

# Probabilistic Control Barrier Functions: Safety in Probability for Discrete-Time Stochastic Systems

Pol Mestres, Blake Werner, Ryan K. Cosner, Aaron D. Ames

**Abstract**—Control systems operating in the real world face countless sources of unpredictable uncertainties. These random disturbances can render deterministic guarantees inapplicable and cause catastrophic safety failures. To overcome this, this paper proposes a method for designing safe controllers for discrete-time stochastic systems that retain probabilistic guarantees of safety. To do this we modify the traditional notion of a *control barrier function* (CBF) to explicitly account for these stochastic uncertainties and call these new modified functions *probabilistic CBFs*. We show that probabilistic CBFs can be used to design controllers that guarantee safety over a finite number of time steps with a prescribed probability. Next, by leveraging various uncertainty quantification methods, such as concentration inequalities, the scenario approach, and conformal prediction, we provide a variety of sufficient conditions that result in computationally tractable controllers with tunable probabilistic guarantees across a plethora of practical scenarios. Finally, we showcase the applicability of our results in simulation and hardware for the control of a quadruped robot.

## I. INTRODUCTION

Motivated by applications in autonomous driving, robotics, and aerospace systems, the topic of safety has recently received a lot of attention in the control theory community. In all of these systems, the design and verification of safe controllers is complicated by the presence of uncertainties in the system, stemming from state estimation errors, imperfect perception, unmodeled system dynamics, among others. Although the study and design of safe controllers that are robust to such sources of uncertainty has received considerable attention in the literature, most existing works assume deterministic and bounded disturbances with known bounds. Such assumptions are uncommon in practice and lead to overly conservative performance. In contrast, this work is motivated by the need to design controllers that maintain safety in the presence of stochastic and potentially unbounded disturbances, conditions that naturally arise when employing techniques such as Kalman filtering or simultaneous localization and mapping (SLAM).

### A. Literature Review

There are a variety of tools to design controllers with safety guarantees, including control barrier functions (CBFs) [1], reachability-based controllers [2], and model predictive control [3]. There has been a substantial effort to adapt these techniques for systems with uncertainty. In the CBF literature, this problem has been extensively studied

PM, BW, AA are with the Department of Mechanical and Civil Engineering, California Institute of Technology, Pasadena, CA 91125, USA. RC is with the Department of Mechanical Engineering, Tufts University, Medford, MA 02155, USA.

for continuous-time systems and deterministic and bounded disturbances [4]–[8]. Other works have explored the problem under various uncertainty models, such as stochastic and possibly unbounded [9]–[13], gaussian process-based [14], [15] or through Wasserstein ambiguity sets [16], [17]. For discrete-time systems, the works [12], [13], [18] leverage the theory of martingales to provide bounds on the probability of exiting the safe set within a certain time horizon. However, the condition on the control input derived in these papers only leverages the first moment of the distribution and can be conservative if more information about the uncertainty is available or not applicable if such first moment is unknown. On the other hand, [19] introduces the risk-sensitive notion of CVaR safety, and a CBF-based condition that can be used to certify it. However, controllers based on CVaR safety can be quite conservative in practice, and are only tractable for a small class of dynamics and safety constraints. Similar risk-aware approaches are taken in [20], [21]. The recent work [22] leverages the scenario approach [23] to design controllers that are safe with a prescribed probability by forcing them to be safe for a number of different scenarios (i.e., uncertainty realizations). Beyond CBF-based approaches, there exist various works in the literature that study the problem of guaranteeing safety constraints for discrete-time systems subject to stochastic uncertainty, particularly in the context of MPC [24]–[26].

### B. Statement of Contributions

We consider the problem of designing safe controllers for discrete-time stochastic systems. Our first contribution is the introduction of a probabilistic notion of CBF which we show can be used to design controllers with probabilistic safety guarantees over a finite time horizon. However, verifying whether a function is a probabilistic CBF is impractical, because it requires knowledge of the uncertainty distribution. In our second contribution we sidestep this difficulty by introducing two different sufficient conditions which guarantee that a function is a probabilistic CBF. These conditions only require knowledge of the moments of the distribution of the value of the CBF at the next time step. The first condition only leverages a known upper bound on the CBF and the first moment of such distribution, whereas the second condition uses its first and second moments. We also study various cases under which these conditions are convex in the control input and are thus amenable to optimization-based control design. Next, we consider the problem of verifying that a function is a probabilistic CBF when moment information is not available but instead we have a dataset consisting of uncertainty realizations. In our third contribution we leverage

concentration inequalities, scenario optimization, and conformal prediction to provide different sufficient conditions that ensure that a function is a probabilistic CBF with high confidence. Finally, we validate our results in both simulation and hardware on quadruped robots.

### C. Notation

We denote by  $\mathbb{N}, \mathbb{R}$  the set of natural and real numbers, respectively. We use bold symbols to represent vectors and matrices and non-bold symbols to represent scalar quantities. For  $N \in \mathbb{N}$ , we let  $[N] = \{1, \dots, N\}$ . Given  $\mathbf{x} \in \mathbb{R}^n$ ,  $\|\mathbf{x}\|$  denotes its Euclidean norm. Given a matrix  $\mathbf{M} \in \mathbb{R}^{n \times n}$ ,  $\text{Tr}(\mathbf{M})$  denotes its trace. For a random variable  $Z$ , we denote by  $\mathbb{E}[Z]$  and  $\text{Var}(Z)$  its expected value and variance, respectively. If  $Z$  is multivariate,  $\text{Cov}(Z)$  denotes its covariance matrix. Given  $\delta > 0$ , and  $k \in \mathbb{N}$  values  $R^{(1)}, \dots, R^{(k)}$  sorted in non-decreasing order,  $\text{Quantile}_{1-\delta}(R^{(1)}, \dots, R^{(k)}, \infty)$  denotes the  $1 - \delta$  quantile of the empirical distribution of the values  $R^{(1)}, \dots, R^{(k)}, \infty$ , which can be equivalently obtained as  $R^{(p)}$ , where  $p = \lceil (k+1)(1-\delta) \rceil$ , and  $\lceil \cdot \rceil$  denotes the ceiling function.

## II. MOTIVATION

Throughout this paper we consider a discrete-time stochastic system of the form:

$$\mathbf{x}_{t+1} = \mathbf{F}(\mathbf{x}_t, \mathbf{u}_t, \mathbf{d}_t), \quad (1)$$

with  $\mathbf{x}_t \in \mathbb{R}^n$  the state,  $\mathbf{u}_t \in \mathbb{R}^m$  the control input, and  $\mathbf{d}_t \in \mathbb{R}^d$  a random variable distributed according to a distribution  $\mathcal{D}$ , i.e.,  $\mathbf{d}_t \sim \mathcal{D}$ . For simplicity, we assume that  $\mathcal{D}$  is state and time independent.

Given a user-prescribed safe set  $\mathcal{C} \subset \mathbb{R}^n$ , defined as the 0-superlevel set of a continuous function  $h$ , (i.e.,  $\mathcal{C} = \{\mathbf{x} \in \mathbb{R}^n : h(\mathbf{x}) \geq 0\}$ ), our goal is to design a state-feedback controller  $\mathbf{k} : \mathbb{R}^n \rightarrow \mathbb{R}^m$  that ensures that the iterates of the closed-loop system obtained from (1) by using  $\mathbf{k}$  remain in  $\mathcal{C}$  at all times. However, the presence of the stochastic disturbance in (1) poses a significant challenge in finding such controller. In fact, this is impossible if the distribution  $\mathcal{D}$  is not supported on a bounded set, because the iterates will exit the set  $\mathcal{C}$  in finite time with probability 1 (cf. [27, Remark 1]). Hence, we are interested in a notion of safety over a finite time horizon and with a prescribed probability, which we formalize next.

