

Safe Probabilistic Planning for Human-Robot Interaction using Conformal Risk Control

Jake Gonzales¹

Kazuki Mizuta²

Karen Leung^{2*}

Lillian J. Ratliff^{1*}

Abstract— In this paper, we present a novel probabilistic safe control framework for human-robot interaction that combines control barrier functions (CBFs) with conformal risk control to provide formal safety guarantees while considering complex human behavior. The approach uses conformal risk control to quantify and control the prediction errors in CBF safety values and establishes formal guarantees on the probability of constraint satisfaction during interaction. We introduce an algorithm that dynamically adjusts the safety margins produced by conformal risk control based on the current interaction context. Through experiments on human-robot navigation scenarios, we demonstrate that our approach significantly reduces collision rates and safety violations as compared to baseline methods while maintaining high success rates in goal-reaching tasks and efficient control. The code, simulations, and other supplementary material can be found on the project website: <https://jakeagonzales.github.io/crc-cbf-website/>.

I. INTRODUCTION

The safe deployment of autonomous robots in human environments, from autonomous driving to service robots, presents fundamental challenges in ensuring safety under the unpredictability of human behavior. In particular, the uncertainty in human behavior is *multimodal*, where there could be multiple possible distinct behaviors (e.g., passing to the left or right of someone), and *history dependent*, where future behaviors depend on interaction history. Common approaches to safety-critical control in the presence of uncertainty make simplifying distributional assumptions (e.g. Gaussian) about the uncertainty and leverage chance constraints to provide probabilistic safety guarantees [1]. While sampling-based methods can handle more complex distributions, they lack formal safety guarantees and can be computationally expensive for real-time applications [2].

A core challenge is in developing robot planning algorithms that can handle the complex uncertainty associated with human behavior without being overly conservative and providing quantifiable confidence in constraint satisfaction. Recent theoretical advancements in statistical verification have provided powerful tools for establishing formal guarantees on complex predictive models, including deep neural networks. In particular, *conformal prediction* [3] offers a distribution-free framework for constructing prediction sets with guaranteed finite-sample coverage properties, and has been extended to the adaptive setting to maintain coverage

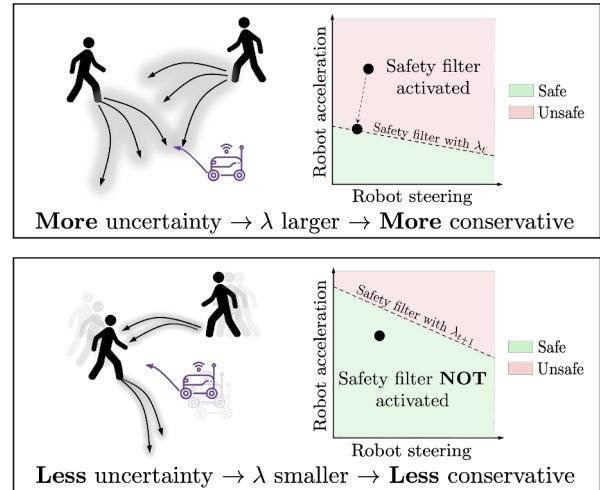


Fig. 1: We develop a risk-aware adaptive safety filter that dynamically adjusts robot conservativeness based on the uncertainty of human behavior. The plots show the robot's control space with red regions indicating unsafe actions. In high-risk scenarios (top), our approach increases the **safety margin parameter λ** , creating more restrictive control constraints. In low-risk scenarios (bottom), λ decreases, allowing less conservative behavior while maintaining safety guarantees.

guarantees under arbitrary distribution shifts using adaptive conformal prediction [4]. This extension is important for robot decision-making in dynamic environments where adaptive conformal prediction allows you to recalibrate prediction sets online, without requiring stationarity assumptions. Conformal Risk Control (CRC) [5] further generalizes this approach by directly controlling arbitrary prediction risks through guarantees on the expected value of user-defined loss functions, making it particularly suitable for safety-critical applications where precise risk quantification is essential.

In this work, we present a principled framework that combines conformal risk control with control-theoretic safety constraints to achieve provable safety guarantees in human-robot interactions. Specifically, we apply CRC to algorithmically tune a *safety margin* for a safety constraint, expressed using Control Barrier Functions (CBFs). We devise a strategy to tune the safety margin to meet a user-desired probabilistic safety specification. We apply this technique to learn a model using offline data to estimate the safety margin conditioned on a given context, thereby resulting in robots that dynamically adjust their conservativeness based on uncertainty. Figure 1 illustrates this concept. The practical implementation of our approach does not rely on any distributional assumptions or worst-case bounds, and is both computationally efficient for real-time control and theoretically grounded.

¹University of Washington, Department of Electrical and Computer Engineering.
²University of Washington, Department of Aeronautics and Astronautics.
 *Equal advising. JG and KM are supported by the UW + Amazon Science Hub Fellowship. KL was supported by a National Science Foundation award under Grant No. 2430686.

Email: {jakegonz, mizuta, ratliff, kymleung}@uw.edu

Statement of contributions. The contributions are multi-fold: **(i)** a novel safe control framework combining CBFs with CRC to provide high-probability safety guarantees for human-robot interactions; **(ii)** rigorous theoretical analysis establishing the connection between CRC-based uncertainty quantification and probabilistic safety guarantees; **(iii)** an assumption-light algorithm for dynamic uncertainty adaptation of safety margins; and **(iv)** a practically implementable and empirically validated approach to handle temporal dependencies in human-robot interactions. To our knowledge, this is the first use of conformal risk control to quantify uncertainty in control-theoretic safety constraints for joint human-robot systems.

A. Related work

Safe planning with CBFs. By enforcing control invariance, CBFs provide theoretical guarantees for safety-critical control [6]. Yet, the practical implementation of CBFs faces challenges. For example, manual construction of CBFs is increasingly difficult for high-dimensional systems with input constraints. This limitation has motivated learning-based approaches to approximate CBFs [7]–[10].

Although CBFs have shown promise in single-agent settings, their application to multi-agent scenarios—particularly with human agents—remains underdeveloped. The inherent stochasticity of human behavior and unknown policies of opponents in multi-agent settings presents fundamental challenges that classical CBF formulations struggle to adequately address. Towards this end, CBFs have been extended to robust formulations [11]–[13] for *a priori* known bounded disturbances, however this is not practical for real-world deployment with complex and unbounded uncertainties due to human interactions or other autonomous agents.

Safe planning with conformal prediction. To address uncertainties in multi-agent settings with minimal assumptions, there has been a line of work on *conformal prediction* in safety critical control. For instance, ignoring feedback effects, recent work [14] uses conformal prediction in a model predictive control formulation to model opponent uncertainty. In a follow-up work [15], the authors leveraged ideas from adaptive conformal prediction to adapt to distribution shift by quantifying uncertainty from past agent observations without requiring stationarity assumptions on agent motion. More closely aligned with our approach, [16] integrates adaptive conformal prediction with CBF constraints in order to safely adapt to uncertainty in multi-robot systems.

