

# Statistical Contraction for Chance-Constrained Trajectory Optimization of Non-Gaussian Stochastic Systems

Rihan Aaron D'Silva and Hiroyasu Tsukamoto

**Abstract**—We present a distribution-free approach to robust trajectory optimization and control of discrete-time, nonlinear, and non-Gaussian stochastic systems, with closed-loop guarantees on chance constraint satisfaction. Our framework employs conformal inference to generate coverage-based confidence sets for the closed-loop dynamics around arbitrary reference trajectories under uncertainty. It thereby constructs a joint nonconformity score to quantify both the validity of contraction (i.e., incremental stability) conditions and the impact of external stochastic disturbance on the closed-loop dynamics, without any distributional assumptions. Via appropriate constraint tightening, chance constraints can be reformulated into tractable, statistically valid deterministic constraints on the reference trajectories. This enables a formal pathway to certify the performance of learning-based motion planners and controllers, such as those with neural contraction metrics, in safety-critical real-world applications. Notably, our statistical guarantees are non-diverging and can be computed with finite samples of the underlying uncertainty, without overly conservative structural priors. We demonstrate our approach in motion planning problems for designing safe, dynamically feasible trajectories in both numerical simulations and hardware experiments.

## I. INTRODUCTION

Generating interpretable, high-performance motion plans for dynamical systems under unstructured uncertainties is central to their deployment in safety-critical applications. As the scope and scale of the tasks entrusted to these systems have expanded significantly in recent years, such planning problems have become increasingly nonlinear and stochastic, with the underlying uncertainties often non-Gaussian. Consequently, data-driven and learning-based approaches have been widely pursued as promising alternatives to classical formulations. As the gap between their empirical performance and theoretical guarantees continues to widen, we take a step back and establish a statistically rigorous, closed-loop perspective to address this challenge.

Among classical model-based approaches, chance-constrained programming [1] provides a natural framework for handling system uncertainties in motion planning. It operates through relaxing hard constraints into probabilistic counterparts that hold with high confidence, yielding less conservative solutions. Since nominal chance constraints are generally nonconvex and often computationally intractable, they may be reformulated into tractable deterministic problems by introducing distributional assumptions or

This work benefited from technical discussions within DARPA's Safe and Assured Foundation Robots for Open eNvironments (SAFRON) Program, under contract number HR0011-25-3-0331. The authors are with the Department of Aerospace Engineering, The Grainger College of Engineering, University of Illinois Urbana-Champaign, Urbana, Illinois 61801, rihanad2@illinois.edu, hiroyasu@illinois.edu.

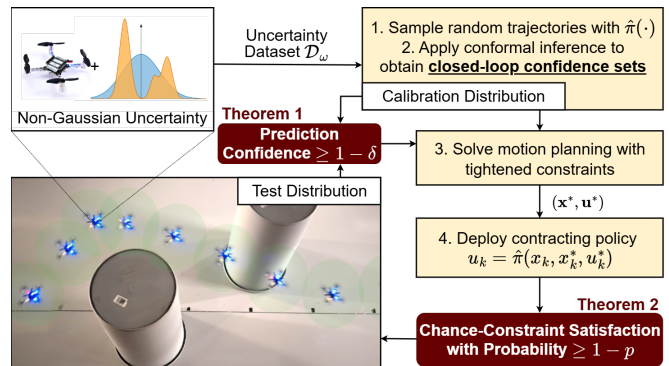


Fig. 1: Illustration of our proposed approach. Given a finite dataset  $\mathcal{D}_w$  of non-Gaussian disturbance samples, from predictions of the control contraction metrics  $\hat{M}$ , tracking policy  $\hat{\pi}(\cdot)$  for contraction rate  $\lambda$ , we infer high probability chance constraint satisfaction guarantees for safe motion planning. Code : <https://github.com/Rihan24/SCC-TrajOpt>.

approximations of underlying uncertainty. For example, existing methods consider additive Gaussian noise [2], [3], which allows propagation of the uncertainty through linear dynamics in closed form, leading to deterministic relaxations of chance constraints.

Alternatively, distributionally robust approaches seek constraint satisfaction across a family of probability distributions belonging to an ambiguity set. Such a set is commonly defined via moment specifications [4] or via distributional metrics [5], [6]. However, extending these ideas beyond linear settings becomes challenging, since uncertainty or ambiguity sets cannot, in general, be propagated analytically through nonlinear dynamics. Several ideas have been proposed to address this issue, e.g., through linearization [7], [8], polynomial chaos expansions [9], or learning-based methods [10], [11]. These methods propagate statistical information through approximated dynamics, albeit at the cost of strict probabilistic guarantees.

Sampling-based approaches address this challenge from a different angle, pursuing formal a priori guarantees achievable purely through data. The scenario approach bypasses explicit distributional modeling by enforcing constraints directly on a finite set of sampled realizations or scenarios [12], [13]. It enables statistically valid inference of probabilistic certificates on constraint satisfaction. Sample average approximation methods aim to, instead, approximate chance constraints using the empirical distribution over the samples [14]. Another line of work is conformal inference, also known as conformal prediction (CP) [15]–[17], which has recently gained attention in control theory as a powerful, yet relatively light-weight tool for distribution-free uncertainty quantification associated with point or time-series

predictions. Given a predictor and a negatively oriented score function to measure the predictor's performance, it exploits the rank statistics of the score to construct a high-confidence bound on the prediction error. CP and its variants have been employed in chance-constrained optimization [18] and a variety of planning and control settings [19]. This includes, but is not limited to, safe planning in unknown dynamic environments [20], [21], deterministic robustification of model-based controllers [22]–[24], uncertainty-aware planning with diffusion models [25], data-driven state estimation [26], prediction region construction for linear systems with non-Gaussian noise [27], conformally robust Lyapunov and barrier function synthesis [28], and deterministic contraction-based planning and control [29].

However, streamlined methods that exploit CP to alleviate distribution mismatches in closed-loop trajectory planning of nonlinear, non-Gaussian stochastic systems remain underexplored. While there are a few related efforts in contraction theory, they often operate under restrictive assumptions. For example, extreme value theory-informed collision checking [30] provides statistical safety with a contracting controller but only under bounded perturbations. Stochastic contraction-based predictive control [31] provides probabilistic reachability but requires an upper bound on the norm of the noise covariance.

This work investigates how the distribution-free concepts in CP can be integrated with the strong structural guarantees of contraction-based robust motion planning [32]–[37]. We specifically consider discrete-time nonlinear stochastic systems, where the uncertainty may be non-Gaussian. We leverage a CP-based constraint tightening method to define, for each uncertainty sample, a joint nonconformity score that quantifies both the predicted distributional mismatch and the incremental stability violations of a learned contraction metric and its associated controller. We show that this approach offers a statistically rigorous characterization of how far the system's true trajectory may deviate from the reference trajectory only with finite data samples. This thereby offers a non-diverging, closed-loop guarantee of generating provably safe and dynamically feasible motion plans. Here, chance constraints are reformulated into their deterministic counterparts, agnostic to how the distribution is predicted or how the metric is constructed, with no strong structural priors. Unlike the scenario-based approach, the problem size does not grow with the number of data samples. We validate the performance of our approach through numerical simulations with the Dubins Car and hardware experiments with the Crazyflie drone, both under non-Gaussian uncertainty. The results indicate that it indeed achieves the closed-loop satisfaction of constraints at the desired probability level.

**Notation :** The set  $\mathbb{I}_{\{a,b\}}$  denotes the collection of integers in  $\{a, b\}$  and  $\mathbb{I}_{\geq 0}$  denotes the set of non-negative integers. We will use bold face letters  $\mathbf{x} = \{x_0, x_1, \dots, x_N\}$ ,  $\mathbf{u} = \{u_0, u_1, \dots, u_{N-1}\}$  and  $\mathbf{w} = \{w_0, w_1, \dots, w_{N-1}\}$  to denote the state, control and disturbance trajectories over a time horizon  $N \in \mathbb{N}$ . Quantile $_{1-\alpha}(\mathcal{P})$  denotes the  $(1 - \alpha)$ th quantile of the distribution  $\mathcal{P}$ . For  $A, B \in \mathbb{R}^{n \times n}$ , we use

$A \succeq B$  to denote positive semi-definiteness of  $A - B$ . For  $x \in \mathbb{R}^n$ , we let  $\|x\|$  denote the Euclidean norm and  $\|x\|_M := \sqrt{x^\top M x}$  to denote the weighted norm with respect to a positive definite matrix  $M$ . For a matrix  $A \in \mathbb{R}^{l \times n}$ ,  $A_i$  denotes the  $i$ th row of  $A$  and for a vector  $b \in \mathbb{R}^n$ ,  $b_i$  denotes the  $i$ th element of  $b$ . For two sets  $A$  and  $B$ ,  $A \oplus B = \{a + b \mid a \in A, b \in B\}$  denotes the Minkowski set addition over the sets. The smallest eigen value of a symmetric matrix  $M$  is denoted by  $\lambda_{\min}(M)$ . We use the bar accent ( $\bar{\mathbf{x}}, \bar{\mathbf{u}}$ ) to denote nominal dynamics and hat notation  $\hat{a}$  to denote learned components.

