

# Sampling and Quantization-Aware Control Barrier Functions for Safety-Critical Control of Cyber-Physical Systems

Luyao Niu, Bhaskar Ramasubramanian, Andrew Clark, and Radha Poovendran

**Abstract**—Safety is critical to a wide range of cyber-physical systems (CPS). Safety violations may damage CPS and cause harm to humans that co-exist in the operating environment. However, it is nontrivial to guarantee safety of complex CPS whose computation and control workload are shifted to the cloud. The reason is that the system states which evolve continuously are *sampled* periodically and *quantized* before being sent to the controller to compute control inputs. Moreover, the controller may operate with finite precision, making the coefficients involved in computation different from those of the actual system. Consequently, the synthesized control inputs to the system may lead to safety violations. In this paper, we study the co-design of quantizer and control inputs for such CPS. We construct a control barrier function (CBF) constraint for the digital controller and analyze how it differs from the CBF constraint formulated using the actual system states and dynamics. We observe that this difference is dependent on the sampling error, quantization error, and error induced by finite precision of the controller. We derive upper bounds of these errors and use the bounds to design a state quantizer. We show that the problem of designing a quantizer can be converted to a facility location problem. We prove the submodularity of the quantizer design problem, and leverage the submodularity property to develop an efficient greedy algorithm to construct the quantizer. Given the quantized states calculated by the quantizer, we modify the CBF constraint used by the controller to synthesize control inputs for the system at each sampling interval. We show that the synthesized inputs guarantee the system safety. We demonstrate the proposed approach using a numerical case study on a batch reactor system.

## I. INTRODUCTION

Safety is critical to multiple application domains of cyber-physical systems (CPS) such as autonomous driving and critical infrastructures [1], [2]. Safety is normally formulated as a forward invariance property of a given safety region [3]. Safety violations may damage the system and cause catastrophic harm to human operators. Control barrier function (CBF)-based approaches [3] are widely-used to synthesize controllers with safety guarantees.

However, the safety guarantees provided by CBF-based controllers may become invalid for complex CPS whose

L. Niu and R. Poovendran are with the Department of Electrical and Computer Engineering, University of Washington, Seattle, WA. Email: {luyaoniu,rp3}@uw.edu

B. Ramasubramanian is with the Electrical and Computer Engineering Program at Western Washington University, Bellingham, WA, USA. Email: ramasub@wwu.edu

A. Clark is with the Department of Electrical and Systems Engineering, Washington University in St. Louis, St. Louis, MO, USA. Email: andrewclark@wustl.edu

This work was supported by the Office of Naval Research grant N0014-23-1-2386, National Science Foundation grants CNS 1941670, CNS 2153136, and Air Force Office of Scientific Research grants FA9550-22-1-0054 and FA9550-23-1-0208.

computing and control workload are shifted to cloud and hence are not co-located with the system [4]–[6]. The reason is that while the system states evolve in a continuous manner, the controller are normally implemented in a digital manner. This *continuous-digital mismatch* leads to the following discrepancies between the system and controller. First, the controller treats the system as a sampled-data system [7]–[9]. The system states are sampled periodically and sent to the controller for computing control inputs, which will be applied to the system in a zero-order hold manner. Moreover, the sampled system states are quantized to accommodate limited network bandwidth and computation resource of the controller [10]–[12]. Finally, the controller may operate with finite precision, and thus the coefficients involved in any computation by the controller (e.g., system dynamics) may not exactly match those of the actual system.

At present, the effects of sampling and quantization have been investigated separately. Safety guarantees for sampled-data systems have been investigated in [7]–[9]. Quantizer design and feedback quantized control have been studied in [10]–[12] to ensure system stability. However, how to design a quantizer and synthesize control inputs for complex CPS that encounter continuous-digital mismatch to guarantee safety have been less studied.

In this paper, we study the problem of designing a quantizer and synthesizing control inputs to guarantee safety of CPS. We formulate the continuous-digital mismatch for such systems by comparing the CBF constraint formulated using the actual system dynamics and states and that formulated using the sampled and quantized states perceived by the controller that operates with finite precision. Here a CBF constraint is an inequality imposed on the control input to guarantee forward invariance of the safety region [3]. The difference between these two CBF constraints captures the effects of sampling and quantization errors, as well as finite precision representation of coefficients. To address these effects, we develop a quantizer for the system and bound the sampling and quantization errors. We modify the CBF constraint used by the controller to incorporate these errors and synthesize control inputs with safety guarantees. To summarize, this paper makes the following contributions.

- We bound the difference between those two CBF constraints. We derive a sufficient condition such that satisfying the CBF constraint formulated for controllers that operate with finite precision using quantized states at each sampling interval implies satisfying the other one formulated using actual system dynamics and states.
- We develop a quantizer for the system to quantize the

sampled states. We show that the quantizer design can be formulated as a submodular maximization problem. We develop a greedy algorithm to efficiently solve the submodular maximization problem and hence the quantizer design.

- We establish the safety guarantee of control inputs synthesized using our formulated CBF constraint for the controller. We empirically evaluate our approach to demonstrate its effectiveness.

The remainder of this paper is organized as follows. Related work is reviewed in Section II. We present the system model and problem formulation in Section III. Our solution approach is described in Section IV. Section V demonstrates the proposed approach using a numerical case study. Section VI concludes this paper.

## II. RELATED WORK

A wide range of methods has been proposed to synthesize control inputs to guarantee safety of CPS. Typical examples include Hamilton-Jacobi-Bellman-Isaccs (HJI) equation [13], mixed-integer program (MIP) [14], and control barrier function (CBF) as well as control Lyapunov function (CLF)-based approaches [1]–[3]. In particular, CBF-based approaches formulate a CBF constraint on the control input such that the safety region remains forward invariant. CBF-based approaches have been extended to systems operating under different scenarios such as in the presence stochastic noises [15], [16]. CBF-based approaches have also been extended to address the sampling errors and applied to sampled-data systems [7]–[9]. Different from our work, these approaches are not applicable to CPS whose controllers are subject to finite precision and quantization errors.

Quantization has been shown to result in performance degradation and lead to unstable system behaviors [17]. To address the effect of quantization, non-uniform quantizers such as logarithm quantizers have been proposed [10]. In [18], the problem of designing state quantizers for control systems is converted to the multicenter problem from computational geometry [19]. This paper develops a similar insight as from [18] to map the problem of designing quantizers to the facility location problem [20]. Different from [18], we show the submodularity property of the quantizer design problem, and develop a greedy algorithm to efficiently construct the quantizer.

In addition to quantizer design, quantized feedback control with stability guarantees have been studied in [10], [21], [22]. However, these approaches are not readily applicable to ensure safety of CPS since they ignore the effects of sampling and finite precision of controllers. In this paper, we characterize the errors induced by sampling and quantization, and develop a CBF-based approach to synthesize control inputs to guarantee system safety.

## III. PROBLEM FORMULATION

We consider a continuous-time control-affine system

$$\dot{x}_t = f(x_t) + g(x_t)u_t \quad (1)$$

where  $x_t \in \mathcal{X} \subseteq \mathbb{R}^n$  is the system state at time  $t \geq 0$  and  $u_t \in \mathcal{U} \subseteq \mathbb{R}^m$  is the input provided by the controller. We assume that the admissible control input set  $\mathcal{U}$  is compact and convex. Specifically, each dimension  $j$  of  $u_t$ , denoted as  $u_t[j]$ , is bounded as  $u_{min}[j] \leq u_t[j] \leq u_{max}[j]$ . We denote  $u_{min}, u_{max} \in \mathbb{R}^m$  as  $u_{min} = [u_{min}[1], \dots, u_{min}[m]]^T$  and  $u_{max} = [u_{max}[1], \dots, u_{max}[m]]^T$ . In Eq. (1),  $f(x_t) \in \mathbb{R}^n$  is a vector-valued function and  $g(x_t) \in \mathbb{R}^{n \times m}$  is a matrix-valued function.

