase the order of the function $V(x)$ to 6 which, however, does not result in significant decrease of the value of γ.

Next we check if a polynomial secondary controller $h_s(x_2)$ of order no higher than 4 can be found to keep the states within the safe set. We again apply the alternating algorithm to check if condition (16) is feasible with the initial guess being $V = V_1(x)$. It turns out that, $h_s(x_2) = -0.31761x_2^3 - 1.2534x_2$ can ensure safety of the closed-loop system (19) with $V = V_2(x) = 0.8881x_1^2 + 2.669x_2^2$.

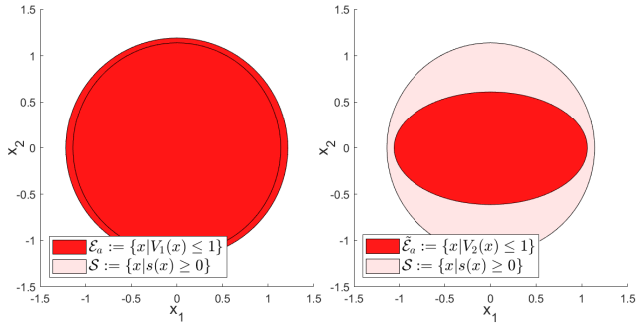

Fig. 2. Plots of the safe set $\mathcal{S}$ and the ellipsoidal over-approximation $\mathcal{E}_a$ of the reachable set $\mathcal{R}_a$ with the primary control only (left) and ellipsoidal over-approximation $\tilde{\mathcal{E}}_a$ of the reachable set $\tilde{\mathcal{R}}_a$ with primary and secondary controls (right).

The plots of the ellipsoidal over-approximations $\mathcal{E}_a$ and $\tilde{\mathcal{E}}_a$ of the respective reachable sets $\mathcal{R}_a$ and $\tilde{\mathcal{R}}_a$ found via the corresponding alternating algorithms and the safe set are given in Fig. 2. It can be seen that, after introducing the secondary controller $h_s(x_2) = -0.31761x_2^3 - 1.2534x_2$ to the closed-loop system, its state x always remains in a subset of the safe set when initialized at the origin.

VII. CONCLUSIONS

In this work, based on invariant set analysis and SOS programming, we provide sufficient conditions for safety verification and control design of a class of polynomial nonlinear systems in the presence of adversarial signals. The conditions are computationally tractable via an alternating algorithm. We show that it is possible to improve the safety performance of a nonlinear system by using a subset of sensors that are attack-free. A numerical simulation on a second order nonlinear system verifies the theoretical result.

There are several possible future research directions to be explored. First, it is interesting to investigate a secondary controller design approach that recovers the performance achieved by the primary controller at least locally while ensuring safety. Another interesting topic would be the analysis of how the choice of sensors characterized by the matrix C_s impacts the performance of the secondary controller.

REFERENCES

[1] A. A. Cárdenas, S. Amin, and S. Sastry, "Research challenges for the security of control systems.," *HotSec*, vol. 5, p. 15, 2008.

[2] Y. Lin, M. S. Chong, and C. Murguia, "Plug-and-play secondary control for safety of LTI systems under attacks," *arXiv preprint arXiv:2212.00593*, 2022. To appear in the proceedings of the IFAC World Congress 2023.

[3] K. Margellos and J. Lygeros, "Hamilton–Jacobi formulation for reach–avoid differential games," *IEEE Transactions on Automatic Control*, vol. 56, no. 8, pp. 1849–1861, 2011.

[4] S. Bansal, M. Chen, S. Herbert, and C. J. Tomlin, "Hamilton-Jacobi reachability: A brief overview and recent advances," in *Proceedings of the IEEE 56th Conference on Decision and Control*, pp. 2242–2253, IEEE, 2017.

[5] F. Blanchini, "Set invariance in control," *Automatica*, vol. 35, no. 11, pp. 1747–1767, 1999.