## II. PROBLEM SETUP & PRELIMINARIES

We consider a class of discrete-time nonlinear systems with additive stochastic process noise, given by,

$$x_{k+1} = f(x_k, u_k) + D(x_k)w_k \quad (1)$$

where  $x_k \in \mathbb{R}^{n_x}$ ,  $u_k \in \mathbb{R}^{n_u}$  and  $w_k \in \mathbb{R}^{n_w}$  are the state, controls and process noise at time  $k \in \mathbb{I}_{\geq 0}$ , respectively. Here  $f$  is a known smooth function and  $D : \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_x \times n_w}$  is a known state-dependent matrix function.

*Assumption 1:* The process noise  $\mathbf{w}$  consists of i.i.d samples from an unknown zero mean probability distribution  $\mathcal{Q}$ , that is,  $w_k \sim \mathcal{Q}$  and  $\mathbb{E}[w_k] = 0$  for all  $k \in \mathbb{I}_{\geq 0}$ . A disturbance dataset is available, given by  $\mathcal{D}_w = \{\mathbf{w}^{(0)}, \mathbf{w}^{(1)}, \dots, \mathbf{w}^{(K-1)}\}$  for  $K \in \mathbb{N}$ , where each  $\mathbf{w}^{(j)}$  represents samples of process noise over the time horizon  $N$ , with  $w_k^{(j)} \sim \mathcal{Q}$  for all  $k \in \mathbb{I}_{[0, N-1]}$ . The support of distribution  $\mathcal{Q}$  can be unbounded.

Given an initial condition  $x_0 = \mathbf{x}(0)$ , we aim to drive the system (1) to a target region  $\mathcal{X}_N \subset \mathbb{R}^{n_x}$  over a time horizon  $N \in \mathbb{N}$ , while satisfying state constraints  $\mathcal{X} \subset \mathbb{R}^{n_x}$  and minimizing given step cost  $c(\cdot)$  and final cost  $c_F(\cdot)$ . This problem can be formulated as a chance-constrained optimal control problem, given by,

*Problem 1:*

$$\begin{aligned} \min_{\mathbf{x}, \mathbf{u}} \quad & \mathbb{E} \left[ c_F(x_N) + \sum_{k=0}^{N-1} c(x_k, u_k) \right] \\ \text{s.t.} \quad & x_{k+1} = f(x_k, u_k) + D(x_k)w_k \quad k \in \mathbb{I}_{[0, N-1]} \\ & \Pr(x_k \in \mathcal{X}) \geq 1 - p \quad k \in \mathbb{I}_{[1, N-1]} \quad (2a) \\ & \Pr(x_N \in \mathcal{X}_N) \geq 1 - p \quad (2b) \\ & x_0 = \mathbf{x}(0) \end{aligned}$$

where  $p \in (0, 1)$  denotes the maximum allowed probability of constraint violation. Problem 1 is generally intractable due to nonconvex distributional constraints on the uncertain states. A tractable deterministic reformulation of the problem becomes challenging since the distribution  $\mathcal{Q}$  is not explicitly known, on top of the difficulties in propagating this uncertainty through the nonlinear dynamics. To tackle this, we will first restrict the controls to time-parametrized state feedback policy  $u_k = \pi_k(x_k)$  and then consider chance constraint satisfaction for the closed-loop dynamics,

$$x_{k+1} = \varphi_\pi(x_k, k) + D(x_k)w_k \quad (3)$$

where  $\varphi_\pi(x, k) := f(x, \pi_k(x))$  denotes the vector field of the nominal closed-loop dynamics.

Our approach makes use of the disturbance dataset  $\mathcal{D}_w$  and is based on contraction theory and finite sample guarantees associated with statistical inference. It designs a suitable policy  $\pi_k(\cdot)$  which is a feasible solution to Problem 1; that is, the closed-loop state  $x_k$  in (3) satisfies constraints  $x_k \in \mathcal{X}$  and  $x_N \in \mathcal{X}_N$  with a confidence of at least  $(1-p)$ . We now briefly introduce the necessary preliminaries.

**Closed-loop Contraction:** Contraction theory offers a set of tools to analyze and control the incremental stability of nonlinear system trajectories with respect to each other [32].

Consider the nominal closed-loop dynamics,  $x_{k+1} = \varphi_\pi(x_k, k)$ . Let  $F_k := \frac{\partial \varphi_\pi(x, k)}{\partial x}$  denote the Jacobian matrix of  $\varphi_\pi$ , and  $\Gamma(a_k, b_k)$  denote the set of piecewise-smooth curves  $c: s \in [0, 1] \rightarrow \mathbb{R}^{n_x}$  connecting  $a_k$  to  $b_k$  for some  $a_k, b_k \in \mathbb{R}^{n_x}$ , with  $c(0) = a_k$  and  $c(1) = b_k$ . For a square matrix function  $\Theta(x, k) \in \mathbb{R}^{n \times n}$  and uniformly positive definite matrix function  $M(x, k) := \Theta(x, k)^\top \Theta(x, k) \in \mathbb{R}^{n \times n}$ , the Riemann energy for defining the distance between  $a_k$  and  $b_k$  is given by [33],

$$E(a_k, b_k) := \min_{c \in \Gamma(a_k, b_k)} \int_0^1 \frac{\partial c(s)}{\partial s}^\top M(c(s), k) \frac{\partial c(s)}{\partial s} ds. \quad (4)$$

Lemma 1 establishes conditions for uniform incremental stability for the nominal closed-loop system [32], [38].

*Lemma 1:* Consider a uniformly bounded matrix function  $M_k := M(x, k)$ , i.e.,  $\underline{m}I \preceq M_k \preceq \bar{m}I$  for some  $\bar{m} \geq \underline{m} > 0$ . If  $F_k^\top M_{k+1} F_k \preceq \lambda M_k$  with  $\lambda \in [0, 1)$  holds for all  $x \in \mathbb{R}^{n_x}$  and  $k \in \mathbb{I}_{\geq 0}$ , then the energy  $E(x_{1,k}, x_{2,k})$  decreases exponentially over the flow of the nominal closed-loop system and satisfies the decay condition  $E(\varphi_\pi(x_{1,k}, k), \varphi_\pi(x_{2,k}, k)) \leq \lambda E(x_{1,k}, x_{2,k})$ , where  $(x_{1,k}, x_{2,k})$  denotes any pair of the solution trajectories. The nominal closed-loop system is then said to be globally contracting with rate  $\lambda$ .

**Challenges in Metric Synthesis:** Control Contraction Metric (CCM) [33] establishes conditions for the existence of a feedback tracking controller such that the closed-loop system is contracting with a given rate. The search for a suitable CCM and corresponding differential feedback gain requires solving infinite-dimensional linear matrix inequalities, which are often solved using finite-dimensional approximations with restrictive assumptions, e.g., sum-of-squares relaxation [33], [39]. A tracking controller is then constructed by computing the geodesics with respect to the calculated CCM. Existing methods propose neural network parameterizations of a contraction metric and its tracking controller [40], [41]. However, the quality of such learned components depends on their structural priors and training, and thus remains primarily empirical. Verifying the validity, in terms of satisfying conditions for incremental stability, becomes crucial for their trustworthy usage.