**Assumption 1.** We assume that each entry of function  $f(x_t)$ , denoted as  $f_j(x_t)$ , is locally Lipschitz with Lipschitz coefficient  $L_{f_j}$ . We also assume that each entry of function  $g(x_t)$ , denoted as  $g_{ij}(x_t)$ , is locally Lipschitz with Lipschitz coefficient  $L_{g_{ij}}$ .

We note that Assumption 1 is commonly made for reachability and safety analysis [1], [23], [24]. We consider that the dynamical system (1) is subject to a safety constraint

$$x_t \in \mathcal{C}, \quad \forall t \geq 0. \quad (2)$$

where the safety set  $\mathcal{C}$  is compact and defined as

$$\mathcal{C} = \{x \in \mathcal{X} : h(x) \geq 0\}, \quad (3)$$

and  $h : \mathcal{X} \rightarrow \mathbb{R}$  is a continuously differentiable function. We say the system in Eq. (1) is safe with respect to  $\mathcal{C}$  if Eq. (2) holds. Since function  $h$  is continuously differentiable, it must be locally Lipschitz. We denote the Lipschitz coefficient of the  $i$ -entry of  $\frac{\partial h}{\partial x}(x)$  as  $L_{dh_i}$  for  $i = 1, \dots, n$ .

In this paper, we consider that the system (1) is controlled by a controller that may not be co-located with the dynamical system [4]–[6]. The controller is designed to ensure safety subject to two constraints. First, the controller may have limited memory and computation power, and thus operates with finite precision. Second, if the controller is not co-located with the dynamical system, the communication channel between the controller and system has limited bandwidth.

To accommodate these two constraints, the state of the system is *sampled* and *quantized* before being sent to the controller. Specifically, the system state  $x_t$  is sampled with a sampling period  $\tau$  as  $x_k^s = x_{k\tau}$  for  $k = 0, 1, \dots$ . The sampled system state  $x_k^s$  is then quantized via a quantizer  $q : \mathbb{R}^n \rightarrow \mathbb{Y}$ , where  $\mathbb{Y} \subset \mathbb{R}^n$  is a finite set. The controller receives the quantized sample state  $x_k^q = q(x_k^s)$  for each  $k = 0, 1, \dots$ . In the remainder of this paper, we refer to  $x_k^s$  and  $x_k^q$  as the *sampled state* and *quantized state*, respectively.

The quantized state  $x_k^q$  is sent to the controller to compute a control input so that system (1) is safe with respect to  $\mathcal{C}$ . Since the controller operates with finite precision, the quantized state  $x_k^q$  may be rounded as  $\tilde{x}_k^q$ . Without loss of generality, we consider that  $x_k^q$  can be represented by the finite precision adopted by the controller. This condition is not restrictive since we can always adapt the design of quantizer to meet this requirement. Similarly, coefficients of functions  $f$ ,  $g$ , and  $h$  are represented using finite precision, leading to functions  $\tilde{f}$ ,  $\tilde{g}$ , and  $\tilde{h}$  used by the controller if they are involved in any computation. The controller invokes a function  $\mu : \mathbb{Y} \rightarrow \mathbb{S}^m$  to compute a control input  $u_k$  for each

sampling period  $[k\tau, (k+1)\tau)$ , where  $\mathbb{S} \subset \mathbb{R}$  is the set of values that can be represented using the precision adopted by the controller. We consider that  $u_k$  can be transmitted to the system without requiring further quantization. The system applies control input  $u_k$  in a zero-order hold manner, i.e.,  $u_t = u_k$  for all  $t \in [k\tau, (k+1)\tau)$ , where  $k = 0, 1, \dots$

We summarize the problem of interest as follows.

**Problem 1.** Consider a dynamical system in the form of Eq. (1). Design a quantizer  $q$  and synthesize a controller  $\mu$  such that the system is safe with respect to set  $\mathcal{C}$  given in Eq. (3).

#### IV. SOLUTION APPROACH

In this section, we first present preliminary background and necessary notations used in our solution approach to Problem 1. We next describe an overview of our solution approach. Then we quantify the error incurred when sampling and quantizing the system states, and leverage the derived error to design a quantizer. We finally present the control synthesis to ensure safety of the system.

##### A. Notations and Preliminary Background

A continuous function  $\alpha : [0, a) \rightarrow [0, \infty)$  belongs to class  $\mathcal{K}$  if it is strictly increasing and  $\alpha(0) = 0$  [24]. A continuous function  $\alpha : [-b, a) \rightarrow (-\infty, \infty)$  belongs to extended class  $\mathcal{K}$  if it is strictly increasing and  $\alpha(0) = 0$  for some  $a, b > 0$ . We use  $x[j]$  to denote the  $j$ -th entry of a vector  $x$ .

Control barrier functions (CBFs) [3] have been widely used to verify safety or synthesize controllers with provable safety guarantees. A CBF is defined as follows.

**Definition 1** (Control Barrier Function (CBF) [3]). Consider the system (1) and a continuously differentiable function  $h : \mathcal{X} \rightarrow \mathbb{R}$ . If there exists an extended class  $\mathcal{K}$  function  $\alpha$  such that for all  $x \in \mathcal{X}$  the following inequality holds

$$\sup_{u \in \mathcal{U}} \left\{ \frac{\partial h}{\partial x}(x)f(x) + \frac{\partial h}{\partial x}(x)g(x)u + \alpha(h(x)) \right\} \geq 0,$$

then function  $h$  is a CBF.

Given Definition 1, one can compute the control input at each time  $t$  using the following quadratic program [3] to guarantee the safety of Eq. (1):

$$\min_{u \in \mathcal{U}} u^T u \quad (4a)$$

$$\text{s.t. } \frac{\partial h}{\partial x}(x)f(x) + \frac{\partial h}{\partial x}(x)g(x)u + \alpha(h(x)) \geq 0 \quad (4b)$$

When there exist no sampling and quantization errors, the control input obtained using Eq. (4) provides the following safety guarantee.

**Lemma 1** ([3]). Given the system (1) and a safe set (3) defined by some continuously differentiable function  $h : \mathcal{X} \rightarrow \mathbb{R}$ , if  $h$  is a CBF defined on  $\mathcal{X}$  and  $x_0 \in \mathcal{C}$ , then the control input obtained by Eq. (4) guarantees that  $x_t \in \mathcal{C}, \forall t \geq 0$ .

As we will show in subsequent sections, when the controller operates with finite precision and there exist sampling and quantization errors, the safety guarantee in Lemma 1 may become invalid.

##### B. Solution Overview

In this paper, we assume that function  $h$  in Eq. (3) is a CBF for system (1). We consider a CBF-based controller. When the controller operates with finite precision and the system states are sampled and quantized, the constraint in Eq. (4b) is represented in the following form to compute control input  $u_k$ :

$$\frac{\partial \tilde{h}}{\partial x}(x_k^q)[\tilde{f}(x_k^q) + \tilde{g}(x_k^q)u_k] + \tilde{\alpha}(\tilde{h}(x_k^q)) \geq 0. \quad (5)$$

We observe that the discrepancies between Eq. (4b) and Eq. (5) are raised by two reasons. First, the actual system state  $x_t$  becomes the quantized state  $x_k^q$ . Furthermore, functions  $f, g, h$ , and  $\alpha$  become  $\tilde{f}, \tilde{g}, \tilde{h}$ , and  $\tilde{\alpha}$  since the controller operates with finite precision. Consequently, a control input  $u_k$  that satisfies Eq. (5) may not necessarily satisfy Eq. (4b).

