Chance-Constrained Sequential Convex Programming for Robust Trajectory Optimization

Thomas Lew, Riccardo Bonalli, Marco Pavone

Abstract— Planning safe trajectories for nonlinear dynamical systems subject to model uncertainty and disturbances is challenging. In this work, we present a novel approach to tackle chance-constrained trajectory planning problems with nonconvex constraints, whereby obstacle avoidance chance constraints are reformulated using the signed distance function. We propose a novel sequential convex programming algorithm and prove that under a discrete time problem formulation, it is guaranteed to converge to a solution satisfying first-order optimality conditions. We demonstrate the approach on an uncertain 6 degrees of freedom spacecraft system and show that the solutions satisfy a given set of chance constraints.

I. INTRODUCTION

Planning safe trajectories for autonomous systems in the presence of uncertainty (e.g., uncertain model parameters, external disturbances) is critical to their use in real-world applications, especially for robotic systems with limited control authority. Accordingly, probabilistic models have become popular to characterize the dynamics of uncertain systems, as they can be derived from data [1], [2], and allow one to quantify and impose constraints on the risk associated with robotic operations in the face of uncertainty [3]. This motivates the development of trajectory optimization algorithms tailored to tackle this class of problems.

The problem of computing a safe minimal cost trajectory under model uncertainty and external random disturbances has been approached from the perspectives of robust control [4] and chance-constrained optimal control. However, considering fixed deterministic bounds on all uncertainties as in robust control can be too conservative or lead to infeasibility (e.g., when uncertainties follow probability distributions with unbounded support). Instead, chance constraints guarantee the satisfaction of constraints for a pre-defined probability level p. For a constraint $g(\boldsymbol{x}) \leq 0$ with a random variable $\boldsymbol{x}$, chance constraints are expressed as

$$\Pr(g(\boldsymbol{x}) \leq 0) \geq p. \tag{1}$$

To handle such constraints, most approaches reformulate (1) as one or multiple deterministic constraints. The first class of methods consists of using indicator functions $\mathcal{I}(\cdot)$:

$$\Pr(g(\boldsymbol{x}) \leq 0) = \mathbb{E}(\mathcal{I}(\boldsymbol{x})), \text{ with } \mathcal{I}(\boldsymbol{x}) = \begin{cases} 1 \text{ if } g(\boldsymbol{x}) \leq 0 \\ 0 \text{ otherwise.} \end{cases} \tag{2}$$
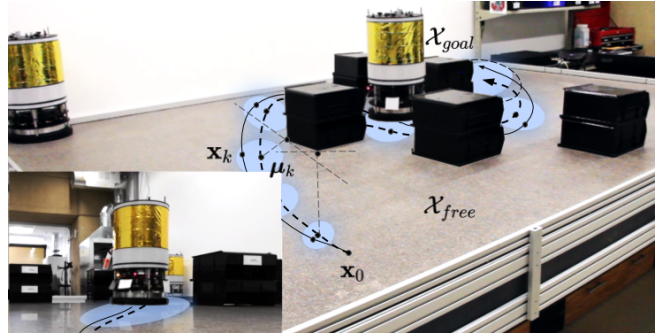
Fig. 1: *Using a chance-constrained problem formulation, the proposed sequential convex programming algorithm efficiently plans safe trajectories despite nonlinear dynamics, nonconvex constraints, model uncertainty, and external disturbances.*

Methods such as the scenario approach [5], the Bernstein approximation [6], [7], and Monte-Carlo-based methods [8], [9] reformulate (1) by sampling $\boldsymbol{x}$ to approximate $\mathbb{E}(\mathcal{I}(\boldsymbol{x}))$, which can be computationally expensive. For discretized state spaces, chance-constrained dynamic programming [10], [11] also leverages (2) to incorporate chance constraints, but such methods are too slow for fast re-planning.

On the other hand, for specific classes of constraints $g(\boldsymbol{x})$ and probability distributions $p(\boldsymbol{x})$, such as affine constraints with radial distributions, equivalent explicit reformulations of (1) exist [12], [13]. Probabilistic confidence sets have been proposed to conservatively reformulate such constraints. [14]. For uncertain systems subject to polytopic constraints, these properties have been employed in chance-constrained trajectory planning [15], [16] and model predictive control [17], [18]. To handle nonconvex obstacles avoidance constraints, it is common to decompose the free space as multiple polytopes, and to solve the resulting problem with mixed integer programming [15], [16], [19]. However, for chance-constrained problems, such methods typically assume linear dynamics, and are slow for a large number of obstacles. A chance-constrained problem formulation is considered in [20], although it is limited to ellipsoidal obstacles. To address these limitations, the method proposed in this work exploits explicit reformulations of chance constraints and signed distance functions to handle obstacles of arbitrary shape.

Reliably and efficiently computing a trajectory satisfying nonlinear dynamics and nonconvex chance constraints is of particular interest. Recent work on sequential convex programming (SCP) has proposed efficient algorithms for trajectory optimization [21]–[26]. By iteratively formulating convex approximations of the original nonconvex optimization problem, this class of algorithms is capable of computing

feasible trajectories in real time. Importantly, recent results provide theoretical guarantees ensuring that the resulting solution satisfies first-order necessary conditions for optimality of the original problem [22]–[26]. However, these works consider deterministic dynamics, penalization of the constraints, and heuristics to facilitate convergence, such as virtual controls and nonsmooth functions which complicate the convergence analysis of the algorithm. In this work, we provide a formulation and algorithm tailored to solve the class of discrete-time chance-constrained trajectory planning problems, as well as a proof of the convergence of our algorithm to a locally optimal point.

The key contributions of this work are as follows:

1) We reformulate the chance-constrained trajectory optimization problem with obstacle avoidance constraints by leveraging signed distance functions to explicitly reformulate chance constraints and probabilistic confidence sets for obstacles of arbitrary shape.

2) We provide a novel sequential convex programming trajectory optimization algorithm with hard enforcement of constraints to solve the chance-constrained problem.

3) We prove the convergence of our algorithm to a point guaranteed to satisfy first-order necessary optimality conditions and all constraints, which is critically enabled by our discrete-time C^1 problem formulation.

This paper is organized as follows. In Section II, we state the chance-constrained optimal control problem. In Section III, we review known approaches to model uncertainty and to reformulate linear chance constraints, In Section IV, these results are leveraged together with the signed distance function to express a deterministic reformulation of the original chance-constrained problem. This problem is convexified in Section V and solved using a novel SCP algorithm. In Section VI, we prove that under mild assumptions, the proposed SCP method is guaranteed to converge to a point satisfying first-order necessary optimality conditions of the nonconvex problem. Finally, the approach is validated on a 6 DoF uncertain spacecraft system in Section VII.

Notations: $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the multivariate normal distribution of mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$, $\chi_n^2(p)$ and $\Phi^{-1}(p)$ the p-th quantiles of the χ^2 distribution with n degrees of freedom and of the inverse cumulative function of $\mathcal{N}(0,1)$, respectively. For $\mathbf{A} \in \mathbb{R}^{n \times m}$, $\mathbf{A}_{i.}$ denotes the i-th row of $\mathbf{A}$ and A_{ij} the j-th element of $\mathbf{A}_{i.}$. For $\mathbf{b} \in \mathbb{R}^n$, b_i denotes the i-th element of $\mathbf{b}$. For a vector-valued function $\boldsymbol{f}(\cdot)$, $\nabla_x \boldsymbol{f}$ denotes the Jacobian matrix of $\boldsymbol{f}$ with respect to $\boldsymbol{x}$.