Conformal prediction methods create confidence regions around other agents’ trajectories by calibrating prediction sets for coverage. As such they lack the ability to directly control the *risk* of safety constraint violations. Instead of targeting fixed coverage, CRC aims to control the expected loss (risk) under a user-specified loss function and potentially adaptive or contextual thresholds. Indeed, our approach addresses this limitation by leveraging CRC to explicitly quantify prediction errors in safety values consequent of uncertainty due to other agent behavior, and encoding this directly into the safety constraints. This provides a more

precise and principled approach to managing the trade-off between safety and performance. Parallel work by [17] proposes a similar formulation that uses conformal decision theory [18] to tune the conservativeness of the CBF constraint. A limitation of this approach, however, is that it only provides guarantees on the *average constraint violation* over time rather than probabilistic safety guarantees for each timestep, arguably a crucial need for safety-critical control involving humans. We establish this theoretical connection and empirically evaluate our methods using human behavior models trained on real-world pedestrian data.

II. PRELIMINARIES

Consider a human-robot joint dynamical system where the goal is to design a control policy for the robot such that the joint state $x(t) \in \mathbb{R}^n$ of the dynamical system remains safe with high probability given the stochastic unknown policies of all other agents.

Let R denote the index of the robot agent, and $H = \{1, \dots, H\}$ be the index set of all agents excluding the robot agent. For any time $t \in \mathbb{R}_+$, define the joint state $x(t) := (x_R(t), x_H(t))$ where $x_R(t) \in \mathcal{X}_R \subseteq \mathbb{R}^{n_R}$ and $x_H(t) \in \mathcal{X}_H \subseteq \mathbb{R}^{n_H}$, denote the robot and human states, respectively, and let $\mathcal{X} := \mathcal{X}_R \times \mathcal{X}_H$ define the joint state space. Analogously, define the joint control input $u(t) := (u_R(t), u_H(t))$ where $u_R(t) \in \mathcal{U}_R \subseteq \mathbb{R}^{m_R}$ and $u_H(t) \in \mathbb{R}^{m_H}$ denote the robot and human control inputs, respectively, and let $\mathcal{U} := \mathcal{U}_R \times \mathcal{U}_H$ be the joint control space. Note that we drop the explicit dependence on t where obvious from context.

Assumption 1: The joint control space \mathcal{U} is non-empty, closed, and convex set. Moreover, there exists strictly positive, finite constants b_u and B_u such that $b_u \mathbb{B} \subseteq \mathcal{U} \subseteq B_u \mathbb{B}$ where $\mathbb{B} = \{u \in \mathbb{R}^m \mid \|u\| = 1\}$ is the unit ball.

This assumption implies the set \mathcal{U} is compact with a non-empty interior; otherwise, we can map \mathcal{U} to a lower dimensional space. Let $n := n_R + n_H$ be the state dimension and $m := m_R + m_H$ be the control input dimension. The dynamical system under consideration is given by

$$\dot{x} = f(x) + Bu, \quad (1)$$

where $B = [B_R \ B_H]$. We make the following regularity assumption.

Assumption 2: The map $f : \mathcal{X} \rightarrow \mathbb{R}^n$ is locally Lipschitz continuous in x so that $x(t)$ is Lipschitz continuous. Let $\mathfrak{L}_x \in [0, \infty)$ be the corresponding Lipschitz constant.

A. Safety

As noted the goal is to give a guarantee on the state remaining safe. To gain some intuition, let us consider a deterministic control-affine system:

$$\dot{x} = f(x) + g(x)u \quad (2)$$

where $x \in \mathcal{X} \subset \mathbb{R}^n$ is the state, and $u \in U \subset \mathbb{R}^m$ is the control input. Let $\mathcal{S} \subseteq \mathcal{X}$ denote the set of *safe states* for the system in (2). One way to characterize such a set is via the 0-superlevel set of a mapping:

$$\mathcal{S} := \{x \in \mathcal{X} \mid h(x) \geq 0\},$$

where $h(\cdot)$ is a continuously differentiable function.

In this setting, the objective is to synthesize a feedback control policy $\pi : x \mapsto u$ ensuring forward invariance of \mathcal{S} under closed-loop dynamics:

$$\dot{x} = f(x) + g(x)\pi(x).$$

That is, trajectories initialized in \mathcal{S} at time t_0 must remain within \mathcal{S} for all $t \geq t_0$. This is known as the safe planning problem. One common approach to solving this problem is to use CBFs [6].

Definition 1: The mapping $h(\cdot)$ is a CBF for the system (2) if there exists an extended class- \mathcal{K}_∞ ¹ function \mathcal{K} such that

$$\sup_{u \in \mathcal{U}} \{L_f h(x) + L_g h(x)u + \mathcal{K}(h(x))\} \geq 0, \quad \forall x \in \mathcal{X}, \quad (3)$$

where $L_f h(x) = \nabla h(x)^\top f(x)$ and $L_g h(x) = \nabla h(x)^\top g(x)$. For a given CBF $h(\cdot)$ and class- \mathcal{K}_∞ function \mathcal{K} , the set of admissible safe control actions is defined as

$$C(x) := \{u \in \mathcal{U} \mid L_f h(x) + L_g h(x)u + \mathcal{K}(h(x)) \geq 0\}. \quad (4)$$

Any locally Lipschitz continuous controller $\pi(x) \in C(x)$ guarantees forward invariance of the safe set \mathcal{S} under the closed-loop dynamics. This motivates the design of a minimally-invasive safety filter:

$$\pi(x) := \min_{u \in \mathcal{U}} \|u - u_{\text{nom}}(x)\|_2^2 \quad (\text{CBF-QP})$$

$$\text{subject to } L_f h(x) + L_g h(x)u + \mathcal{K}(h(x)) \geq 0,$$

where $u_{\text{nom}}(x)$ represents a control input from a potentially unsafe nominal controller. Given the control-affine structure of our system, $C(x)$ forms a polyhedral set, reducing (CBF-QP) to a convex quadratic program.

In the setting we consider, however, there are two additional practically motivated challenges:

- **Discretization:** Agent controllers generally operate in discrete time due to sampling. Therefore we must discretize the dynamics and synthesize a robust safe policy.
- **Uncertainty:** The human's behavior is unknown *a priori* to the robot. This means they do not have a deterministic model of the control input u_H , and must learn it from data.