**Conformal Prediction (CP):** CP is a statistical tool for uncertainty quantification that leverages rank statistics of *exchangeable* random variables. It uses nonconformity scores to construct prediction intervals that are valid up to the user-defined probability level while being distribution agnostic [19], [42]. CP is also model-agnostic in na-

ture and holds out when the nonconformity scores result from a function composition, e.g., a neural network. Let  $S^{(1)}, \dots, S^{(K)}, S^{(K+1)}$  be a collection of  $(K+1)$  nonconformity scores, where  $\mathcal{S} = \{S^{(1)}, \dots, S^{(K)}\}$  are scores drawn from the calibration distribution and  $S^{(K+1)}$  corresponds to the test sample. In robotic applications, the assumption of exchangeability can be too restrictive for inference on the test sample, due to possible distribution shifts with respect to the calibration dataset. This paper thus employs a CP variant called Weighted Conformal Prediction (W-CP) [17], which relaxes exchangeability by reweighting calibration samples to better match the test distribution. Let  $\{\tilde{w}_j\}_{j=1}^K$  be a set of positive weights with  $\tilde{w}_j \in [0, 1]$  for all  $j \in \mathbb{I}_{[1, K]}$ . Define the weighted quantile as,

$$q_{1-\alpha}(\mathcal{S}) := \text{Quantile}_{1-\alpha} \left( \sum_{i=1}^K \tilde{w}_i \delta_{S^{(i)}} + \tilde{w}_\infty \delta_\infty \right)$$

where  $\tilde{w}_j = \tilde{w}_j / (\sum_{i=1}^K \tilde{w}_i + 1)$ ,  $\tilde{w}_\infty = 1 / (\sum_{i=1}^K \tilde{w}_i + 1)$  and  $\delta_a$  denotes the point mass probability at  $a \in \mathbb{R} \cup \{-\infty, +\infty\}$ .

*Lemma 2:* [17] Let  $d_{\text{TV}}(\mathcal{P}_1, \mathcal{P}_2)$  be the total variation distance between distributions  $\mathcal{P}_1$  and  $\mathcal{P}_2$ . For any  $\alpha \in (0, 1)$ ,

$$\Pr(S^{(K+1)} \leq q_{1-\alpha}(\mathcal{S})) \geq 1 - \bar{\alpha}$$

where  $\bar{\alpha} = \alpha + \sum_{i=1}^K \tilde{w}_i d_{\text{TV}}(\bar{\mathcal{S}}, \bar{\mathcal{S}}_i)$ . Here  $\bar{\mathcal{S}} = \mathcal{S} \cup S^{(K+1)}$  and  $\bar{\mathcal{S}}_i$  is the set  $\bar{\mathcal{S}}$  with scores  $S^{(i)}$  and  $S^{(K+1)}$  swapped for each  $i = 1, \dots, K$ , and  $\Pr$  denotes the marginal probability measure over randomness in draws of the scores in  $\mathcal{S}$ .

*Remark 1:* The vanilla CP guarantee from [42] is obtained for exchangeable scores by setting  $\tilde{w}_1 = \dots = \tilde{w}_K = 1$ . Assuming  $\tilde{w}_1 S^{(1)} \leq \dots \leq \tilde{w}_K S^{(K)}$ , if  $K \geq \lceil (1-\alpha)(K+1) \rceil$ , one can pick  $q_{1-\alpha}(\mathcal{S}) = \tilde{w}_{j_\alpha} S^{(j_\alpha)}$  where  $j_\alpha = \lceil (1-\alpha)(K+1) \rceil$ . The weights are chosen so as to minimize the coverage gap  $\sum_{i=1}^K \tilde{w}_i d_{\text{TV}}(\mathcal{S}, \mathcal{S}_i)$  with higher weights assigned to data points that resemble the test distribution more closely [17], [22].

### III. UNCERTAINTY QUANTIFICATION & CONTRACTION

We aim to construct the following confidence sets for the closed-loop system (3) under a suitable policy  $\pi(\cdot)$ , with finite samples of the noise distribution  $\mathcal{D}_w$ .

*Definition 1:* Let  $x \in \mathbb{R}^{n_x}$  be a random vector and  $M$  be a positive definite matrix. For  $\mu \in \mathbb{R}^{n_x}$ ,  $\mathcal{B}^\theta(\mu, M) := \{\tilde{x} \mid \|\tilde{x} - \mu\|_{M^{-1}} \leq 1\}$  is an ellipsoidal confidence set of probability level  $\theta \in (0, 1)$  for  $x$  if  $\Pr(x \in \mathcal{B}^\theta(\mu, M)) \geq \theta$ .

We call the state and control trajectory pair  $(\bar{x}, \bar{u})$  as a target trajectory if they are forward complete solutions of the nominal dynamics  $\bar{x}_{k+1} = f(\bar{x}_k, \bar{u}_k)$  in (1). Let us consider neural network predictions for a state-independent CCM  $\hat{M}_k$  and the tracking policy  $\hat{\pi}(\cdot, \bar{x}, \bar{u})$ , trained to make the nominal closed-loop system controlled by  $u_k = \hat{\pi}(x_k, \bar{x}_k, \bar{u}_k)$  contracting with rate  $\lambda \in [0, 1)$  under the metric defined through  $\hat{M}_k$ . We will consider the target tracking feedback policy  $u_k = \pi_k(x_k) = \hat{\pi}(x_k, \bar{x}_k, \bar{u}_k)$ , thereby reformulating Problem 1 into a deterministic search over the space of target trajectories  $(\bar{x}, \bar{u})$ . Here, we aim to quantify the epistemic uncertainty associated with these

learned components and their contribution to the uncertainty of the closed-loop dynamics, along with the effect of stochastic noise during deployment. We will perform a calibration step to evaluate the controller  $\hat{\pi}$  over random tracking scenarios in the presence of noise characterized by the disturbance dataset  $\mathcal{D}_w$ . This will be used to show that for a suitable choice of nonconformity score, we can obtain meaningful predictive inference on the confidence set of the closed-loop state in (1) over the time horizon via Lemma 2. First, we make the following assumption,

*Assumption 2:*  $\hat{\pi}(x, x, u) = u \quad \forall x \in \mathbb{R}^{n_x}$  and  $u \in \mathbb{R}^{n_u}$ . This assumption ensures that the target trajectory  $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$  is a valid solution to the nominal closed loop system. Let  $\hat{M}_k = \Theta_k^\top \Theta_k$  and define,

$$\Delta_V(x_k, \bar{x}_k, \bar{u}_k) := \max \left\{ 0, -\lambda \left\| \hat{\Theta}_k(x_k - \bar{x}_k) \right\| + \left\| \hat{\Theta}_{k+1}(\varphi_{\hat{\pi}}(x_k) - \varphi_{\hat{\pi}}(\bar{x}_k)) \right\| \right\} \quad (5)$$

where  $\varphi_{\hat{\pi}}(x_k) := f(x_k, \hat{\pi}(x_k, \bar{x}_k, \bar{u}_k))$ . Note that under a state-independent and time indexed metric defined through  $\hat{M}_k$ , the geodesic is equivalent to a straight line connecting the two points and the Riemann energy  $E(x_k, \bar{x}_k)$  in (4) reduces to to the weighted norm  $\|x_k - \bar{x}_k\|_{\hat{M}_k}$ . Hence, given  $(\bar{x}_k, \bar{u}_k)$ , equation (5) denotes the residuals of the decay condition in (1) at time  $k$ . Let  $\Omega$  denote the distribution of  $N$  length state and control target trajectories with initial state  $\bar{x}_0 = \mathbf{x}(0)$ . By taking a random sample  $(\bar{\mathbf{x}}^{(j)}, \bar{\mathbf{u}}^{(j)}) \sim \Omega$  for every  $j$ -th sample  $\mathbf{w}^{(j)}$  from dataset  $\mathcal{D}_w$ , one can generate  $K$  distinct closed-loop realizations  $\mathbf{x}^{(j)}$  of the system (1) under the feedback policy  $u_k = \hat{\pi}(x_k, \bar{x}_k^{(j)}, \bar{u}_k^{(j)})$ , i.e.,

$$x_{k+1}^{(j)} = f(x_k^{(j)}, \hat{\pi}(x_k^{(j)}, \bar{x}_k^{(j)}, \bar{u}_k^{(j)})) + D(x_k^{(j)})w_k^{(j)}$$

with initial state  $x_0^{(j)} = \mathbf{x}(0)$  for all  $j \in \mathbb{I}_{[1, K]}$ . Now for each  $j \in \mathbb{I}_{[1, K]}$  and  $k \in \mathbb{I}_{[1, N]}$ , define the nonconformity score as follows,

$$S_k^{(j)} := \sum_{i=0}^{k-1} \lambda^i \left( \Delta_V(x_i^{(j)}, \bar{x}_i^{(j)}, \bar{u}_i^{(j)}) + \left\| \hat{\Theta}_{i+1} D(x_i^{(j)}) w_i^{(j)} \right\| \right) \quad (6)$$

