

Safety-Critical Control With Limited Information

Maison Clouâtre^{ID}, *Graduate Student Member, IEEE*, Makhin Thitsa^{ID}, Wesley Kinney^{ID},
Andrea Conti^{ID}, *Fellow, IEEE*, and Moe Z. Win^{ID}, *Fellow, IEEE*

Abstract—This letter explores safety-critical control of nonlinear systems in settings where a finite-rate communication channel stands in the path of state feedback. We show that the mere existence of a nominally safe control law (certified by an exponential barrier function) suffices to provide safe control in these limited-information settings. We introduce the notion of “safety escape time”, the minimum time a system takes to become unsafe in the absence of actuation. The results complement the existing literature on stabilizing control with limited information and represent a step towards a complete understanding of safety-critical control with limited information.

Index Terms—Safety escape time, safety, control with limited information, barrier function, networked control.

I. INTRODUCTION

NETWORKED SYSTEMS enable ubiquitous communication and control—while presenting new challenges and design constraints. In the theory of networked systems, tremendous effort has been poured into tackling the problem of control with limited information [1], [2]. As connected technologies become ever more prevalent, safely controlling [3], [4], [5] networked systems is of paramount interest. However, in the vast literature on control with limited information (e.g., see [6] and the references therein), little attention has been paid to the problem of safety.

Stabilizing control over finite capacity channels was pioneered around the turn of the 21st century [7], [8], [9]. After initially focusing on linear systems, the theory was expanded to encompass stabilizing nonlinear systems under communication constraints and with quantized inputs [10], [11], [12]. Researchers also developed state estimation [13], [14] and

distributed filtering strategies [15], [16], [17], [18] both in the presence of limited information. An exposition on control of networked systems is given in [19]. Inspired by this literature on stabilizing control with limited information, an interesting open question is how to achieve safety-critical control with limited information. This letter provides an initial answer to this question in the scenario where a finite-rate communication channel stands in the path of state feedback. Whereas traditional feedback paths are assumed perfect, a finite-rate feedback channel limits the state information available to the controller. Exploiting a topological robustness property that naturally occurs in barrier functions, we show that the mere existence of a nominally safe control (whose safety is certified by an exponential barrier function) guarantees the existence of a safe control in the limited-information setting (referred to as a “limited-information safety control”). The key contributions of this letter can be summarized as follows, we

- establish how topological robustness can be used to certify the safety of a system with only an estimate of its state;
- introduce the notion of “safety escape time”, the time it takes for an unactuated system to go unsafe, as well as derive a lower bound on the safety escape time; and
- construct a class of limited-information safety controllers that guarantee the safe operation of a dynamical system over a finite-rate state feedback channel.

The results established herein leverage the approach to limited-information observers developed in [11], [14], [20] and are inspired by exponential barrier function-based control methodologies from the traditional control setting [4].

Notation: Vectors and matrices are denoted by bold lowercase and uppercase letters, respectively. The norm $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}$ is taken to be the ∞ -norm. The closed n -ball of radius $r > 0$ centered at \mathbf{z} is denoted $\mathcal{B}(\mathbf{z}, r) \triangleq \{\mathbf{y} \in \mathbb{R}^n : \|\mathbf{y} - \mathbf{z}\| \leq r\}$. The diameter of a bounded set $\mathcal{C} \subset \mathbb{R}^n$ is denoted $\text{diam}(\mathcal{C}) = \sup \{\|\mathbf{x} - \mathbf{y}\| : \mathbf{x}, \mathbf{y} \in \mathcal{C}\}$. Given a continuous function $f : \mathbb{R}^{n_1} \rightarrow \mathbb{R}^{n_2}$ and a compact set $\mathcal{D} \subset \mathbb{R}^{n_1}$, the maximum of $\|f(\mathbf{x})\|$ over all $\mathbf{x} \in \mathcal{D}$ is denoted $B_f(\mathcal{D})$. Given locally Lipschitz f , the Lipschitz constant (in terms of the metric induced by $\|\cdot\|$) of f restricted to the compact set \mathcal{D} is denoted $C_f(\mathcal{D})$. The Lie derivative of $h : \mathbb{R}^n \rightarrow \mathbb{R}$ along $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is denoted $L_f h(\mathbf{x})$. The set of all non-negative integers is denoted \mathbb{N} . The time derivative is denoted with a dot as in $\dot{\mathbf{x}}$. The symbol \mathcal{U} denotes the set of all bounded piecewise continuous functions from $[0, \infty)$ into \mathbb{R}^m , i.e., each $\mathbf{u}(\cdot) \in \mathcal{U}$ has a finite number of discontinuities on any subinterval $[a, b] \subset [0, \infty)$.

Manuscript received 8 March 2024; revised 15 May 2024; accepted 29 May 2024. Date of publication 17 June 2024; date of current version 24 October 2024. The fundamental research described in this letter was supported, in part, by the National Science Foundation under Grant CNS-2148251 and by federal agency and industry partners in the RINGS program. Recommended by Senior Editor S. Dey. (Corresponding author: Moe Z. Win.)

Maison Clouâtre is with the Wireless Information and Network Sciences Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139 USA.

Makhin Thitsa and Wesley Kinney are with the Department of Electrical and Computer Engineering, Mercer University, Macon, GA 31207 USA.

Andrea Conti is with the Department of Engineering and CNIT, University of Ferrara, 44122 Ferrara, Italy.

Moe Z. Win is with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: moewin@mit.edu).