B. Discretization & Robust Control Barrier Function

We discretize the continuous-time dynamics given in (1). Let $t_k = k \cdot \Delta T$ for $k \in \mathbb{N}$ denote sampling instants with fixed timestep $\Delta T \in \mathbb{R}_{>0}$. Under this time discretization, the controller operates using a zero-order hold control law: for each $k \in \mathbb{N}$, set

$$u(t) = u_k, \quad \forall t \in [t_k, t_{k+1}). \quad (5)$$

Let $x_k := (x_R(t_k), x_H(t_k))$ be the corresponding sampled state². Thus, discrete-time joint dynamics are give by

$$x_{k+1} = f_d(x_k) + B_{R,d}u_{R,k} + B_{H,d}u_{H,k} := F_d(x_k, u_k), \quad (6)$$

where $B_{j,d} := B_j \cdot \Delta T$ for $j \in \{R, H\}$. The discrete-time dynamics $f_d(x_k)$ are obtained through standard numerical integration of $f(x(t))$ over the time interval $[t_k, t_{k+1}]$ under

¹An extended class- \mathcal{K}_∞ function $\mathcal{K} : \mathbb{R} \rightarrow \mathbb{R}$ is continuous, strictly increasing, satisfies $\mathcal{K}(0) = 0$, and $\mathcal{K}(r) \rightarrow \infty$ as $r \rightarrow \infty$.

²Under a zero-order hold control law as in (5) with compact set \mathcal{U} , uniqueness of the maximal closed-loop solution x_k is guaranteed [19].

the zero-order hold assumption on the control inputs. The discrete-time dynamics $f_d(x_k)$ are implicitly defined by this integral using Gronwall's inequality and the Lipschitz continuity of f .

To ensure safety despite the time discretization, we construct a *robust CBF* [20] and corresponding safe controller set. We show first in the deterministic setting that the zero-order hold controller remains safe on every interval $[t_k, t_{k+1}]$ as long as $u_k \in C(x_k)$ and $x_k \in \mathcal{S}$. To this end, we make the following regularity assumptions.

Assumption 3: The mappings $h \in C^1$ and $\mathcal{K} \circ h$ are \mathfrak{L}_h and $\mathfrak{L}_{\mathcal{K} \circ h}$ Lipschitz continuous, respectively.

Observe that $L_f h(x)$ and $L_g h(x)$ are, respectively, $\mathfrak{L}_h \mathfrak{L}_f$ and $\mathfrak{L}_h \mathfrak{L}_g$ Lipschitz continuous under Assumption 2.

Now, define the robust safe control set—which we will use to select the control at each time t_k —as follows: for the current state and agent action $(x(t_k), u_H(t_k))$, the robust safe set for the robot is given by

$$\tilde{C}(x(t_k), u_H(t_k)) = \{u_R \in \mathcal{U}_R \mid L_f h(x(t_k)) + L_g h(x(t_k))u + \mathcal{K}(h(x(t_k))) \geq \eta, u = (u_R, u_H(t_k))\}$$

where $\eta := \Delta T \cdot \mathfrak{L}_x (\mathfrak{L}_h \mathfrak{L}_g R_u + \mathfrak{L}_h \mathfrak{L}_f + \mathfrak{L}_{\mathcal{K} \circ h})$ is a robustness margin accounting for discretization.

Lemma 1 (Robust CBF): Consider the interval $[t_k, t_{k+1}]$ where $\Delta T = t_{k+1} - t_k$. Suppose that the agent is employing a zero-order hold policy, and that at each discrete time step k , the proposed algorithm selects $u_{R,k} \in \tilde{C}(x(t_k), u_H(t_k))$ for the current state and agent action. Further, suppose that the robot uses a fixed control input $u_R(t) \equiv u_{R,k}$ on $[t_k, t_{k+1}]$ and that the current state is safe—i.e., $x(t_k) \in \mathcal{S}$. Then $x(t) \in \mathcal{S}$ for all $t \in [t_k, t_{k+1}]$.

The proof of the preceding lemma is given in Appendix C. Observe that this lemma provides a safety guarantee for the entire discretized time interval.

In the deterministic setting—i.e., where the human behavior policy is known to the robot—we could easily modify the optimization problem in (CBF-QP) by replacing the control barrier function constraint with the robust version defined above: namely,

$$\pi(x(t_k)) := \min_{u_R} \|u_R - u_{\text{nom}}(x(t_k))\|_2^2 \quad (\text{RCBF-QP})$$

$$\text{subject to } u_R \in \tilde{C}(x(t_k), u_H(t_k))$$

However, in practice, such knowledge would not be available to the robot.

C. Stochastic Human Behavior with Probabilistic Safety

Now that we have tackled the time discretization, we can deal with the second challenge on the stochasticity of the human policy. As noted above, the human's policy is unknown *a priori*. Indeed, since the human's actions are stochastic, we cannot optimize (RCBF-QP) since the robot does not have $u_H(t_k)$. Instead, we aim to ensure this condition holds with high probability.

The following lemma establishes a probabilistic safety guarantee by imposing constraints on the robots control actions given a stochastic policy $P(u_H(t_k) | x_{0:k})$ for the human agents, where $x_{0:k} := (x(t_0), \dots, x(t_k))$.

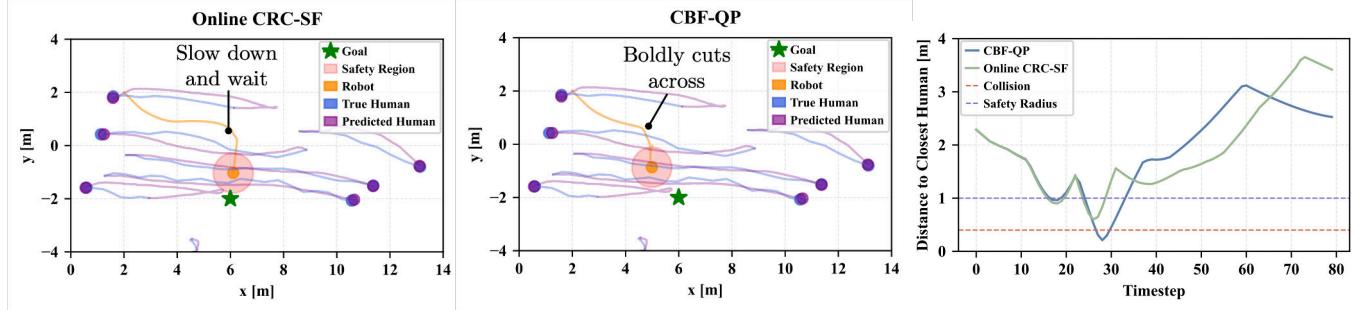


Fig. 2: **Multi-Agent Scenario.** Example of the human crowd simulation setup. The CRC Safety Filter maintains the probabilistic safety guarantees while the standard CBF exhibits unsafe behavior. The safety margin λ adapts based on human prediction uncertainty, ensuring larger distances are maintained during uncertain interactions.