II. PROBLEM DEFINITION

The problem of trajectory optimization consists of computing a sequence of control inputs $\boldsymbol{u}_k$, for a discrete time dynamical system $\boldsymbol{x}_{k+1} = \boldsymbol{f}(\boldsymbol{x}_k, \boldsymbol{u}_k, \boldsymbol{w}_k)$ with $\boldsymbol{x}_k$ the state and $\boldsymbol{w}_k$ random parameters and disturbances, whose solution $\boldsymbol{x}_k$ connects an initial state $\boldsymbol{x}(0) \in \mathbb{R}^n$ to a goal region $\mathcal{X}_{\text{goal}} \subset \mathbb{R}^n$, minimizes given final $l_f(\cdot)$ and step $l(\cdot)$ costs, and satisfies obstacle avoidance constraints $\boldsymbol{x}_k \in \mathcal{X}_{\text{free}} \subset \mathbb{R}^n$, input constraints $\boldsymbol{u}_k \in \mathcal{U} \subset \mathbb{R}^m$ and linear

structural constraints $\mathbf{H}\boldsymbol{x}_k \leq \mathbf{h}$ with $\mathbf{H} \in \mathbb{R}^{l \times n}$, $\mathbf{h} \in \mathbb{R}^l$, such as bounds on translational and angular speed. Due to uncertainty in the dynamics, we embed this problem in the framework of stochastic optimal control and enforce state and input constraints as chance constraints. The resulting optimal control problem (OCP) can be written as follows:

Problem 1. Original Chance-Constrained OCP

$$\min_{\boldsymbol{x}, \boldsymbol{u}} \ \mathbb{E}\left\{ l_f(\boldsymbol{x}_N) + \sum_{k=0}^{N-1} l(\boldsymbol{x}_k, \boldsymbol{u}_k, \boldsymbol{w}_k) \right\} \tag{3a}$$

$$\text{s.t.} \ \ \boldsymbol{x}_{k+1} = \boldsymbol{f}(\boldsymbol{x}_k, \boldsymbol{u}_k, \boldsymbol{w}_k) \qquad k=0\ldots N-1 \tag{3b}$$

$$\Pr(\mathbf{H}\boldsymbol{x}_k \leq \mathbf{h}) \geq p_x \qquad k=1\ldots N \tag{3c}$$

$$\Pr(\boldsymbol{x}_k \in \mathcal{X}_{\text{free}}) \geq p_x \qquad k=1\ldots N-1 \tag{3d}$$

$$\Pr(\boldsymbol{u}_k \in \mathcal{U}) \geq p_u \qquad k=0\ldots N-1 \tag{3e}$$

$$\Pr(\boldsymbol{x}_N \in \mathcal{X}_{\text{goal}}) \geq p_x \tag{3f}$$

$$\boldsymbol{x}_0 = \boldsymbol{x}(0), \tag{3g}$$

where $k \in \{0, \cdots, N\}$, with N the optimization horizon, $\mathcal{X}_{\text{free}}$ the safe region, $\mathcal{X}_{\text{goal}} \subset \mathcal{X}_{\text{free}}$ the goal region, $\mathcal{U}$ the admissible input set, p_x and p_u the probability thresholds for the state and input chance constraints. We assume that $\boldsymbol{f} \in C^2$, $\mathcal{X}_{\text{goal}}$ and $\mathcal{U}$ are polytopes and that $\mathcal{X}_{\text{free}} = \mathbb{R}^n \backslash \mathcal{X}_{\text{obs}}$, $\mathcal{X}_{\text{obs}} = \bigcup_{i=1}^M \mathcal{O}_i$ where $\mathcal{O}_i \subset \mathbb{R}^n$ are closed convex sets. Nonconvex obstacle avoidance constraints (3d) are separated from linear inequality state constraints (3c) to enable the user to specify different probability thresholds for them.

As the uncertainty of the state trajectory grows over time, planning an open-loop trajectory can lead to a conservative reformulation of chance constraints which does not account for feedback controllers typically used to track the planned trajectory. Therefore, we develop our formulation using a control policy $\boldsymbol{u}_k = \boldsymbol{\nu}_k + \mathbf{K}(\boldsymbol{x}_k)(\boldsymbol{x}_k - \boldsymbol{\mu}_k)$, where $\mathbf{K}(\boldsymbol{x})$ is a pre-defined state feedback controller gain (e.g., a linear-quadratic regulator). Due to the uncertainty in $\boldsymbol{x}_k$, $\boldsymbol{u}_k$ is also uncertain, which motivates the use of chance constraints in (3e). This notation is general, as the user may set $\mathbf{K}(\boldsymbol{x}) = \mathbf{0}$ to plan for an open-loop trajectory instead.

III. PRELIMINARIES

To solve Problem 1, our strategy entails reformulating chance constraints as deterministic constraints. To do so, as it is common in the literature [20], [27] , we (approximately) represent the probability distributions for $\boldsymbol{x}_k$ and $\boldsymbol{u}_k$ as Gaussian distributions. Specifically, given a nominal control input trajectory $(\boldsymbol{\nu}_k)_{k=0}^{N-1}$, the two moments $(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)_{k=0}^N$ used to approximate the Gaussian distribution of the state trajectory can be computed using different methods, such as the sigma point transform or first order Taylor expansions. For the latter, given $\boldsymbol{\mu}_0 = \boldsymbol{x}(0)$ and $\boldsymbol{\Sigma}_0 = \mathbf{0}_{n^2}$ and assuming $\boldsymbol{x}_k, \boldsymbol{w}_k$ are independent, the two moments of $\boldsymbol{x}_{k+1}$ can be recursively computed as

$$\boldsymbol{\mu}_{k+1} = \boldsymbol{f}(\boldsymbol{\mu}_k, \boldsymbol{\nu}_k, \mathbb{E}(\boldsymbol{w}_k)) \tag{4a}$$

$$\boldsymbol{\Sigma}_{k+1} = \nabla_x \boldsymbol{f}_{\text{cl}} \cdot \boldsymbol{\Sigma}_k \cdot \nabla_x \boldsymbol{f}_{\text{cl}}^T + \nabla_w \boldsymbol{f} \cdot \boldsymbol{\Sigma}_w \cdot \nabla_w \boldsymbol{f}^T, \tag{4b}$$

where $\nabla_x \boldsymbol{f}_{\mathrm{cl}} = (\nabla_x \boldsymbol{f} + \nabla_u \boldsymbol{f} \cdot \mathbf{K})$, $\boldsymbol{\Sigma}_w = \mathrm{Var}(\boldsymbol{w}_k)$, and $\mathbf{K}$ and all gradients are evaluated at $(\boldsymbol{\mu}_k, \boldsymbol{\nu}_k, \mathbb{E}(\boldsymbol{w}_k))$. Due to space constraints, we refer the reader to the Appendix[1] for derivations. While this approximation might in general be quite coarse, it is indeed common, works well in practice (as confirmed by our experiments), and allows us to reformulate an otherwise intractable stochastic optimal control problem as a deterministic one (see Section IV).

Most importantly, given this Gaussian approximation, affine chance constraints can be rigorously reformulated as deterministic constraints in two ways: (1) exactly, using reformulations for radial distributions, and (2) conservatively, via ellipsoidal confidence sets. We discuss the relative merits of both approaches in Remark 2 in Section IV.

First, consider $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$, where $\mathbf{b}$ follows a radial probability distribution [12], $c \in \mathbb{R}$ and $p \in (0.5, 1)$. Then, affine chance constraints $\Pr(\mathbf{a}^T \mathbf{b} \leq c) \geq p$ can be reformulated as $\mathbf{a}^T \bar{\mathbf{b}} + \kappa(p)\sqrt{\mathbf{a}^T \boldsymbol{\Sigma}_b \mathbf{a}} \leq c$, where $\kappa(p) > 0$ depends on the distribution of $\mathbf{b}$ [12]. In particular, if $\mathbf{b} \sim \mathcal{N}(\bar{\mathbf{b}}, \boldsymbol{\Sigma}_b)$, then

$$\Pr(\mathbf{a}^T \mathbf{b} \leq c) \geq p \iff \mathbf{a}^T \bar{\mathbf{b}} + \Phi^{-1}(p)\sqrt{\mathbf{a}^T \boldsymbol{\Sigma}_b \mathbf{a}} \leq c, \quad (5)$$

where $\Phi^{-1}(p)$ denotes the p-th quantile of the inverse cumulative function of the standard normal distribution [13].

The other method to reformulate chance constraints involves confidence sets, characterizing a region in which a random variable lies for a given probability p. In particular, ellipsoidal confidence sets are defined as follows:

Definition III.1. (Ellipsoidal Confidence Set)[2]
An ellipsoidal set $\mathcal{B}^p(\boldsymbol{\mu}, \mathbf{Q})$, $\mathbf{Q} \in \mathbb{R}^{n \times n}$, $\mathbf{Q} \succ 0$, defined as

$$\mathcal{B}^p(\boldsymbol{\mu}, \mathbf{Q}) := \left\{ \boldsymbol{x} \mid (\boldsymbol{x} - \boldsymbol{\mu})^T \mathbf{Q}^{-1}(\boldsymbol{x} - \boldsymbol{\mu}) \leq 1 \right\}, \quad (6)$$

is an ellipsoidal confidence set of probability level p for $\boldsymbol{x}$ if

$$\Pr(\boldsymbol{x} \in \mathcal{B}^p(\boldsymbol{\mu}, \mathbf{Q})) \geq p. \quad (7)$$

Remark 1. For a Gaussian-distributed random variable $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\mathcal{B}^p(\boldsymbol{\mu}, \chi_n^2(p)\boldsymbol{\Sigma})$ is an ellipsoidal confidence set of probability level p for $\boldsymbol{x}$.