Digital Object Identifier 10.1109/LCSYS.2024.3415477

II. PROBLEM FORMULATION

Consider an input-affine nonlinear system with state $\mathbf{x} \in \mathbb{R}^n$ and input $\mathbf{u}(\cdot) \in \mathcal{U}$. The system's dynamics are given by

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\mathbf{u}, \quad \mathbf{x}(0) = \mathbf{x}_0 \quad (1)$$

where $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ and $\mathbf{g} : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times m}$ are locally Lipschitz. The goal of a safety-critical controller is to choose a control law that ensures the state $\mathbf{x}(t)$ remains “safe”. We define safety using a continuously differentiable scalar test function $h : \mathbb{R}^n \rightarrow \mathbb{R}$. If $h(\mathbf{x}) \geq 0$ then the state is safe; otherwise the state is unsafe. The set of all safe states is denoted $\mathcal{C} \triangleq \{\mathbf{x} \in \mathbb{R}^n : h(\mathbf{x}) \geq 0\}$ and is known as the “safe set”. Notice that this definition of safety is common in the literature [5]. It is useful to define the notation

$$\mathcal{C}_{\alpha,\beta} \triangleq \{\mathbf{x} \in \mathbb{R}^n : h(\mathbf{x}) \in [\alpha, \beta]\} \quad (2a)$$

$$\mathcal{C}_\alpha \triangleq \{\mathbf{x} \in \mathbb{R}^n : h(\mathbf{x}) \geq \alpha\} \quad (2b)$$

for $\beta \geq \alpha \geq 0$. In general, the goal of safety-critical control is to design feedback control laws that render the safe set forward-invariant. Henceforth, the following assumption about the safe set \mathcal{C} will be made.

Assumption 1: The safe set \mathcal{C} is bounded (i.e., compact since \mathcal{C} is closed by definition) and has non-empty interior (i.e., there exists $\mathbf{x} \in \mathcal{C}$ with the property $h(\mathbf{x}) > 0$). \square

Because $\mathbf{u}(\cdot)$ is bounded piecewise continuous, the solution¹ $\mathbf{x}(t)$ to (1) is unique over some maximal time interval $\mathcal{I} = [0, b)$ with $b > 0$ [21]. For now, suppose that $\hat{\mathbf{x}}(t)$ is an estimate of the state $\mathbf{x}(t)$ over the same maximal time interval \mathcal{I} and that $\mathbf{e}(t) \triangleq \hat{\mathbf{x}}(t) - \mathbf{x}(t)$ is the state estimation error for all $t \in \mathcal{I}$. Since the function h is continuously differentiable, it is Lipschitz on the compact set \mathcal{C} . Denote the Lipschitz constant of h on \mathcal{C} by $C_h(\mathcal{C})$. The results established in this letter are based on the following lemma regarding the topological robustness of barrier functions. This lemma is likely also of interest in non-limited information settings.

Lemma 1: Let $\mathbf{u}(\cdot)$ be any bounded piecewise continuous function and let $\mathbf{x}_0 \in \mathcal{C}$. If the inequality

$$h(\hat{\mathbf{x}}(t)) > C_h(\mathcal{C})\|\mathbf{e}(t)\| \quad (3)$$

holds for all time $t \in \mathcal{I}$, then the solution $\mathbf{x}(t)$ to (1) is safe for all time $t \in \mathcal{I} = [0, \infty)$. \square

Proof: Due to (3) and the fact that $\|\mathbf{e}(t)\| \geq 0$, the state estimate $\hat{\mathbf{x}}(t)$ is always safe. Suppose that the system's state $\mathbf{x}(t)$ goes unsafe during the maximal interval of existence \mathcal{I} . By continuity, there exists some time $\tau \geq 0$ such that $h(\mathbf{x}(\tau)) = 0$. Since $\mathbf{x}(\tau) \in \mathcal{C}$, the Lipschitz condition reveals the inequality

$$h(\hat{\mathbf{x}}(\tau)) = |h(\hat{\mathbf{x}}(\tau)) - h(\mathbf{x}(\tau))| \leq C_h(\mathcal{C})\|\mathbf{e}(\tau)\|$$

which contradicts the lemma's hypothesis. Therefore $\mathbf{x}(t)$ is safe for all $t \in \mathcal{I}$. On the other hand, if the maximal interval of existence is finite (i.e., the solution has finite escape time), then the proof of [22, Th. 3.3, p. 94] shows that there must be a time $\tau \in \mathcal{I}$ where $\mathbf{x}(\tau) \notin \mathcal{C}$. This is a contradiction; therefore, the maximal interval of existence for the unique solution $\mathbf{x}(t)$ to (1) is $[0, \infty)$. \square

¹The solutions to the initial value problem (1) are defined in the sense of Carathéodory [21].

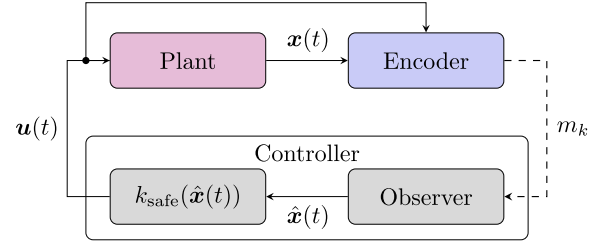


Fig. 1. Block diagram of a limited-information safety controller. Encoded messages m_k describing the plant's state $\mathbf{x}(t)$ are sent to the controller over a finite-rate channel (dashed line) at each transmission time t_k . The controller consists of two parts: an observer and a nominally safe control law $k_{\text{safe}}(\cdot)$. The observer constructs a state estimate $\hat{\mathbf{x}}(t)$ using the messages received from the channel and the controller applies the control $\mathbf{u}(t) = k_{\text{safe}}(\hat{\mathbf{x}}(t))$ to the input of the plant.

The lemma states that, if the state estimate is far enough into the interior of the safe set with respect to the estimation error, then the system is itself safe. For arbitrary choice of $\eta > 1$, the prior lemma proves that the inequality

$$h(\hat{\mathbf{x}}(t)) \geq \eta C_h(\mathcal{C})\|\mathbf{e}(t)\| \quad (4)$$

certifies the safety of system (1). This style of safety certificate will be useful in our subsequent analysis.

The scenario of interest in this letter is that in which state measurements are sent to the controller over a finite-rate communication channel, as shown in Figure 1. In particular, we will construct a general class of controllers that, under appropriate conditions, can guarantee safety in this scenario. These controllers will consist of two parts: a state observer and a nominally safe control law $k_{\text{safe}} : \mathcal{C} \rightarrow \mathbb{R}^m$. Further details regarding each of these two components are forthcoming. Scenarios similar to that depicted in Figure 1 have been previously studied in seminal works on stabilizing control over finite-capacity channels, including [9]. See [19, p. 55] for a study of observers located prior to the communication channel.