**Definition 1.** (Probabilistic safety over a finite time horizon): *Let  $\epsilon \in (0, 1)$ , and  $H \in \mathbb{N}$ . We say that the controller  $\mathbf{k} : \mathbb{R}^n \rightarrow \mathbb{R}^m$  is  $\epsilon$ -safe over a time horizon  $H$  if for any  $\mathbf{x}_0 \in \mathcal{C}$ , the iterates of (1) under  $\mathbf{u}_t = \mathbf{k}(\mathbf{x}_t)$  are such that*

$$\mathbb{P}\left(\bigcap_{t=0}^H \{\mathbf{x}_t \in \mathcal{C}\}\right) \geq 1 - \epsilon.$$

Next we introduce the notion of probabilistic CBF, which will prove to be a useful tool to certify probabilistic safety over a finite time horizon.

**Definition 2.** (Probabilistic CBF): *Let  $\delta \in (0, 1)$ . The function  $h : \mathbb{R}^n \rightarrow \mathbb{R}$  is a  $\delta$ -probabilistic CBF if there exists*

$\alpha \in [0, 1]$  such that for each  $\mathbf{x} \in \mathcal{C}$  there exists  $\mathbf{u} \in \mathbb{R}^m$  such that

$$\mathbb{P}\left(h(\mathbf{F}(\mathbf{x}, \mathbf{u}, \mathbf{d})) \geq \alpha h(\mathbf{x})\right) \geq 1 - \delta, \quad (2)$$

where  $\mathbf{d} \sim \mathcal{D}$ . We note that Definition 2 is simply a probabilistic version of the standard CBF condition for discrete-time systems [28]. In the sequel, whenever the parameter  $\delta$  is clear from the context, we refer to a  $\delta$ -probabilistic CBF simply as a probabilistic CBF. We also fix  $\alpha \in [0, 1]$ ,  $\delta \in (0, 1)$  and let  $\Delta h : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^d \rightarrow \mathbb{R}$  be defined as  $\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d}) = h(\mathbf{F}(\mathbf{x}, \mathbf{u}, \mathbf{d})) - \alpha h(\mathbf{x})$ . The following result shows that probabilistic CBFs can be used to verify the notion of probabilistic safety over a given time horizon in Definition 1.

**Proposition 1.** (Probabilistic guarantees over a finite time horizon): *Let  $\mathbf{x}_0 \in \mathcal{C}$ ,  $H \in \mathbb{N}$ , and  $\mathbf{k} : \mathbb{R}^n \rightarrow \mathbb{R}^m$  be such that  $\mathbf{u} = \mathbf{k}(\mathbf{x})$  satisfies (2) for each  $\mathbf{x} \in \mathcal{C}$ . Then, for the system (1) under  $\mathbf{u}_t = \mathbf{k}(\mathbf{x}_t)$ , we have*

$$\mathbb{P}\left(\bigcap_{t=0}^H \{\mathbf{x}_t \in \mathcal{C}\}\right) \geq (1 - \delta)^H. \quad (3)$$

In particular, for any  $\epsilon > 0$ , if

$$\delta \leq 1 - (1 - \epsilon)^{\frac{1}{H}}, \quad (4)$$

then  $\mathbf{k}$  is  $\epsilon$ -safe over a time horizon  $H$ .

*Proof.* Define the event  $\Gamma_t := \bigcap_{s=0}^t \{\mathbf{x}_s \in \mathcal{C}\}$ . By the law of conditional probabilities [29, page 63]

$$\mathbb{P}\left(\bigcap_{t=0}^H \{\mathbf{x}_t \in \mathcal{C}\}\right) = \prod_{t=1}^H \mathbb{P}\left(\mathbf{x}_t \in \mathcal{C} | \Gamma_{t-1}\right). \quad (5)$$

Note that for each  $t \in [H]$ ,

$$\begin{aligned} \mathbb{P}\left(\mathbf{x}_t \in \mathcal{C} | \Gamma_{t-1}\right) &= \\ \mathbb{P}\left(h(\mathbf{F}(\mathbf{x}_{t-1}, \mathbf{k}(\mathbf{x}_{t-1}), \mathbf{d}_{t-1})) \geq 0 | \Gamma_{t-1}\right) &\geq \\ \mathbb{P}\left(h(\mathbf{F}(\mathbf{x}_{t-1}, \mathbf{k}(\mathbf{x}_{t-1}), \mathbf{d}_{t-1})) \geq \alpha h(\mathbf{x}_{t-1}) | \Gamma_{t-1}\right) &\geq 1 - \delta, \end{aligned}$$

The first inequality follows from the fact that if  $\mathbf{x}_{t-1} \in \mathcal{C}$ , then  $h(\mathbf{x}_{t-1}) \geq 0$ , and  $h(\mathbf{F}(\mathbf{x}_{t-1}, \mathbf{k}(\mathbf{x}_{t-1}), \mathbf{d}_{t-1})) \geq \alpha h(\mathbf{x}_{t-1})$  implies that  $h(\mathbf{F}(\mathbf{x}_{t-1}, \mathbf{k}(\mathbf{x}_{t-1}), \mathbf{d}_{t-1})) \geq 0$ . The last inequality follows from (2). Hence, by (5) we have that (3) holds. By taking  $\delta$  to satisfy (4), it follows that  $(1 - \delta)^H \geq 1 - \epsilon$ , and  $\mathbf{k}$  is  $\epsilon$ -safe over a time horizon  $H$ .  $\square$

Proposition 1 shows that controllers obtained from the probabilistic CBF condition (2) can be used to verify the notion of probabilistic safety introduced in Definition 1. Unfortunately, verifying that  $h$  is a probabilistic CBF as per Definition 2 is impractical because it requires computing a probability over the distribution  $\mathcal{D}$ , which is often unknown. In the following sections, we devise various sufficient conditions that can be used to verify Definition 2 without requiring full knowledge of the distribution  $\mathcal{D}$ .

### III. VERIFICATION OF PROBABILISTIC CBFs USING KNOWN MOMENTS

In practice, although one generally does not know the full distribution  $\mathcal{D}$ , it is common to have information about some of its moments. In this section we study how to leverage this information to verify that  $h$  is a probabilistic CBF. Our first result establishes a sufficient condition for that to hold by imposing a constraint on the expected value of  $\Delta h$ .

**Proposition 2.** (Markov-based condition): *Suppose that there exists  $b > 0$  such that for all  $\mathbf{x} \in \mathcal{C}$  and  $\mathbf{u} \in \mathbb{R}^m$ ,  $\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d}) \leq b$  almost surely. Furthermore, suppose that for all  $\mathbf{x} \in \mathcal{C}$ , there exists  $\mathbf{u} \in \mathbb{R}^m$  such that*

$$\mathbb{E}[\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d})] \geq b(1 - \delta). \quad (6)$$

*Then,  $h$  is a  $\delta$ -probabilistic CBF.*

*Proof.* Note that by assumption, for each  $\mathbf{x} \in \mathcal{C}$  there exists  $\mathbf{u} \in \mathbb{R}^m$  and  $b > 0$  such that the random variable  $b - \Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d})$  is positive almost surely. By Markov's inequality [30, Theorem 5.11],

$$\begin{aligned} \mathbb{P}(\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d}) \leq 0) &= \mathbb{P}(b - \Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d}) \geq b) \\ &\leq \frac{\mathbb{E}[b - \Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d})]}{b}. \end{aligned}$$

Now, from (6) we have that  $\mathbb{E}[b - \nabla h(\mathbf{x}, \mathbf{u}, \mathbf{d})] \leq b\delta$  and hence we have for each  $\mathbf{x} \in \mathcal{C}$  there exists  $\mathbf{u} \in \mathbb{R}^m$  such that  $\mathbb{P}(\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d}) \leq 0) \leq \delta$ , from where it follows that  $h$  is a  $\delta$ -probabilistic CBF.  $\square$

Since (6) is derived through Markov's inequality, we refer to it as the **Markov-based condition**. Proposition 2 shows that even if the distribution  $\mathcal{D}$  is unknown, the expected values of the random variable  $\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d})$  can be used to verify that  $h$  is a probabilistic CBF. Note also that one simple condition that ensures that  $\Delta h$  is uniformly upper bounded is if  $h$  is uniformly upper bounded.