and collect the scores  $\mathcal{S}_k := \{S_k^{(j)}\}_{j=1}^K$ . This will be our calibration dataset at time  $k$ . Note that the nonconformity score defined in equation (6) maps a target trajectory  $(\bar{\mathbf{x}}^{(j)}, \bar{\mathbf{u}}^{(j)})$  and noise realization  $\mathbf{w}^{(j)}$  to a positive real number. Hence, the calibration distribution can be identified with  $(\Omega, \mathcal{Q}^N)$ . Let  $u_k = \hat{\pi}(x_k, \bar{x}_k^{(K+1)}, \bar{u}_k^{(K+1)})$  denote the control policy during deployment. The corresponding target trajectory  $(\bar{\mathbf{x}}^{(K+1)}, \bar{\mathbf{u}}^{(K+1)}) \sim \Xi$  is obtained by solving Problem 1, as to be described in Section IV, and not known apriori. Here,  $\Xi$  denotes the distribution of target trajectories that are feasible solutions to the proposed optimization problem. We propose the following result for the test score  $S_k^{(K+1)}$  obtained from the test distribution  $(\Xi, \mathcal{Q}^N)$ .

*Theorem 1:* Suppose that the Assumptions 1 and 2 hold. Consider a sample target trajectory  $(\bar{\mathbf{x}}, \bar{\mathbf{u}}) \sim \Xi$ . Let  $\bar{w}_j$  denote the normalized weights for  $j \in \mathbb{I}_{[1, K]}$ . Suppose also that the coverage gap  $\sum_{i=1}^{K+1} \bar{w}_i d_{\text{TV}}(\bar{\mathcal{S}}_k, \bar{\mathcal{S}}_{k,i})$  is upper

bounded by  $\bar{\delta}$ , where  $0 \leq \bar{\delta} < \delta$  for some  $\delta \in (0, 1)$  and for all  $k \in \mathbb{I}_{[1, N]}$ . Then  $\mathcal{B}^{1-\delta}(\bar{x}_k, W_k)$  is a confidence set with probability level  $(1 - \delta)$  for the state  $x_k$  in (1) at time  $k \in \mathbb{I}_{[1, N]}$  under the feedback policy  $u_k = \hat{\pi}(x_k, \bar{x}_k, \bar{u}_k)$ , where  $W_k := C_k^2 \hat{M}_k^{-1}$  and  $C_k := q_{1-\delta+\bar{\delta}}(\mathcal{S}_k)$ .

*Proof:* Define  $V_k := V(x_k, \bar{x}_k) = \|\hat{\Theta}_k(x_k - \bar{x}_k)\|$ . Then we have,

$$\begin{aligned} V_{k+1} &= \left\| \hat{\Theta}_{k+1}(x_{k+1} - \bar{x}_{k+1}) \right\| \\ &\leq \left\| \hat{\Theta}_{k+1}(\varphi_{\pi}(x_k) - \bar{x}_{k+1}) \right\| + \left\| \hat{\Theta}_{k+1} D(x_k) w_k \right\| \\ &\stackrel{(5)}{\leq} \lambda V_k + \Delta_V(x_k, \bar{x}_k, \bar{u}_k) + \left\| \hat{\Theta}_{k+1} D(x_k) w_k \right\| \\ &\leq \lambda^{k+1} V_0 + \sum_{i=0}^k \left( \Delta_V(x_i, \bar{x}_i, \bar{u}_i) + \left\| \hat{\Theta}_{i+1} D(x_i) w_i \right\| \right) \lambda^i \\ &\stackrel{(6)}{=} S_{k+1}^{(K+1)}. \end{aligned}$$

Note that  $V_k = \|x_k - \bar{x}_k\|_{\hat{M}_k}$  and  $V_0 = 0$  by construction. Evaluate the quantile  $C_k = q_{1-\delta+\bar{\delta}}(\mathcal{S}_k)$  (see Remark 1) for each  $k \in \mathbb{I}_{[1, N]}$ . Applying Lemma 2 and taking the complement, we obtain,

$$\Pr(V_k > C_k) \leq \delta - \bar{\delta} + \sum_{i=1}^K \bar{w}_i d_{\text{TV}}(\bar{\mathcal{S}}_k, \bar{\mathcal{S}}_{k,i}) \leq \delta.$$

Taking a complement again gives the desired claim.  $\blacksquare$

*Remark 2:* Note that  $W_k = C_k^2 \hat{M}_k^{-1}$  dictates the size and shape of the prediction set of  $x_k$ , analogous to that of covariance in Gaussian-based [43] methods. Our approach, as in existing contraction-based methods (§4.2, [39]) allows us to control the associated conservativeness up to a certain degree by an appropriate design choice of contraction metric bounds  $\underline{m}, \bar{m}$  and the contraction rate  $\lambda$  in equation (5).

*Remark 3:* Lemma 2 provides a way to handle the loss of exchangeability caused by distribution shift between simulated rollouts and real-world deployment. To generalize the setup in Lemma 2,  $D_w$  may be constructed using samples from an approximate noise distribution  $\hat{\mathcal{Q}}$  instead of the actual noise distribution  $\mathcal{Q}$  (see Section V-B). In such cases, the application of Lemma 2 and subsequently Theorem 1 will assume the gap between the calibration distribution  $(\Omega, \hat{\mathcal{Q}}^N)$  and the unknown test distribution  $(\Xi, \mathcal{Q}^N)$  can be sufficiently controlled by the weights. We could also update  $\delta$  of Lemma 2 using adaptive CP [44].

Although we do not consider closed-loop control constraints in our approach, we can obtain such guarantees by imposing some structure on the learned controller  $\hat{\pi}$ , as we show in the following result.

*Corollary 1:* Let  $\mathcal{B}^{1-\delta}(\bar{x}_k, W_k)$  be a confidence set for  $x_k$  at probability level  $(1 - \delta)$ . Suppose  $u_k = \hat{\pi}(x_k, \bar{x}_k, \bar{u}_k) = k(x_k, \bar{x}_k) + \bar{u}_k$  and Assumption 2 holds, i.e.,  $k(x, x) = 0$  for all  $x \in \mathbb{R}^{n_x}$ . If  $k(\cdot)$  is Lipschitz with Lipschitz constant  $L$ , then  $\mathcal{B}^{1-\delta}(\bar{u}_k, Z_k)$  is a confidence set for  $u_k$  at probability level  $(1 - \delta)$ , where  $Z_k = \frac{L^2}{\lambda_{\min}(W_k^{-1})} I$

*Proof:* We have the following,

$$\begin{aligned} \|u_k - \bar{u}_k\| &= \|k(x_k, \bar{x}_k)\| = \|k(x_k, \bar{x}_k) - k(\bar{x}_k, \bar{x}_k)\| \\ &\leq L\|x_k - \bar{x}_k\| \leq L \frac{\|x_k - \bar{x}_k\|_{W_k^{-1}}}{\sqrt{\lambda_{\min}(W_k^{-1})}}. \end{aligned}$$

From Definition 1, we get  $\|x_k - \bar{x}_k\|_{W_k^{-1}} \leq 1$  with probability at least  $(1 - \delta)$ , which implies,  $\Pr(\|u_k - \bar{u}_k\| \leq \frac{L}{\sqrt{\lambda_{\min}(W_k^{-1})}}) \geq 1 - \delta$ . Then the result follows also from Definition 1. ■

#### IV. DETERMINISTIC REFORMULATION

We will use confidence sets constructed using Theorem 1 to reformulate chance-constraints in (2a), (2b) into deterministic constraints on the target trajectory  $(\bar{x}, \bar{u})$  via suitable constraint tightening. Essentially, we aim to find a tightened state constraint set  $\mathcal{X}_k^c$  such that  $\mathcal{X}_k^c \oplus \mathcal{B}^{1-\delta}(\bar{x}_k, W_k) \subseteq \mathcal{X}$  for all  $k \in \mathbb{I}_{[0, N]}$ , whereby restricting  $\bar{x}_k \in \mathcal{X}_k^c$  ensures  $x_k \in \mathcal{X}$  with probability at least  $(1 - \delta)$ . We will use the approach from [7] to construct suitable convex over approximations of the set  $\mathcal{X}_k^c$  for particular classes of sets  $\mathcal{X}$ .