A. Safety Escape Time

The system (1) may become unsafe over time if no actuation is applied. After all, this is why control is needed to ensure safety. Henceforth the following assumption is made.

Assumption 2: The initial state \mathbf{x}_0 of the system lies in the interior of the safe set \mathcal{C} . Therefore, there exists some $\delta > 0$ such that $h(\mathbf{x}_0) \geq \delta$. \square

The particular value of δ need not be known by the controllers developed here, which is further clarified in Section III. We are interested in the minimum time for a system whose initial state satisfies $h(\mathbf{x}_0) = \delta > 0$ to become unsafe when no actuation is applied. We refer to this time as the “safety escape time”. Let $\xi(\mathbf{x}_0, t)$ be the solution (on a maximal time interval \mathcal{I}) of

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}), \quad \mathbf{x}(0) = \mathbf{x}_0 \quad (5)$$

at time $t \in \mathcal{I}$. The safety escape time is defined as

$$T_s(\delta) \triangleq \inf \{t > 0 : \mathbf{x}_0 \in \mathbb{R}^n, h(\mathbf{x}_0) = \delta, h(\xi(\mathbf{x}_0, t)) < 0\}. \quad (6)$$

Following the same logic as the proof of Lemma 1, if the interval of existence \mathcal{I} of the solution $\xi(\mathbf{x}_0, \cdot)$ is finite, then there indeed exists some time $t \in \mathcal{I}$ such that $h(\xi(\mathbf{x}_0, t)) < 0$.

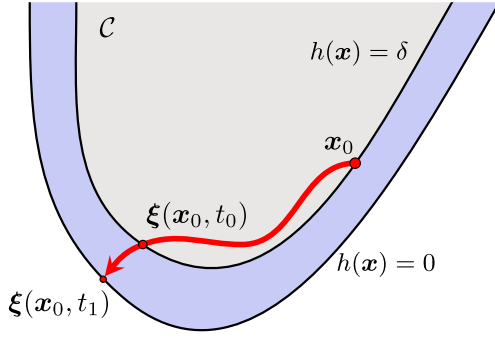


Fig. 2. Visualization of the trajectory $\xi(x_0, \cdot)$ studied in the proof of Theorem 1. The gray area denotes the set $\mathcal{C}_\delta = \{x : h(x) > \delta\}$ and the lilac area represents the set $\mathcal{C}_{0,\delta} = \{x : h(x) \in [0, \delta]\}$; the safe set $\mathcal{C} = \{x : h(x) \geq 0\}$ is the union of these two areas. The system starts at time $t = 0$ in the state x_0 with $h(x_0) = \delta$. At time t_0 , the trajectory is at point $\xi(x_0, t_0)$ which also satisfies $h(\xi(x_0, t_0)) = \delta$. For time t in the interval $[t_0, t_1]$ the solution $\xi(x_0, t)$ lies entirely within the compact set $\mathcal{C}_{0,\delta}$. At time t_1 the solution reaches the boundary of the safe set and $h(\xi(x_0, t_1)) = 0$.

In the uninteresting case where the set on the right side of (6) is empty, the system initiating from x_0 with $h(x_0) = \delta$ is safe for all time with zero actuation and the safety escape time is defined to be $T_s(\delta) \triangleq +\infty$.

Calculating the safety escape time would be difficult in practice for arbitrary nonlinear systems; however, the next theorem will provide a useful lower bound. Recall that $B_f(\mathcal{D})$ is the maximum of $\|f(x)\|$ over all x in a compact set $\mathcal{D} \subset \mathbb{R}^n$. Given $\delta > 0$, the set $\mathcal{C}_{0,\delta} \subseteq \mathcal{C}$ may be expressed as $\mathcal{C}_{0,\delta} = h^{-1}([0, \delta])$. It follows that $\mathcal{C}_{0,\delta}$ is closed because h is continuous, and $\mathcal{C}_{0,\delta}$ is bounded because it is contained in \mathcal{C} . By the Heine-Borel theorem, $\mathcal{C}_{0,\delta}$ is compact. Let f and h be non-trivial so that $C_h(\mathcal{C}_{0,\delta})$ and $B_f(\mathcal{C}_{0,\delta})$ are non-zero.

Theorem 1: The safety escape time is lower bounded by a positive constant as given by

$$T_s(\delta) \geq \frac{\delta}{C_h(\mathcal{C}_{0,\delta}) B_f(\mathcal{C}_{0,\delta})}. \quad (7)$$

□

Proof: Consider x_0 with $h(x_0) = \delta$ such that the solution $\xi(x_0, \cdot)$ to (5) goes unsafe. By continuity of this solution and continuity of h , there exists some time $t_1 > 0$ where

$$h(\xi(x_0, t_1)) = 0. \quad (8)$$

Choose t_1 to be the smallest time for which (8) occurs. Due to continuity, there exists a time t_0 with $0 \leq t_0 < t_1$ so that $\xi(x_0, t) \in \mathcal{C}_{0,\delta}$ for all $t \in [t_0, t_1]$. This is depicted in Figure 2. The width of the interval $[t_0, t_1]$ provides a lower bound on the safety escape time.

By Lipschitz continuity of h on the compact set $\mathcal{C}_{0,\delta}$,

$$h(\xi(x_0, t_1)) \geq h(\xi(x_0, t_0)) - C_h(\mathcal{C}_{0,\delta}) \|\xi(x_0, t_1) - \xi(x_0, t_0)\| \quad (9)$$

and therefore

$$\|\xi(x_0, t_1) - \xi(x_0, t_0)\| \geq \frac{\delta}{C_h(\mathcal{C}_{0,\delta})} \quad (10)$$

as $h(\xi(x_0, t_0)) = \delta$ and $h(\xi(x_0, t_1)) = 0$. Applying the integral form of the mean value theorem [23, p. 247] to the

function $\|f(\xi(x_0, \cdot))\|$ on the compact interval $[t_0, t_1]$ reveals that there exists a time $t' \in (t_0, t_1)$ such that

$$\|\xi(x_0, t_1) - \xi(x_0, t_0)\| \leq (t_1 - t_0) \|f(\xi(x_0, t'))\|. \quad (11)$$