In practice, in order to design a controller that ensures probabilistic safety, we are interested in including input constraints such as (6) as constraints of an optimization problem. However, as it is also the case for deterministic discrete-time CBFs [28], the condition (6) is generally not convex in  $\mathbf{u}$ , which compromises the computational tractability of such optimization-based control design. The following result leverages the convexity properties of  $h$  to derive another set of conditions for which the convexity properties with respect to  $\mathbf{u}$  are easier to analyze.

**Corollary 3.** (Markov-based condition under convexity assumptions): *Suppose that there exists  $b > 0$  such that for all  $\mathbf{x} \in \mathcal{C}$  and  $\mathbf{u} \in \mathbb{R}^m$ ,  $\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d}) \leq b$  almost surely. Further assume that one of the following conditions holds:*

- $h$  is convex and for all  $\mathbf{x} \in \mathcal{C}$  there exists  $\mathbf{u} \in \mathbb{R}^m$  such that

$$\tilde{\Delta}h(\mathbf{x}, \mathbf{u}) := h(\mathbb{E}[\mathbf{F}(\mathbf{x}, \mathbf{u}, \mathbf{d})]) - \alpha h(\mathbf{x}) \geq b(1 - \delta); \quad (7)$$

- $h$  is concave,  $\lambda > 0$  satisfies  $\sup_{\mathbf{x} \in \mathbb{R}^n} \|\nabla^2 h(\mathbf{x})\| \leq \lambda$  and for all  $\mathbf{x} \in \mathcal{C}$ , there exists  $\mathbf{u} \in \mathbb{R}^m$  such that

$$\hat{\Delta}h(\mathbf{x}, \mathbf{u}) := \tilde{\Delta}h(\mathbf{x}, \mathbf{u}) - \frac{\lambda}{2} \text{Tr}(\text{Cov}(\mathbf{F}(\mathbf{x}, \mathbf{u}, \mathbf{d}))) \geq b(1 - \delta), \quad (8)$$

*Then,  $h$  is a  $\delta$ -probabilistic CBF.*

*Proof.* First, if  $h$  is convex, by Jensen's inequality, we have  $h(\mathbb{E}[\mathbf{F}(\mathbf{x}, \mathbf{u}, \mathbf{d})]) \leq \mathbb{E}[h(\mathbf{F}(\mathbf{x}, \mathbf{u}, \mathbf{d}))]$ . Therefore, the satisfaction of (7) implies that (6) holds. By Proposition 2, this means that  $h$  is a  $\delta$ -probabilistic CBF. Second, if  $h$  is concave, by [18, Lemma 1] (which is based on a result from [31]), we have that

$$\begin{aligned} \mathbb{E}[h(\mathbf{F}(\mathbf{x}, \mathbf{u}, \mathbf{d}))] &\geq \\ h(\mathbb{E}[\mathbf{F}(\mathbf{x}, \mathbf{u}, \mathbf{d})]) - \frac{\lambda}{2} \text{Tr}(\text{Cov}(\mathbf{F}(\mathbf{x}, \mathbf{u}, \mathbf{d}))). \end{aligned} \quad (9)$$

Hence, the satisfaction of (8) implies that (6) holds. By Proposition 2, this means that  $h$  is a  $\delta$ -probabilistic CBF.  $\square$

The following remark comments on the convexity properties with respect to  $\mathbf{u}$  of the conditions (7) and (8).

**Remark 1.** (Convexity properties of constraints (7), (8)): *For general  $\mathbf{F}$  and  $h$ , conditions (7) and (8) are not convex in  $\mathbf{u}$ . In fact, even if  $\mathbf{F}$  is affine in  $\mathbf{u}$ , and  $h$  is convex, (7) is not convex in  $\mathbf{u}$ . However, in some cases, such as when  $h$  is quadratic, (7) is a quadratic concave constraint in  $\mathbf{u}$ . Although it can not be included as a constraint in a convex program, there exist very efficient heuristics to solve non-convex quadratically constrained quadratic programs (QCQPs) [32]. On the other hand, (8) is convex in  $\mathbf{u}$  if  $\mathbf{F}$  is affine in  $\mathbf{u}$  and  $\mathbf{d}$  is additive, (i.e.  $\mathbf{F}(\mathbf{x}, \mathbf{u}, \mathbf{d}) = F_1(\mathbf{x}) + F_2\mathbf{u} + \mathbf{d}$ , for  $F_1 : \mathbb{R}^n \rightarrow \mathbb{R}^n$  and  $F_2 \in \mathbb{R}^{n \times m}$ ) and  $h$  is concave. Additionally, if  $\text{Trace}(\text{Cov}(h(\mathbf{F}(\mathbf{x}, \mathbf{u}, \mathbf{d})))) \leq c$  for some  $c > 0$ ,  $h$  is concave and  $\mathbf{F}$  is affine in  $\mathbf{u}$ , then the condition that we get after replacing  $\frac{\lambda}{2} \text{Tr}(\text{Cov}(\mathbf{F}(\mathbf{x}, \mathbf{u}, \mathbf{d})))$  by  $\frac{\lambda c}{2}$  in (8) is convex in  $\mathbf{u}$ .* •

The following result provides another sufficient condition for  $h$  to be a probabilistic CBF. In this case, the condition does not require a known upper bound for  $\Delta h$  but assumes knowledge of its variance (i.e., its second moment).

**Proposition 4.** (Cantelli-based condition): *Suppose that for each  $\mathbf{x} \in \mathcal{C}$  there exists  $\mathbf{u} \in \mathbb{R}^m$  such that*

$$\mathbb{E}[\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d})] - \sqrt{\text{Var}(\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d})) \frac{1 - \delta}{\delta}} \geq 0. \quad (10)$$

*Then,  $h$  is a  $\delta$ -probabilistic CBF.*

*Proof.* By Cantelli's inequality [33], we have that for any  $\mathbf{x} \in \mathbb{R}^n$ ,  $\mathbf{u} \in \mathbb{R}^m$ , and  $\kappa \geq 0$ ,

$$\begin{aligned} \mathbb{P}(\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d}) \leq \mathbb{E}[\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d})] - \kappa) \\ \leq \frac{\text{Var}(\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d}))}{\text{Var}(\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d})) + \kappa^2} \end{aligned}$$

By taking  $\kappa = \kappa_{\mathbf{x}, \mathbf{u}} := \sqrt{\text{Var}(\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d})) \frac{1-\delta}{\delta}}$ , we have

$$\mathbb{P}(\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d}) \leq \mathbb{E}[\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d})] - \kappa_{\mathbf{x}, \mathbf{u}}) \leq \delta.$$

Now, for each  $\mathbf{x} \in \mathcal{C}$ , select  $\mathbf{u} \in \mathbb{R}^m$  so that (10) holds. It follows that

$$\begin{aligned} \mathbb{P}(\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d}) \leq 0) &\leq \\ \mathbb{P}(\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d}) \leq \mathbb{E}[\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d})] - \kappa_{\mathbf{x}, \mathbf{u}}) &\leq \delta. \end{aligned}$$

Hence,  $\mathbb{P}(\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d}) \geq 0) \geq 1 - \delta$ , and  $h$  is a  $\delta$ -probabilistic CBF.  $\square$

Since (10) is derived through Cantelli's inequality, we refer to it as the **Cantelli-based condition**. Proposition 4 shows that even if the distribution  $\mathcal{D}$  is unknown, if the expected value and variance of  $\Delta h$  are known, then one can verify that  $h$  is a probabilistic CBF. The following result is an analogue of Corollary 3 for the Cantelli-based condition.