**Example 1.** Consider an LTI system  $\dot{x}_t = Ax_t + Bu_t$ , where  $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$  and  $B = [0, 1]^T$ . Suppose the safety set is defined as  $\mathcal{C} = \{x : h(x) = 1.01 - \|x\|_2 \geq 0\}$  and the controller uses 8-bit floats. When  $x = [0.8, 0.6]^T$  and  $\alpha(h(x)) = 10h(x)$ , one can verify that although there exists a feasible control input  $u = -0.716 \in [-1, 1]$  that satisfies Eq. (4b), the solution  $u = -0.3 \in [-1, 1]$  to Eq. (5) fails to guarantee safety.

To address the discrepancy between (4b) and Eq. (5), for any state  $x_t$  within any sampling interval  $[k\tau, (k+1)\tau)$ , quantized state  $x_k^q = q(x_k^s)$ , and control input  $u_k \in \mathcal{U}$ , we define the following error

$$e(x_t, x_k^q, u_k) = \frac{\partial \tilde{h}}{\partial x}(x_k^q)[\tilde{f}(x_k^q) + \tilde{g}(x_k^q)u_k] - \frac{\partial h}{\partial x}(x_t)[f(x_t) + g(x_t)u_k]. \quad (6)$$

Using Eq. (6), the following relation holds for any state  $x_t, t \in [k\tau, (k+1)\tau)$ , quantized state  $x_k^q = q(x_k^s)$ , and control input  $u_k \in \mathcal{U}$

$$\begin{aligned} & \frac{\partial h}{\partial x}(x_t)[f(x_t) + g(x_t)u_k] \\ &= \frac{\partial \tilde{h}}{\partial x}(x_k^q)[\tilde{f}(x_k^q) + \tilde{g}(x_k^q)u_k] - e(x_t, x_k^q, u_k) \\ &\geq \frac{\partial \tilde{h}}{\partial x}(x_k^q)[\tilde{f}(x_k^q) + \tilde{g}(x_k^q)u_k] - \max_{\substack{x_t, t \in [k\tau, (k+1)\tau) \\ x_k^q = q(x_k^s), u_k \in \mathcal{U}}} e(x_t, x_k^q, u_k). \end{aligned}$$

Therefore, if there exist some extended class  $\mathcal{K}$  functions  $\alpha(h(x_t))$  and  $\tilde{\alpha}(\tilde{h}(x_k^q))$  satisfying  $\alpha(h(x_t)) \geq \tilde{\alpha}(\tilde{h}(x_k^q))$ , then

$$\begin{aligned} & \frac{\partial h}{\partial x}(x_t)[f(x_t) + g(x_t)u_k] + \alpha(h(x_t)) \\ &\geq \frac{\partial \tilde{h}}{\partial x}(x_k^q)[\tilde{f}(x_k^q) + \tilde{g}(x_k^q)u_k] \\ &\quad + \tilde{\alpha}(\tilde{h}(x_k^q)) - \max_{\substack{x_t, t \in [k\tau, (k+1)\tau) \\ x_k^q = q(x_k^s), u_k \in \mathcal{U}}} e(x_t, x_k^q, u_k). \quad (7) \end{aligned}$$

Based on Eq. (7), we modify the CBF constraint used by the controller as follows:

$$\frac{\partial \tilde{h}}{\partial x}(x_k^q)[\tilde{f}(x_k^q) + \tilde{g}(x_k^q)u_k] + \tilde{\alpha}(\tilde{h}(x_k^q)) - \max_{\substack{x_t, t \in [k\tau, (k+1)\tau] \\ x_k^q = q(x_k^s), u_k \in \mathcal{U}}} e(x_t, x_k^q, u_k) \geq 0. \quad (8)$$

We have that if a control input  $u_k$  satisfies Eq. (8), then Eq. (4b) must hold using  $u_k$ , and thus the system is safe according to Lemma 1. We note that evaluating the modified CBF constraint in Eq. (8) and computing control input  $u_k$  requires quantifying and bounding the error  $e(x_t, x_k^q, u_k)$  in Eq. (6). In the next subsection, we will bound this error and use this bound to design quantizer  $q$ .

### C. Quantification of Sampling and Quantization Errors

In this subsection, we characterize and bound the error  $e(x_t, x_k^q, u_k)$  defined in Eq. (6). We first analyze the deviation between the quantized state  $x_k^q$  and actual system state  $x_t$  within one sampling period in the following lemma.

**Lemma 2.** *Let  $\Theta$  and  $\beta$  be given as*

$$\theta(u) = \sqrt{\sum_{j=1}^n \left( L_{f_j} + \sum_{s=1}^m L_{g_{j,s}} |u[s]| \right)^2}, \quad \Theta = \max_{u \in \mathcal{U}} \theta(u), \quad (9)$$

$$\beta = \sup_{x \in \mathcal{C}, u \in \mathcal{U}} (\|f(x) + g(x)u\|_2). \quad (10)$$

Consider a sampling interval  $[k\tau, (k+1)\tau)$ . Let  $\delta(x_k^q) = \|x_k^s - x_k^q\|_2$  be the quantization error incurred by quantizer  $q$  satisfying  $x_k^q = q(x_k^s)$ . If  $x_t$  is within a neighborhood of  $x_k^s$  such that  $f$  and  $g$  are Lipschitz over this neighborhood, then

$$\|x_k^q - x_t\|_2 \leq \frac{\|\beta\|_2}{\Theta} (e^{\Theta\tau} - 1) + \delta(x_k^q)$$

holds for all  $t \in [k\tau, (k+1)\tau)$ .

*Proof.* Consider the quantized state  $x_k^q = q(x_k^s)$  and actual state  $x_t$  for  $t \in [k\tau, (k+1)\tau)$ . We have that

$$\begin{aligned} \|x_k^q - x_t\|_2 &= \|x_k^q - x_k^s + x_k^s - x_t\|_2 \\ &\leq \|x_k^q - x_k^s\|_2 + \|x_k^s - x_t\|_2 \\ &\leq \delta(x_k^q) + \frac{\|\beta\|_2}{\Theta} (e^{\Theta\tau} - 1) \end{aligned}$$

where the first inequality holds by triangle inequality, and the second inequality holds by [7, Proposition 1] and the definition of  $\delta(x_k^q)$ .  $\square$

Lemma 2 shows that  $\|x_k^q - x_t\|_2$  can be bounded by the error  $\delta(x_k^q)$  introduced by quantization and the error  $\frac{\|\beta\|_2}{\Theta} (e^{\Theta\tau} - 1)$  introduced by sampling. For any quantized state  $x_k^q$ , we denote  $\mathcal{B}(x_k^q)$  as

$$\mathcal{B}(x_k^q) = \{x : \|x_k^q - x\|_2 \leq \delta(x_k^q) + \frac{\|\beta\|_2}{\Theta} (e^{\Theta\tau} - 1)\}. \quad (11)$$

In what follows, we decompose the error  $e(x_t, x_k^q, u_k)$  in Eq. (6) into those induced by sampling, quantization, and finite precision of controller. Specifically, we have

$$\begin{aligned} e(x_t, x_k^q, u_k) &= \frac{\partial \tilde{h}}{\partial x}(x_k^q)[\tilde{f}(x_k^q) + \tilde{g}(x_k^q)u_k] - \frac{\partial h}{\partial x}(x_t)[f(x_t) + g(x_t)u_k] \\ &= \frac{\partial \tilde{h}}{\partial x}(x_k^q)[\tilde{f}(x_k^q) + \tilde{g}(x_k^q)u_k] - \frac{\partial h}{\partial x}(x_t)[f(x_k^q) + g(x_k^q)u_k] \\ &\quad + \frac{\partial h}{\partial x}(x_t)[f(x_k^q) + g(x_k^q)u_k] - \frac{\partial h}{\partial x}(x_t)[f(x_t) + g(x_t)u_k] \\ &= \left[ \frac{\partial \tilde{h}}{\partial x}(x_k^q) - \frac{\partial \tilde{h}}{\partial x}(x_t) \right] [\tilde{f}(x_k^q) + \tilde{g}(x_k^q)u_k] \\ &\quad + \left[ \frac{\partial \tilde{h}}{\partial x}(x_t) - \frac{\partial h}{\partial x}(x_t) \right] [\tilde{f}(x_k^q) + \tilde{g}(x_k^q)u_k] \\ &\quad + \frac{\partial h}{\partial x}(x_t) [\tilde{f}(x_k^q) + \tilde{g}(x_k^q)u_k - f(x_k^q) - g(x_k^q)u_k] \\ &\quad + \frac{\partial h}{\partial x}(x_t) [f(x_k^q) - f(x_t) + g(x_k^q)u_k - g(x_t)u_k] \end{aligned} \quad (12)$$

where the first equality holds by the definition in Eq. (6), the second equality holds by subtracting and adding  $\frac{\partial h}{\partial x}(x_t)[f(x_k^q) + g(x_k^q)u_k]$ , and the last equality holds by subtracting and adding  $\frac{\partial \tilde{h}}{\partial x}(x_t)[\tilde{f}(x_k^q) + \tilde{g}(x_k^q)u_k]$ .