Since $\xi(x_0, t) \in \mathcal{C}_{0,\delta}$ for all $t \in [t_0, t_1]$,

$$\|\xi(x_0, t_1) - \xi(x_0, t_0)\| \leq (t_1 - t_0) B_f(\mathcal{C}_{0,\delta}). \quad (12)$$

Combining this with (10) shows that

$$(t_1 - t_0) \geq \frac{\delta}{C_h(\mathcal{C}_{0,\delta}) B_f(\mathcal{C}_{0,\delta})}. \quad (13)$$

In the worst case scenario $t_0 = 0$, and the earliest time t_1 that $h(\xi(x_0, t_1)) = 0$ can occur is $t_1 = \delta / (C_h(\mathcal{C}_{0,\delta}) B_f(\mathcal{C}_{0,\delta}))$. This proves the theorem. □

An alternative version of this theorem can be stated without using the subset $\mathcal{C}_{0,\delta}$ of \mathcal{C} . While it is a looser bound in general, the constants involved may be easier to calculate.

Corollary 2: The safety escape time is lower bounded by a positive constant as given by

$$T_s(\delta) \geq \frac{\delta}{C_h(\mathcal{C}) B_f(\mathcal{C})}. \quad (14)$$

□

Proof: Since $\mathcal{C}_{0,\delta} \subseteq \mathcal{C}$, it follows that $C_h(\mathcal{C}_{0,\delta}) \leq C_h(\mathcal{C})$ and $B_f(\mathcal{C}_{0,\delta}) \leq B_f(\mathcal{C})$. The proof follows by Theorem 1. □

The limited-information safety controllers proposed in this letter consist of a transient *estimation-only phase*. This phase begins at time $t = 0$. For the duration of this phase no actuation is applied to the system, and the goal of this phase is to quickly compute a state estimate which is precise enough to select a safe control law. This phase cannot take too long otherwise the system will go unsafe. It is for this reason that we have defined the safety escape time $T_s(\delta)$.

Remark 1: Theorem 1 provides a bound on the time that the controller is allowed to “do nothing”, i.e., apply zero actuation, before the system becomes unsafe. As such, we informally refer to this result as the “do nothing theorem.” □

B. Limited-Information State Observer

This section describes the limited-information state observer considered in this letter. This observer is a variant, adapted specifically to the input-affine dynamics (1), of the existing approach for both linear and nonlinear state observers over finite-rate channels such as [11], [14], [20]. Recall the system’s block diagram from Figure 1. The encoder measures the system’s state $x(t)$ and transmits messages to the observer over a finite-rate channel. Specifically, at each transmission time $t_k = kT$, where $k \in \mathbb{N}$ and $T > 0$, the encoder transmits a symbol $m_k \in \{1, 2, \dots, M\} \triangleq \mathcal{M}$ over the channel. The set \mathcal{M} is known as the coding alphabet and T is known as the transmission period. The job of the observer is to decode the received messages and to use them in computing an estimate $\hat{x}(t)$ of the system’s state $x(t)$. This estimate is called the state of the observer. At each time t_k the observer is allowed to update its state abruptly (discontinuously) based on messages received from the encoder. However, for $t \in (t_k, t_{k+1})$ the observer’s state is updated according to the system’s dynamics:

$$\dot{\hat{x}} = f(\hat{x}) + g(\hat{x})u, \quad \hat{x}(t_k) = \hat{x}_k \quad (15)$$

where \hat{x}_k is the state at the update time t_k (the latter equation in (15) plays the role of the “initial condition” at time t_k).

The control input applied to (1) will be a continuous function $u(t) \triangleq k_{\text{safe}}(\hat{x}(t))$ of the observer’s state $\hat{x}(t)$, and hence $u(\cdot) \in \mathcal{U}$. Suppose that the control ensures the observer’s state $\hat{x}(t)$ is confined to some compact set $\mathcal{D} \subset \mathbb{R}^n$ over the time period of interest. Note that this does not imply that \mathcal{D} is a subset of \mathcal{C} . Since $k_{\text{safe}}(\cdot)$ is continuous and \mathcal{D} is bounded, there exists some $B_{k_{\text{safe}}}(\mathcal{D}) > 0$ such that $\|k_{\text{safe}}(z)\| \leq B_{k_{\text{safe}}}(\mathcal{D})$ for all $z \in \mathcal{D}$. Let $C_f(\mathcal{D})$ and $C_g(\mathcal{D})$ denote the Lipschitz coefficients of f and g , respectively, on \mathcal{D} . Define

$$G_a \triangleq \exp\{[C_f(\mathcal{D}) + C_g(\mathcal{D}) B_{k_{\text{safe}}}(\mathcal{D})] T\}, \text{ and} \quad (16a)$$

$$G_e \triangleq \exp\{C_f(\mathcal{D}) T\}. \quad (16b)$$

Between transmission times t_k and t_{k+1} the observer error’s growth is bounded by $\|e(t_k)\| \leq G_a \|e(t_{k-1})\|$ (resp. $\|e(t_k)\| \leq G_e \|e(t_{k-1})\|$) if actuation (resp. no actuation) is applied to the system.² This is a direct result of Grönwall’s inequality. Here G_a and G_e are known as “growth bounds.”