Alternatively to (5), if $\mathcal{B}^p(\bar{\mathbf{b}}, \mathbf{Q}_b)$ is an ellipsoidal confidence set of probability level p for $\mathbf{b}$, then [18]

$$\Pr(\mathbf{a}^T \mathbf{b} \leq c) \geq p \impliedby \mathbf{a}^T \bar{\mathbf{b}} + \sqrt{\mathbf{a}^T \mathbf{Q}_b \mathbf{a}} \leq c. \quad (8)$$

From Remark 1, if $\mathbf{b} \sim \mathcal{N}(\bar{\mathbf{b}}, \boldsymbol{\Sigma}_b)$ then $\mathbf{Q}_b = \chi_n^2(p)\boldsymbol{\Sigma}_b$. Since $\chi_n(p) \geq \Phi^{-1}(p) \; \forall n$, (5) is less conservative than (8) for a single chance constraint. However, when considering multiple constraints such as (3c), using (8) can be less conservative as we will discuss in the following section.

[1]The Appendix is available at `asl.stanford.edu/wp-content/papercite-data/pdf/Lew.Bonalli.Pavone.ECC20.pdf`.
[2]Definition III.1 is equivalent to [14, Lemma 2], but with $\boldsymbol{\mu} \neq \mathbf{0}$. For $\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, it is obtained by applying [14, Lemma 2] to $(\boldsymbol{x} - \boldsymbol{\mu})$ with $\mathbf{Q} = \chi_n^2(p)\boldsymbol{\Sigma}$.

IV. DETERMINISTIC PROBLEM REFORMULATION

To reformulate Problem 1 as a deterministic problem with optimization variables $(\boldsymbol{\mu}_k, \boldsymbol{\nu}_k)_k$, we exploit the Gaussian approximation and present a method to conservatively reformulate obstacle avoidance chance constraints by leveraging signed distance functions.

A. Cost and Dynamics

As in [17], we use an approximation of the expected cost along the nominal trajectory $(\boldsymbol{\mu}_k, \boldsymbol{\nu}_k)_k$ as

$$\mathbb{E}\{l(\boldsymbol{x}_k, \boldsymbol{u}_k, \boldsymbol{w}_k)\} \approx l(\boldsymbol{\mu}_k, \boldsymbol{\nu}_k, \mathbb{E}\{\boldsymbol{w}_k\}) := \tilde{l}(\boldsymbol{\mu}_k, \boldsymbol{\nu}_k) \quad (9)$$
$$\mathbb{E}\{l_f(\boldsymbol{x}_N)\} \approx l_f(\boldsymbol{\mu}_N) := \tilde{l}_f(\boldsymbol{\mu}_N). \quad (10)$$

This mean-equivalent reformulation could be replaced with different approaches, e.g., using a first-order approximation or risk metrics [3]. Choosing a cost function is not the focus of this paper and we leave the reader choose a problem-dependent cost. For the dynamics, we use (4) to reformulate the state trajectory as a function of $(\boldsymbol{\mu}_k, \boldsymbol{\nu}_k)_k$ only. Note that our proposed trajectory optimization algorithm can leverage other uncertainty propagation techniques, e.g., [18].

B. Polytopic Chance Constraints

The linear state chance constraints in (3c) can be conservatively rewritten using (a) $(\Pr(\bigwedge_i A_i) \geq p) \equiv (\Pr(\bigvee_i \bar{A}_i) \leq 1 - p)$, where $A_i \in \{0, 1\}$ denotes a random event, and (b) Boole's inequality $(\Pr(\bigvee_i A_i) \leq \sum_i \Pr(A_i))$, as

$$\Pr(\mathbf{H}\boldsymbol{x}_k \leq \mathbf{h}) \geq p_x \overset{(a)}{\iff} \Pr\left(\bigvee_{i=1}^{l} \mathbf{H}_{i\cdot}\boldsymbol{x}_k > h_i \right) \leq 1 - p_x$$
$$\overset{(b)}{\impliedby} \sum_{i=1}^{l} \Pr\left(\mathbf{H}_{i\cdot}\boldsymbol{x}_k > h_i \right) \leq 1 - p_x. \quad (11)$$

Using a uniform risk allocation with $\delta_x^l = \frac{1-p_x}{l}$, a conservative condition for (11) is given as

$$(11) \impliedby \Pr(\mathbf{H}_{i\cdot}\boldsymbol{x}_k > h_i) \leq \delta_x^l, \qquad i = 1, \ldots, l$$
$$\iff \Pr(\mathbf{H}_{i\cdot}\boldsymbol{x}_k \leq h_i) \geq 1 - \delta_x^l, \quad i = 1, \ldots, l. \quad (12)$$

Finally, exploiting the approximation $\boldsymbol{x}_k \sim \mathcal{N}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ and using (5), (12) can be equivalently rewritten with l deterministic constraints as

$$\mathbf{H}_{i\cdot}\boldsymbol{\mu}_k + \Phi^{-1}(1-\delta_x^l)\sqrt{\mathbf{H}_{i\cdot}\boldsymbol{\Sigma}_k\mathbf{H}_{i\cdot}^T} \leq \mathbf{h}_i, \; i = 1, \ldots, l. \quad (13)$$

Alternatively, assuming that $\mathcal{B}^{p_x}(\boldsymbol{\mu}_k, \mathbf{Q}_k)$ is an ellipsoidal confidence set of probability level p_x for $\boldsymbol{x}_k$ and using (8), (3c) can be directly reformulated as l deterministic constraints (without the use of Boole's inequality) as

$$\mathbf{H}_{i\cdot}\boldsymbol{\mu}_k + \sqrt{\mathbf{H}_{i\cdot}\mathbf{Q}_k\mathbf{H}_{i\cdot}^T} \leq h_i, \quad i = 1, \ldots, l. \quad (14)$$

Remark 2. For a given statespace dimension n, there exists a critical number of linear constraints l such that it is preferable to use (14) than (13). For the Gaussian distribution, such values are reported in Appendix IX-F.

The same approach is used to reformulate control input chance constraints in (3e). Let $\mathcal{U} = \{\boldsymbol{u} \mid \mathbf{G}^u \boldsymbol{u} \leq \mathbf{b}^u\}$ with

$\mathbf{G}^u \in \mathbb{R}^{r \times m}$, $\mathbf{b}^u \in \mathbb{R}^r$, and $\delta_u = \frac{1-p_u}{r}$. Since $\boldsymbol{u} = \boldsymbol{\nu} + \mathbf{K}(\boldsymbol{x} - \boldsymbol{\mu})$, we approximate $\boldsymbol{u}_k \sim \mathcal{N}(\boldsymbol{\nu}_k, \mathbf{K}\boldsymbol{\Sigma}_k\mathbf{K})$ and obtain

$$\Pr(\boldsymbol{u}_k \in \mathcal{U}) \geq p_u \iff \Pr(\mathbf{G}^u\boldsymbol{u}_k \leq \mathbf{b}^u) \geq p_u, \quad i = 1...r$$
$$\impliedby \mathbf{G}^u_{i\cdot}\boldsymbol{\nu}_k + \Phi^{-1}(1-\delta_u)\sqrt{\mathbf{G}^u_{i\cdot}\mathbf{K}\boldsymbol{\Sigma}_k\mathbf{K}^T\mathbf{G}^{uT}_{i\cdot}} \leq \mathbf{b}_i, \quad (15)$$

where (15) holds $\forall i = 1, ..., r$, and $\mathbf{K} = \mathbf{K}(\boldsymbol{\mu}_k)$. Similarly, if $\mathcal{B}^{p_u}(\boldsymbol{\nu}_k, \mathbf{Q}^u_k)$ satisfies Definition III.1 for $\boldsymbol{u}_k$, (3e) can be reformulated as $\mathbf{G}^u_{i\cdot}\boldsymbol{\nu}_k + \sqrt{\mathbf{G}^u_{i\cdot}\mathbf{Q}^u_k\mathbf{G}^{uT}_{i\cdot}} \leq \mathbf{b}_i, i = 1, ..., r$.

Finally, with $\mathcal{X}_{\text{goal}} = \{\boldsymbol{x} \mid \mathbf{P}\boldsymbol{x} \leq \mathbf{p}\}$, $\mathbf{P} \in \mathbb{R}^{q \times n}, \mathbf{p} \in \mathbb{R}^q$, and $\delta_N = \frac{1-p_x}{q}$, the goal region constraint $\Pr(\boldsymbol{x}_N \in \mathcal{X}_{\text{goal}}) \geq p_x$ in (3f) is reformulated similarly using either (13) or (14).