By inspecting Eq. (12), we have that the first term encodes the error induced when evaluating  $\frac{\partial \tilde{h}}{\partial x}(x)$  using the quantized state  $x_k^q$  instead of the actual state  $x_t$ . The second term depends on the error induced by the finite precision representation of  $\frac{\partial \tilde{h}}{\partial x}(x_t)$ . The third term is dependent on the error induced by using finite precision to represent the system dynamics in Eq. (1). The last term captures the difference between the system dynamics evaluated using the actual state  $x_t$  and quantized state  $x_k^q$ .

We denote the last two terms of  $e(x_t, x_k^q, u_k)$  as follows:

$$\begin{aligned} \epsilon(x_t, x_k^q, u_t) &= \frac{\partial h}{\partial x}(x_t) [\tilde{f}(x_k^q) + \tilde{g}(x_k^q)u_k - f(x_k^q) \\ &\quad - g(x_k^q)u_k] + \frac{\partial h}{\partial x}(x_t) [f(x_k^q) - f(x_t) + g(x_k^q)u_k - g(x_t)u_k]. \end{aligned} \quad (13)$$

Let  $\Delta_g(x_k^q) = \|\tilde{g}(x_k^q) - g(x_k^q)\|_F$  be the Frobenius norm of matrix  $\tilde{g}(x_k^q) - g(x_k^q)$  and  $\Delta_f(x_k^q) = \|\tilde{f}(x_k^q) - f(x_k^q)\|_2$ . We further define  $\gamma(x_k^q) := \sup_{x \in \mathcal{B}(x_k^q)} \|\frac{\partial h}{\partial x}(x)\|_2$ . Then the error in Eq. (13) can be bounded as follows.

**Proposition 1.** *The error in Eq. (13) is bounded as*

$$\|\epsilon(x_t, x_k^q, u_k)\|_2 \leq E(x_k^q), \quad (14)$$

where  $E(x_k^q)$  is given as

$$\begin{aligned} E(x_k^q) &= \gamma(x_k^q) \left[ \Delta_f(x_k^q) + \Delta_g(x_k^q)U \right. \\ &\quad \left. + \left( \sqrt{\sum_{i=1}^n L_{f_i}^2} + \sqrt{\sum_{i=1}^n \sum_{j=1}^m L_{g_{ij}}^2} U \right) \left( \frac{\|\beta\|_2}{\Theta} (e^{\Theta\tau} - 1) + \delta(x_k^q) \right) \right] \end{aligned} \quad (15)$$

where  $U$  is the upper bound of  $\|u_k\|_2$ , i.e.,  $\|u_k\|_2 \leq U$ .

*Proof.* The proof is deferred to Appendix.  $\square$

Using Proposition 1 and Eq. (12), a sufficient condition for Eq. (8) to hold is as follows:

$$\begin{aligned}
& \frac{\partial \tilde{h}}{\partial x}(x_k^q)[\tilde{f}(x_k^q) + \tilde{g}(x_k^q)u_k] + \tilde{\alpha}(\tilde{h}(x_k^q)) \\
& \quad - \max_{\substack{x_t, t \in [k\tau, (k+1)\tau] \\ x_k^q = q(x_k^*), u_k \in \mathcal{U}}} e(x_t, x_k^q, u_k) \\
& \geq \frac{\partial \tilde{h}}{\partial x}(x_k^q)[\tilde{f}(x_k^q) + \tilde{g}(x_k^q)u_k] + \tilde{\alpha}(\tilde{h}(x_k^q)) - E(x_k^q) \\
& \quad - \left[ \frac{\partial \tilde{h}}{\partial x}(x_k^q) - \frac{\partial \tilde{h}}{\partial x}(x_t) \right] [\tilde{f}(x_k^q) + \tilde{g}(x_k^q)u_k] \\
& \quad - \left[ \frac{\partial \tilde{h}}{\partial x}(x_t) - \frac{\partial \tilde{h}}{\partial x}(x_k^q) \right] [\tilde{f}(x_k^q) + \tilde{g}(x_k^q)u_k] \\
& = \frac{\partial h}{\partial x}(x_t)[\tilde{f}(x_k^q) + \tilde{g}(x_k^q)u_k] + \tilde{\alpha}(\tilde{h}(x_k^q)) - E(x_k^q) \\
& \geq 0. \tag{16}
\end{aligned}$$

Based on Eq. (8), we note that any control input  $u_k$  that satisfies Eq. (16) ensures that Eq. (4b) holds, and hence system (1) is safe. The existence of such a control input  $u_k$  satisfying Eq. (16) relies on the quantized state  $x_k^q$  and design of quantizer  $q$ , which will be described in the next subsection.

#### D. Quantizer Design

In this subsection, we describe how to design a quantizer  $q$  to guarantee safety. We have that a control input  $u_k$  that satisfies Eq. (16) must exist if the following inequality holds

$$\begin{aligned}
& \frac{\partial h}{\partial x}(x_t)\tilde{f}(x_k^q) + \left[ \frac{\partial h}{\partial x}(x_t)\tilde{g}(x_k^q) \right]^+ u_{max} \\
& + \left[ \frac{\partial h}{\partial x}(x_t)\tilde{g}(x_k^q) \right]^- u_{min} + \tilde{\alpha}(\tilde{h}(x_k^q)) - E(x_k^q) \geq 0 \tag{17}
\end{aligned}$$

for any system state  $x_t \in \mathcal{B}(x_k^q)$  within sampling interval  $[k\tau, (k+1)\tau)$ , where  $[v]^+$  and  $[v]^-$  replaces the non-negative and negative entries of a vector  $v$  with zero, respectively. We observe that the satisfaction of Eq. (17) depends on the quantized state  $x_k^q$ . Therefore, Eq. (17) can be leveraged as a design criterion of quantizer design.

Inspired by [18], we cast the problem of quantizer design as a facility location problem [20]. Specifically, a facility location problem focuses on selecting locations of facilities to optimize certain performance metrics such as average distance to demanding points and maximal coverage of all facilities. Here, a quantized state can be viewed as a facility and all other states  $x_t \in \mathcal{C}$  can be interpreted as demanding points that need to be covered by at least one facility. For any state  $x_t \in \mathcal{C}$ , it can be covered by a quantized state  $x_k^q$  if Eq. (17) is satisfied. We denote the finite set of quantized states  $\mathcal{V}$ . Then the problem of designing a quantizer can be formulated as follows:

$$\min_{\mathcal{V}} |\mathcal{V}| \tag{18a}$$

$$\text{s.t. } \mathcal{C} \subseteq \cup_{x' \in \mathcal{V}} \mathcal{D}(x') \tag{18b}$$

where  $|\mathcal{V}|$  is the cardinality of  $\mathcal{V}$  and set  $\mathcal{D}(x')$  is defined as

$$\mathcal{D}(x') = \{x : x \in \mathcal{B}(x') \text{ and Eq. (17) holds}\} \cap \mathcal{C}. \tag{19}$$

The objective function in Eq. (18a) minimizes the cardinality of set  $\mathcal{V}$ , i.e., the number of quantized states. The constraint in Eq. (18b) guarantees the quantizer to be capable of quantizing all states in the safety set  $\mathcal{C}$ . In the context of facility location problem, Eq. (18a) minimizes the number of facilities and Eq. (18b) ensures all demanding points are covered by at least one facility.