Lemma 2: Suppose the joint control space \mathcal{U} satisfies Assumption 1, the system (6) satisfies Assumptions 2 and 3, and both the robot and humans employ zero-order hold control policies. Fix a confidence level $\gamma \in (0, 1)$. Suppose $x(t_k) \in \mathcal{S}$ and the robot selects $u_R(t_k)$ such that

$$\Pr(u_R(t_k) \in \tilde{\mathcal{C}}(x(t_k), u_H(t_k))) \geq 1 - \gamma, \quad (7)$$

where the probability is taken with respect to the distribution of human actions $u_H(t_k) \sim P(u_H(t_k) | x_{0:k})$. Then, the joint state remains safe for all $t \in [t_k, t_{k+1}]$ —i.e.,

$$\Pr(\{x(t) \in \mathcal{S} \ \forall t \in [t_k, t_{k+1}]\}) \geq 1 - \gamma. \quad (8)$$

This lemma trivially follows from Lemma 1 which implies that if $u_R(t_k) \in \tilde{\mathcal{C}}(x(t_k), u_H(t_k))$ and $x(t_k) \in \mathcal{S}$, then $x(t) \in \mathcal{S}$ for all $t \in [t_k, t_{k+1}]$.

Given the safety guarantee implied by Lemma 2, we modify the (RCBF-QP) to incorporate this high probability guarantee: namely,

$$\pi(x(t_k)) := \min_{u_R} \|u_R - u_{\text{nom}}(x(t_k))\|_2^2 \quad (\text{SRCBF-QP})$$

$$\text{subject to } \Pr(u_R(t_k) \in \tilde{\mathcal{C}}(x(t_k), u_H(t_k))) \geq 1 - \gamma$$

This optimization problem now has a chance constraint.

III. ROBUST AND SAFE CONTROL POLICY SYNTHESIS

In this section, we describe a method to address the probabilistic safety constraint in (SRCBF-QP). We reformulate the CBF condition using CRC to quantify the risk associated with the uncertainty in human’s stochastic policy.

A. Safety Barrier Certificates

Following [21], we define two safety barrier certificates, the deterministic $\mathcal{B}(x_k, u_k)$ using ground truth states and the predicted $\hat{\mathcal{B}}(\hat{x}_k, \hat{u}_k)$ using estimated states:

$$\mathcal{B}(x_k, u_k) = L_f h(x_k) + L_g h(x_k) u_k + \mathcal{K}(h(x_k)) - \eta, \quad (9)$$

$$\hat{\mathcal{B}}(\hat{x}_k, \hat{u}_k) = L_f \hat{h}(\hat{x}_k) + L_g \hat{h}(\hat{x}_k) \hat{u}_k + \mathcal{K}(\hat{h}(\hat{x}_k)) - \eta,$$

where $u_k = (u_{R,k}, u_{H,k})$ and $\hat{u}_k = (\hat{u}_{R,k}, \hat{u}_{H,k})$. For safety, we require $\mathcal{B}(x_k, u_k) \geq 0$, but can only compute $\hat{\mathcal{B}}(\hat{x}_k, \hat{u}_k)$ during planning.

Assumption 4 (Bounded barrier prediction error): There exists a finite constant $\xi \in \mathbb{R}_{>0}$ such that, for all time steps $k \in \mathbb{N}$, the difference between the true barrier certificate $\mathcal{B}(x_k, u_k)$ and the predicted barrier certificate $\hat{\mathcal{B}}(\hat{x}_k, \hat{u}_k)$ is bounded $\sup |\mathcal{B}(x_k, u_k) - \hat{\mathcal{B}}(\hat{x}_k, \hat{u}_k)| \leq \xi$. This bound holds for all joint states at each timestep k .

Remark 1: While this assumption is used in the theoretical analysis to establish statistically valid guarantees, the proposed method does not require explicit knowledge of this bound in practice since CRC implicitly learns the appropriate margins from data.

B. Safety Margin Synthesis via Conformal Risk Control

To address the chance constraint in (SRCBF-QP), we leverage nonexchangeable CRC [22], a statistical verification technique that provides a flexible framework for quantifying prediction errors in dynamic settings. Theoretical guarantees for classical conformal uncertainty quantification methods hold under the assumption that the data is exchangeable³ [23]. Data from controlled dynamical systems is not exchangeable since the current state depends on the past state and control inputs. Non-exchangeable CRC accounts for these dependencies by essentially reweighting the conformal prediction sets. For brevity, we will refer to nonexchangeable CRC simply as *CRC* from here on out.

To produce safety guarantees using the estimated barrier certificate, we need to bound the estimation error. The procedure we develop to do this uses a *safety margin* parameter that is selected based on the historical estimates of the safety barrier certificates as we describe below.

Let $\lambda \in \Lambda \subseteq \mathbb{R}_{>0}$ be a *safety margin* parameter for constructing C_λ that characterizes the set of safe robot control inputs under uncertainty:

$$C_\lambda(\hat{x}_k, \hat{u}_{H,k}) = \{u_{R,k} \in \mathcal{U}_R \mid \hat{\mathcal{B}}(\hat{x}_k, \hat{u}_k) - \lambda_k \geq 0\}. \quad (10)$$

For any $\lambda, \lambda' \in \Lambda$, we have $\lambda \leq \lambda' \implies C_{\lambda'}(\cdot) \subseteq C_\lambda(\cdot)$, i.e., larger λ yields a more restrictive control space. The safety margin essentially creates a buffer that accounts for prediction uncertainty: as uncertainty increases, λ increases to ensure safety by further restricting the set of permissible controls, directly capturing the risk associated with the human’s stochastic policy.

1) *Conformal Risk Control Safety Filter:* Now, we present a risk-aware safety filter that enforces probabilistic safety constraints through a quadratic program formulation. At each discrete timestep k , the CRC safety filter computes a control input that minimally deviates from a nominal robot policy

³Exchangeability is a slight relaxation of the well-known condition of independent and identically distributed data.

while ensuring safety with high probability:

$$\pi(x(t_k)) := \min_{u_R} \|u_R - u_{\text{nom}}(x(t_k))\|_2^2 \quad (\text{CRC-SF})$$

$$\text{subject to } \Pr(u_R(t_k) \in C_\lambda(\hat{x}_k, \hat{u}_{H,k})) \geq 1 - \gamma$$

where $u_{\text{nom}}(x(t_k))$ is the nominal robot control and $C_\lambda(\hat{x}_k, \hat{u}_{H,k})$ is the set of safe robot controls parameterized by safety margin λ . We develop a method to update λ online based on interaction context (Section IV). This formulation provides single-timestep safety guarantees under human behavioral uncertainty (Theorem 1).