**Corollary 5.** (Use of convexity properties for Cantelli-based condition): Let  $\tilde{\Delta}h$  and  $\hat{\Delta}h$  be defined as in Corollary 3 and assume that one of the following conditions holds:

- $h$  is convex and for all  $\mathbf{x} \in \mathcal{C}$  there exists  $\mathbf{u} \in \mathbb{R}^m$  such that

$$\tilde{\Delta}h(\mathbf{x}, \mathbf{u}) \geq \sqrt{\text{Var}(\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d})) \frac{1-\delta}{\delta}}; \quad (11)$$

- $h$  is concave,  $\lambda > 0$  satisfies  $\sup_{\mathbf{x} \in \mathbb{R}^n} \|\nabla^2 h(\mathbf{x})\| \leq \lambda$  and for all  $\mathbf{x} \in \mathcal{C}$ , there exists  $\mathbf{u} \in \mathbb{R}^m$  such that

$$\hat{\Delta}h(\mathbf{x}, \mathbf{u}) \geq \sqrt{\text{Var}(\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d})) \frac{1-\delta}{\delta}}. \quad (12)$$

Then,  $h$  is a  $\delta$ -probabilistic CBF.

The proof of Corollary 5 follows an argument analogous to that of Corollary 3.

**Remark 2.** (Convexity of the associated constraints for known mean and variance): Comments similar to those made in Remark 1 follow regarding conditions (10), (11), (12). However, in this case the presence of the term  $\text{Var}(\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d}))$  can further complicate the convexity guarantees for (11), (12). Nonetheless, in the case where  $\mathbf{d}$  is additive and  $h$  is affine,  $\text{Var}(\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d})) = \text{Var}(h(\mathbf{d}))$ , and (11), (12) have the same dependence in  $\mathbf{u}$  as the one discussed in Remark 1. Additionally, in the case where  $\Delta h$  is affine in  $\mathbf{u}$  (10) is a second order cone (SOC) constraint, which is convex and can be easily used in a second-order cone program (SOCP) as in [14], [34]. •

The results in Propositions 2 and 4, leverage knowledge of the first and second moments of the distribution of  $\Delta h$  to provide bounds on the probability of its tails. Although it is possible to derive similar (possibly tighter) results when knowledge of higher moments is available (cf. [35]–[37] for concentration inequalities using higher moments), this usually leads to conditions with complicated dependencies on the control input  $\mathbf{u}$  and that are computationally not amenable for optimization-based control.

#### IV. VERIFICATION OF PROBABILISTIC CBFs USING DATA

Section III established a set of results that show how to verify that a function is a probabilistic CBF by leveraging moments of the uncertainty distribution. In this section, we take an alternative approach that relies on the availability of samples of the uncertainty distribution, rather than knowledge of its moments.

However, since the samples themselves are random, in general it is not possible to exactly certify that  $h$  is a probabilistic CBF. Instead, one can only do that with a given confidence. The following definition formalizes this idea.

**Definition 3.** (Probabilistic CBF with a given confidence): Let  $\delta \in (0, 1)$  and  $\beta \in (0, 1)$ . Let  $\mathbf{D}_N = \{\mathbf{d}^{(i)}\}_{i=1}^N$  be  $N \in \mathbb{N}$  independent identically distributed samples with  $\mathbf{d}^{(i)} \sim \mathcal{D}$ . The function  $h : \mathbb{R}^n \rightarrow \mathbb{R}$  is a  $\delta$ -probabilistic CBF with confidence  $1 - \beta$  if for each  $\mathbf{x} \in \mathcal{C}$ , there exists  $\mathbf{u} \in \mathbb{R}^m$  (dependent on the dataset  $\mathbf{D}_N$ ) such that

$$\mathbb{P}_N(\mathbb{P}(\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d}) \geq 0) \geq 1 - \delta) \geq 1 - \beta, \quad (13)$$

where  $\mathbb{P}_N$  denotes the probability with respect to  $\mathbf{D}_N$ .

The following result shows that the notion of CBF in Definition 3 can be used to verify that safety holds over a finite time horizon with a given probability over the stochasticity of (1) and a given confidence over  $\mathbf{D}_N$ .

**Proposition 6.** (Safety over a finite time horizon for probabilistic CBFs with given confidence): Let  $\epsilon \in (0, 1)$  and  $\tilde{\beta} \in (0, 1)$ . Let  $\mathbf{D}_N = \{\mathbf{d}^{(i)}\}_{i=1}^N$  be  $N$  independent identically distributed samples with  $\mathbf{d}^{(i)} \sim \mathcal{D}$ . Let  $\delta \in (0, 1)$ ,  $\beta \in (0, 1)$  be such that (4) and  $\beta \leq \frac{\tilde{\beta}}{H}$  hold. Further assume that  $h$  is a  $\delta$ -probabilistic CBF with confidence  $1 - \beta$ . Let  $\mathbf{k} : \mathbb{R}^n \rightarrow \mathbb{R}^m$  be such that  $\mathbf{u}_t = \mathbf{k}(\mathbf{x}_t)$  satisfies (13) for each  $\mathbf{x}_t \in \mathcal{C}$ . Then,  $\mathbf{k}$  is  $\epsilon$ -safe over a horizon  $H$  with confidence  $1 - \tilde{\beta}$ , i.e., for any  $\mathbf{x}_0 \in \mathcal{C}$ , the iterates of (1) under  $\mathbf{k}$  satisfy

$$\mathbb{P}_N\left(\mathbb{P}\left(\bigcap_{t=0}^H \{\mathbf{x}_t \in \mathcal{C}\}\right) \geq 1 - \epsilon\right) \geq 1 - \tilde{\beta}. \quad (14)$$

*Proof.* As shown in the proof of Proposition 1 and using the same notation, for any dataset  $\mathbf{D}_N$ , we have

$$\begin{aligned} \mathbb{P}\left(\bigcap_{t=0}^H \{\mathbf{x}_t \in \mathcal{C}\}\right) &\geq \\ \geq \prod_{t=1}^H \mathbb{P}\left(\Delta h(\mathbf{x}_{t-1}, \mathbf{k}(\mathbf{x}_{t-1}), \mathbf{d}_{t-1}) \geq 0 | \Gamma_{t-1}\right). \end{aligned}$$

Hence, if we let  $A_N$  be the event

$$\prod_{t=1}^H \mathbb{P}\left(\Delta h(\mathbf{x}_{t-1}, \mathbf{k}(\mathbf{x}_{t-1}), \mathbf{d}_{t-1}) \geq 0 | \Gamma_{t-1}\right) \geq 1 - \epsilon,$$

(defined by the randomness of the dataset  $\mathbf{D}_N$ ) we have

$$\mathbb{P}_N\left(\mathbb{P}\left(\bigcap_{t=0}^H \{\mathbf{x}_t \in \mathcal{C}\}\right) \geq 1 - \epsilon\right) \geq \mathbb{P}_N(A_N). \quad (15)$$