We note that directly solving Eq. (18) is challenging because of the following reasons. First, the safety set  $\mathcal{C}$  is an infinite and uncountable set while we can only choose a discrete finite set  $\mathcal{V}$ . Furthermore, it is difficult to evaluate the function  $\text{Vol}(\cup_{x' \in \mathcal{V}} \mathcal{D}(x'))$ . We address these challenges by first relaxing the constrained optimization problem as an unconstrained optimization problem, and then developing a sampling-based approach to derive an approximate solution.

We relax the constrained optimization in Eq. (18) as the following unconstrained problem

$$\min_{\mathcal{V} \subset \mathcal{C}} |\mathcal{V}| + \lambda(\text{Vol}(\mathcal{C}) - \text{Vol}(\cup_{x' \in \mathcal{V}} \mathcal{D}(x'))), \tag{20}$$

where  $\text{Vol}(\cdot)$  denotes the volume of a set and  $\lambda > 0$  is a hyper-parameter. The term  $\lambda(\text{Vol}(\mathcal{C}) - \text{Vol}(\cup_{x' \in \mathcal{V}} \mathcal{D}(x')))$  converts constraint (18b) to a penalty scaled by  $\lambda$ .

We then develop a sampling-based approach to approximately solve Eq. (20). We first sample a countable finite set of states, denoted as  $\mathcal{S}$ , from the safety set  $\mathcal{C}$ . We constrain  $\mathcal{V}$  as  $\mathcal{V} \subset \mathcal{S}$ . Given the samples in  $\mathcal{S}$ , we approximate the function  $\text{Vol}(\mathcal{D}(x'))$  by counting the number of samples in  $\mathcal{S}$  that satisfies Eq. (19). We characterize this sampling-based approach as follows.

**Proposition 2.** *Let  $\mathcal{V}^*$  be the optimal solution to Eq. (18) and  $\mathcal{V}$  be the solution to Eq. (20) obtained using the sampling-based approach. For any  $\varepsilon > 0$  and  $\phi \in [0, 1]$ , if  $|\mathcal{S}| \geq -\frac{1}{\varepsilon} \ln(\phi)$ , then with probability at least  $1 - \phi$ ,  $|\text{Vol}(\cup_{x' \in \mathcal{V}} \mathcal{D}(x')) - \text{Vol}(\cup_{x' \in \mathcal{V}^*} \mathcal{D}(x'))| \leq \varepsilon$  holds.*

*Proof.* Suppose that there exists a set of states  $\mathcal{T} \subset \mathcal{C}$  defined as  $\mathcal{T} = \mathcal{C} \setminus \cup_{x' \in \mathcal{V}} \mathcal{D}(x')$ . When  $\mathcal{T} \neq \emptyset$ , we have that  $\text{Vol}(\mathcal{C}) - \text{Vol}(\cup_{x' \in \mathcal{V}} \mathcal{D}(x')) > 0$ . We denote the probability mass of  $\mathcal{T}$  as  $\varepsilon$ . Then we have that

$$\begin{aligned}
& |\text{Vol}(\cup_{x' \in \mathcal{V}} \mathcal{D}(x')) - \text{Vol}(\cup_{x' \in \mathcal{V}^*} \mathcal{D}(x'))| \\
& = |\text{Vol}(\mathcal{C}) - \text{Vol}(\cup_{x' \in \mathcal{V}} \mathcal{D}(x')) \\
& \quad - (\text{Vol}(\mathcal{C}) - \text{Vol}(\cup_{x' \in \mathcal{V}^*} \mathcal{D}(x')))| \tag{21a}
\end{aligned}$$

$$\leq 0 + |\text{Vol}(\mathcal{C}) - \text{Vol}(\cup_{x' \in \mathcal{V}} \mathcal{D}(x'))| \tag{21b}$$

where Eq. (21a) holds by adding and subtracting  $\text{Vol}(\mathcal{C})$ , and Eq. (21b) follows from Eq. (19), triangle inequality, and the fact that  $\mathcal{V}^*$  is the optimal solution to Eq. (18).

When the samples in  $\mathcal{S}$  are drawn independently, we have that the probability that  $\mathcal{T} \cap \mathcal{S} = \emptyset$  is  $(1 - \varepsilon)^{|\mathcal{S}|} \leq e^{-\varepsilon|\mathcal{S}|}$ , where the inequality holds by the fact that  $1 - a \leq e^{-a}$ . Therefore, if  $e^{-\varepsilon|\mathcal{S}|} \leq \phi$ , we have that  $|\mathcal{S}| \geq -\frac{1}{\varepsilon} \ln(\phi)$ , completing our proof.  $\square$

We note that implementing the sampling-based approach is still computationally expensive due to the combinatorial nature of Eq. (20). In what follows, we characterize the formulation in Eq. (20) by showing its equivalence to a submodular maximization problem, which leads to an efficient greedy algorithm to obtain an approximate solution.

**Theorem 1.** *The objective function in Eq. (20) is monotone supermodular in  $\mathcal{V}$ .*

*Proof.* We consider  $\mathcal{V} \subset \mathcal{V}' \subset \mathcal{C}$ , where  $\mathcal{V}$  and  $\mathcal{V}'$  are finite. We first prove the monotonicity of Eq. (20). Note that  $\text{Vol}(\mathcal{D}(x')) \geq 0$ . Therefore,  $\text{Vol}(\cup_{x' \in \mathcal{V}} \mathcal{D}(x')) \leq \text{Vol}(\cup_{x' \in \mathcal{V}'} \mathcal{D}(x'))$  must hold for any  $\mathcal{V} \subset \mathcal{V}'$ . Hence, we have that

$$\text{Vol}(\mathcal{C}) - \text{Vol}(\cup_{x' \in \mathcal{V}} \mathcal{D}(x')) \geq \text{Vol}(\mathcal{C}) - \text{Vol}(\cup_{x' \in \mathcal{V}'} \mathcal{D}(x')),$$

and Eq. (20) is monotone nonincreasing in  $\mathcal{V}$ .

In what follows, we prove the supermodularity of Eq. (20). Let  $x'' \notin \mathcal{V}'$ . We consider two possible cases. First, if  $\mathcal{D}(x'') \cap \mathcal{D}(x') = \emptyset$  for all  $x' \in \mathcal{V}'$ , then

$$\begin{aligned} & \text{Vol}(\cup_{x' \in \mathcal{V} \cup \{x''\}} \mathcal{D}(x')) - \text{Vol}(\cup_{x' \in \mathcal{V}} \mathcal{D}(x')) \\ &= \text{Vol}(\cup_{x' \in \mathcal{V} \cup \{x''\}} \mathcal{D}(x')) - \text{Vol}(\cup_{x' \in \mathcal{V}'} \mathcal{D}(x')). \end{aligned} \quad (22)$$

In the second case, we consider that there exists some  $y \in \mathcal{V}'$  such that  $\mathcal{D}(x'') \cap \mathcal{D}(y) \neq \emptyset$ . Note that  $y$  may or may not belong to set  $\mathcal{V}$ . Hence, we have that

$$\begin{aligned} & \text{Vol}(\cup_{x' \in \mathcal{V} \cup \{x''\}} \mathcal{D}(x')) - \text{Vol}(\cup_{x' \in \mathcal{V}} \mathcal{D}(x')) \\ & \geq \text{Vol}(\cup_{x' \in \mathcal{V}' \cup \{x''\}} \mathcal{D}(x')) - \text{Vol}(\cup_{x' \in \mathcal{V}'} \mathcal{D}(x')). \end{aligned} \quad (23)$$

Combining Eq. (22) and (23), we have that

$$\begin{aligned} & \text{Vol}(\cup_{x' \in \mathcal{V} \cup \{x''\}} \mathcal{D}(x')) - \text{Vol}(\cup_{x' \in \mathcal{V}} \mathcal{D}(x')) \\ & \geq \text{Vol}(\cup_{x' \in \mathcal{V}' \cup \{x''\}} \mathcal{D}(x')) - \text{Vol}(\cup_{x' \in \mathcal{V}'} \mathcal{D}(x')) \end{aligned}$$

holds for all  $\mathcal{V} \subset \mathcal{V}'$ . Hence,  $\text{Vol}(\cup_{x' \in \mathcal{V}} \mathcal{D}(x'))$  is submodular in  $\mathcal{V}$ , indicating that  $-\text{Vol}(\cup_{x' \in \mathcal{V}} \mathcal{D}(x'))$  is supermodular.