The encoding strategy and observer updating strategy at the transmission times are as follows. The encoder and decoder start by considering a common safe reference point $r \in \mathcal{C}$. Note that $\|r - x_0\| \leq \text{diam}(\mathcal{C})$. The encoder divides $\mathcal{B}(r, \text{diam}(\mathcal{C}))$ into M equal hypercubes and transmits the symbol $j \in \{1, 2, \dots, M\}$ associated with the cube containing $x(t_0)$. When the symbol j is received, the observer knows that the true state lies in cube j and defines $\hat{x}(t_0)$ to be the center of said cube. The state estimation error is then bounded by $\|\hat{x}(t_0) - x_0\| \leq \text{diam}(\mathcal{C})/M^{1/n} \triangleq \hat{e}_0$. Using this initial condition, one can iteratively define an upper bound \hat{e}_k on the observer error at each step $k \in \mathbb{N}$. For the first few transmissions, say until step $l \in \mathbb{N}$, no actuation is applied to the system. Exact details regarding time step l are provided in Section III. During this *estimation-only phase*, the observer knows with certainty that $x(t_k) \in \mathcal{B}(\hat{x}(t_{k-1}), G_e \hat{e}_{k-1})$ at each transmission time t_k .³ Let $\hat{x}(t_k^-)$ be the observer’s state just before time t_k . The encoder divides $\mathcal{B}(\hat{x}(t_k^-), G_e \hat{e}_{k-1})$ into M cubes of equal volume, as it did at time t_0 , and transmits the symbol associated with the cube containing $x(t_k)$. When the symbol j is received, the observer defines $\hat{x}(t_k)$ to be the center of cube j . As a result, the error bound at each step is

$$\|e(t_k)\| \leq \frac{G_e}{M^{1/n}} \hat{e}_{k-1} \triangleq \hat{e}_k. \quad (17)$$

For $k \geq l$, the same estimation strategy is performed by replacing the growth bound G_e with the larger quantity G_a which accounts for the effects of actuation:

$$\|e(t_k)\| \leq \frac{G_a}{M^{1/n}} \hat{e}_{k-1} \triangleq \hat{e}_k. \quad (18)$$

If T and M are chosen such that $G_a < M^{1/n}$ (and therefore $G_e < M^{1/n}$), then (17) and (18) will produce a sequence $\{\|e(t_k)\|\}_{k \in \mathbb{N}}$ which is exponentially convergent to zero. Further details, including how to confine $x(t)$ and $\hat{x}(t)$ to an appropriate compact set \mathcal{D} , are provided below.

²More precise bounds can be given if f , g , and u are continuously differentiable functions of the state x , e.g., see [14].

³It is not the focus of this letter to consider disturbances in the dynamics, which could put $x(t_k)$ outside the ball $\mathcal{B}(\hat{x}(t_{k-1}), G_e \hat{e}_{k-1})$.

Algorithm 1: Limited-Information Safety Controller—State Observer and Control Parameter Updates

Initialize: $\hat{x}(t_0^-) \triangleq r \in \mathcal{C}$, $\mathcal{D} \triangleq \mathcal{B}(r, 2 \text{diam}(\mathcal{C}))$,
 $\hat{e}_{-1} \triangleq \text{diam}(\mathcal{C})$, $\delta_{-1} \triangleq \eta C_h(\mathcal{C}) G_a \text{diam}(\mathcal{C})$, $G \triangleq 1$;
for $k = 0, 1, 2, \dots$ **do**
 Divide $\mathcal{B}(\hat{x}(t_k^-), G \hat{e}_{k-1})$ into M equal hypercubes
 and let $\hat{x}(t_k)$ equal the center of the hypercube that
 $x(t_k)$ falls into;
 Set $\hat{e}_k = G \hat{e}_{k-1} / M^{1/n}$ and $\delta_k = G \delta_{k-1} / M^{1/n}$;
 if $h(\hat{x}(t_k)) < \delta_k$ **then**
 Set $G = G_e$; \triangleright *estimation-only phase*
 Set $u(t) = 0$ for $t \in [t_k, t_{k+1})$;
 else
 Set $G = G_a$; \triangleright *estimation-actuation phase*
 Set $u(t) = k_{\text{safe}}(\hat{x}(t))$ for $t \in [t_k, t_{k+1})$;
 end
 Simulate $\hat{x}(t)$ over the interval $[t_k, t_{k+1})$ according
 to (1) whilst applying $u(t)$ to the real system;
end

III. LIMITED-INFORMATION SAFETY CONTROLLERS

In this section we will show how safety can be ensured using a nominally safe controller that is certified by an exponential barrier certificate, and we will present a class of separation-principle style limited-information safety controllers. As depicted in Figure 1, these controllers operate by cascading the observer described in Section II-B with a nominally safe control law. The proposed methodology is summarized in Algorithm 1. This algorithm consists of two sequential phases. The first is an *estimation-only phase* during which no actuation is applied to the system. The goal during this time is to sufficiently reduce the state estimation error so that state estimate $\hat{x}(t)$ may be used to select a control law which will keep the true state $x(t)$ safe. This phase should last no longer than the safety escape time described in Section II-A—as waiting any longer to apply actuation could result in the system going unsafe. The controller will apply actuation if the condition

$$h(\hat{x}(t_l)) \geq \delta_l \quad (19)$$

is satisfied for some transmission time t_l , where δ_k for $k \in \mathbb{N}$ is defined to be

$$\delta_k \triangleq \eta C_h(\mathcal{C}) G_a \hat{e}_k. \quad (20)$$

It will be shown that T and M can be chosen such that the *estimation-actuation phase* begins within the safety escape time $T_s(\delta)$. Once the *estimation-actuation phase* begins, the controller uses a control law $u(t) = k_{\text{safe}}(\hat{x}(t))$ which is assumed to satisfy the exponential barrier condition

$$L_f h(\hat{x}(t)) + L_g h(\hat{x}(t)) k_{\text{safe}}(\hat{x}(t)) \geq -\sigma h(\hat{x}(t)) \quad (21)$$

for some $\sigma \geq 0$ over each time interval $t \in [t_k, t_{k+1})$, $k \geq l$. The control law is piecewise continuous: it is continuous between transmission times but is allowed to update discontinuously each time the observer updates.