2) *Optimizing the Safety Margin Parameter:* To select the safety margin parameter, we define a loss function $\mathcal{L} : \Lambda \rightarrow [0, B]$ that quantifies the barrier prediction error:

$$\mathcal{L}(\lambda) = \max\{0, |\mathcal{B}(x_k, u_k) - \hat{\mathcal{B}}(\hat{x}_k, \hat{u}_k)| - \lambda\}. \quad (11)$$

This loss function is bounded under Assumption 4; this is a technical condition that is needed for the theoretical analysis of the CRC method. Further, this loss function quantifies the amount by which the prediction error exceeds the safety margin λ , and satisfies the key properties required by the CRC framework: (i) non-negativity, (ii) boundedness by Assumption 4, and (iii) monotonically non-increasing with respect to λ .

To account for temporal dependencies in human-robot interaction, we apply nonexchangeable CRC using geometrically decaying weights. The optimal safety margin $\hat{\lambda}_k$ is chosen as the smallest value that ensures the weighted risk remains below α with an appropriate correction term:

$$\hat{\lambda}_k = \inf \left\{ \lambda : \frac{1}{n_w + 1} \sum_{i=1}^{n_k} w_i \mathcal{L}_i(\lambda) + \frac{B}{n_w + 1} \leq \alpha \right\}, \quad (12)$$

where $w_i = \rho^{n_k+1-i}$ are geometrically decaying weights with $\rho \in (0, 1)$, n_k is the size of the calibration set (detailed in Section IV), and $n_w = \sum_{i=1}^{n_k} w_i$. This formulation ensures

$$\mathbb{E}[\mathcal{L}(\hat{\lambda}_k)] \leq \alpha + \beta, \quad (13)$$

where $\alpha \in (0, 1)$ is a user-specified risk level. Here β represents the *total variation distance* that quantifies the degree of non-exchangeability in the data. As discussed in [22], in practice β is difficult to compute, but with a proper choice of weights and gradual distribution shift such as that in time series data, β is likely to be small.

C. High Probability Safety Guarantees

This formulation leads to the main theoretical results. First, we establish that CRC guarantees on the barrier prediction error translate to probabilistic bounds on the difference between true and predicted barrier values.

Lemma 3 (Barrier Value Concentration): Given a risk threshold $\alpha \in (0, 1)$ and confidence level $\gamma \in (0, 1)$, if we compute $\hat{\lambda}$ using non-exchangeable CRC to satisfy $\mathbb{E}[\mathcal{L}(\hat{\lambda})] \leq \alpha + \beta$, then setting $\epsilon = (\alpha + \beta)/\gamma$ gives

$$\Pr(|\mathcal{B}(x_k, u_k) - \hat{\mathcal{B}}(\hat{x}_k, \hat{u}_k)| \leq \hat{\lambda}_k + \epsilon) \geq 1 - \gamma. \quad (14)$$

The proof of Lemma 3 is contained in the Appendix. Building on this result, we establish our safety guarantee.

Algorithm 1 Offline Safety Margin Calibration using CRC

```

1: Input: Risk threshold  $\alpha$ , human policy  $P(u_H | x_{0:k})$ , nominal robot policy  $\pi_{\text{nom}}$ , discount factor  $\rho$ , trajectories per batch  $K$ , total trajectories  $M$ , horizon  $N$ 
2: Initialize: Training dataset  $\mathcal{T} \leftarrow \emptyset$ 
3: for  $b = 1$  to  $M/K$  do  $\triangleright$  Each batch uses  $K$  trajectories to compute one set of  $\lambda$  values
4:    $\mathcal{D}_b \leftarrow \emptyset$   $\triangleright$  Calibration data for batch  $b$ 
5:   for  $j = 1$  to  $K$  do  $\triangleright$  Interaction trajectories
6:     Initialize states:  $x_{R,0}, x_{H,0}$ 
7:     for  $k = 0$  to  $N - 1$  do  $\triangleright$  Data Collection
8:       Sample human action:  $\hat{u}_{H,k} \sim P(u_H | x_{0:k})$ 
9:       Set  $u_{R,k} = \pi_R(x_k, \hat{u}_{H,k})$  solving (RCBF-QP)
10:      Update states:  $x_{k+1} = F_d(x_k, u_k)$  via (6)
11:      Compute barrier certificates as in (9)
12:       $\mathcal{D}_b \leftarrow \mathcal{D}_b \cup \{(\mathcal{B}_k, \hat{\mathcal{B}}_k)\}$ 
13:   for  $k = 0$  to  $N - 1$  do  $\triangleright$  Risk Calibration
14:      $\mathcal{D}_{b,k} \leftarrow \{(\mathcal{B}_i, \hat{\mathcal{B}}_i, k_i) \in \mathcal{D}_b \mid k \leq k_i \leq N - 1\}$ 
15:      $n_k \leftarrow |\mathcal{D}_{b,k}| = K \times (N - k)$ 
16:      $w_i \leftarrow \rho^{n_k+1-i}$  for  $i = 1, \dots, n_k$ 
17:     Compute  $\hat{\lambda}_{b,k}$  using (12)
18:     Extract feature vector  $\phi_{b,k}$  for each  $K$  trajectories
19:      $\mathcal{T} \leftarrow \mathcal{T} \cup \{(\phi_{b,k}, \hat{\lambda}_{b,k})\}$ 
20: return  $\mathcal{T}$ 

```

Theorem 1 (CRC-CBF Safety Guarantee): Consider the human-robot system (6) with barrier certificates defined in (9). Given a confidence level $\gamma \in (0, 1)$ and risk threshold $\alpha \in (0, 1)$, if we have the CRC guarantee such that $\hat{\lambda}$ satisfies $\mathbb{E}[\mathcal{L}(\hat{\lambda})] \leq \alpha + \beta$ and set $\epsilon = (\alpha + \beta)/\gamma$, then the prediction set defining the safe set of control inputs under uncertainty,

$$C_\lambda(\hat{x}_k, \hat{u}_{H,k}) = \{u_R \in \mathcal{U}_R \mid \hat{\mathcal{B}}(\hat{x}_k, \hat{u}_k) - (\hat{\lambda}_k + \epsilon) \geq 0\}, \quad (15)$$

ensures that $\Pr(h(x_{k+1}) \geq 0) \geq 1 - \gamma$ holds.

The proof of the above theorem can be found in the Appendix. This result shows that by bounding prediction errors on barrier functions using CRC, we can construct control constraints that guarantee safety with high probability.