Noting that  $|\mathcal{V}|$  is a modular function in  $\mathcal{V}$ , we have that Eq. (20) is supermodular in  $\mathcal{V}$  by definition.  $\square$

---

**Algorithm 1** Greedy Algorithm for Quantizer Design

---

- 1: Initialize  $\kappa \leftarrow 0$ ,  $\mathcal{V}^0 \leftarrow \emptyset$ ,  $\eta > 0$ , and
  - 2: **repeat**
  - 3:    $x^* = \text{argmax}_{x'' \in \mathcal{S} \setminus \mathcal{V}^\kappa} \{ \text{Vol}(\cup_{x' \in \mathcal{V}^\kappa \cup \{x''\}} \mathcal{D}(x')) - \text{Vol}(\cup_{x' \in \mathcal{V}^\kappa} \mathcal{D}(x')) \}$
  - 4:    $\mathcal{V}^{\kappa+1} \leftarrow \mathcal{V}^\kappa \cup \{x^*\}$
  - 5:    $\kappa \leftarrow \kappa + 1$
  - 6: **until**  $\text{Vol}(\cup_{x' \in \mathcal{V}^\kappa \cup \{x^*\}} \mathcal{D}(x')) - \text{Vol}(\cup_{x' \in \mathcal{V}^\kappa} \mathcal{D}(x')) < \eta$
  - 7: **return**  $\mathcal{V}^{\kappa-1}$
- 

If a function  $f(\cdot)$  is supermodular, then  $-f(\cdot)$  is submodular. Hence, Theorem 1 indicates that the optimization problem in Eq. (20) is a submodular maximization problem, which is NP-hard to solve [25]. Given Theorem 1 and the sampling-based approach, we develop a greedy algorithm presented in Algorithm 1 to efficiently compute  $\mathcal{V}$ . The

algorithm initializes the iteration index  $\kappa$  as zero and set  $\mathcal{V}^0$  at iteration zero as an empty set. Then the algorithm greedily searches for a state  $x'' \in \mathcal{S} \setminus \mathcal{V}^\kappa$  that maximizes  $\text{Vol}(\cup_{x' \in \mathcal{V}^\kappa \cup \{x''\}} \mathcal{D}(x')) - \text{Vol}(\cup_{x' \in \mathcal{V}^\kappa} \mathcal{D}(x'))$ . This step repeats until there is no such  $x'' \in \mathcal{S} \setminus \mathcal{V}^\kappa$  such that  $\text{Vol}(\cup_{x' \in \mathcal{V}^\kappa \cup \{x''\}} \mathcal{D}(x')) - \text{Vol}(\cup_{x' \in \mathcal{V}^\kappa} \mathcal{D}(x')) \geq \eta$ , where  $\eta$  is a tunable parameter. As shown in [25], the greedy algorithm provides us a provable optimality guarantee.

Let  $\mathcal{V}$  be the set of quantized states given by Algorithm 1. Then the quantizer works as follows. For any state  $x_t \in \mathcal{C}$ , the quantizer finds a state  $x \in \mathcal{S}$  as  $x = \text{argmin}_{x' \in \mathcal{S}} \|x' - x_t\|_2$ . Let  $x_k^q$  be the quantized state associated with  $x$ , i.e.,  $x \in \mathcal{D}(x_k^q)$ . Then state  $x_t$  is quantized as  $x_k^q$ .

### E. Control Synthesis and Safety Guarantee

In what follows, we describe how to compute a control input  $u_k$  given the quantized state  $x_k^q$  for each sampling interval. After receiving the the quantized state  $x_k^q$ , the controller computes a control input  $u_k \in \mathcal{U}$  such that the following constraint is satisfied

$$\frac{\partial \tilde{h}}{\partial x}(x_k^q)[\tilde{f}(x_k^q) + \tilde{g}(x_k^q)u_k] + \tilde{\alpha}(\tilde{h}(x_k^q)) - W(x_k^q, u_k) \geq 0, \quad (24)$$

where

$$\begin{aligned} W(x_k^q, u_k) = & E(x_k^q) + \left[ \sqrt{\sum_{i=1}^n L_{dh_i}^2} \left( \frac{\|\beta\|_2}{\Theta} (e^{\Theta\tau} - 1) \right) \right. \\ & \left. + \delta(x_k^q) + \Delta_{dh}(x_k^q) \right] \|\tilde{f}(x_k^q) + \tilde{g}(x_k^q)u_k\|_2 \end{aligned} \quad (25)$$

and  $\Delta_{dh}(x_k^q) = \sup_{x_t \in \mathcal{B}(x_k^q)} \left\| \frac{\partial \tilde{h}}{\partial x}(x_t) - \frac{\partial \tilde{h}}{\partial x}(x_k^q) \right\|_2$ . Note that constraint (24) can be converted to a quadratic constraint in  $u_k$  when  $x_k^q$  is given. Thus the controller can compute the control input by solving a quadratically constrained quadratic program. The synthesized control input  $u_k$  provides the following safety guarantee.

**Theorem 2.** *Suppose  $\mathcal{V}$  satisfies constraint (18b). Assume that there exist some extended class  $\mathcal{K}$  function  $\alpha(h(x_t))$  such that  $\alpha(h(x_t)) \geq \tilde{\alpha}(\tilde{h}(x_k^q))$  for all  $x_t \in \mathcal{D}(x_k^q)$ . If  $x_0, x_0^q \in \mathcal{C}$  and there exists a control input  $u_k$  at each sampling interval satisfying Eq. (24), then the system given in Eq. (1) is safe.*

*Proof.* We prove the theorem by showing that  $u_k$  guarantees that Eq. (4b) holds true. Based on Lemma 2 and the fact that  $\tilde{h}$  is continuously differentiable, we have that

$$\begin{aligned} & \left[ \frac{\partial \tilde{h}}{\partial x}(x_k^q) - \frac{\partial \tilde{h}}{\partial x}(x_t) \right] [\tilde{f}(x_k^q) + \tilde{g}(x_k^q)u_k] \leq \sqrt{\sum_{i=1}^n L_{dh_i}^2} \\ & \cdot \left( \frac{\|\beta\|_2}{\Theta} (e^{\Theta\tau} - 1) + \delta(x_k^q) \right) \|\tilde{f}(x_k^q) + \tilde{g}(x_k^q)u_k\|_2. \end{aligned} \quad (26)$$

Furthermore, the definition of  $\Delta_{dh}(x_k^q)$  yields that

$$\begin{aligned} & \left[ \frac{\partial \tilde{h}}{\partial x}(x_t) - \frac{\partial h}{\partial x}(x_t) \right] [\tilde{f}(x_k^q) + \tilde{g}(x_k^q)u_k] \\ & \leq \Delta_{dh}(x_k^q) \|\tilde{f}(x_k^q) + \tilde{g}(x_k^q)u_k\|_2. \end{aligned} \quad (27)$$