Now note that if the dataset  $\mathbf{D}_N$  is such that the event  $B_{t,N}$ , defined as

$$\mathbb{P}\left(\Delta h(\mathbf{x}_{t-1}, \mathbf{k}(\mathbf{x}_{t-1}), \mathbf{d}_{t-1}) \geq 0 | \Gamma_{t-1}\right) \geq (1 - \epsilon)^{\frac{1}{H}},$$

holds for all  $t \in [H]$ , then  $A_N$  also holds. Therefore,

$$\mathbb{P}_N(A_N) \geq \mathbb{P}_N\left(\bigcap_{t=1}^H B_{t,N}\right) \geq \sum_{t=1}^H \mathbb{P}_N(B_{t,N}) - (H-1), \quad (16)$$

where the last inequality follows from Fréchet's inequality [38]. Now, since  $1 - \delta \geq (1 - \epsilon)^{\frac{1}{H}}$ , if the event  $C_{t,N}$ , defined as

$$\mathbb{P}\left(\Delta h(\mathbf{x}_{t-1}, \mathbf{k}(\mathbf{x}_{t-1}), \mathbf{d}_{t-1}) \geq 0 | \Gamma_{t-1}\right) \geq 1 - \delta,$$

holds, then  $B_{t,N}$  also holds. Therefore,  $\mathbb{P}_N(B_{t,N}) \geq \mathbb{P}_N(C_{t,N})$ . Now, since  $h$  is a  $\delta$ -probabilistic CBF with confidence  $\beta$ , we have  $\mathbb{P}_N(C_{t,N}) \geq 1 - \beta$  for all  $t \in [H]$ , and (15), (16) we have

$$\mathbb{P}_N\left(\mathbb{P}\left(\bigcap_{t=0}^H \{\mathbf{x}_t \in \mathcal{C}\}\right) \geq 1 - \epsilon\right) \geq 1 - H\beta.$$

Now the result follows from the fact that  $\beta \leq \frac{\tilde{\beta}}{H}$ .  $\square$

In the rest of the section, we fix  $\delta, \beta \in (0, 1)$ , and consider a dataset  $\mathbf{D}_N$  of  $N \in \mathbb{N}$  samples as defined in Proposition 6. The following result establishes a sufficient condition for  $h$  to be a probabilistic CBF with a given confidence in the case where the random variable  $\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d})$  is bounded.

**Proposition 7.** (Hoeffding-based condition): *Suppose that there exist constants  $a, b, \epsilon > 0$  with  $\beta \geq 2 \exp\{-\frac{2N\epsilon^2}{(b-a)^2}\}$  and such that for each  $\mathbf{x} \in \mathcal{C}$  and  $\mathbf{u} \in \mathbb{R}^m$ ,  $a \leq \Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d}) \leq b$  almost surely and*

$$\frac{1}{N} \sum_{i=1}^N \Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d}^{(i)}) \geq \epsilon + b(1 - \delta), \quad (17)$$

*Then,  $h$  is a  $\delta$ -probabilistic CBF with confidence  $1 - \beta$ .*

*Proof.* Let  $\mathbf{x} \in \mathcal{C}$  and pick  $\mathbf{u} \in \mathbb{R}^m$  so that (17) holds. Note that such  $\mathbf{u}$  is dependent on the dataset  $\mathbf{D}_N$ . Now, we have

$$\begin{aligned} \mathbb{P}_N\left(\left|\mathbb{E}[\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d})] - \frac{1}{N} \sum_{i=1}^N \Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d}^{(i)})\right| \leq \epsilon\right) \\ \geq 1 - 2 \exp\left\{-\frac{2N\epsilon^2}{(b-a)^2}\right\} \geq 1 - \beta, \end{aligned}$$

where in the first inequality we have used Hoeffding's inequality [39], and in the last inequality we have used  $\beta \geq 2 \exp\{-\frac{2N\epsilon^2}{(b-a)^2}\}$ . This, together with (17), implies that

$$\mathbb{P}_N\left(\mathbb{E}[\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d})] \geq b(1 - \delta)\right) \geq 1 - \beta.$$

Now, the result follows from Proposition 2.  $\square$

Since (17) is derived through Hoeffding's inequality, we refer to it as the **Hoeffding-based condition**.

**Remark 3.** (Conditions of Proposition 7): *A simple scenario in which  $a, b$  as in Proposition 7 exist is if  $h$  is uniformly bounded (which we can always assume without loss of generality). Moreover,  $\beta \geq 2 \exp\{-\frac{2N\epsilon^2}{(b-a)^2}\}$  holds if  $N$  is sufficiently large. We note that we can also state a version of Proposition 7 in which the  $a, b, \epsilon, N$  are dependent on  $\mathbf{x}$  and  $\mathbf{u}$  (instead of being uniform for all  $\mathbf{x} \in \mathcal{C}$  and  $\mathbf{u} \in \mathbb{R}^m$ ). Although the assumption that  $a, b, \epsilon, N$  are uniform across  $\mathbf{x}$  and  $\mathbf{u}$  might add some conservatism, it leads to a simpler dependency on  $\mathbf{u}$  in (17). In fact, condition (17) is convex in  $\mathbf{u}$  under assumptions such as the ones in Remark 1.*  $\bullet$

Next we introduce another result that certifies that  $h$  is a probabilistic CBF with a given confidence. In this case, the result is distribution-free (meaning it holds without any assumptions on the uncertainty distribution) and uses the theory of the scenario approach [23].

**Proposition 8.** (Scenario approach-based condition): *Suppose that for each  $\mathbf{x} \in \mathcal{C}$ , there exists  $\mathbf{u} \in \mathbb{R}^m$  such that*

$$\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d}^{(i)}) \geq 0, \quad i \in [N]. \quad (18)$$

*Further suppose that for each  $\mathbf{x} \in \mathcal{C}$  and  $\mathbf{d} \in \mathbb{R}^d$ , the function  $\mathbf{u} \rightarrow -\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d})$  is convex and  $N$  satisfies*

$$\sum_{i=0}^{d-1} \binom{N}{i} \delta^i (1 - \delta)^{N-i} \leq \beta. \quad (19)$$

*Then,  $h$  is a  $\delta$ -probabilistic CBF with confidence  $1 - \beta$ .*

*Proof.* Let  $\mathbf{c} \in \mathbb{R}^m$  and take  $\mathbf{u}_x \in \mathbb{R}^m$  as a minimizer of the optimization problem

$$\min_{\mathbf{u} \in \mathbb{R}^m} \mathbf{c}^\top \mathbf{u}, \quad \text{s.t. } \Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d}^{(i)}) \geq 0, \quad i \in [N], \quad (20)$$

which is feasible by assumption. Now, by [40, Theorem 2.4] we have that for each  $\mathbf{x} \in \mathcal{C}$ ,

$$\mathbb{P}_N\left(\mathbb{P}(\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d}) \leq 0) > \delta\right) \leq \beta.$$

Equivalently,

$$\mathbb{P}_N\left(\mathbb{P}(\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d}) \leq 0) \leq \delta\right) \geq 1 - \beta. \quad (21)$$

Now, note that for any  $\mathbf{D}_N$ , if the event  $X_N$  defined as  $\mathbb{P}(\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d}) \leq 0) \leq \delta$  holds, then the event  $Y_N$ , defined as  $\mathbb{P}(\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d}) \geq 0) \geq 1 - \delta$  also holds. Hence,  $\mathbb{P}_N(Y_N) \geq \mathbb{P}_N(X_N)$ , and the result follows from (21).  $\square$

Since condition (18) is derived through the scenario approach, we refer to it as the **scenario-based condition**. The following remark discusses the assumptions in Proposition 8.