[6] S. Prajna, A. Jadbabaie, and G. J. Pappas, "A framework for worst-case and stochastic safety verification using barrier certificates," *IEEE Transactions on Automatic Control*, vol. 52, no. 8, pp. 1415–1428, 2007.

[7] A. D. Ames, X. Xu, J. W. Grizzle, and P. Tabuada, "Control barrier function based quadratic programs for safety critical systems," *IEEE Transactions on Automatic Control*, vol. 62, no. 8, pp. 3861–3876, 2017.

[8] S. Coogan and M. Arcak, "A dissipativity approach to safety verification for interconnected systems," *IEEE Transactions on Automatic Control*, vol. 60, no. 6, pp. 1722–1727, 2014.

[9] U. Topcu, A. K. Packard, P. Seiler, and G. J. Balas, "Robust region-of-attraction estimation," *IEEE Transactions on Automatic Control*, vol. 55, no. 1, pp. 137–142, 2010.

[10] A. Papachristodoulou and S. Prajna, "Robust stability analysis of nonlinear hybrid systems," *IEEE Transactions on Automatic Control*, vol. 54, no. 5, pp. 1035–1041, 2009.

[11] X. Tan, W. S. Cortez, and D. V. Dimarogonas, "High-order barrier functions: Robustness, safety, and performance-critical control," *IEEE Transactions on Automatic Control*, vol. 67, no. 6, pp. 3021–3028, 2021.

[12] W. Xiao, C. Belta, and C. G. Cassandras, "Adaptive control barrier functions," *IEEE Transactions on Automatic Control*, vol. 67, no. 5, pp. 2267–2281, 2021.

[13] P. Parillo, "Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization," *PhD thesis, California Institute of Technology*, 2000.

[14] Z. Jarvis-Wloszek, R. Feeley, W. Tan, K. Sun, and A. Packard, "Some controls applications of sum of squares programming," in *Proceedings of the IEEE 42nd Conference on Decision and Control*, vol. 5, pp. 4676–4681, IEEE, 2003.

[15] M. Jones and M. M. Peet, "Using SOS and sublevel set volume minimization for estimation of forward reachable sets," *IFAC-PapersOnLine*, vol. 52, no. 16, pp. 484–489, 2019.

[16] A. Teixeira, I. Shames, H. Sandberg, and K. H. Johansson, "A secure control framework for resource-limited adversaries," *Automatica*, vol. 51, pp. 135–148, 2015.

[17] G. Blekherman, P. A. Parrilo, and R. R. Thomas, *Semidefinite optimization and convex algebraic geometry*. SIAM, 2012.

[18] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear matrix inequalities in system and control theory*. SIAM, 1994.

[19] J. Bochnak, M. Coste, and M.-F. Roy, *Géométrie algébrique réelle*, vol. 12. Springer Science & Business Media, 1987.

[20] Z. W. Jarvis-Wloszek, *Lyapunov based analysis and controller synthesis for polynomial systems using sum-of-squares optimization*. PhD thesis, the University of California, Berkeley, 2003.

[21] J.-P. Aubin, A. M. Bayen, and P. Saint-Pierre, *Viability theory: new directions*. Springer Science & Business Media, 2011.

[22] H. Yin, M. Arcak, A. Packard, and P. Seiler, "Backward reachability for polynomial systems on a finite horizon," *IEEE Transactions on Automatic Control*, vol. 66, no. 12, pp. 6025–6032, 2021.

[23] K. S. Schweidel, H. Yin, S. W. Smith, and M. Arcak, "Safe-by-design planner–tracker synthesis with a hierarchy of system models," *Annual Reviews in Control*, 2022.

[24] A. Papachristodoulou, J. Anderson, G. Valmorbida, S. Prajna, P. Seiler, P. Parrilo, M. M. Peet, and D. Jagt, "SOSTOOLS version 4.00 sum of squares optimization toolbox for MATLAB," *arXiv preprint arXiv:1310.4716*, 2013.

[25] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optimization Methods and Software*, vol. 11, no. 1-4, pp. 625–653, 1999.