Using Proposition 1 and Eq. (26) - (27), we bound the error  $e(x_t, x_k^q, u_k)$  as

$$\|e(x_t, x_k^q, u_k)\|_2 \leq W(x_k^q, u_k). \quad (28)$$

We can thus rewrite Eq. (8) as

$$\begin{aligned} & \frac{\partial \tilde{h}}{\partial x}(x_k^q) [\tilde{f}(x_k^q) + \tilde{g}(x_k^q)u_k] + \tilde{\alpha}(\tilde{h}(x_k^q)) \\ & - \max_{\substack{x_t, t \in [k\tau, (k+1)\tau] \\ x_k^q = q(x_k^s), u_k \in \mathcal{U}}} e(x_t, x_k^q, u_k) \\ & \geq \frac{\partial \tilde{h}}{\partial x}(x_k^q) [\tilde{f}(x_k^q) + \tilde{g}(x_k^q)u_k] + \tilde{\alpha}(\tilde{h}(x_k^q)) - W(x_k^q). \end{aligned} \quad (29)$$

Using Eq. (7) and  $\alpha(h(x_t)) \geq \tilde{\alpha}(\tilde{h}(x_k^q))$  holds, Eq. (4b) must hold when  $u_k$  satisfies Eq. (24). Therefore, the system is safe when  $x_0 \in \mathcal{C}$  according to Lemma 1.  $\square$

## V. NUMERICAL CASE STUDY

Consider a batch reactor system [26] whose dynamics are given as  $\dot{x}_t = Ax_t + Bu_t$ , where

$$A = \begin{bmatrix} 1.38 & 0.2077 & 6.715 & 5.676 \\ 0.5814 & 4.29 & 0 & 0.675 \\ 1.067 & 4.273 & 6.654 & 5.893 \\ 0.048 & 4.273 & 1.343 & 2.104 \end{bmatrix},$$

$$B = \begin{bmatrix} 0 & 0 \\ 5.679 & 0 \\ 1.136 & -3.146 \\ 1.136 & 0 \end{bmatrix}.$$

We consider a safety constraint requiring the system to stay within a safety set  $\mathcal{C} = \{x : h(x) := 25 - x[3]^2 \geq 0\}$ . The dynamical system uses sampling period  $\tau = 0.1$ . The system adopts a logarithm quantizer [10] defined as

$$q(v) = \begin{cases} l_i, & \text{if } \frac{1}{1+\rho}l_i < v \leq \frac{1}{1-\rho}l_i, v > 0 \\ 0, & \text{if } v = 0 \\ -q(-v) & \text{if } v < 0 \end{cases}, \quad (30)$$

where  $\rho = \frac{1-\hat{\rho}}{1+\hat{\rho}}$ ,  $\hat{\rho} \in (0, 1)$ , and  $\mathcal{L} = \{\pm l_i, l_i = \hat{\rho}^i l_0, i = \pm 1, \pm 2, \dots\} \cup \{0\}$  is the set of quantization intervals.

We set the initial state  $x_0 = [2, 1, -1, 1]^T$  and  $\hat{\rho} = 0.5$ . When sampling and quantization error are ignored, we depict the trajectory of  $x[3]$  in Fig. 1 using dashed line. We have that the system violates the safety constraint at the sampling interval  $k = 6$ . Particularly, the system state becomes  $x_{0.6} = [17.5952, 0.1452, 5.4564, 7.2049]^T$  using a CBF-based controller. We observe that  $x[3] > 5$ , and thus the safety constraint is violated.

We next simulate our approach. We set  $\eta = 1$  and uniformly sample 500 states as set  $\mathcal{S}$  to compute the set of quantized states  $\mathcal{V}$ . The controller uses Eq. (24) to compute the control input at each sampling interval. We simulate

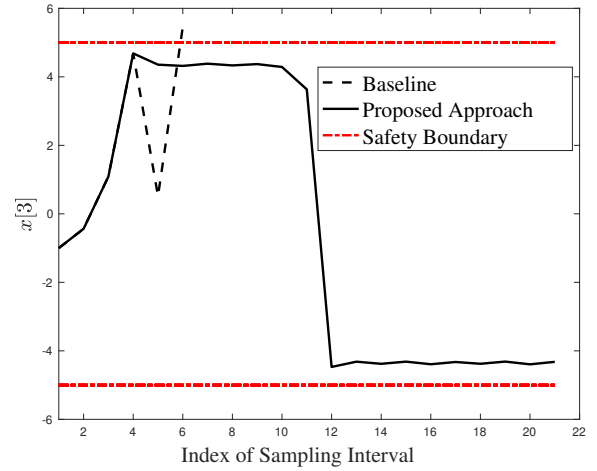


Fig. 1: This figure presents the trajectories of  $x[3]$  obtained using a CBF-based controller that ignores sampling and quantization error (dashed curve) and using our proposed approach (solid curve). The boundaries of the safety set  $\mathcal{C} = \{x : 25 - x[3]^2 \geq 0\}$  are shown using red lines. The CBF-based controller violates the safety constraint at sampling interval 6, while our approach guarantees safety.

the trajectory of  $x[3]$  using the synthesized control input, as shown by the solid curve in Figure 1. We observe that state  $x[3] \in [-5, 5]$  for all sampling intervals, and hence the system is safe.

## VI. CONCLUSION AND FUTURE WORK

In this paper, we studied the co-design of quantizer and control input for cyber-physical systems to guarantee safety. We formulated a control barrier function (CBF) constraint for the controller that operates with finite precision using quantized states at each sampling interval. We compared this CBF constraint with the one formulated using actual system dynamics and states, and derived an upper bound for their difference. We leveraged the upper bound to develop a quantizer for the system to quantize the states. We synthesized a control input for the controller with finite precision using the quantized states at each sampling interval. We proved that the synthesized control inputs guaranteed that the system to be safe. We demonstrated the proposed approach using a numerical case study. Our future work will investigate the co-design of quantizer and controller for hybrid and interconnected CPS.

## REFERENCES

- [1] L. Wang, A. D. Ames, and M. Egerstedt, "Safety barrier certificates for collisions-free multirobot systems," *IEEE Transactions on Robotics*, vol. 33, no. 3, pp. 661–674, 2017.
- [2] M. H. Cohen and C. Belta, "Approximate optimal control for safety-critical systems with control barrier functions," in *IEEE Conference on Decision and Control*. IEEE, 2020, pp. 2062–2067.
- [3] 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, 2016.