**Remark 4.** (Assumptions of Proposition 8): *As mentioned in [23, Theorem 1], a sufficient condition for (19) to hold is  $N \geq \frac{2}{\delta} \left( \ln \frac{1}{\beta} + d \right)$ . Also, although Proposition 8 requires  $-\Delta_\alpha$  to be convex in  $\mathbf{u}$ , the recent paper [41] shows that a similar result can be obtained in the non-convex setting.*  $\bullet$

Another method to establish a distribution-free guarantee that  $h$  is a probabilistic CBF with a given confidence is through the theory of conformal prediction [42]. This is leveraged in the following result.

**Proposition 9.** (Conformal prediction-based condition): Suppose that for each  $\mathbf{x} \in \mathcal{C}$ , there exists  $\mathbf{u}$  such that

$$Q_N = \text{Quantile}_{1-\delta+\sqrt{\frac{\log(1/\beta)}{2N}}}(R^{(1)}, \dots, R^{(N)}, \infty) \leq 0, \quad (22)$$

with  $R^{(i)} = -\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d}^{(i)})$ . Then,  $h$  is a  $\delta$ -probabilistic CBF with confidence  $1 - \beta$ .

*Proof.* Let  $\mathbf{x} \in \mathcal{C}$  and pick  $\mathbf{u} \in \mathbb{R}^m$  so that (22) holds. Note that such  $\mathbf{u}$  is dependent on the dataset  $\mathbf{D}_N$ . By [42, Proposition 2a], we have

$$\mathbb{P}_N(\mathbb{P}(-\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d}) \leq Q_N) \geq 1 - \delta) \geq 1 - \beta. \quad (23)$$

Note that if  $\mathbf{d} \in \mathbb{R}^d$  satisfies  $-\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d}) \leq Q_N$ , then we have  $-\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d}) \leq 0$  by (22). Hence, for any fixed dataset  $\mathbf{D}_N$  and  $\mathbf{u}$  satisfying (22), we have

$$\mathbb{P}(-\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d}) \leq Q_N) \leq \mathbb{P}(-\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d}) \leq 0).$$

Therefore, if the event  $\bar{X}_N$  defined as  $\mathbb{P}(-\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d}) \leq Q_N) \geq 1 - \delta$  holds, then the event  $\bar{Y}_N$  defined as  $\mathbb{P}(-\Delta h(\mathbf{x}, \mathbf{u}, \mathbf{d}) \leq 0) \geq 1 - \delta$  also holds. Hence,  $\mathbb{P}_N(\bar{Y}_N) \geq \mathbb{P}_N(\bar{X}_N)$  and by (23) this implies that  $h$  is a  $\delta$ -probabilistic CBF with confidence  $1 - \beta$ .  $\square$

Since condition (22) is derived through conformal prediction, we refer to it as the **conformal prediction-based condition**. We note that as shown in [43, Section 5], the quantile constraint (22) can be reformulated using mixed-integer programming so that it can be used in optimization-based control design with off-the-shelf solvers.

## V. EXPERIMENTAL VALIDATION

In this section we validate our theoretical results by considering the problem of controlling a quadrupedal robot traversing a narrow corridor, inspired by [27, Section 5.D]. We consider the following discrete-time dynamics

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \Delta t \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} v_k^x \\ v_k^y \\ \omega_k \end{pmatrix} + \mathbf{d}_k, \quad (24)$$

where  $\mathbf{x}_k = [x, y, \theta]^\top$  and  $\Delta t > 0$ . We let  $h(\mathbf{x}) = 0.5^2 - y^2$ .

### A. Simulation

In simulation, we model the uncertainties in the terrain through a Gaussian disturbance with standard deviation  $\sigma$  acting only in the  $y$  direction, i.e.,  $\mathbf{d}_k = [0, d_k, 0]$ , with  $d_k \sim \mathcal{N}(0, \sigma)$ . We consider a time horizon of 20 steps and take  $\delta = 0.1$ ,  $\beta = 0.01$ ,  $\alpha = 0.01$  and  $\Delta t = 0.1$ . We consider a nominal controller  $\mathbf{k}_{\text{nom}}(\mathbf{x}) = [0.2, 0, -\theta]$ . For each of the different conditions —Markov (in the form of (8)), Cantelli, Hoeffding, scenario approach, and conformal prediction—, we compute the control input at every state by solving an optimization problem that finds the closest input to the nominal controller that satisfies the corresponding condition. We take generate datasets of size

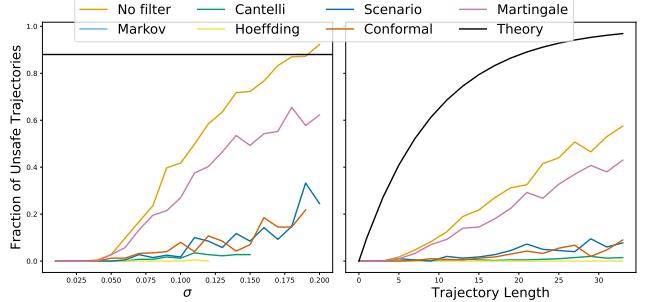


Fig. 1: (left) Number of unsafe trajectories for different values of the disturbance variance  $\sigma$ . (right) Number of unsafe trajectories for different values of the trajectory length. No filter refers to the implementation of the nominal controller, whereas Markov, Cantelli, Hoeffding, Scenario, and Conformal refer to the approaches in Propositions 2, 7, 8, 9, respectively.

3252, 113, and 300 for the Hoeffding, scenario, and conformal approaches, respectively. For the Markov, Hoeffding and scenario approach-based condition, the optimization problem at every state is a convex quadratically constrained quadratic program (QCQP), which we solve using the `cvxpy` library in Python. The Cantelli-based condition leads to a non-convex problem, which we solve using the trust region solver of the `minimize` function in the `SciPy` library in Python. Using the reformulation in [43, Section 5], the conformal prediction-based method leads to a mixed-integer quadratically constrained quadratic program, which we solve using the `gurobipy` library in Python. We implement each of the approaches (along with the nominal controller and the approach based on [27], which we refer to as the Martingale approach) in 400 different trajectories with initial condition  $\mathbf{x}_0 = [0, 0, 0]$ .

Figure 1 shows the number of unsafe trajectories for different values of  $\sigma$  (and fixed horizon 20)) and different time horizon values (and fixed  $\sigma$ ). For a fixed time horizon of 20 Proposition 1 ensures that the probability that the trajectory is safe over 20 steps is at least  $(1-\delta)^{20} \approx 0.12$ , and hence the probability that the trajectory is unsafe over these 20 steps is at most 0.88. (for the approaches in Section IV, this is with confidence  $1 - \beta \cdot 20 = 0.8$ ). As can be seen in Figure 1, all the approaches considered in this paper lead to a fraction of unsafe trajectories significantly lower than the theoretical bound, which suggests that the conditions might be conservative. The source of this conservatism can be attributed to the limited knowledge of the uncertainty distribution, the use of Jensen's inequality (cf. Corollaries 3 and 5), and the fact that these results do not consider full trajectory information, and instead derive probabilistic bounds only based on a condition at every state.

Figure 2 shows 400 different trajectories generated with the different approaches for  $\sigma = 0.06$  and trajectory length 20. Note that we do not plot trajectories from the Markov approach because it is infeasible for such value of  $\sigma$ .