IV. ALGORITHMIC DETAILS

In this section, we present the algorithmic details for implementing the proposed probabilistic safety framework. We begin with a brief high-level overview of the procedure. The proposed method first trains a stochastic human behavioral policy $P(u_H | x_{0:k})$ on real-world pedestrian data. The interaction history $x_{0:k} = \{x_0, x_1, \dots, x_k\}$ captures the joint state evolution of all agents in the scene, enabling the human policy to account for multi-agent dynamics through observed interactions. While our theoretical framework supports conditioning on robot control actions—through the development of robot-aware human behavior models as in [24]—our current implementation is trained on pedestrian datasets where robot actions are not present [25]. Specifically, our human behavioral model takes as input the positions and velocities of all agents in the scene over the interaction history and outputs a sequence of predicted human controls, following

Algorithm 2 CRC Safety Filter with Online Updates

```

1: Input: Initial states  $(x_{R,0}, x_{H,0})$ , stochastic human policy
    $P(u_H | x_{0:k})$ , LSTM model  $\Phi$ , timestep  $\Delta T$ , horizon  $N$ 
2: Output: Safe robot trajectory  $\{x_{R,k}\}_{k=0}^N$ 
3: Sample human trajectories:  $\{\hat{x}_{H,k}\}_{k=0}^N \sim P(u_H | x_{0:k})$ 
4: Compute predicted human states  $\{\hat{x}_{H,k}\}_{k=0}^N$  by integrating
   human dynamics
5: for  $k = 0$  to  $N - 1$  do
6:   Compute  $\hat{B}_k$  using  $(\hat{x}_{H,k}, \hat{u}_{H,k})$  via (9)
7:   Extract feature vector  $\phi_k$ 
8:   Update safety margin:  $\tilde{\lambda}_k = \Phi(\phi_k)$ 
9:   Solve (CRC-SF) to obtain  $u_{R,k}$ 
10:  Apply control  $u_{R,k}$  to the robot system
11: return  $\{x_{R,k}\}_{k=0}^N$ 

```

approaches similar to [26]. The learned human policy captures relative spatial information (e.g., distance to nearest agents, relative velocities) that enables generalization from the pedestrian-only training data to human-robot interaction scenarios in our simulated environments.

The proposed method requires the specification of an *a priori* nominal robot control policy π_{nom} that governs robot behavior during the offline calibration phase. The nominal policy can be any reasonable control strategy chosen by the user—for instance, a goal-reaching controller or any existing deterministic safe control method. The nominal policy serves as the robot’s baseline behavior around which we calibrate our uncertainty quantification and safety margin adaptation.

In an *offline phase*, we use the learned human policy together with the nominal robot policy to generate interaction trajectories $\tau = \{(x_k, u_k)\}_{k=0}^{N-1}$ that capture the joint evolution of human-robot state-action pairs over a finite horizon N . For each trajectory, we compute both the true barrier certificate values B_k using ground truth human actions and the predicted values \hat{B}_k using actions sampled from our behavior model. Then, we apply CRC to the collected barrier certificate pairs $\{(B_k, \hat{B}_k)\}$ to compute the optimal safety margins $\{\hat{\lambda}_k\}_{k=0}^{N-1}$. We then train a predictive model that maps the current context in the scene to the appropriate safety margins, with details outlined later in this section. In the *online phase*, we use this model to dynamically adapt λ , enabling risk-adaptive control.

Computing Optimal Safety Margin. We now detail the offline CRC calibration procedure for computing optimal safety margins as outlined in Algorithm 1. We adopt the geometric weight decay approach from [22] to handle temporal dependencies through re-weighting the loss function, ensuring statistically valid risk control guarantees.

We generate M interaction trajectories by simulating the stochastic human policy with a nominal robot policy π_{nom} . In our implementation, π_{nom} solves the robust CBF optimization problem (RCBF-QP) where $u_{\text{nom}}(x_k)$ is computed using a goal-reaching control Lyapunov function. We process the M trajectories in batches of size K , where each batch computes safety margins $\{\hat{\lambda}_{b,k}\}_{k=0}^{N-1}$. Each safety margin is computed

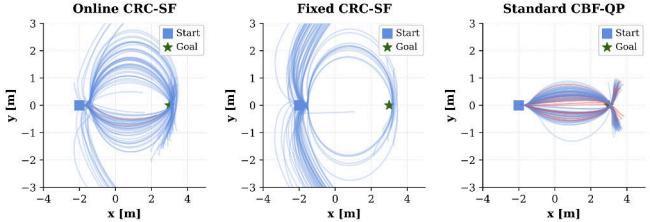


Fig. 3: Trajectories generated from 100 single-agent head-on interactions with each method. Red trajectories indicate collisions.

by collecting all barrier certificate data from timesteps k through $N-1$ across the K trajectories in that batch, yielding $n_k = K \times (N - k)$ total samples. We apply exponentially decaying weights $w_i = \rho^{n_k + 1 - i}$ to emphasize more recent timesteps, then solve (12) to compute the optimal safety margin. We then use these safety margins as ground truth labels to train a predictive model to update λ online.

Online Updates. To enable dynamic safety margin adaptation during deployment, we train a 2-layer LSTM network with 32 hidden units, denoted Φ , that takes as input a feature vector ϕ concatenating the robot state x_R , human state \hat{x}_H , Euclidean distance between agents $\|p_R - p_H\|$, predicted barrier value \hat{B} , and timestep k . We choose an LSTM to capture temporal dependencies in the interaction sequences, though simpler models such as MLPs could also be used. The LSTM is trained using MSE loss between the ground truth safety margins from the offline calibration phase and the predicted values from our model, which achieves low error on the training and validation data. During online deployment, we perform inference at each timestep k to update $\tilde{\lambda}_k = \Phi(\phi_k)$ in (CRC-SF), allowing the robot to dynamically restrict the set of admissible controls in (10) based on the uncertainty in the stochastic human policy.

V. EXPERIMENTAL RESULTS

To demonstrate the effectiveness of our method, we evaluate CRC-SF through single-agent and multi-agent experiments designed to answer the following research questions:

RQ1. How does CRC-SF compare in terms of safety and performance to baseline methods?

RQ2. What impact does the online adaptation of the safety margin have on balancing safety and performance compared to fixed safety margins?

RQ3. How does our approach perform in complex multi-agent scenarios where human behavior is modeled using a prediction model learned from real-world pedestrian data?

A. Experimental setup

Choosing the risk threshold α . To demonstrate the application of safety-critical control using our framework, we choose $\alpha = 0.01$ for all of the experiments. According to our theoretical analysis, this should roughly relate to $\gamma = 0.99$, seeking to provide high probabilistic safety guarantees.