Let  $\mathcal{X}$  be defined by polytopic and non-convex obstacle avoidance constraints on the state,  $\mathcal{X} = \mathcal{X}_{\text{poly}} \cap \mathcal{X}_{\text{free}}$ , where  $\mathcal{X}_{\text{poly}} = \{x \in \mathbb{R}^{n_x} \mid Ax \leq b\}$ ,  $A \in \mathbb{R}^{l \times n_x}$ ,  $b \in \mathbb{R}^l$ , and  $\mathcal{X}_{\text{free}} = \mathbb{R}^{n_x} \setminus \cup_{i=1}^J \mathcal{O}_i$ . Here  $\mathcal{O}_i \subset \mathbb{R}^{n_x}$ ,  $i \in \mathbb{I}_{[1, J]}$  are closed convex sets representing  $J$  obstacles. First, we split the joint chance constraints  $\Pr(x_k \in \mathcal{X}_{\text{poly}} \cap \mathcal{X}_{\text{free}}) \geq 1 - p$  into single chance-constraints using the Bonferroni inequality [9] with equal risk allocation accordingly,

$$\Pr(x_k \in \mathcal{X}_{\text{poly}}) \geq 1 - \bar{p} \quad (7)$$

$$\Pr(x_k \in \mathcal{X}_{\text{free}}) \geq 1 - \bar{p} \quad (8)$$

where  $\bar{p} = p/2$ . The above result is obtained by applying Boole's inequality over the complement of the event  $x_k \in \mathcal{X}_{\text{poly}} \cap \mathcal{X}_{\text{free}}$ . Let  $\mathcal{B}^{1-\bar{p}}(\hat{x}_k, W_k)$  be a confidence set for  $x_k$  with probability level  $(1 - \bar{p})$ . Then the polytopic chance constraint may be conservatively reformulated as  $l$  deterministic constraints [7],

$$(7) \Leftrightarrow A_i \bar{x}_k + \sqrt{A_i W_k A_i^\top} \leq b_i \quad i \in \mathbb{I}_{[1, l]}. \quad (9)$$

Here the inequality condition in (9) is the robust counterpart to the condition  $A_i x_k \leq b_i$  over the confidence set  $\mathcal{B}^{1-\bar{p}}(\hat{x}_k, W_k)$  [45]. For the obstacle avoidance constraints, we will use the signed distance function  $d_i : \mathcal{X} \rightarrow \mathbb{R}$  that returns the shortest distance from point  $x_k$  to the boundary  $\partial \mathcal{O}_i$  of the set  $\mathcal{O}_i$  [7]. It is defined as,

$$d_i(x_k) := \inf_{y \in \mathcal{O}_i} \|x_k - y\| - \inf_{z \notin \mathcal{O}_i} \|x_k - z\|$$

where  $d_i(x_k) \geq 0 \implies x_k \notin \mathcal{O}_i$ . Define,

$$n_i(x) = \frac{x - \check{x}_i}{d_i(x)}, \quad \check{x}_i = \arg \min_{y \in \partial \mathcal{O}_i} \|x - y\|.$$

Then, we can conservatively reformulate obstacle avoidance chance constraint (8) into  $J$  deterministic constraints as,

$$(8) \Leftrightarrow d_i(\bar{x}_k) - \sqrt{n_i^\top W_k n_i} \geq 0, \quad i \in \mathbb{I}_{[1, J]}. \quad (10)$$

For more details, refer [7]. Let  $\mathcal{X}_N = \{x \mid Hx \leq h, H \in \mathbb{R}^{r \times n_x}, h \in \mathbb{R}^r\}$ . The deterministic reformulation of (2b) then follows the same approach as in (9) with a confidence set of probability level  $(1 - p)$ , which here gives  $r$  deterministic constraints.

Now consider the following relaxed deterministic reformulation of Problem 1 for some  $\delta \in (0, 1)$ .

*Problem 2:*

$$\min_{\bar{x}, \bar{u}} c_F(\bar{x}_N) + \sum_{k=0}^{N-1} c(\bar{x}_k, \bar{u}_k)$$

$$\text{s.t. } \bar{x}_{k+1} = f(\bar{x}_k, \bar{u}_k), \quad k \in \mathbb{I}_{[0, N-1]}$$

$$A_i \bar{x}_k + \bar{\eta}_k(\delta) \|A_i\|_{M_k^{-1}} \leq b_i, \quad i \in \mathbb{I}_{[1, l]}, k \in \mathbb{I}_{[1, N-1]}$$

$$d_i(\bar{x}_k) - \bar{\eta}_k(\delta) \|n_i\|_{M_k^{-1}} \geq 0, \quad i \in \mathbb{I}_{[1, J]}, k \in \mathbb{I}_{[1, N-1]}$$

$$H_i \bar{x}_N + \eta(\delta) \|H_i\|_{M_N^{-1}} \leq h_i, \quad i \in \mathbb{I}_{[1, r]}$$

$$\bar{x}_0 = \mathbf{x}(0)$$

where,  $\bar{\eta}_k(\delta) = q_{1-\frac{\delta}{2}}(\mathcal{S}_k)$  and  $\eta(\delta) = q_{1-\delta}(\mathcal{S}_N)$ . We then have the following result,

*Theorem 2:* Let  $(\mathbf{x}^*, \mathbf{u}^*)$  be the solution to Problem 2 for  $\delta \in (0, 1)$ , if it exists. Suppose that Assumptions 1 and 2 hold and let the coverage gap be upper bounded by  $\bar{\delta}$  for all  $k \in \mathbb{I}_{[0, N]}$ . Then the control policy  $u_k = \hat{\pi}(x_k, x_k^*, u_k^*)$  is a feasible solution to Problem 1 if  $\delta = p - 2\bar{\delta}$ .

*Proof:* For each  $k \in \mathbb{I}_{[1, N-1]}$ , set  $W_k = \bar{\eta}_k(\delta)^2 M_k^{-1}$  and  $W_N = \eta(\delta)^2 M_N^{-1}$ . From Theorem 1,  $\mathcal{B}^{1-\frac{\delta}{2}-\bar{\delta}}(x_k^*, W_k)$  and  $\mathcal{B}^{1-\delta-\bar{\delta}}(x_N^*, W_N)$  are confidence sets for  $x_k$ ,  $k \in \mathbb{I}_{[0, N-1]}$ , and  $x_N$  in system (1), respectively, under the policy  $u_k = \hat{\pi}(x_k, x_k^*, u_k^*)$ . Since  $\mathbf{x}^*, \mathbf{u}^*$  satisfies the tightened constraints in Problem 2, from (9) and (10), we get  $\Pr(x_k \in \mathcal{X}) \geq 1 - \delta - 2\bar{\delta} = 1 - p$  thus satisfying constraint (2a). Similarly, we obtain  $\Pr(x_N \in \mathcal{X}_N) \geq 1 - \delta - \bar{\delta} = 1 - p + \bar{\delta} \geq 1 - p$ , which satisfies (2b). ■

We note that Theorem 2 quantifies the effect of stochastic noise, the quality of predicted contraction metrics and controller, and the conservativeness demanded by the coverage gap between the calibration and test distributions, in one unified way. Under the assumption of the coverage gap bound  $\bar{\delta}$ , it provides a suitable method to ensure chance-constraint satisfaction up to the desired probability level when using the learned controller  $\hat{\pi}$ . In Section V, we validate the same through both numerical and hardware experiments. We note that, practically, it is not possible to know the coverage gap bound  $\bar{\delta}$  a priori, making our guarantees of the *a posteriori* type [18]. To address this, one may update the upper bound by means of verification or re-estimate the quantile online using adaptive CP [24].

#### V. CASE STUDIES

We present our approach with the Dubins car model in numerical simulation and with the Crazyflie quadcopter in the real world. The code for all experiments is made available at <https://github.com/Rihan24/SCC-TrajOpt>. In both experiments, we show that the confidence sets constructed using our approach from finite noise samples are valid up to the prescribed probability level, and the proposed chance

constraints are satisfied. We remark that the main goal of this section is the empirical validation of the formal guarantees established by our theorems, rather than demonstrating the superiority of our approach over all existing data-driven/learning-based motion planning and control methods. For example, even in the case of a poorly learned control policy with a suboptimally designed contraction metric, our theorems still provide a means to formally quantify their performance without relying on structural or distributional priors. For this reason, we compare against only a limited set of existing approaches in the following. These serve solely to illustrate our claims within the constraints of the page limit.