Finally, Table I shows the average execution times for the

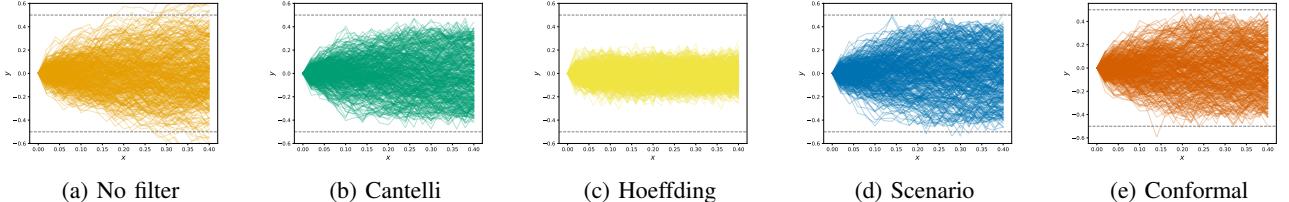


Fig. 2: Evolution of 400 different trajectories with initial condition at  $\mathbf{x}_0 = [0, 0, 0]$  and  $\sigma = 0.06$  and trajectory length 20 for different methods introduced in the paper, along with the nominal controller (left).

	Markov	Cantelli	Hoeffding	Scenario	Conformal	Martingale
$T_{ex}$	3.47	3.78	3.89	79.96	5.65	3.46
$\sigma_0$	0.03	0.16	0.13	0.21	0.2	0.25

TABLE I: Average time it takes to compute the controller ( $T_{ex}$ ) in milliseconds and smallest value of  $\sigma$  that leads to infeasibility  $\sigma_0$  for the different approaches.

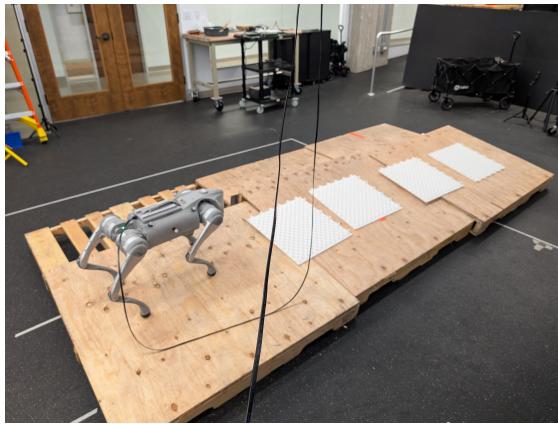


Fig. 3: Experimental Setup

different approaches as well as the smallest value of  $\sigma$  that leads to infeasibility of the corresponding condition.

The choice of which of the different approaches should be used depends on a variety of factors. Firstly, on what information is available about the uncertainty distribution. Secondly, on the desired level of conservativeness with respect to the safety constraints. Figure 1 and Table I show that there is a tradeoff between the empirical number of safety violations and the smallest variance that leads to infeasibility. Approaches yielding less safety violations generally lead to infeasible conditions for smaller values of  $\sigma$ . Thirdly, another point to consider is the computational time (cf. Table I). Although the martingale approach shows a larger value of  $\sigma_0$  in Table I, it fails to meet the prescribed safety guarantees that the other approaches successfully satisfy.

### B. Hardware

To validate the effectiveness of our approach beyond simulation, we deploy it on a quadruped robot navigating a challenging obstacle course designed to induce substantial disturbances. The course is 3m long and 1m wide, with a laterally slanted wooden ramp covered in slippery plates (Fig. 3). Our system is a Unitree GO2 quadruped, which

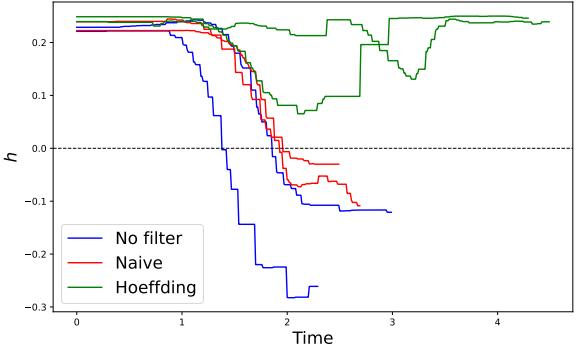


Fig. 4: Evolution of the value of  $h$  for two different trajectories for each of the different approaches considered in the hardware experiment.

we assume follows dynamics 24. The control objective is to traverse the course with a nominal forward velocity of 1 m/s<sup>1</sup>. We compare three controllers: (i) the nominal controller with no safety filter, (ii) a naive CBF-based filter, and (iii) our Hoeffding-based probabilistic CBF filter. For the latter, we collect disturbance data from 5000 rollout steps on the course to characterize the uncertainty. We use the same  $h$  as in simulation, and estimate the global state of the robot by an external OptiTrack system. Results are shown in Fig. 4. We plot the trajectories induced by the nominal controller, the naive filter (i.e., the controller that is computed at each state as the smallest deviation from the nominal satisfying the discrete-time CBF condition assuming the disturbance is identically zero), and the Hoeffding-based approach. The nominal controller and the naive filter consistently violate the safety constraint as the sloped terrain pushes the robot laterally out of the safe set. In contrast, our Hoeffding-based filter substantially reduces the frequency and magnitude of constraint violations, correctly accounting for the stochastic disturbance and succeeding on 4/5 of the tested rollouts.

### VI. CONCLUSION

We have introduced probabilistic CBFs, which can be used to design controllers that satisfy safety constraints with a prescribed probability over a finite time horizon. We have proposed a number of different sufficient conditions to verify that a function is a probabilistic CBF. The first set of conditions leverage knowledge of the moments

<sup>1</sup>Video of the experiment: <https://youtu.be/Lu9tl-teSLc>

of the uncertainty distribution, whereas the second set of conditions utilize realizations of the uncertainty. We have studied the computational tractability of such conditions for its use in optimization-based control. In particular, we have implemented such controllers on a quadruped, both in simulation and hardware. In future work, we plan to find less conservative conditions to ensure that a function is a probabilistic CBF and bridge the gap between the theoretical probabilistic safety guarantees and the empirical results. The results in this paper also open the door to the design of safe controllers in systems with state estimation errors induced by techniques such as the Kalman Filter or SLAM.

#### ACKNOWLEDGMENTS

The authors thank Ryan M. Bena for his help with the quadruped hardware experiment.