The main result of this section is stated below. The significance of the Theorem 2 is that the existence of a

nominally safe control law, certified by an exponential barrier function, guarantees the existence of a limited-information safety controller. This is proved by showing that Algorithm 1 is one particular safe controller. Notably in Algorithm 1, δ does not need to be known. Rather, a limited-information safety controller with given T and M can be certified to keep safe all system trajectories with $h(\mathbf{x}_0) \geq \delta$ down to a particular value of δ given by

$$\check{\delta} \triangleq \min_{k \in \mathbb{N}} \left\{ (1 + \eta G_a) C_h(\mathcal{C}) \left(\frac{G_e}{M^{1/n}} \right)^k \frac{\text{diam}(\mathcal{C})}{M^{1/n}} + (C_h(\mathcal{C}) B_f(\mathcal{C}) T) k \right\}. \quad (22)$$

Theorem 2: Let the transmission period T and size of the coding alphabet M satisfy

$$(e^{-\sigma T} \eta - 1) M^{\frac{1}{n}} \geq \eta G_a. \quad (23)$$

Then, any trajectory of (1) originating from $\mathcal{C}_{\check{\delta}}$ will remain in \mathcal{C} for all time under the Algorithm 1. \square

Note that, by monotonicity, for any $\delta > 0$ there exists T and M so that $\check{\delta} < \delta$ and (23) is satisfied. Regarding the hypotheses of Theorem 2, $\mathbf{x}_0 \in \mathcal{C}_{\check{\delta}}$ ensures that the *estimation-only phase* terminates before the safety escape time. Inequality (23) ensures that the state estimation error converges exponentially and that, once Algorithm 1 enters the *estimation-actuation phase*, it remains in this for all time. The proof of Theorem 2 will be given after the presentation of preliminary lemmata. The following lemma proves that Algorithm 1 enters the *estimation-actuation phase* before the system (1) goes unsafe due to the free evolution of the *estimation-only phase*. It also proves that $\hat{\mathbf{x}}(t)$ is confined to the compact set $\mathcal{D} \triangleq \mathcal{B}(\mathbf{r}, 2 \text{diam}(\mathcal{C}))$ up until that point, recalling from Section II-B that \mathbf{r} is the initial reference point of the observer.

Lemma 2: Let $l' \in \mathbb{N}$ be the minimizing argument of (22). If inequality (23) holds and $h(\mathbf{x}_0) = \delta \geq \check{\delta}$ then:

- 1) the time $t_{l'} = l'T$ is strictly less than the safety escape time $T_s(\delta)$;
- 2) throughout the initial *estimation-only phase* of Algorithm 1, $\hat{\mathbf{x}}(t)$ lies in \mathcal{D} ; and,
- 3) the algorithm enters the *estimation-actuation phase* at some time $t_l \leq t_{l'}$, i.e., within the safety escape time. \square

Proof: One can see from (22) that $t_{l'} < \delta / (C_h(\mathcal{C}) B_f(\mathcal{C}))$. By Theorem 1 one may conclude $t_{l'} < T_s(\delta)$. This proves the first part of the lemma.

Next, manipulating (23) and basic algebra reveal that $G_a/M^{1/n} < 1$. Because $G_e < G_a$ it follows that $G_e/M^{1/n} < 1$. If the algorithm immediately enters the *estimation-actuation phase* then the second and third parts of the lemma are vacuously true, so consider the contrary. The algorithm selects $\hat{\mathbf{x}}_0$ which is contained in $\mathcal{D} = \mathcal{B}(\mathbf{r}, 2 \text{diam}(\mathcal{C}))$. Suppose for contradiction that $\hat{\mathbf{x}}(t)$ leaves \mathcal{D} within the time interval $[t_0, t_1]$ over which $\mathbf{x}(t) \in \mathcal{C}$. By continuity there exists some time $\tau \in [t_0, t_1]$ such that $\|\hat{\mathbf{x}}(\tau) - \mathbf{r}\| = 2 \text{diam}(\mathcal{C})$. However, an application of Grönwall's inequality reveals

$$\begin{aligned} \|\mathbf{x}(\tau) - \hat{\mathbf{x}}(\tau)\| &\leq e^{C_f(\mathcal{D})(\tau-t_0)} \|\mathbf{x}(t_0) - \hat{\mathbf{x}}(t_0)\| \\ &\leq \frac{G_e}{M^{1/n}} \text{diam}(\mathcal{C}) < \text{diam}(\mathcal{C}). \end{aligned} \quad (24)$$

On the other hand, because $\mathbf{x}(\tau)$ is still safe it follows that $\|\mathbf{x}(\tau) - \mathbf{r}\| \leq \text{diam}(\mathcal{C})$. The triangle inequality gives

$$\begin{aligned} \|\hat{\mathbf{x}}(\tau) - \mathbf{r}\| &\leq \|\hat{\mathbf{x}}(\tau) - \mathbf{x}(\tau)\| + \|\mathbf{x}(\tau) - \mathbf{r}\| \\ &< 2 \text{diam}(\mathcal{C}) \end{aligned} \quad (25)$$

which is a contradiction. Therefore $\hat{\mathbf{x}}(t) \in \mathcal{D}$ for all time $t \in [t_0, t_1]$. The same argument shows that $\hat{\mathbf{x}}(t)$ lies in \mathcal{D} throughout the initial *estimation-only phase* so long as t is within the safety escape time. This point will be addressed next, which will complete the proof of the lemma's second and third parts.

As previously stated, the point of the assumption $\mathbf{x}_0 \in \mathcal{C}_{\check{\delta}}$ is to ensure the algorithm exits the *estimation-only phase* within the safety escape time. We already know that $t_{l'} < T_s(\delta)$. If the algorithm remains in the estimation phase up to time $t_{l'}$, the proof of Theorem 1 can be modified to show

$$h(\mathbf{x}(t_{l'})) \geq \delta - C_h(\mathcal{C}) B_f(\mathcal{C}) t_{l'}. \quad (26)$$

Combining this with (22) and the lemma in the Appendix proves

$$h(\hat{\mathbf{x}}(t_{l'})) \geq \eta C_h(\mathcal{C}) G_a \hat{e}_{l'} = \delta_{l'}. \quad (27)$$

Therefore, by at least time step $t_{l'} < T_s(\delta)$ the condition necessary for Algorithm 1 to enter the *estimation-actuation phase* is satisfied. This completes the proof of the lemma's second and third parts. \square

Algorithm 1 applies actuation once $h(\hat{\mathbf{x}}(t_l)) \geq \delta_l$ for some $l \in \mathbb{N}$. Next it is shown that this inequality is satisfied at each subsequent $k \geq l$ so long as $\mathbf{x}(t) \in \mathcal{D}$ (which happens, in particular, if $\mathbf{x}(t)$ remains safe).