C. Obstacle Avoidance Constraints

Consider M obstacles $\mathcal{O}_i \subset \mathbb{R}^n$ and the unsafe and safe sets $\mathcal{X}_{\text{obs}} = \bigcup_{i=1}^M \mathcal{O}_i$ and $\mathcal{X}_{\text{free}} = \mathbb{R}^n \setminus \mathcal{X}_{\text{obs}}$. Then,

$$\Pr(\boldsymbol{x}_k \in \mathcal{X}_{\text{free}}) \geq p_x \iff \Pr\left(\bigwedge_{i=1}^M \boldsymbol{x}_k \notin \mathcal{O}_i\right) \geq p_x. \quad (16)$$

Following similar derivations as before with a uniform risk allocation $\delta_x := \frac{1-p_x}{M}$, a sufficient condition for (16) is

$$(16) \overset{(a)}{\iff} \Pr\left(\bigvee_{i=1}^M \boldsymbol{x}_k \in \mathcal{O}_i\right) \leq 1 - p_x \quad (17)$$
$$\overset{(b)}{\impliedby} \sum_{i=1}^M \Pr(\boldsymbol{x}_k \in \mathcal{O}_i) \leq 1 - p_x \impliedby \bigwedge_{i=1}^M \Pr(\boldsymbol{x}_k \in \mathcal{O}_i) \leq \delta_x$$
$$\iff \Pr(\boldsymbol{x}_k \notin \mathcal{O}_i) \geq 1 - \delta_x, \quad i = 1, ..., M. \quad (18)$$

For general nonconvex sets $\mathcal{O}_i$ and probability distributions for $\boldsymbol{x}_k$, the constraints above are nonconvex. Instead of using conservative polytopic approximations of the obstacles as in [19] and to avoid the use of mixed integer programming, we leverage the signed distance function $d_i : \mathcal{X} \to \mathbb{R}$: a nonlinear function which returns the shortest distance from a point $\boldsymbol{x}$ to the boundary $\partial\mathcal{O}_i$ of a set $\mathcal{O}_i$ [21]. It is defined as

$$d_i(\boldsymbol{x}) = \inf_{\mathbf{y} \in \mathcal{O}_i} \|\boldsymbol{x} - \mathbf{y}\|_2 - \inf_{\boldsymbol{z} \notin \mathcal{O}_i} \|\boldsymbol{x} - \boldsymbol{z}\|_2. \quad (19)$$

States out of obstacles return positive signed distances, i.e., $\boldsymbol{x} \notin \mathcal{O}_i \iff d_i(\boldsymbol{x}) \geq 0$. Also, note that $\nabla d_i(\boldsymbol{x})$ is given as

$$\mathbf{n}_i(\boldsymbol{x}) = \frac{\boldsymbol{x} - \boldsymbol{x}^i_{\text{obs}}}{d_i(\boldsymbol{x})}, \text{ with } \boldsymbol{x}^i_{\text{obs}} = \underset{\boldsymbol{x}_{\text{obs}} \in \partial\mathcal{O}_i}{\arg\min} \|\boldsymbol{x}_{\text{obs}} - \boldsymbol{x}\|_2, \quad (20)$$

where $\partial\mathcal{O}_i$ denotes the boundary of $\mathcal{O}_i$. Compared to [26] which uses projection operators, the signed distance function is differentiable almost everywhere which facilitates the convergence analysis of our algorithm and enables its initialization with infeasible trajectories. Further, although $d_i(\cdot)$ is nonconvex, it is possible to formulate a convex conservative approximation of each individual obstacle avoidance constraint $\boldsymbol{x}_k \notin \mathcal{O}_i$ by simply linearizing the constraint $d_i(\boldsymbol{x}) \geq 0$. Indeed, define $\boldsymbol{x}^j_k \in \mathbb{R}^n$, $d^j_{ik} = d_i(\boldsymbol{x}^j_k)$ and $\mathbf{n}^j_{ik} = \mathbf{n}_i(\boldsymbol{x}^j_k)$. Then, given $p \in (0, 1)$,

$$d^j_{ik} + \mathbf{n}^{jT}_{ik}(\boldsymbol{x}_k - \boldsymbol{x}^j_k) \geq 0 \implies \boldsymbol{x}_k \notin \mathcal{O}_i, \text{ and} \quad (21)$$

$$\Pr\left(d^j_{ik} + \mathbf{n}^{jT}_{ik}(\boldsymbol{x}_k - \boldsymbol{x}^j_k) \geq 0\right) \geq p \implies \Pr(\boldsymbol{x}_k \notin \mathcal{O}_i) \geq p. \quad (22)$$

This result is generalized to the free set $\mathcal{X}_{\text{free}}$ and summarized in the following proposition:

Proposition 1. Define $\boldsymbol{x}, \boldsymbol{x}^j \in \mathbb{R}^n$, $\mathcal{X}_{\text{free}} = \mathbb{R}^n \setminus \mathcal{X}_{\text{obs}}$, $d_{\min} \geq 0$ a safety margin, $d^j_i = d_i(\boldsymbol{x}^j)$, and $\mathbf{n}^j_i = (\boldsymbol{x}^j - \boldsymbol{x}_{\text{obs}})/d^j_i$. If $\mathcal{X}_{\text{obs}} = \bigcup_{i=1}^M \mathcal{O}_i$, and $\mathcal{O}_i \subset \mathbb{R}^n$ are closed convex sets, then

$$d^j_i + \mathbf{n}^{jT}_i(\boldsymbol{x} - \boldsymbol{x}^j) \geq d_{\min}, \ i = 1, ..., M \implies \boldsymbol{x} \in \mathcal{X}_{\text{free}}. \quad (23)$$

Further, for $p \in (0, 1)$, and any probability distribution of $\boldsymbol{x}$,

$$\Pr\left(d^j_i + \mathbf{n}^{jT}_i(\boldsymbol{x} - \boldsymbol{x}^j) \geq d_{\min}, \ i = 1, ..., M\right) \geq p \quad (24)$$
$$\implies \Pr(\boldsymbol{x} \in \mathcal{X}_{\text{free}}) \geq p.$$

Proof. Due to space constraints, we refer the reader to the Appendix for the full proof of the first fact.

For the second, define the indicator functions $\mathcal{I}_l(\boldsymbol{x})$ and $\mathcal{I}(\boldsymbol{x})$, such that $\mathcal{I}_l(\boldsymbol{x}) = 1$ if $d^j_i + \mathbf{n}^{jT}_i(\boldsymbol{x} - \boldsymbol{x}^j) \geq d_{\min}$, $\mathcal{I}(\boldsymbol{x}) = 1$ if $\boldsymbol{x} \in \mathcal{X}_{\text{free}}$, and $\mathcal{I}_l(\boldsymbol{x}) = 0, \mathcal{I}(\boldsymbol{x}) = 0$ otherwise. Using (2), it holds that $(24) \iff \mathbb{E}(\mathcal{I}_l(\boldsymbol{x})) \geq p$ and $\Pr(\boldsymbol{x} \in \mathcal{X}_{\text{free}}) \geq p \iff \mathbb{E}(\mathcal{I}(\boldsymbol{x})) \geq p$. From the definitions of $\mathbb{E}(\cdot)$, $\mathcal{I}_l(\boldsymbol{x})$, $\mathcal{I}(\boldsymbol{x})$ and due to Proposition 1, $\mathbb{E}(\mathcal{I}(\boldsymbol{x})) \geq \mathbb{E}(\mathcal{I}_l(\boldsymbol{x}))$. Hence, $\mathbb{E}(\mathcal{I}_l(\boldsymbol{x})) \geq p \implies \mathbb{E}(\mathcal{I}(\boldsymbol{x}))) \geq p$, from which Proposition 1 follows. $\square$

Importantly, (21) holds for any $\boldsymbol{x}^j_k$ and is linear in $\boldsymbol{x}_k$. Thus, using the procedure used to obtain (13), each chance constraint in (18) can be conservatively reformulated as

$$d^j_{ik} + \mathbf{n}^{jT}_{ik}(\boldsymbol{\mu}_k - \boldsymbol{x}^j_k) - \Phi^{-1}(1-\delta_x)\sqrt{\mathbf{n}^{jT}_{ik}\boldsymbol{\Sigma}_k\mathbf{n}^j_{ik}} \geq 0. \quad (25)$$

As Proposition 1 holds $\forall \boldsymbol{x}^j_k$, we set $\boldsymbol{x}^j_k = \boldsymbol{\mu}_k$ and obtain

$$(18) \impliedby d_i(\boldsymbol{\mu}_k) - \Phi^{-1}(1-\delta_x)\sqrt{\mathbf{n}^T_i\boldsymbol{\Sigma}_k\mathbf{n}_i} \geq 0, \quad (26)$$

with $i = 1, ..., M$ and $\mathbf{n}_i = \mathbf{n}_i(\boldsymbol{\mu}_k)$.