#### REFERENCES

- [1] A. D. Ames, S. Coogan, M. Egerstedt, G. Notomista, K. Sreenath, and P. Tabuada, "Control barrier functions: theory and applications," in *European Control Conference*, Naples, Italy, 2019, pp. 3420–3431.
- [2] S. Bansal, M. Chen, S. Herbert, and C. J. Tomlin, "Hamilton-Jacobi Reachability: A Brief Overview and Recent Advances," in *IEEE Conf. on Decision and Control*, Melbourne, Australia, Dec. 2017, pp. 2242–2253.
- [3] J. B. Rawlings, D. Q. Mayne, and M. M. Diehl, *Model Predictive Control: Theory, Computation, and Design*. Nob Hill Publishing, 2017.
- [4] S. Kolathaya and A. D. Ames, "Input-to-state safety with control barrier functions," *IEEE Control Systems Letters*, vol. 3, no. 1, pp. 108–113, 2019.
- [5] A. Anil, A. J. Taylor, C. R. He, G. Orosz, and A. D. Ames, "Safe controller synthesis with tunable input-to-state safe control barrier functions," *IEEE Control Systems Letters*, vol. 6, pp. 908–913, 2022.
- [6] M. Jankovic, "Robust control barrier functions for constrained stabilization of nonlinear systems," *Automatica*, vol. 96, pp. 359–367, 2018.
- [7] A. Anil, T. G. Molnar, A. D. Ames, and G. Orosz, "Generalizing robust control barrier functions from a controller design perspective," *IEEE Open Journal of Control Systems*, vol. 4, pp. 54–69, 2025.
- [8] W. Wand and X. Xu, "Disturbance observer-based robust control barrier functions," in *American Control Conference*, 2023, pp. 3681–3687.
- [9] A. Clark, "Control barrier functions for stochastic systems," *Automatica*, vol. 130, p. 109688, 2021.
- [10] O. So and C. Fan, "Comment(s) on "control barrier functions for stochastic systems","" *Automatica*, vol. 180, p. 112185, 2025.
- [11] S. Prajna, A. Jadbabaie, and G. J. Pappas, "Stochastic safety verification using barrier certificates," in *IEEE Conf. on Decision and Control*, Paradise Island, Bahamas, Dec. 2004, pp. 929–934.
- [12] C. Santoyo, M. Dutreix, and S. Coogan, "Verification and control for finite-time safety of stochastic systems via barrier functions," in *2019 IEEE Conference on Control Technology and Applications (CCTA)*, 2019, pp. 712–717.
- [13] J. Steinhardt and R. Tedrake, "Finite-time regional verification of stochastic non-linear systems," *The International Journal of Robotics Research*, vol. 31, no. 7, pp. 901–923, 2012.
- [14] F. Castañeda, J. J. Choi, B. Zhang, C. J. Tomlin, and K. Sreenath, "Pointwise feasibility of Gaussian process-based safety-critical control under model uncertainty," in *IEEE Conf. on Decision and Control*, Austin, Texas, USA, 2021, pp. 6762–6769.
- [15] K. Long, V. Dhiman, M. Leok, J. Cortés, and N. Atanasov, "Safe control synthesis with uncertain dynamics and constraints," *IEEE Robotics and Automation Letters*, vol. 7, no. 3, pp. 7295–7302, 2022.
- [16] K. Long, Y. Yi, J. Cortés, and N. Atanasov, "Safe and stable control synthesis for uncertain system models via distributionally robust optimization," in *American Control Conference*, San Diego, California, Jun. 2023, pp. 4651–4658.
- [17] P. Mestres, K. Long, N. Atanasov, and J. Cortés, "Feasibility analysis and regularity characterization of distributionally robust safe stabilizing controllers," *IEEE Control Systems Letters*, vol. 8, pp. 91–96, 2024.
- [18] R. K. Cosner, P. Culbertson, and A. D. Ames, "Bounding Stochastic Safety: Leveraging Freedman's Inequality with Discrete-Time Control Barrier Functions," *IEE Control Systems Letters*, vol. 8, pp. 1937–1942, 2024.
- [19] M. Ahmadi, X. Xiong, and A. D. Ames, "Risk-Averse Control via CVaR Barrier Functions: Application to Bipedal Robot Locomotion," *IEEE Control Systems Letters*, vol. 6, pp. 878–883, 2022.
- [20] M. Black, G. Fainekos, B. Hoxha, D. Prokhorov, and D. Panagou, "Safety under uncertainty: tight bounds with risk-aware control barrier functions," in *IEEE International Conference on Robotics and Automation (ICRA)*, London, UK, 2023, pp. 12 686–12 692.
- [21] M. Vahs, C. Pek, and J. Tumova, "Belief control barrier functions for risk-aware control," *IEEE Robotics and Automation Letters*, vol. 8, no. 12, pp. 8565–8572, 2023.
- [22] A. A. D. Nascimento, A. Papachristodoulou, and K. Margellos, "Probabilistically safe controllers based on control barrier functions and scenario model predictive control," in *2024 IEEE 63rd Conference on Decision and Control (CDC)*, 2024, pp. 1814–1819.
- [23] M. C. Campi, S. Garatti, and M. Prandini, "The scenario approach for systems and control design," *Annual Reviews in Control*, vol. 32, no. 2, pp. 149–157, 2009.
- [24] A. Z. S. Pfrommer, T. Gautam and S. Sojoudi, "Safe reinforcement learning with chance-constrained model predictive control," in *4th Annual Learning for Dynamics and Control Conference, Proceedings of Machine Learning Research*, vol. 168, 2022, pp. 291–303.
- [25] K. P. Wabersich, L. Hewing, A. Carron, and M. N. Zeilinger, "Probabilistic model predictive safety certification for learning-based control," *IEEE Transactions on Automatic Control*, vol. 67, no. 1, pp. 176–188, 2022.
- [26] M. Farina, L. Giulioni, and R. Scattolini, "Stochastic linear Model Predictive Control with chance constraints - A review," *Journal of Process Control*, vol. 44, pp. 53–67, 2016.
- [27] R. K. Cosner, P. Culbertson, A. J. Taylor, and A. D. Ames, "Robust safety under stochastic uncertainty with discrete-time control barrier functions," in *Robotics: Science and Systems*, 2023.
- [28] A. Agrawal and K. Sreenath, "Discrete control barrier functions for safety-critical control of discrete systems with application to bipedal robot navigation," in *Robotics: Science and Systems*, 2017.
- [29] S. M. Ross, *A first course in probability, eighth edition*. Macmillan, 2010.
- [30] A. Klenke, *Probability Theory*. New York: Springer, 2013.
- [31] R. A. Becker, "The variance drain and Jensen's inequality," 2012.
- [32] J. Park and S. Boyd, "General heuristics for nonconvex quadratically constrained quadratic programming," *arXiv preprint arXiv:1703.07870v2*, 2017.
- [33] F. P. Cantelli, "Sui confini della probabilità," in *Atti del Congresso Internazionale dei Matematici 6*, 1928, pp. 47–59.
- [34] P. Mestres and J. Cortés, "Feasibility and regularity analysis of safe stabilizing controllers under uncertainty," *Automatica*, vol. 167, p. 111800, 2024.
- [35] K. Isii, "Inequalities of the types of Chebyshev and Cramér-Rao and mathematical programming," *Annals of the Institute of Statistical Mathematics*, vol. 16, pp. 277–293, 1964.
- [36] L. Vandenberghe, S. Boyd, and K. Comanor, "Generalized Chebyshev Bounds via Semidefinite Programming," *SIAM Review*, vol. 49, no. 1, pp. 52–64, 2007.
- [37] B. B. Bhattacharyya, "One sided Chebyshev inequality when the first four moments are known," *Communications in Statistics - Theory and Methods*, vol. 16, no. 9, pp. 2789–2791, 1987.
- [38] M. Fréchet, "Généralisations du théorème des probabilités totales," *Fundamenta Mathematicae*, vol. 1, no. 25, pp. 379–387, 1935.
- [39] W. Hoeffding, "Probability inequalities for sums of bounded random variables," *Journal of the American Statistical Association*, vol. 58, no. 301, pp. 13–30, 1963.
- [40] M. C. Campi and S. Garatti, "The exact feasibility of randomized solutions of uncertain convex programs," *SIAM Journal on Optimization*, vol. 19, no. 2, pp. 1211–1230, 2008.
- [41] S. Garatti and M. C. Campi, "Non-convex scenario optimization," *Mathematical Programming, Series A*, vol. 209, pp. 557–608, 2024.
- [42] V. Vovk, "Conditional validity of inductive conformal predictors," in *Asian Conference on Machine Learning*, 2012, pp. 475–490.
- [43] Y. Zhao, Z. Yu, M. Sesia, J. V. Deshmukh, and L. Lindemann, "Conformal predictive programming for chance constrained optimization," *arXiv preprint arXiv:2402.07407*, 2024.