- [4] G. Cai, J. Dias, and L. Seneviratne, "A survey of small-scale unmanned aerial vehicles: Recent advances and future development trends," *Unmanned Systems*, vol. 2, no. 02, pp. 175–199, 2014.
- [5] M. Lv, D. Wang, Z. Peng, L. Liu, and H. Wang, "Event-triggered neural network control of autonomous surface vehicles over wireless network," *Science China Information Sciences*, vol. 63, no. 5, p. 150205, 2020.
- [6] X.-M. Zhang, Q.-L. Han, X. Ge, D. Ding, L. Ding, D. Yue, and C. Peng, "Networked control systems: A survey of trends and techniques," *IEEE/CAA Journal of Automatica Sinica*, vol. 7, no. 1, pp. 1–17, 2019.
- [7] L. Niu, H. Zhang, and A. Clark, "Safety-critical control synthesis for unknown sampled-data systems via control barrier functions," in *IEEE Conference on Decision and Control*, 2021, pp. 6806–6813.
- [8] J. Breeden, K. Garg, and D. Panagou, "Control barrier functions in sampled-data systems," *arXiv preprint arXiv:2103.03677*, 2021.
- [9] A. Singletary, Y. Chen, and A. D. Ames, "Control barrier functions for sampled-data systems with input delays," in *IEEE Conference on Decision and Control*. IEEE, 2020, pp. 804–809.
- [10] N. Elia and S. K. Mitter, "Stabilization of linear systems with limited information," *IEEE Transactions on Automatic Control*, vol. 46, no. 9, pp. 1384–1400, 2001.
- [11] H. Gao and T. Chen, "A new approach to quantized feedback control systems," *Automatica*, vol. 44, no. 2, pp. 534–542, 2008.
- [12] M. Wakaiki, A. Cetinkaya, and H. Ishii, "Stabilization of networked control systems under DoS attacks and output quantization," *IEEE Transactions on Automatic Control*, vol. 65, no. 8, pp. 3560–3575, 2019.
- [13] C. Tomlin, G. J. Pappas, and S. Sastry, "Conflict resolution for air traffic management: A study in multiagent hybrid systems," *IEEE Transactions on Automatic Control*, vol. 43, no. 4, pp. 509–521, 1998.
- [14] D. Mellinger, A. Kushleyev, and V. Kumar, "Mixed-integer quadratic program trajectory generation for heterogeneous quadrotor teams," in *IEEE International Conference on Robotics and Automation*. IEEE, 2012, pp. 477–483.
- [15] A. Clark, "Control barrier functions for stochastic systems," *Automatica*, vol. 130, p. 109688, 2021.
- [16] C. Santoyo, M. Dutreix, and S. Coogan, "A barrier function approach to finite-time stochastic system verification and control," *Automatica*, vol. 125, p. 109439, 2021.
- [17] R. Miller, M. Mousa, and A. Michel, "Quantization and overflow effects in digital implementations of linear dynamic controllers," *IEEE Transactions on Automatic Control*, vol. 33, no. 7, pp. 698–704, 1988.
- [18] F. Bullo and D. Liberzon, "Quantized control via locational optimization," *IEEE Transactions Automatic Control*, vol. 51, no. 1, pp. 2–13, 2006.
- [19] M. De Berg, *Computational geometry: Algorithms and applications*. Springer Science & Business Media, 2000.
- [20] A. Okabe and A. Suzuki, "Locational optimization problems solved through Voronoi diagrams," *European Journal of Operational research*, vol. 98, no. 3, pp. 445–456, 1997.
- [21] R. W. Brockett and D. Liberzon, "Quantized feedback stabilization of linear systems," *IEEE Transactions on Automatic Control*, vol. 45, no. 7, pp. 1279–1289, 2000.
- [22] I. R. Petersen and A. V. Savkin, "Multi-rate stabilization of multivariable discrete-time linear systems via a limited capacity communication channel," in *IEEE Conference on Decision and Control*, 2001, pp. 304–309.
- [23] A. Taylor, A. Singletary, Y. Yue, and A. Ames, "Learning for safety-critical control with control barrier functions," in *Learning for Dynamics and Control*. PMLR, 2020, pp. 708–717.
- [24] H. K. Khalil, *Nonlinear Systems*. Prentice hall Upper Saddle River, NJ, 2002, vol. 3.
- [25] A. Krause and D. Golovin, "Submodular function maximization." *Tractability*, vol. 3, no. 71-104, p. 3, 2014.
- [26] W. P. Heemels, K. H. Johansson, and P. Tabuada, "An introduction to event-triggered and self-triggered control," in *IEEE Conference on Decision and Control*. IEEE, 2012, pp. 3270–3285.

## APPENDIX

In this section, we present the proof of Proposition 1. We first consider the first term,  $\frac{\partial h}{\partial x}(x_t)[\tilde{f}(x_k^q) + \tilde{g}(x_k^q)u_k -$

$f(x_k^q) - g(x_k^q)u_k]$ , in  $\epsilon(x_t, x_k^q, u_k)$ . We have

$$\begin{aligned} & \left\| \frac{\partial h}{\partial x}(x_t)[\tilde{f}(x_k^q) + \tilde{g}(x_k^q)u_k - f(x_k^q) - g(x_k^q)u_k] \right\|_2 \\ & \leq \Delta_f(x_k^q) \left\| \frac{\partial h}{\partial x}(x_t) \right\|_2 + \Delta_g(x_k^q) \left\| \frac{\partial h}{\partial x}(x_t) \right\|_2 \|u_k\|_2, \end{aligned} \quad (31)$$

where the inequality holds by Cauchy-Schwarz inequality and the definitions of  $\Delta_f(x_k^q)$  and  $\Delta_g(x_k^q)$ . Here  $\Delta_f(x_k^q)$  and  $\Delta_g(x_k^q)$  capture the error induced by finite precision when representing  $f(x_k^q)$  and  $g(x_k^q)$ , respectively.

We next consider the second term  $\frac{\partial h}{\partial x}(x_t)[f(x_k^q) - f(x_t) + g(x_k^q)u_k - g(x_t)u_k]$  in  $\epsilon(x_t, x_k^q, u_k)$ . We have that

$$\begin{aligned} & \left\| \frac{\partial h}{\partial x}(x_t)[f(x_k^q) - f(x_t) + g(x_k^q)u_k - g(x_t)u_k] \right\|_2 \quad (32a) \\ & \leq \left\| \frac{\partial h}{\partial x}(x_t)[f(x_k^q) - f(x_t)] \right\|_2 \\ & \quad + \left\| \frac{\partial h}{\partial x}(x_t)[g(x_k^q)u_k - g(x_t)u_k] \right\|_2 \quad (32b) \end{aligned}$$

$$\begin{aligned} & \leq \sqrt{\sum_{i=1}^n L_{f_i}^2} \left\| \frac{\partial h}{\partial x}(x_t) \right\|_2 \|x_k^q - x_t\|_2 \\ & \quad + \sqrt{\sum_{i=1}^n \sum_{j=1}^m L_{g_{ij}}^2} \left\| \frac{\partial h}{\partial x}(x_t) \right\|_2 \|x_k^q - x_t\|_2 \|u_k\|_2, \end{aligned} \quad (32c)$$

where the first inequality holds by triangle inequality, and the second inequality holds by Cauchy-Schwarz inequality and Assumption 1.

Using triangle inequality, we then have that

$$\begin{aligned} & \|\epsilon(x_t, x_k^q, u_k)\|_2 \leq \Delta_f(x_k^q) \left\| \frac{\partial h}{\partial x}(x_t) \right\|_2 \\ & + \Delta_g(x_k^q) \left\| \frac{\partial h}{\partial x}(x_t) \right\|_2 \|u_k\|_2 + \sqrt{\sum_{i=1}^n L_{f_i}^2} \left\| \frac{\partial h}{\partial x}(x_t) \right\|_2 \|x_k^q - x_t\|_2 \\ & \quad + \sqrt{\sum_{i=1}^n \sum_{j=1}^m L_{g_{ij}}^2} \left\| \frac{\partial h}{\partial x}(x_t) \right\|_2 \|x_k^q - x_t\|_2 \|u_k\|_2. \end{aligned} \quad (33)$$

Using Lemma 2, we have that

$$\|x_k^q - x_t\|_2 \leq \frac{\|\beta\|_2}{\Theta} (e^{\Theta\tau} - 1) + \delta(x_k^q). \quad (34)$$

Moreover, for any state  $x_k^q$ , we have that

$$\left\| \frac{\partial h}{\partial x}(x_t) \right\|_2 \leq \gamma(x_k^q) := \sup_{x \in \mathcal{B}(x_k^q)} \left\| \frac{\partial h}{\partial x}(x) \right\|_2. \quad (35)$$

We further note that  $\|u_k\|_2 \leq U$ . Then Eq. (33) can be written as

$$\begin{aligned} & \|\epsilon(x_t, x_k^q, u_k)\|_2 \leq \gamma(x_k^q) \left[ \Delta_f(x_k^q) + \Delta_g(x_k^q)U \right. \\ & \left. + \left( \sqrt{\sum_{i=1}^n L_{f_i}^2} + \sqrt{\sum_{i=1}^n \sum_{j=1}^m L_{g_{ij}}^2} U \right) \left( \frac{\|\beta\|_2}{\Theta} (e^{\Theta\tau} - 1) + \delta(x_k^q) \right) \right], \end{aligned}$$

which completes our proof.