Lemma 3: Assume that T and M satisfy the assumptions of Theorem 2. If for some time t_k inequality (19) holds, then inequality (19) also holds at time t_{k+1} so long as $\mathbf{x}(t) \in \mathcal{D}$ for all $t \in [t_k, t_{k+1}]$. \square

Proof: Suppose that for some $k \in \mathbb{N}$ the condition (19) is satisfied so that Algorithm 1 chooses to apply actuation over the interval $[t_k, t_{k+1}]$. Applying Grönwall's inequality to (21) reveals that

$$h(\hat{\mathbf{x}}(t)) \geq e^{-\sigma(t-t_k)} h(\hat{\mathbf{x}}(t_k)) \quad (28)$$

for all t in the interval $[t_k, t_{k+1}]$. Therefore, the state of the observer right before updating at time t_{k+1} satisfies $h(\hat{\mathbf{x}}(t_{k+1}^-)) \geq e^{-\sigma T} h(\hat{\mathbf{x}}(t_k)) > 0$. The goal of the proof is to show that the new state estimate $\hat{\mathbf{x}}(t_{k+1})$ chosen by the observer is safe. For contradiction, suppose that the observer chooses an unsafe state. By the lemma's assumption on the true state $\mathbf{x}(t)$, the observer chooses $\hat{\mathbf{x}}(t_{k+1})$ from the ball $\mathcal{B}(\hat{\mathbf{x}}(t_{k+1}^-), G_a \hat{e}_k)$. Since $h(\hat{\mathbf{x}}(t_{k+1}^-)) > 0$, $h(\hat{\mathbf{x}}(t_{k+1})) < 0$, and h is continuous, there exists some $\mathbf{z} \in \mathcal{B}(\hat{\mathbf{x}}(t_{k+1}^-), G_a \hat{e}_k)$ such that $h(\mathbf{z}) = 0$ and $\|\mathbf{z} - \hat{\mathbf{x}}(t_{k+1}^-)\| \leq G_a \hat{e}_k$. Note that \mathbf{z} is safe. By $C_h(\mathcal{C})$ -Lipschitz continuity of h on \mathcal{C} ,

$$0 = h(\mathbf{z}) \geq h(\hat{\mathbf{x}}(t_{k+1}^-)) - C_h(\mathcal{C}) \|\mathbf{z} - \hat{\mathbf{x}}(t_{k+1}^-)\|$$

$$\geq h(\hat{\mathbf{x}}(t_{k+1}^-)) - C_h(\mathcal{C}) G_a \hat{e}_k$$

$$\text{Eq. (28)} \rightarrow \geq e^{-\sigma T} h(\hat{\mathbf{x}}(t_k)) - C_h(\mathcal{C}) G_a \hat{e}_k$$

$$\text{Eq. (19), (20)} \rightarrow \geq (e^{-\sigma T} \eta - 1) C_h(\mathcal{C}) G_a \hat{e}_k$$

$$\text{Eq. (23)} \rightarrow \geq \eta \left(\frac{G_a}{M^{1/n}} \right) C_h(\mathcal{C}) G_a \|e(t_k)\|. \quad (29)$$

Therefore the state estimation error is zero and $\hat{\mathbf{x}}(t_{k+1}^-) = \hat{\mathbf{x}}(t_{k+1})$ since $M^{1/n}$ is odd. This gives the contradiction $h(\hat{\mathbf{x}}(t_{k+1})) > 0$. We have shown that at time t_{k+1} the state estimate $\hat{\mathbf{x}}(t_{k+1})$ chosen by the observer is safe. Since $\hat{\mathbf{x}}(t_{k+1}) \in \mathcal{C}$, the same Lipschitz continuity argument shows

$$h(\hat{\mathbf{x}}(t_{k+1})) \geq \eta C_h(\mathcal{C}) G_a \hat{e}_{k+1} = \delta_{k+1}. \quad (30)$$

Therefore, at time step $k + 1$, Algorithm 1 remains in the *estimation-actuation phase*, and so on. \square

Proof of Theorem 2: To this point it has been shown that the controller enters the *estimation-actuation phase* by time t_l . Since time t_l is smaller than the safety escape time $T_s(\delta)$, the system is safe up until t_l despite the lack of actuation. Let t_l be the time step in which actuation is first applied. We need only to show that $\mathbf{x}(t) \in \mathcal{C}$ for all $t \in [t_l, t_{l+1})$, and then the theorem follows from Lemma 3 and induction on $k \geq l$. Suppose that $\mathbf{x}(t)$ leaves the safe set within the time interval of interest. Then there exists a first time $\tau \in [t_l, t_{l+1})$ such that $h(\mathbf{x}(\tau)) = 0$. Up to that point,

$$\|\mathbf{e}(\tau)\| \leq G_a \hat{e}_l. \quad (31)$$

On the other hand, in light of equations (20), (23), and (28), the inequality

$$h(\hat{\mathbf{x}}(\tau)) \geq e^{-\sigma T} \eta C_h(\mathcal{C}) G_a \hat{e}_l > C_h(\mathcal{C}) G_a \hat{e}_l \quad (32)$$

holds for all $t \in [t_l, t_{l+1})$. Combining this inequality with the error bound (18) and applying Lemma 1 yields a contradiction. Thus, the true state $\mathbf{x}(t)$ remains safe for all time $t \in [t_l, t_{l+1})$. \square

IV. CONCLUSION

This letter explored safety-critical control of nonlinear systems over finite-rate state feedback channels. We showed that the existence of a nominally safe control law, certified by an exponential barrier function, suffices to provide safe control in limited-information scenarios. We introduced the notion of safety escape time and derived a lower bound on this time, enabling the construction of a class of limited-information safety controllers. This letter represents a starting point for a theory of safety-critical control with limited information.