Similarly, if $\mathcal{B}^{p_x}(\boldsymbol{\mu}_k, \mathbf{Q}_k)$ satisfies Definition III.1 for $\boldsymbol{x}_k$, (16) can be conservatively reformulated with Proposition 1 as

$$(16) \impliedby d_i(\boldsymbol{\mu}_k) - \sqrt{\mathbf{n}^T_i\mathbf{Q}_k\mathbf{n}_i} \geq 0, \quad i = 1, ..., M. \quad (27)$$

As mentioned in Remark 2 and shown in Figure 2, there exists obstacle configurations for which (27) is less conservative than (26).
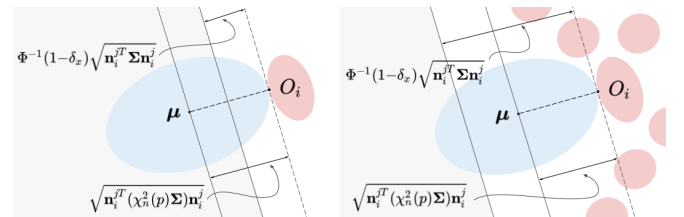

Fig. 2: Consider $\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. For many obstacles or linear chance constraints, using ellipsoidal confidence sets (Def. III.1, in blue) rather than risk allocation and the exact reformulation of each chance constraint (Eq. (26), feasible set in gray) is less conservative.

D. Nonlinear Deterministic Problem

Using the deterministic reformulations defined above, the chance-constrained optimal control problem (OCP) can be conservatively reformulated as a deterministic problem with optimization variables $(\boldsymbol{\mu}, \boldsymbol{\nu}) := (\boldsymbol{\mu}_k, \boldsymbol{\nu}_k)_{k=0}^N$ as

Problem 2. Chance-Constrained OCP (**CC-OCP**)

$$\min_{\boldsymbol{\mu}, \boldsymbol{\nu}} \quad \tilde{l}_f(\boldsymbol{\mu}_N) + \sum_{k=0}^{N-1} \tilde{l}_k(\boldsymbol{\mu}_k, \boldsymbol{\nu}_k)$$

$$\text{s.t.} \quad \boldsymbol{\mu}_{k+1} = \tilde{f}(\boldsymbol{\mu}_k, \boldsymbol{\nu}_k), \quad \boldsymbol{\mu}_0 = \boldsymbol{x}(0), \quad \boldsymbol{\Sigma}_{k+1}(\boldsymbol{\mu}, \boldsymbol{\nu}) = (4b), \quad \boldsymbol{\Sigma}_0 = \mathbf{0}_{n^2},$$

$$\mathbf{H}_{i\cdot} \boldsymbol{\mu}_k + \Phi_{\delta_x^l}^{-1} \|\mathbf{H}_{i\cdot}\|_{\boldsymbol{\Sigma}_k} \leq h_i, \quad d_i(\boldsymbol{\mu}_k) - \Phi_{\delta_x}^{-1} \|\mathbf{n}_{ik}\|_{\boldsymbol{\Sigma}_k} \geq 0,$$

$$\mathbf{G}_i^u \boldsymbol{\nu}_k + \Phi_{\delta_u}^{-1} \|\mathbf{G}_{i\cdot}^u \mathbf{K}\|_{\boldsymbol{\Sigma}_k} \leq b_i, \quad \mathbf{P}_{i\cdot} \boldsymbol{\mu}_N + \Phi_{\delta_N}^{-1} \|\mathbf{P}_{i\cdot}\|_{\boldsymbol{\Sigma}_N} \leq p_i,$$

where the indices i and j are defined separately for each constraint in this section, $\Phi_\delta^{-1} := \Phi^{-1}(1-\delta)$, $\tilde{f}(\boldsymbol{\mu}_k, \boldsymbol{\nu}_k) = f(\boldsymbol{\mu}_k, \boldsymbol{\nu}_k, \mathbb{E}(\boldsymbol{w}_k)))$ and $\|\mathbf{a}\|_{\boldsymbol{\Sigma}} := \sqrt{\mathbf{a}^T \boldsymbol{\Sigma} \mathbf{a}}$ for conciseness. We stress that instead of using the reformulation for Gaussian distributions with Φ_δ^{-1} with (5) as in this formulation, ellipsoidal confidence sets can be used with (8), as specified in Remark 2.

V. Convex Problem Reformulation and Sequential Convex Programming

We provide solutions for (**CC-OCP**) by leveraging sequential convex programming (SCP). Specifically, in Algorithm 1, we provide Chance-Constrained SCP (**CC-SCP**), a version of [23, Algorithm 1] designed to solve (**CC-OCP**). By leveraging the discrete time and chance constraints setting in (**CC-OCP**), we are able to prove that, when convergence is achieved, (**CC-SCP**) finds a point that 1) is a feasible point for (**CC-OCP**), and in particular, it satisfies state constraints, and 2) it satisfies strong necessary conditions for optimality for (**CC-OCP**).

The working principle of (**CC-SCP**) consists of successively linearizing the costs and nonconvex constraints, seeking a solution of the original problem through a sequence of convex problems. Given the solution $(\boldsymbol{\mu}^j, \boldsymbol{\nu}^j)$ from the convexified problem at iteration j, the convex approximation of (**CC-OCP**) at the current iteration $(j+1)$ is described next.

First, we approximate the cost terms $\tilde{l}_f(\cdot), \tilde{l}_k(\cdot)$ around $\boldsymbol{z}^j = (\boldsymbol{\mu}^j, \boldsymbol{\nu}^j)$ using a first order Taylor approximation and denote each linearized term as $l_f^j(\boldsymbol{\mu}_N)$ and $l_k^j(\boldsymbol{\mu}_k, \boldsymbol{\nu}_k)$. Next, all nonconvex constraints, including the dynamics and chance constraints, are also linearized around $(\boldsymbol{\mu}^j, \boldsymbol{\nu}^j)$.

To avoid artificial unboundedness [25], where the solution of the linearized problem may lie far from the linearization trajectory $(\boldsymbol{\mu}^j, \boldsymbol{\nu}^j)$, it is necessary to add trust region constraints $\|\boldsymbol{\mu}_k - \boldsymbol{\mu}_k^j\|^2 \leq \Delta_j$, $\|\boldsymbol{\nu}_k - \boldsymbol{\nu}_k^j\|^2 \leq \Delta_j$, where $\Delta_j \in [0, \Delta_0]$, $\Delta_0 > 0$ is the trust region radius. Also, to provide theoretical convergence guarantees, we penalize this constraint using a convex C^1 approximation $\varphi(t)$ for $\max\{t, 0\}$ satisfying $\frac{d\varphi}{dt}(t) = 0$ for $t \leq 0$.

Therefore, the convexified chance-constrained optimal control problem at iteration $(j+1)$ is defined as follows:

Problem 3. Chance-Constrained Lin. OCP (**CC-LOCP**)

$$\min_{\boldsymbol{\mu}, \boldsymbol{\nu}} \quad l_f^j(\boldsymbol{\mu}_N) + \sum_{k=0}^{N-1} l_k^j(\boldsymbol{\mu}_k, \boldsymbol{\nu}_k) + \tag{29a}$$
$$\omega_j \left(\sum_{k=1}^{N} \varphi(\|\boldsymbol{\mu}_k - \boldsymbol{\mu}_k^j\|^2 - \Delta_j) + \sum_{k=0}^{N-1} \varphi(\|\boldsymbol{\nu}_k - \boldsymbol{\nu}_k^j\|^2 - \Delta_j) \right)$$

s.t.   Linearized Version of the Constraints in (**CC-OCP**).

where $\omega_j \geq 0$ is a penalization weight. The complete problem formulation is expressed in the Appendix.

Another key component of (**CC-SCP**) is the computation of an accuracy ratio ρ to quantify whether the solution to (**CC-LOCP**) is close to a solution of (**CC-OCP**): where $|\Delta l_k^{j+1}|, \|\Delta f_k^{j+1}\|, \|\Delta g_k^{j+1}\|$ quantify the differences between the true and linearized costs, dynamics and constraints, respectively, normalized by the respective linearized expressions as in [23].

(**CC-SCP**), outlined in Algorithm 1, consists of solving a sequence of (**CC-LOCP**) (line 3) to obtain a solution to (**CC-OCP**). If the solution lies outside the trust region (line 4) or is deemed not accurate with respect to the nonlinear problem (lines 5-6), it is rejected and weights are updated accordingly (lines 7-8, 12-14) to ensure that the next solution satisfies these conditions. Otherwise, each solution is accepted and used to obtain the next (**CC-LOCP**). The algorithm terminates once it converges or if it cannot find a feasible solution (line 2). At line 15, Δ is shrunk by a coefficient α to satisfy an assumption necessary for the theoretical convergence result presented in the next section, although this step appeared to have no impact in practice.