Baselines. We compare our approach Online CRC-SF against three baselines: (i) CBF-QP which does not consider uncertainty in human behavior. This baseline provides a direct comparison to show the benefits of incorporating

human behavioral uncertainty into the safety constraints. (ii) Fixed CRC-CF which uses a *fixed* λ value computed from considering the entire future trajectory in the calibration phase. This baseline method will help validate the use of an online update scheme to adjust the safety margin to account for different contexts. (iii) MPPI [2], a sampling-based method that can handle complex uncertainty distributions, but does not provide formal guarantees.

Metrics. We evaluate performance using five key metrics: (i) collision rate, measuring the percentage of simulations where the robot collides with a human; (ii) safety constraint violation rate, measuring instances when the robot violates the safe control constraint but not necessarily resulting in collision; (iii) goal-reaching success rate, capturing the percentage of simulations where the robot reaches within a small neighborhood of its target position within the allotted time horizon; (iv) control effort, quantifying the magnitude of control inputs applied during the trajectory execution; and (v) control smoothness, assessing the rate of change in control inputs over time.

B. Single-agent scenario

We study a challenging head-on interaction between a human and robot, illustrated in Figure 3. The human randomly chooses to pass either left or right of the robot, requiring the robot to adapt its decisions in real-time. The scenario mimics a corridor situation where the robot must pass the human safely while minimizing deviation from its path to efficiently reach the opposite side.

Simulation setup. Both the robot and human are modeled using the dynamically-extended unicycle model with state, $x = [p_x, p_y, \theta, v]^\top$ where (p_x, p_y) represents the 2D position, θ is the heading angle, and v is the velocity. The control inputs $u = [u_1, u_2]^\top$ represent steering and acceleration, bounded by $u \in [-0.3, 0.3] \times [-1, 1]$. The safety constraint is defined using a barrier function that maintains a minimum squared distance between the agents $h(x_R, x_H) = \|p_R - p_H\|_2^2 - R^2$ where $R = 1.0\text{m}$ is the relative safety radius. The class- \mathcal{K}_∞ function is a linear function $\mathcal{K}_{\text{CBF}}(r) = r$.

We set $N = 8.0\text{s}$ (80 steps at $\Delta T = 0.1$). The human’s behavior is simulated using a stochastic policy that generates noisy control inputs around a nominal goal-reaching trajectory, with bounded Gaussian noise ($\sigma = 1.2$, clipped to ± 0.5) added to create noticeable uncertainty in the human’s future states. We ran 100 test instances of the simulation.

Results. Metrics from this single-agent experiment are summarized in Table I (bottom). This single-agent experiment offers insight into **RQ1** and **RQ2**. We observe that our Online CRC-SF achieves the best safety-efficiency trade-off. CBF-QP and Fixed CRC-SF, are either very efficient but not safe, or vice versa. Both CRC-SF approaches perform well regarding the safety metrics, which is to be expected, but for Fixed CRC-SF, we see that it is very conservative where safety violations are close to zero percent, but has the worst efficiency by a significantly large margin. While Online CRC-SF offers a “close-second” across all metrics;

Method	Multi-Agent Scenario				
	Coll. (%) \downarrow	Safety Viol.(%) \downarrow	Goal (%) \uparrow	Control Effort \downarrow	Control Smooth. \downarrow
CBF-QP	38.8	81.4	96.4	0.30	0.12
Fixed CRC-SF	10.2	49.0	31.6	0.53	0.23
Online CRC-SF	3.0	53.2	84.8	0.43	0.19
MPPI	34.4	81.2	77.6	0.65	0.46
Single-Agent Scenario					
CBF-QP	16.0	57.0	100.0	0.31	0.12
Fixed CRC-SF	0.0	1.0	14.0	1.35	0.30
Online CRC-SF	2.0	15.0	78.0	0.59	0.20

TABLE I: Performance comparison between multi-agent and single-agent human-robot navigation scenarios. The multi-agent results represent averages across five different test configurations (each with distinct initial and goal positions), with each configuration run 100 times. The single-agent results represent averages from 100 head-on simulations.

its safety and performance are significantly more similar to the best performing method in each metric than the worst.

C. Multi-agent scenario

We evaluate our approach using real-world pedestrian trajectory data recorded from crowd interactions [25].

Human Behavior Model. We adapt prior work [27] to develop a diffusion model trained on real-world pedestrian trajectory datasets. This enables us to sample multiple possible human action sequences conditioned on the interaction history. Human states are then obtained by integrating these predicted controls using known single-integrator dynamics [24], [26]. This probabilistic formulation allows the robot to reason about the distribution of possible future human states and actions, which is essential for our safety-critical control approach. By sampling from this learned distribution, we can evaluate the probabilistic safety constraints and, crucially, dynamically adapt the robot’s level of conservativeness based on the uncertainty in human behavior predictions.

CBF construction. Rather than enforcing constraints with respect to every human in the environment, which could be computationally intractable, we define a neighborhood around the robot within which safety guarantees must be maintained. Given a robot state x_R , we consider the set of humans $\mathcal{N}(x_R) = \{x_H^i : \|p_R - p_H^i\| \leq R_N\}$ where $R_N = 3\text{m}$ is the neighborhood radius. For each human i in this neighborhood, we construct a corresponding joint barrier function: $h_i(x_R, x_H^i) = \|p_R - p_H^i\|_2^2 - R^2$.

Test scenarios. We performed experiments on five different scenarios with different initial and goal positions for the robot. For each scenario, we ran 100 different trials to account for the stochastic nature of the prediction model. An example of a multi-agent scenario is shown in Figure 2.

D. Discussion

Qualitative results. Figure 2 visualizes the trajectories from Online CRC-SF and CBF-QP. We observe that since Online CRC-SF method adapts its safety margin based on context, it proactively “slows down and waits” for a pedestrian to pass as it starts to turn and move perpendicularly to the direction of motion (i.e., towards the bottom of the plot). However,

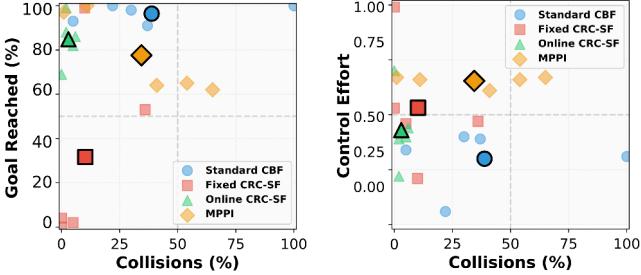


Fig. 4: Efficiency versus safety results across five test scenarios for the multi-agent setting. Darker markers with a black outline indicate the mean value across the five test scenarios. Online CRC-SF achieves the best safety-efficiency trade-off while maintaining more consistent performance across scenarios.

CBF-QP, which does not account for uncertainty, “boldly cuts across” collides with an oncoming pedestrian. The distance to the closest human over these interactions is shown in Figure 2 (right). While Online CRC-SF occasionally violates the safety radius, it consistently recovers without collisions, unlike CBF-QP—suggesting future work to strengthen guarantees during temporary constraint violations.