APPENDIX

Lemma 4: If for some time $k \in \mathbb{N}$ the inequality

$$h(\mathbf{x}(t_k)) \geq (1 + \eta G_a) C_h(\mathcal{C}) \hat{e}_k \quad (33)$$

holds then $h(\hat{\mathbf{x}}(t_k)) \geq \eta C_h(\mathcal{C}) G_a \hat{e}_k$ also holds. \square

Proof: Inequality (33) ensures $\mathbf{x}(t_k) \in \mathcal{C}$. The Lipschitz-style argument that has been used throughout this letter can be used to show that $\hat{\mathbf{x}}(t_k) \in \mathcal{C}$ also. By Lipschitz continuity of h on \mathcal{C}

$$\begin{aligned} h(\hat{\mathbf{x}}(t_k)) &\geq h(\mathbf{x}(t_k)) - C_h(\mathcal{C}) \|\mathbf{z} - \mathbf{x}(t_k)\| \\ &\geq h(\mathbf{x}(t_k)) - C_h(\mathcal{C}) \hat{e}_k \end{aligned}$$

and hence $h(\hat{\mathbf{x}}(t_k)) \geq \eta C_h(\mathcal{C}) G_a \hat{e}_k$ as given in the lemma's statement. \square

ACKNOWLEDGMENT

The authors would like to celebrate the life and legacy of Sanjoy K. Mitter—a pioneer, mentor, and friend. His work on control with limited information was impetus for much research, including that contained in this letter.

REFERENCES

- [1] N. Elia and S. K. Mitter, "Stabilization of linear systems with limited information," *IEEE Trans. Autom. Control*, vol. 46, no. 9, pp. 1384–1400, Sep. 2001.
- [2] D. Liberzon, "Nonlinear control with limited information," *Commun. Inf. Syst.*, vol. 9, no. 1, pp. 41–58, 2009.
- [3] S. Prajna and A. Rantzer, "On the necessity of barrier certificates," *IFAC Proc. Vol.*, vol. 38, no. 1, pp. 526–531, 2005.
- [4] Q. Nguyen and K. Sreenath, "Exponential control barrier functions for enforcing high relative-degree safety-critical constraints," in *Proc. IEEE Am. Control Conf.*, 2016, pp. 322–328.
- [5] A. D. Ames, S. Coogan, M. Egerstedt, G. Notomista, K. Sreenath, and P. Tabuada, "Control barrier functions: Theory and applications," in *Proc. 18th Eur. Control Conf. (ECC)*, 2019, pp. 3420–3431.
- [6] M. Franceschetti, M. J. Khojasteh, and M. Z. Win, "The many facets of information in networked estimation and control," *Annu. Rev. Control, Robot., Auton. Syst.*, vol. 6, no. 1, pp. 233–259, 2023.
- [7] V. S. Borkar and S. K. Mitter, "LQG control with communication constraints," in *Communications, Computation, Control, and Signal Processing: A Tribute to Thomas Kailath*, A. Paulraj, V. Roychowdhury, and C. D. Schaper, Eds., New York, NY, USA: Springer, 1997, pp. 365–373.
- [8] W. S. Wong and R. W. Brockett, "Systems with finite communication bandwidth constraints II: Stabilization with limited information feedback," *IEEE Trans. Autom. Control*, vol. 44, no. 5, pp. 1049–1053, May 1999.
- [9] S. Tatikonda and S. K. Mitter, "Control under communication constraints," *IEEE Trans. Autom. Control*, vol. 49, no. 7, pp. 1056–1068, Jul. 2004.
- [10] D. Nešić and A. R. Teel, "Input-output stability properties of networked control systems," *IEEE Trans. Autom. Control*, vol. 49, no. 10, pp. 1650–1667, Oct. 2004.
- [11] D. Liberzon and J. P. Hespanha, "Stabilization of nonlinear systems with limited information feedback," *IEEE Trans. Autom. Control*, vol. 50, no. 6, pp. 910–915, Jun. 2005.
- [12] W. M. H. Heemels, A. R. Teel, N. Van de Wouw, and D. Nešić, "Networked control systems with communication constraints: Tradeoffs between transmission intervals, delays and performance," *IEEE Trans. Autom. Control*, vol. 55, no. 8, pp. 1781–1796, 2010.
- [13] W. S. Wong and R. W. Brockett, "Systems with finite communication bandwidth constraints I: State estimation problems," *IEEE Trans. Autom. Control*, vol. 42, no. 9, pp. 1294–1298, Sep. 1997.
- [14] D. Liberzon and S. Mitra, "Entropy and minimal data rates for state estimation and model detection," *IEEE Trans. Autom. Control*, vol. 63, no. 10, pp. 3330–3344, Oct. 2018.
- [15] Z. Liu, A. Conti, S. K. Mitter, and M. Z. Win, "Continuous-time distributed filtering with sensing and communication constraints," *IEEE J. Sel. Areas Inf. Theory*, vol. 4, pp. 667–681, Dec. 2023.
- [16] S. Kar, J. M. F. Moura, and H. V. Poor, "Distributed linear parameter estimation: Asymptotically efficient adaptive strategies," *SIAM J. Control Optim.*, vol. 51, no. 3, pp. 2200–2229, 2013.
- [17] Z. Liu, A. Conti, S. K. Mitter, and M. Z. Win, "Communication-efficient distributed learning over networks—Part I: Sufficient conditions for accuracy," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 4, pp. 1081–1101, Apr. 2023.
- [18] Z. Liu, A. Conti, S. K. Mitter, and M. Z. Win, "Communication-efficient distributed learning over networks—Part II: Necessary conditions for accuracy," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 4, pp. 1102–1119, Apr. 2023.
- [19] E. Garcia, P. J. Antsaklis, and L. A. Montestruque, *Model-Based Control of Networked Systems*. Cham, Switzerland: Springer, 2014.
- [20] D. Liberzon, "On stabilization of linear systems with limited information," *IEEE Trans. Autom. Control*, vol. 48, no. 2, pp. 304–307, Feb. 2003.
- [21] F. Fillipov, *Differential equations with discontinuous righthand sides*, Dordrecht, Netherlands: Springer, 1988.
- [22] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Hoboken, NJ, USA: Patience Hall, 2002.
- [23] D. Zwillinger and A. Jeffrey, *Table of Integrals, Series, and Products*, 7th ed. Amsterdam, The Netherlands: Elsevier, 2007.