We leverage a neural network parameterized CCM  $M(x; \theta_m)$  and its controller  $\pi(x, x^*, u^*; \theta_u)$  from the implementation in [41], which satisfies Assumption 2 by design. We consider  $\hat{M}_k = M = \hat{M}(x(0); \theta_m)$  as the predicted matrices for approximately constructing contraction metrics for all  $k \in \mathbb{I}_{[1, N]}$ . The discrete-time contraction rate  $\lambda \in [0, 1)$  is obtained from the chosen continuous time counterpart  $\gamma > 0$  in [41] using  $\lambda = \sqrt{(1 - 2\gamma\Delta t)(\underline{m}/\bar{m})}$  [37], where  $\Delta t$  denotes the sampling rate for the discrete-time system. We note that the quantiles  $\bar{\eta}_k(\delta)$  quickly converge to  $\eta(\delta)$  in a few time steps in both experiments for our choices of  $\lambda$ . Hence, we conservatively set  $\bar{\eta}_k(\delta) = \eta(\delta)$  for all  $k \in \mathbb{I}_{[1, N-1]}$ , which, along with the choice of constant contraction metric, results in confidence sets of the same size over the time horizon. We consider step cost and final cost as  $c(x_k, u_k) = u_k^\top R u_k$  and  $c_F = 0$ , respectively, where  $R$  is some positive definite weight matrix. The optimization in Problem 2 is posed and solved using the CasADi interface [46] with the NLP solver IPOPT [47].

**Choice of  $\Omega$ :** For both experiments, we follow the approach from [41] to sample suitable target trajectories. We first parametrize the target control trajectory  $\bar{u}$  as a linear combination of fixed sinusoidal basis functions with a predefined set of frequencies. We then sample i.i.d. reference control trajectories by randomly sampling the weight for each frequency component from a normal distribution. The initial state  $\bar{x}_0$  is set to  $x(0)$ . We then generate the corresponding target state trajectory  $\bar{x}$  by propagating  $\bar{x}_0$  through the nominal dynamics  $\bar{x}_{k+1} = f(\bar{x}_k, \bar{u}_k)$ .

In both simulation and hardware experiments, the distribution shift between the calibrated and test distributions was minimal and well-accommodated by the slight conservativeness associated with our method. Therefore, using vanilla CP with  $\delta = p$  gave us satisfactory results. For online motion planning, one may consider receding horizon alternatives of our approach and experiments.

### A. Numerical Simulations

We consider the Dubin's car model from [41] with 4 state variables  $x := [p_x, p_y, \theta, v]$  where  $p_x, p_y$  are the positions,  $\theta$  is the heading angle and  $v$  is the velocity. The control inputs are  $u := [\omega, a]$ , where  $\omega$  and  $a$  are angular velocity and linear acceleration respectively. The dynamics is given by  $\dot{x} = f(x, u) = [v \cos(\theta), v \sin(\theta), \omega, a]^\top$ . We discretize this via a zero-order hold on the controls and add noise  $w$  to obtain the

discrete-time dynamics  $x_{k+1} = x_k + \Delta t f(x_k, u_k) + D w_k$ . We will consider the following two cases,

- (i) Uniform noise  $w_k \sim U[w_{\min}, w_{\max}]$  over the hyper-box defined by  $w_{\min} = [-\sigma/5, -\sigma/5, -\sigma, -\sigma]^\top$  and  $w_{\max} = [\sigma/5, \sigma/5, \sigma, \sigma]^\top$ . We set  $\sigma = 0.15$  and  $D = \text{diag}([1, 1, 1, 1])$ .
- (ii) Zero mean 3-component Gaussian mixture noise with density  $\sum_{i=1}^3 \psi_i \mathcal{N}(\mu_i, \Sigma_i)$  with  $\mu_1 = -0.5 \cdot \mathbf{1}_4, \mu_2 = 0 \cdot \mathbf{1}_4, \mu_3 = 0.5 \cdot \mathbf{1}_4, \Sigma_i = (0.05)^2 \cdot I$  for  $i = 1, 2, 3$ ,  $\psi_1, \psi_3 = 0.1, \psi_2 = 0.8$  and  $D = \text{diag}([0, 0, 1, 1])$ , where  $\mathbf{1}_4 = [1, 1, 1, 1]^\top$ . Here  $\mathcal{N}$  is the multivariate Gaussian density.

The noise parameters are chosen so as to cater to the problem dimensions. We consider the motion planning problem from [9], with  $x(0) = [0, 0.4, 0, 0]^\top$ , goal region loosely centered at  $x(N) = [10, 0.4, 0, 0]^\top$  and collision avoidance with a wall and obstacle given by  $\mathcal{X}_{\text{poly}} = \{p_y \mid p_y \leq 2\}$  and obstacle  $\mathcal{O} = \{p_{x,y} \mid \|p_{x,y} - p_{\text{obs}}\| \leq 1.2\}$  respectively, where  $p_{x,y} = [p_x, p_y]^\top, p_{\text{obs}} = [0, 5]^\top$ . We set  $R = 0.1 \cdot I$ , failure probability  $\delta = p = 0.1$ , sampling rate  $\Delta t = 0.05$  and time horizon to  $N = 200$  steps. We set  $\underline{m} = 0.5, \bar{m} = 10$  and  $\gamma = 1$ , which corresponds to a discrete-time contraction rate of  $\lambda = 0.21$ . The disturbance dataset  $\mathcal{D}_w$  is constructed with  $K = 20$  by sampling from the respective noise distributions. The quantile  $\eta(\delta)$  comes to 0.301 and 0.574 for cases (i) and (ii), respectively. We solve Problem 2 to get  $(x^*, u^*)$ , which takes on average 1.67 seconds, and run 200 closed-loop realizations under policy  $u_k = \hat{\pi}(x_k, x_k^*, u_k^*)$ , as shown in Figure 2, to verify satisfaction of chance constraints. We note that all chance constraints are satisfied at the failure probability level  $p = 0.1$ . In particular, the empirical failure probabilities obtained are  $\max_k \{\Pr(x_k \notin \mathcal{X})\} = 0\%$  for case (i) and  $\max_k \{\Pr(x_k \notin \mathcal{X})\} = 1.5\%$  for case (ii).

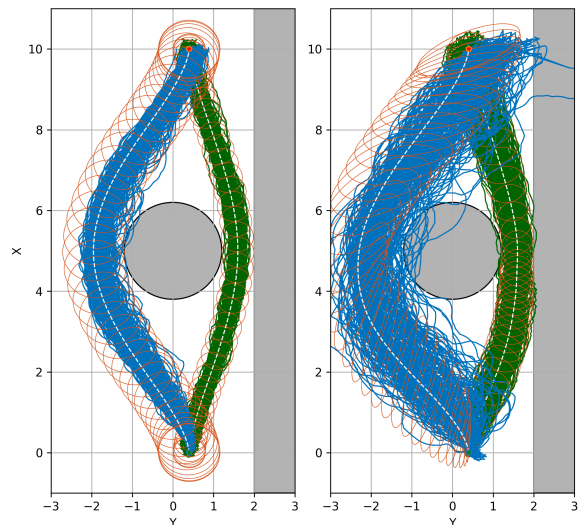


Fig. 2: Closed-loop realizations for the Dubin's car system with our approach (shown on the left) and linearization approach with Gaussian approximation of noise and LQR feedback (shown on the right). Case (i) Uniform noise and Case (ii) Gaussian mixture noise are shown in green and blue, respectively. The corresponding target trajectory  $x^*$  is shown in dashed white lines, and the respective confidence sets are shown in orange.

We compare our approach to the linearization baseline from [7], [8], shown in Figure 2, with Gaussian approx-

imation of the noise and a linear quadratic regulator (LQR) feedback policy. The Gaussian distribution is characterized by MLE estimates of the mean and covariance from samples of the respective non-Gaussian noise distributions, and the LQR feedback gain is computed using linearized closed-loop dynamics at each state. The uncertainty of the corresponding closed-loop state at time  $k$  is characterized as a Gaussian  $x_k \sim \mathcal{N}(\mu_k, \Sigma_k)$ , where  $\mu_k, \Sigma_k$  are obtained by mean and covariance propagation through the linearized dynamics. We use constraint tightening based on ellipsoidal confidence sets  $\mathcal{B}^{1-\frac{p}{2}}(\mu_k, \mathcal{X}_{n_x}^2(1-\frac{p}{2})\Sigma_k)$  for  $x_k$  [7], where  $\mathcal{X}_{n_x}^2(1-\alpha)$  denotes the  $(1-\alpha)$ -th quantile of the  $\chi^2$  distribution with  $n_x$  degrees of freedom. Solving this tightened problem takes on average 103.8 seconds. From Figure 2, we see that Gaussian-based methods perform poorly in accurately capturing the uncertainty associated with the state over the time horizon. The empirical failure probabilities obtained are  $\max_k \{\Pr(x_k \notin \mathcal{X})\} = 10\%$  for case (i) and  $\max_k \{\Pr(x_k \notin \mathcal{X})\} = 20.5\%$  for case (ii). This can be attributed to heavier tail characteristics in the noise and nonlinearity in the dynamics.