Algorithm 1 Chance-Constrained SCP (**CC-SCP**)

Input: Initial guess $(\boldsymbol{\mu}^0, \boldsymbol{\nu}^0)$
Parameters: $\Delta_0, \beta_{\text{fail}}, \beta_{\text{succ}}, \rho_0, \rho_1, \omega_0, \omega_{\max}, \gamma_{\text{fail}}, \alpha$
Output: Solution $(\boldsymbol{\mu}, \boldsymbol{\nu})$

1: $j \leftarrow 0$
2: **while** $(\boldsymbol{\mu}^{j+1}, \boldsymbol{\nu}^{j+1}) \neq (\boldsymbol{\mu}^j, \boldsymbol{\nu}^j)$ **and** $\omega_j \leq \omega_{\max}$ **do**
3:     Solve (**CC-LOCP**)$_{j+1}$ for $(\boldsymbol{\mu}^{j+1}, \boldsymbol{\nu}^{j+1})$
4:     **if** $\|\boldsymbol{\mu}^{j+1} - \boldsymbol{\mu}^j\|^2 \leq \Delta_j$ **and** $\|\boldsymbol{\nu}^{j+1} - \boldsymbol{\nu}^j\|^2 \leq \Delta_j^u$ **then**
5:         Compute model accuracy ratio ρ_j
6:         **if** $\rho_j > \rho_1$ **then**
7:             Reject solution $(\boldsymbol{\mu}^{j+1}, \boldsymbol{\nu}^{j+1})$
8:             $\Delta_{j+1} \leftarrow \beta_{\text{fail}} \Delta_j$
9:         **else**
10:            Accept solution $(\boldsymbol{\mu}^{j+1}, \boldsymbol{\nu}^{j+1})$
11:            $\Delta_{j+1} \leftarrow \begin{cases} \min\{\beta_{\text{succ}} \Delta_j, \Delta_0\} & \text{if } \rho_j < \rho_0 \\ \Delta_j & \text{if } \rho_j \geq \rho_0 \end{cases}$
12:     **else**
13:         Reject solution $(\boldsymbol{\mu}^{j+1}, \boldsymbol{\nu}^{j+1})$
14:         $\omega_{j+1} \leftarrow \gamma_{\text{fail}} \cdot \omega_j$
15:     $\Delta_{j+1} \leftarrow \alpha \Delta_{j+1}$, with $\alpha \approx 1, \alpha < 1$
16:     $j \leftarrow j + 1$
17: **return** $(\boldsymbol{\mu}^{j+1}, \boldsymbol{\nu}^{j+1})$

VI. CONVERGENCE GUARANTEES

(**CC-SCP**) enjoys theoretical guarantees of convergence. Specifically, the main takeaway is that, when convergence is achieved, (**CC-SCP**) returns a feasible point for (**CC-OCP**) satisfying first-order necessary conditions for optimality, i.e., the Karush–Kuhn–Tucker (KKT) conditions associated to (**CC-OCP**). In other words, (**CC-SCP**) leads to points that are necessarily local optima for the nonconvex problem (**CC-OCP**). In this section, we state and prove these theoretical guarantees of convergence. It is worth pointing out that our analysis does not rely on the specific formulation of (**CC-OCP**), but, under mild assumptions, it rather applies to a broader class of nonlinear optimization problems.

A. Reformulation of the Main Problem

Denote $z = (\mu_1, \ldots, \mu_N, \nu_0, \ldots, \nu_{N-1})$. Then, (**CC-OCP**) takes the form of the following nonconvex problem:

Problem 4. Nonlinear/Nonconvex Formulation

$$\min_{z} \quad c(z) \quad \text{s.t.} \quad h(z) = 0, \ g(z) \leq 0,$$

where $c: \mathbb{R}^{N(n+m)} \to \mathbb{R}$, $h: \mathbb{R}^{N(n+m)} \to \mathbb{R}^r$, $g: \mathbb{R}^{N(n+m)} \to \mathbb{R}^s$, where integers $r, s > 0$ sum up over all constraints in formulation (2) and the inequality above holds componentwise.

In practice, (**CC-SCP**) successively solves linearized versions of Problem 4, defined as (**CC-LOCP**). Under the formalism adopted above, these problems can be written as the following family of convex problems:

Problem 5. Convexified Formulation at Iteration $(j + 1)$

$$\min_{z} \quad \left(c(z^j) + \nabla c(z^j)(z - z^j) \right) + \varphi(\|z - z^j\|^2 - \Delta_j)$$
$$\text{s.t.} \quad h(z^j) + \nabla h(z^j)(z - z^j) = 0$$
$$g(z^j) + \nabla g(z^j)(z - z^j) \leq 0$$

where z^j denotes the unique solution to Problem 5 at the previous iteration j, around which Problem 4 was linearized, ∇h and ∇g are the Jacobians of h and g respectively, and with the notation $\|z - z^j\|^2 - \Delta_j$, gathers all the contributions $\|\mu_k - \mu_k^j\|^2 - \Delta_j$, $\|\nu_k - \nu_k^j\|^2 - \Delta_j$ in formulation (3). Recall that the scalar function $\varphi(t)$ is a C^1, convex approximation for $\max\{t, 0\}$ with $\frac{d\varphi}{dt}(t) = 0$ for $t \leq 0$.

B. Necessary Conditions for Optimality

Definition VI.1. A point z satisfies the Linear Independence Constraint Qualification (LICQ) for a nonlinear, nonconvex formulation as in Problem 4 if, for every $i \in A \subseteq \{1, \ldots, s\}$ for which $g_i(z) = 0$, the vectors $\nabla(g_i)(z)$, $i \in A$ and $\nabla(h_1)(z), \ldots, \nabla(h_r)(z)$ are linearly independent.

Definition VI.2. A point z satisfies the KKT conditions for a nonlinear, nonconvex formulation as in Problem 4 if there exists a multiplier $(\alpha, \lambda, \zeta) \in \mathbb{R}^{1+r+s}$ satisfying

$$\alpha \nabla c(z) + \lambda^\top \nabla h(z) + \zeta^\top \nabla g(z) = 0, \tag{32a}$$
$$h(z) = 0, \quad g(z) \leq 0, \quad (\alpha, \lambda, \zeta) \neq 0, \tag{32b}$$
$$\zeta_i g_i(z) = 0, \quad \zeta_i \geq 0, \quad i = 1, \ldots, s. \tag{32c}$$

Below, we recall the classical result concerning first-order necessary conditions for optimality (see, e.g., [28]):

Theorem 1. *Let z^* be an optimum for Problem 4 for which LICQ holds. Then, z^* satisfies the KKT conditions related to Problem 4 with multiplier (α, λ, ζ) satisfying $\alpha = 1$.*

The previous theorem provides the existence of multipliers (α, λ, ζ) for which $\alpha \neq 0$. As a consequence, this might prevent from having a control over the norm of the multiplier, that is, when considering a family of optimization problems (in our case, represented by the family of convexified Problems 5) it might happen that the family of associated multipliers (when they exist) is not bounded. Having bounded families of multipliers will be crucial to prove our theoretical result. For this, below we derive slightly different necessary conditions for optimality tailored to our framework:

Corollary 1. *Let z^* be an optimum for Problem 4 for which LICQ holds. Then, z^* satisfies the KKT conditions related to Problem 4 with multiplier satisfying $\|(\alpha, \lambda, \zeta)\| = 1$, $\alpha \geq 0$.*

Proof. Thanks to Theorem 1, z^* satisfies the KKT conditions related to Problem 4 with multiplier $(\tilde{\alpha}, \tilde{\lambda}, \tilde{\zeta})$ satisfying $\tilde{\alpha} = 1$. Since $(\tilde{\alpha}, \tilde{\lambda}, \tilde{\zeta}) \neq 0$, the conclusion follows by taking $(\alpha, \lambda, \zeta) := (\tilde{\alpha}, \tilde{\lambda}, \tilde{\zeta})/\|(\tilde{\alpha}, \tilde{\lambda}, \tilde{\zeta})\|$. $\square$

C. Assumptions and Proof of the Main Result

The proof of our main result makes use of the following set of assumptions, which we list below. We appropriately comment these assumptions and their validity in the context of our main problem (**CC-OCP**).