Quantitative results. Metrics from the multi-agent experiments are summarized in Table I (top). In answering **RQ3**, we see that even with this more challenging multi-agent setting and the use of a diffusion-based human behavior prediction model, we find that Online CRC-SF performs either the best or a close-second. These results reaffirm the **RQ1** and **RQ2** takeaways from the single-agent experiments. Figure 4 visualizes this safety-efficiency trade-off for each of the five test scenarios. Not only does Online CRC-SF offer the most desirable safety-efficiency trade-off, but its performance is also consistent across the five test scenarios whereas other methods exhibited higher variance.

VI. CONCLUSIONS

We propose a safe probabilistic robot control framework for interacting with multiple human agents. The proposed safe controller can dynamically adjust its conservativeness based on the uncertainty on future human behaviors. We achieve this by leveraging techniques from CRC to form an online estimate of a safety margin for a CBF constraint that will achieve the desired probabilistic safe constraint. The experiments on synthetic and real-world interaction data indicate that the approach achieves the best balance between safety and efficiency compared to existing CBF methods. Moreover, the dynamic adjustment of the safety margin naturally induces proactive behaviors in situations with more uncertainty.

REFERENCES

- [1] H. Zhu and J. Alonso-Mora, “Chance-constrained collision avoidance for MAVs in dynamic environments,” *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 776–783, 2019.
- [2] G. Williams, A. Aldrich, and E. A. Theodorou, “Model predictive path integral control using covariance variable importance sampling,” 2015.
- [3] V. Vovk, A. Gammerman, and G. Shafer, *Algorithmic Learning in a Random World*, 2005.
- [4] I. Gibbs and E. J. Candès, “Conformal inference for online prediction with arbitrary distribution shifts,” *J. Machine Learning Research*, vol. 25, pp. 1–36, 2022.
- [5] A. N. Angelopoulos, S. Bates, A. Fisch, L. Lei, and T. Schuster, “Conformal risk control,” *Int. Conf. Learning Representation*, 2024.
- [6] A. Ames, S. D. Coogan, M. Egerstedt, G. Notomista, K. Sreenath, and P. Tabuada, “Control barrier functions: Theory and applications,” *European Control Conference*, pp. 3420–3431, 2019.
- [7] M. Saveriano and D. Lee, “Learning barrier functions for constrained motion planning with dynamical systems,” in *IEEE/RSJ Int. Conf. on Intelligent Robots & Systems*, 2019.
- [8] L. Lindemann, H. Hu, A. Robey, H. Zhang, D. V. Dimarogonas, S. Tu, and N. Matni, “Learning hybrid control barrier functions from data,” in *Conference on Robot Learning*, 2020.
- [9] A. Robey, H. Hu, L. Lindemann, H. Zhang, D. V. Dimarogonas, S. Tu, and N. Matni, “Learning control barrier functions from expert demonstrations,” in *Proc. IEEE Conf. on Decision and Control*, 2020, pp. 3717–3724.
- [10] O. So, Z. Serlin, M. Mann, J. Gonzales, K. Rutledge, N. Roy, and C. Fan, “How to train your neural control barrier function: Learning safety filters for complex input-constrained systems,” in *Proc. IEEE Conf. on Robotics and Automation*, 2024.
- [11] E. Das, S. X. Wei, and J. W. Burdick, “Robust Control Barrier Functions with Uncertainty Estimation,” *IEEE Tran. Automatic Contr.*, 2025.
- [12] X. Xu, P. Tabuada, J. W. Grizzle, and A. Ames, “Robustness of control barrier functions for safety critical control,” in *IFAC Conf. Analysis and Design of Hybrid Systems*, 2016.
- [13] R. 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.
- [14] L. Lindemann, M. Cleaveland, G. Shim, and G. J. Pappas, “Safe planning in dynamic environments using conformal prediction,” *IEEE Robotics and Automation Letters*, vol. 8, no. 8, pp. 5116–5123, 2023.
- [15] A. Dixit, L. Lindemann, S. Wei, M. Cleaveland, G. J. Pappas, and J. W. Burdick, “Adaptive conformal prediction for motion planning among dynamic agents,” *Conf. Learning for Dynamics and Control*, 2023.
- [16] H. Zhou, Y. Zhang, and W. Luo, “Safety-critical control with uncertainty quantification using adaptive conformal prediction,” *American Control Conf.*, 2024.
- [17] S. Huriot and H. Sibai, “Safe decentralized multi-agent control using black-box predictors, conformal decision policies, and control barrier functions,” *arXiv arxiv:preprint:2409.18862*, 2024.
- [18] J. Lekeufack, A. N. Angelopoulos, A. Bajcsy, M. I. Jordan, and J. Malik, “Conformal decision theory: Safe autonomous decisions from imperfect predictions,” in *Proc. IEEE Conf. on Robotics and Automation*, 2024.
- [19] E. Sontag, “Mathematical control theory (second edition),” 1998.
- [20] S. Dean, A. Taylor, R. Cosner, B. Recht, and A. Ames, “Guaranteeing Safety of Learned Perception Modules via Measurement-Robust Control Barrier Functions,” in *Proc. Conf. Robot Learning*, 2020.
- [21] L. Wang, A. D. Ames, and M. Egerstedt, “Safety barrier certificates for collisions-free multirobot systems,” *IEEE Trans. Robotics*, vol. 33, no. 3, pp. 661–674, 2017.
- [22] A. Farinhas, C. Zerva, D. T. Ulmer, and A. Martins, “Non-exchangeable conformal risk control,” in *Int. Conf. Learning Representations*, 2024.
- [23] V. Vovk, A. Gammerman, and G. Shafer, *Algorithmic Learning in a Random World*, 01 2005.
- [24] E. Schmerling, K. Leung, W. Vollprecht, and M. Pavone, “Multimodal Probabilistic Model-Based Planning for Human-Robot Interaction,” *IEEE Int. Conf. Robotics and Automation*, 2018.
- [25] S. Pellegrini, A. Ess, Schindler, and L. Gool, “You’ll Never Walk Alone: Modeling Social Behavior for Multi-Target Tracking,” in *IEEE Int. Conf. Computer Vision*, 2009.
- [26] T. Salzmann, B. Ivanovic, P. Chakravarty, and M. Pavone, “RON++: Dynamically-Feasible Forecasts With Heterogeneous Data,” in *European Conf. Computer Vision*, 2020.
- [27] K. Mizuta and K. Leung, “CoBL-Diffusion: Diffusion-based conditional robot planning in dynamic environments using control barrier and lyapunov functions,” in *IEEE/RSJ Int. Conf. on Intelligent Robots & Systems*, 2024.