### B. Hardware Experiments

We implement our approach on the **Crazyflie 2.1** drone for safe motion planning through a static cluttered environment. We consider the quadcopter model and controller implementation from [39], [41]. Consider state  $x := [p_x, p_y, p_z, \dot{p}_x, \dot{p}_y, \dot{p}_z, f, \phi, \theta, \psi]^\top$ , where the positions  $p_x, p_y, p_z$  are expressed in the inertial world frame of the drone,  $\phi, \theta, \psi$  represent Euler angles as per the XYZ rotation sequence, both defined according to the North-East-Down frame convention. Here,  $f$  is the net mass-normalized force by the motors. Control inputs are chosen as  $u = [\dot{f}, \dot{\phi}, \dot{\theta}, \dot{\psi}]^\top$ . The system dynamics are given by  $\dot{x} = [\dot{p}_x, \dot{p}_y, \dot{p}_z, -f \sin(\theta), f \cos(\theta) \sin(\phi), g - f \cos(\theta) \cos(\phi), \dot{f}, \dot{\phi}, \dot{\theta}, \dot{\psi}]^\top$ , where  $g = 9.81$ . The contraction metric defined through  $\hat{M}$  and the corresponding tracking controller  $\hat{\pi}$  are obtained for  $\underline{m} = 0.5, \bar{m} = 25$  and  $\gamma = 0.8$ . We deploy the controller offboard on the Crazyflie at 100Hz, with **Qualisys** motion capture feeding pose estimates onto Crazyflie's internal EKF. We remark that the controller  $\hat{\pi}$  requires only a single forward pass to compute control input, and hence can even be deployed onboard the Crazyflie firmware. At each sampling instant, throttle and angle setpoints computed from  $\hat{\pi}$  are passed onto the low-level built-in PID controller onboard the Crazyflie firmware. First, thrust and angle setpoints  $[f_c, \phi_c, \theta_c]$  are obtained by integrating over the control input. The respective throttle command  $\tau_c$  is computed using  $f_c$  via an iterative proportional feedback update on the previous throttle value [39], which allows us to bypass the complex thrust to throttle mapping. The gain of this thrust controller is set to  $K_p = 400$ . The low-level controller, which is assumed to run at a higher rate, takes in throttle and angle setpoints and computes the downstream motor PWM commands.

We consider the problem of collision-free motion planning for the Crazyflie drone over a time period of 7.5 seconds

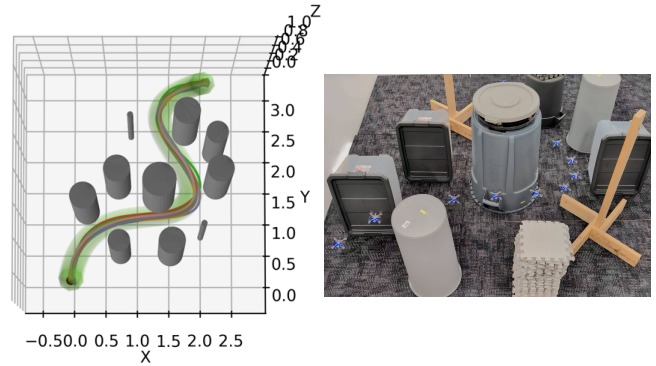


Fig. 3: Left: Closed-loop realizations of the Crazyflie drone using our approach. The 3D confidence set  $\mathcal{B}^{1-\delta}(\hat{x}_k, \eta(\delta)^2 M^{-1})$  for the positions is shown in green. Right: One iteration of Crazyflie trajectory in the cluttered environment.

in a test arena cluttered with obstacles. We set the failure probability level to  $\delta = p = 0.1$  and  $R = 0.01 \cdot I$ . The obstacles are posed as cylindrical bodies for formulating the optimization problem, as shown in Figure 3. The yaw remains uncontrolled in our experiments. To construct the disturbance dataset  $\mathcal{D}_w$ , sampling from the unknown test distribution, that is, by collecting the flight data over the time horizon for  $K \geq \lceil (1-p)(K+1) \rceil$  runs, is not practical and generalizable. Instead, we consider a Gaussian MLE estimate of the non-gaussian process noise, calculated using the formula  $w_k = x_{k+1} - f_{DT}(x_k, u_k)$  from arbitrary trial runs of the Crazyflie drone, and sample from it to generate the calibration dataset  $\mathcal{D}_w$  with  $K = 20$ . Here  $f_{DT}$  denotes the discretized dynamics similar to that in Section V-A with  $D = I$ . We confirm that the MLE estimate is approximately zero mean. Setting  $\Delta t = 0.01$  seconds,  $\lambda = 0.14$  and using approach from Sections III and IV, we get  $\eta(\delta) = 0.08$ . We solve Problem 2 by warm starting solver with a sample state trajectory in the obstacle free region, and deploy the control policy  $u_k = \hat{\pi}(x_k, x_k^*, u_k^*)$  over the time horizon. We run the tracking policy for 15 iterations, as shown in Figure 3. We observe that the drone stays within the defined confidence sets throughout the time horizon in all iterations, which implies that it achieves the prescribed chance constraint.

## VI. CONCLUSION

In this work, we proposed a novel method for chance-constrained trajectory optimization and control for nonlinear systems subject to non-Gaussian stochastic uncertainty. It quantifies distribution mismatch and system contraction via finite samples of the uncertainty with minimal distributional and structural assumptions. Using conformal prediction and contraction theory, we established formal guarantees for chance-constraint tightening, enabling deterministic planning with rigorous performance assurances. For example, our approach can ensure safety under non-Gaussian uncertainty even when learned representations of contraction metrics are employed, thereby bridging the gap in providing strict guarantees for learning-based motion planning. We validated our approach with both simulation and hardware experiments, showing that it achieves the posed chance constraints up to the desired probability level, without being too conservative.