Assumption 1. *Functions c, h and g are C^1.*

This assumption is easily satisfied in the context of (**CC-OCP**). Indeed, all deterministic cost and dynamics are taken to be C^2, therefore, the covariance matrices Σ_k are C^1. Moreover, all original deterministic constraints can be chosen to be C^1, including the signed distance function which is C^∞ when the obstacles are either walls or discs (for which the differentiability holds everywhere but at their centers; this also extends to ellipsoids). A proof of this fact is provided in the Appendix. Finally, it is worth noting that this assumption can always be enforced because any Lipschitz function can be approximated by a smooth function with a precision that is selected by the user (see, e.g., [29]).

Assumption 2. *At each iteration j, Problem 5 has a solution z^j. Moreover, z^j satisfies LICQ related to Problem 5. Finally, the family of solutions z^j, $j \in \mathbb{N}$ is bounded.*

This assumption is classic in convex optimization and easily satisfied in the context of the linearized Problems (**CC-LOCP**). In particular, since the constraints in formulation (29) define a convex and closed feasible region, every Problem (**CC-LOCP**) has at least one solution as soon as the feasible region is not empty and either the convexified version of the cost goes to infinity at infinity or the feasible region is bounded. The latter is satisfied

in all cases considered in this work since the workspace is compact and we set polytopic control input constraints, which also enforce the family of solutions for (**CC-LOCP**) to be bounded. Finally, LICQ is a standard assumption in optimization and it is satisfied by almost any optimization problem (see, e.g., [30, Theorem 1]).

Assumption 3. *Set* $\Delta z^{j+1} := z^{j+1} - z^j$. *We require that*

$$\exists M > 0 \ \ s.t. \ \ \forall j > M, \quad \|\Delta z^{j+1}\| < \|\Delta z^j\|. \quad (33)$$

Assumption 3 plays a mere technical role in the proof, providing the existence of a converging subsequence of solutions z^j's (we can weaken (33) by considering a subsequence). The assumption above is satisfied thanks to step 15 of (**CC-SCP**). It is worth noting that Assumption 3 does not force the entire sequence $(z^j)_{j \in \mathbb{N}}$ to converge. As a counterexample, consider the nonconvergent sequence $(\log(i))_{i \in \mathbb{N}}$ which clearly satisfies relation (33).

Our convergence result consists of proving that, under Assumptions 1, 2 and 3, solving a sequence of Problem 5 yields a solution satisfying the KKT conditions related to Problem 4. To do so, we extract a converging subsequence of solutions for Problems 5, then retrieve a multiplier for Problem 4 by extracting a convergent subsequence of multipliers related to Problems 5, which is possible by Corollary 1. Importantly, this implies that the algorithm does not retrieve trivial, noninformative multipliers for the original formulation.

Theorem 2 (Convergence guarantees). *Assume that Assumptions 1, 2, 3 hold and consider the family* $(z^j)_{j \in \mathbb{N}}$ *where* z^j *is solution of Problem 5 at iteration* j. *The following holds:*

1) *If there exists an iteration* $\bar{j}$ *such that for every* $j \geq \bar{j}$ *it holds* $z^j = z^{\bar{j}}$, *then* $z^{\bar{j}}$ *is a feasible point satisfying the KKT conditions related to Problem 4, with* $\alpha = 1$.
2) *Assume that* $(z^j)_{j \in \mathbb{N}}$ *is an infinite sequence of solution for Problems 5. Then, there exists a subsequence* $(z^{j_\ell})_{\ell \in \mathbb{N}}$ *that converges to a point* $\bar{z}$ *satisfying the KKT conditions related to Problem 4, with* $\alpha \geq 0$.

The convergence behavior of (**CC-SCP**) is completely described by Theorem 2 when the parameter ω takes values in $[\omega_0, \omega_{\max}]$. Due to numerical errors, in practice we are always led to consider the second case of Theorem 2, where usually the entire sequence $(z^j)_{j \in \mathbb{N}}$ converge. It is then useful to state that limiting points satisfy necessary conditions for optimality related to the original nonlinear, nonconvex formulation. Importantly, since every feasible point for Problem 4 satisfies LICQ with probability one (see [30, Theorem 1]), the point $\bar{z}$ whose existence is guaranteed by Theorem 2 is a candidate optimum for Problem 4.

Proof. The proof makes use of classical arguments in optimization. Due to space constraints, we refer to the Appendix.

In the following section, we demonstrate our approach on an uncertain 6 degrees of freedom spacecraft system navigating on-board the International Space Station (ISS).

VII. RESULTS: 6 DoF FREE-FLYER SPACECRAFT

We consider the guidance problem of a free-flyer robot, whose state is given by $\boldsymbol{x} := [\mathbf{p}, \mathbf{v}, \mathbf{q}, \boldsymbol{\omega}] \in \mathbb{R}^{13}$ and its control inputs are $\boldsymbol{u} := [\mathbf{F}, \mathbf{M}] \in \mathbb{R}^6$. We further define

$$\mathbf{S}(\boldsymbol{\omega}) := \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix}, \ \ \boldsymbol{\Omega}(\boldsymbol{\omega}) := \begin{bmatrix} 0 & -\omega_x & -\omega_y & -\omega_z \\ \omega_x & 0 & \omega_z & -\omega_y \\ \omega_y & -\omega_z & 0 & \omega_x \\ \omega_z & \omega_y & -\omega_x & 0 \end{bmatrix}. \quad (34)$$

The continuous time dynamics of the system $\dot{\boldsymbol{x}} = \mathbf{f}_t(\cdot)$ are

$$\dot{\mathbf{p}} = \mathbf{v}, \ m\dot{\mathbf{v}} = \mathbf{F}, \ \dot{\mathbf{q}} = \frac{1}{2}\boldsymbol{\Omega}(\boldsymbol{\omega})\mathbf{q}, \ \mathbf{J}\dot{\boldsymbol{\omega}} = \mathbf{M} - \mathbf{S}(\boldsymbol{\omega})\mathbf{J}\boldsymbol{\omega}. \quad (35)$$

Using a zero-order hold on the controls, we discretize (35)

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \Delta t \cdot \mathbf{f}_t(\boldsymbol{x}_k, \boldsymbol{\nu}_k, m, \mathbf{J}) + \boldsymbol{\epsilon}_k, \quad (36)$$

where $\boldsymbol{\epsilon}_k \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_\epsilon)$ are i.i.d. disturbances. As in [31], m and $J = \mathrm{diag}([J_x, J_y, J_z])$ are distributed according to

$$m \sim \mathcal{N}(7.2, (3.2)^2), \quad J_x, J_y, J_z \sim \mathcal{N}(0.07, (0.015)^2). \quad (37)$$

We consider the problem of planning a trajectory on-board the ISS from $(\mathbf{p}_{\text{init}}, \mathbf{q}_{\text{init}}) = ([9.2, 0, 5], \sqrt{1/3} \cdot [1, 0, 1, 1])$ to $(\mathbf{p}_g, \mathbf{q}_g) = ([11.3, 6, 4.5], [-0.5, 0.5, -0.5, 0.5])$, with zero initial and final velocities $\mathbf{v}, \boldsymbol{\omega}$ for $p = 90\%$. The ISS is defined using 26 polytopes (not shown in Fig. 3) and obstacles of arbitrary shape are placed in the station. Since the total number of obstacles is 30 for this scenario, obstacle avoidance chance constraints are reformulated using ellipsoidal confidence sets with (26) rather than (25), which would actually cause the problem to be infeasible. Linear chance constraints are used to enforce $|v_i| < v_m$, $|\omega_i| < \omega_m$, $|F_i| < F_m$ and $|M_i| < M_m$, with $(v_m, \omega_m, F_m, M_m) = (0.4, 0.8, 0.7, 0.1)$ corresponding to hardware limits of the spacecraft. A linear-quadratic regulator is used to reduce the growth of uncertainty, computed using the linearized dynamics at each state.
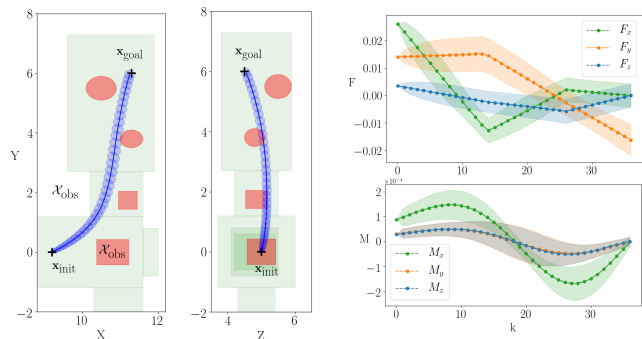

Fig. 3: *Left: projections of the trajectory computed by (CC-SCP), avoiding all spherical and polygonal obstacles on-board the ISS. Right: 90% confidence intervals of the control input trajectory.*

The algorithm is implemented in Python[3], and OSQP [32] is used to solve each convex problem. (**CC-SCP**) is initialized with an infeasible straight line trajectory from $\boldsymbol{x}_{\text{init}}$ to $\boldsymbol{x}_g$. Nevertheless, convergence is achieved after 8 iterations. To verify the satisfaction of chance constraints, 10'000 Monte-Carlo simulations are run. All chance constraints are satisfied at their threshold $p = 90\%$. In particular,

---

[3]The code is available at github.com/StanfordASL/ccscp.