## REFERENCES

- [1] A. Prékopa, *Stochastic Programming*, 1st ed., ser. Mathematics and Its Applications. Springer Dordrecht, 1995, vol. 324.
- [2] K. Okamoto, M. Goldshtein, and P. Tsiotras, “Optimal covariance control for stochastic systems under chance constraints,” *IEEE Control Systems Letters*, vol. 2, no. 2, pp. 266–271, 2018.
- [3] L. Blackmore, M. Ono, and B. C. Williams, “Chance-constrained optimal path planning with obstacles,” *IEEE Transactions on Robotics*, vol. 27, no. 6, pp. 1080–1094, 2011.
- [4] G. C. Calafiore and L. E. Ghaoui, “On Distributionally Robust Chance-Constrained Linear Programs,” *Journal of Optimization Theory and Applications*, vol. 130, pp. 1–22, 2006.
- [5] A. Dixit, M. Ahmadi, and J. W. Burdick, “Distributionally robust model predictive control with total variation distance,” *IEEE Control Systems Letters*, vol. 6, pp. 3325–3330, 2022.
- [6] J. Coulson, J. Lygeros, and F. Dörfler, “Distributionally robust chance constrained data-enabled predictive control,” *IEEE Transactions on Automatic Control*, vol. 67, no. 7, pp. 3289–3304, 2021.
- [7] T. Lew, R. Bonalli, and M. Pavone, “Chance-Constrained Sequential Convex Programming for Robust Trajectory Optimization,” in *IEEE European Control Conference (ECC)*, 2020.
- [8] L. Hewing, J. Kabzan, and M. N. Zeilinger, “Cautious model predictive control using gaussian process regression,” *IEEE Transactions on Control Systems Technology*, vol. 28, no. 6, p. 2736–2743, 2019.
- [9] Y. K. Nakka and S.-J. Chung, “Trajectory optimization of chance-constrained nonlinear stochastic systems for motion planning under uncertainty,” *IEEE Transactions on Robotics*, vol. 39, no. 1, pp. 203–222, 2022.
- [10] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. The MIT Press, 11 2005.
- [11] D. Sun, M. J. Khojasteh, S. Shekhar, and C. Fan, “Uncertainty-aware safe exploratory planning using Gaussian process and neural control contraction metric,” arXiv:2105.06567, 2021.
- [12] G. Calafiore and M. C. Campi, “Uncertain convex programs: Randomized solutions and confidence levels,” *Mathematical Programming*, vol. 102, p. 25–46, 2005.
- [13] G. Schildbach, L. Fagiano, C. Frei, and M. Morari, “The scenario approach for stochastic model predictive control with bounds on closed-loop constraint violations,” *Automatica*, vol. 50, no. 12, p. 3009–3018, 2014.
- [14] A. Hakobyan, G. C. Kim, and I. Yang, “Risk-aware motion planning and control using CVaR-constrained optimization,” *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 3924–3931, 2019.
- [15] V. Vovk, A. Gammernan, and C. Saunders, “Machine-learning applications of algorithmic randomness,” in *16th Int. Conf. Mach. Learn.*, 1999, p. 444–453.
- [16] V. Vovk, A. Gammernan, and G. Shafer, *Algorithmic Learning in a Random World*. Springer, 2022.
- [17] R. F. Barber, E. J. Candès, A. Ramdas, and R. J. Tibshirani, “Conformal prediction beyond exchangeability,” arXiv:2202.13415, 2023.
- [18] Y. Zhao, X. Yu, M. Sesia, J. V. Deshmukh, and L. Lindemann, “Conformal Predictive Programming for Chance Constrained Optimization,” arXiv:2402.07407, 2025.
- [19] L. Lindemann, Y. Zhao, X. Yu, G. J. Pappas, and J. V. Deshmukh, “Formal Verification and Control with Conformal Prediction,” arXiv:2409.00536, 2024.
- [20] L. Lindemann, M. Cleaveland, G. Shim, and G. J. Pappas, “Safe planning in dynamic environments using conformal prediction,” *IEEE Robot. Automat. Lett.*, vol. 8, no. 8, pp. 5116–5123, 2023.
- [21] A. Dixit, L. Lindemann, S. X. Wei, M. Cleaveland, G. J. Pappas, and J. W. Burdick, “Adaptive conformal prediction for motion planning among dynamic agents,” in *Proc. PMLR LADC*, 2023, pp. 300–314.
- [22] K. Y. Chee, T. C. Silva, M. A. Hsieh, and G. J. Pappas, “Uncertainty quantification and robustification of model-based controllers using conformal prediction,” in *Proc. PMLR LADC*, 2024, pp. 528–540.
- [23] K. Y. Chee, M. A. Hsieh, and G. J. Pappas, “Uncertainty quantification for learning-based mpc using weighted conformal prediction,” in *IEEE Conference on Decision and Control (CDC)*, 2023, pp. 342–349.
- [24] H. Zhou, Y. Zhang, and W. Luo, “Safety-critical control with uncertainty quantification using adaptive conformal prediction,” in *American Control Conference (ACC)*, 2024, pp. 574–580.
- [25] J. Sun, Y. Jiang, J. Qiu, P. Nobel, M. J. Kochenderfer, and M. Schwager, “Conformal prediction for uncertainty-aware planning with diffusion dynamics model,” in *NeurIPS*, vol. 36, 2023, pp. 80324–80337.
- [26] S. Yang, G. J. Pappas, R. Mangharam, and L. Lindemann, “Safe perception-based control under stochastic sensor uncertainty using conformal prediction,” in *IEEE Conference on Decision and Control (CDC)*, 2023, pp. 6072–6078.
- [27] E. E. Vlahakis, L. Lindemann, P. Sopasakis, and D. V. Dimarogonas, “Conformal prediction for distribution-free optimal control of linear stochastic systems,” *IEEE Control Systems Letters*, vol. 8, p. 2835–2840, 2024.
- [28] T.-W. Hsu and H. Tsukamoto, “Statistical guarantees in data-driven nonlinear control: Conformal robustness for stability and safety,” *IEEE Control Systems Letters*, vol. 9, p. 997–1002, 2025.
- [29] S. Wei, M. Ornik, and H. Tsukamoto, “Conformal contraction for robust nonlinear control with distribution-free uncertainty quantification,” *IEEE Conference on Decision and Control preprint arXiv:2507.13613*, Dec. 2025. [Online]. Available: <https://arxiv.org/pdf/2507.13613>
- [30] C. Knuth, G. Chou, J. Reese, and J. Moore, “Statistical Safety and Robustness Guarantees for Feedback Motion Planning of Unknown Underactuated Stochastic Systems,” *IEEE International Conference on Robotics and Automation (ICRA)*, 2023.
- [31] J. Köhler and M. N. Zeilinger, “Predictive control for nonlinear stochastic systems: Closed-loop guarantees with unbounded noise,” *IEEE Transactions on Automatic Control*, pp. 1–16, 2025.
- [32] J.-J. E. Slotine and W. Lohmiller, “On contraction analysis for nonlinear systems,” *Automatica*, vol. 34, no. 6, p. 683–696, 1998.
- [33] I. R. Manchester and J.-J. E. Slotine, “Control contraction metrics: Convex and intrinsic criteria for nonlinear feedback design,” *IEEE Trans. Autom. Control*, vol. 62, no. 6, pp. 3046–3053, Jun. 2017.
- [34] S. Singh, A. Majumdar, J.-J. E. Slotine, and M. Pavone, “Robust online motion planning via contraction theory and convex optimization,” in *IEEE International Conference on Robotics and Automation (ICRA)*, 2017, pp. 5883–5890.
- [35] B. T. Lopez, J.-J. E. Slotine, and J. P. How, “Dynamic tube MPC for nonlinear systems,” in *Proc. Amer. Control Conf.*, 2019, pp. 1655–1662.
- [36] A. Lakshmanan, A. Gahlawat, and N. Hovakimyan, “Safe feedback motion planning: A contraction theory and  $\mathcal{L}_1$ -adaptive control based approach,” in *IEEE CDC*, 2020, pp. 1578–1583.
- [37] H. Tsukamoto, S.-J. Chung, and J.-J. E. Slotine, “Contraction theory for nonlinear stability analysis and learning-based control: A tutorial overview,” *Annual Reviews in Control*, vol. 52, p. 1–14, 2021. [Online]. Available: <https://arxiv.org/abs/2110.00675>
- [38] R. M. Lai Wei and J. Bao, “Discrete-time contraction-based control of nonlinear systems with parametric uncertainties using neural networks,” arXiv:2105.05432, May 2021.
- [39] S. Singh, B. Landry, A. Majumdar, J.-J. Slotine, and M. Pavone, “Robust feedback motion planning via contraction theory,” *International Journal of Robotics Research*, vol. 42, no. 9, 2019.
- [40] H. Tsukamoto and S.-J. Chung, “Neural contraction metrics for robust estimation and control: A convex optimization approach,” *IEEE Control Systems Letters*, vol. 5, no. 1, p. 211–216, 2020.
- [41] D. Sun, S. Jha, and C. Fan, “Learning certified control using contraction metric,” in *Conference on Robot Learning*, ser. Proceedings of Machine Learning Research, vol. 155. PMLR, Nov. 2021, pp. 1519–1539.
- [42] G. Shafer and V. Vovk, “A tutorial on conformal prediction,” *Journal of Machine Learning Research*, vol. 9, p. 371–421, 2008.
- [43] L. Hewing, K. P. Wabersich, and M. N. Zeilinger, “Recursively feasible stochastic model predictive control using indirect feedback,” *Automatica*, vol. 119, 2020.
- [44] R. J. Tibshirani, R. Foygel Barber, E. Candès, and A. Ramdas, “Conformal prediction under covariate shift,” *NeurIPS*, vol. 32, 2019.
- [45] T. Koller, F. Berkenkamp, M. Turchetta, and A. Krause, “Learning-based model predictive control for safe exploration,” in *Proceedings of the IEEE Conference on Decision and Control (CDC)*, 2018.
- [46] J. A. E. Andersson, J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl, “CasADi: A software framework for nonlinear optimization and optimal control,” *Mathematical Programming Computation*, vol. 11, p. 1–36, 2019.
- [47] A. Wächter and L. T. Biegler, “On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming,” *Mathematical Programming*, vol. 106, no. 1, p. 25–57, 2006.