$\min_k\{\Pr(\boldsymbol{x}_k \in \mathcal{X}_{\text{free}})\} = 95.4\%$, indicating that the Gaussian parameterization of the trajectory distribution is reasonable. In comparison, running the algorithm without considering uncertainty yields $\min_k\{\Pr(\boldsymbol{x}_k \in \mathcal{X}_{\text{free}})\} = 60.5\%$.

An alternative approach avoiding a chance-constrained problem formulation consists of using buffers to tighten constraints. To show that this deterministic approach can be difficult to tune in practice and can yield higher cost trajectories, we run experiments on 10 different scenarios with obstacles sampled randomly. For each scenario, we first run (**CC-SCP**) with $p = 95\%$, which we compare with the variant neglecting uncertainty but using 6 different buffer radius to inflate all obstacles. As shown in Figure 4, choosing an appropriate clearance can lead to conservative trajectories or violate chance constraints. Importantly, its value for each constraint is problem dependent and unknown a-priori.
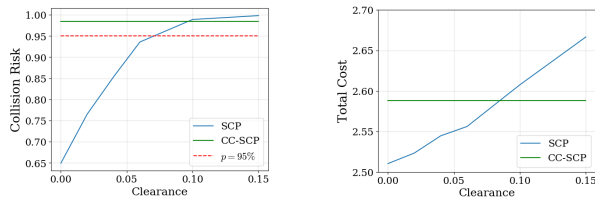

*Fig. 4: Results comparing (**CC-SCP**) with the mean-equivalent SCP approach with obstacle padding.*

VIII. Conclusion

In this paper, we presented a trajectory optimization algorithm based on sequential convex programming designed to solve chance-constrained planning problems. The proposed method is independent of the shape of obstacles and is guaranteed to converge to a solution satisfying first-order necessary optimality conditions.

Future work includes leveraging different uncertainty propagation schemes to theoretically guarantee the satisfaction of chance constraints, and exploring risk allocation methods to further reduce the conservatism of our approach.

IX. Acknowledgement

References

[1] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*, 1st ed. MIT Press, 2006.

[2] J. Harrison, A. Sharma, and M. Pavone, "Meta-learning priors for efficient online bayesian regression," in *Workshop on Algorithmic Foundations of Robotics*, 2018.

[3] A. Majumdar and M. Pavone, "How should a robot assess risk? Towards an axiomatic theory of risk in robotics," in *Int. Symp. on Robotics Research*, 2017.

[4] A. Bemporad and M. Morari, "Robust model predictive control: A survey," in *Robustness in identification and control*, 1999.

[5] G. Calafiore and M. Campi, "Uncertain convex programs: randomized solutions and confidence levels," *Mathematical Programming*, vol. 102, no. 1, pp. 25–46, 2005.

[6] N. A. and S. A., "Convex approximations of chance constrained programs," *SIAM Journal on Optimization*, vol. 17, no. 4, pp. 969–996, 2006.

[7] Z. Zhao and M. Kumar, "Split-bernstein approach to chance-constrained optimal control," *AIAA Journal of Guidance, Control, and Dynamics*, vol. 40, no. 11, pp. 2782–2795, 2017.

[8] L. J. Hong, Y. Yang, and L. Zhang, "Sequential convex approximations to joint chance constrained programs: A monte carlo approach," *Operations Research*, vol. 59, no. 3, pp. 617–630, 2011.

[9] L. Janson, E. Schmerling, and M. Pavone, "Monte Carlo motion planning for robot trajectory optimization under uncertainty," in *Int. Symp. on Robotics Research*, 2015.

[10] M. Ono, M. Pavone, Y. Kuwata, and J. Balaram, "Chance-constrained dynamic programming with application to risk-aware robotic space exploration," *Autonomous Robots*, vol. 39, no. 4, 2015.

[11] T. Lew, F. Lyck, and G. Müller, "Chance-constrained optimal altitude control of a rocket," *European Conf. for Aeronautics and Space Sciences*, 2019.

[12] G. C. Calafiore and L. E. Ghaoui, "On Distributionally Robust Chance-Constrained Linear Programs," *Journal of Optimization Theory & Applications*, vol. 130, no. 1, pp. 1–22, 2006.

[13] A. Prékopa, *Stochastic progamming*, 1st ed. Springer Netherlands, 1995.

[14] L. Hewing, A. Carron, K. P. Wabersich, and M. N. Zeilinger, "On a correspondence between probabilistic and robust invariant sets for linear systems," in *European Control Conference*, 2018.

[15] 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.

[16] D. Strawser and B. Williams, "Approximate Branch and Bound for Fast, Risk-Bound Stochastic Path Planning," in *Proc. IEEE Conf. on Robotics and Automation*, 2018.

[17] L. Hewing and M. N. Zeilinger, "Stochastic Model Predictive Control for Linear Systems using Probabilistic Reachable Sets," in *Proc. IEEE Conf. on Decision and Control*, 2018.

[18] T. Koller, F. Berkenkamp, M. Turchetta, and A. Krause, "Learning-based model predictive control for safe exploration," in *Proc. IEEE Conf. on Decision and Control*, 2018.

[19] B. Landry, R. Deits, P. R. Florence, and R. Tedrake, "Aggressive quadrotor flight through cluttered environments using mixed integer programming," in *Proc. IEEE Conf. on Robotics and Automation*, 2016.

[20] 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.

[21] J. Schulman, Y. Duan, J. Ho, A. Lee, I. Awwal, H. Bradlow, J. Pan, S. Patil, K. Goldberg, and P. Abbeel, "Motion planning with sequential convex optimization and convex collision checking," *Int. Journal of Robotics Research*, vol. 33, no. 9, pp. 1251–1270, 2014.

[22] R. Bonalli, A. Bylard, A. Cauligi, T. Lew, and M. Pavone, "Trajectory optimization on manifolds: A theoretically-guaranteed embedded sequential convex programming approach," in *Robotics: Science and Systems*, 2019.

[23] R. Bonalli, A. Cauligi, A. Bylard, and M. Pavone, "GuSTO: guaranteed sequential trajectory optimization via sequential convex programming," in *Proc. IEEE Conf. on Robotics and Automation*, 2019.

[24] X. Liu and P. Lu, "Solving nonconvex optimal control problems by convex optimization," *AIAA Journal of Guidance, Control, and Dynamics*, vol. 37, no. 3, pp. 750 – 765, 2014.

[25] Y. Mao, M. Szmuk, and B. Açikmeşe, "Successive convexification of non-convex optimal control problems and its convergence properties," in *Proc. IEEE Conf. on Decision and Control*, 2016.

[26] Y. Mao, D. Dueri, M. Szmuk, and B. Açikmeşe, "Successive Convexification of Non-Convex Optimal Control Problems with State Constraints," *IFAC-Papers Online*, vol. 50, no. 1, pp. 4063–4069, 2017.

[27] S. Patil, J. van den Berg, and R. Alterovitz, "Estimating probability of collision for safe motion planning under Gaussian motion and sensing uncertainty," in *Proc. IEEE Conf. on Robotics and Automation*, 2012.

[28] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed. Springer New York, 2006.

[29] M. Hirsch, *Differential Topology*, 1st ed. Springer, 1994.

[30] J. E. Spingarn and R. T. Rockafellar, "The Generic Nature of Optimality Conditions in Nonlinear Programming," *Mathematics of Operations Research*, vol. 4, no. 4, pp. 425–430, 1979.

[31] M. Ekal and R. Ventura, "On the accuracy of inertial parameter estimation of a free-flying robot while grasping an object," *Journal of Intelligent & Robotic Systems*, 2019.

[32] B. Stellato, G. Banjac, P. Goulart, A. Bemporad, and S. Boyd, "OSQP: An operator splitting solver for quadratic programs," 2017, Available at https://arxiv.org/abs/1711.08013.

[33] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge Univ. Press, 2004.

[34] A. A. Ahmadi, "Convex and Conic Optimization," Lecture notes, Spring 2016.

APPENDIX

A. Proof of Theorem 2 (Convergence guarantees)

Proof. The first assertion is a particular case of the second one, where Theorem 1 is used instead of its Corollary 1. Therefore, we only need to prove the second assertion.

From the Bolzano-Weierstrass theorem, the boundedness assumption in Assumption 2 allows us to extract a subsequence $(z^{j_\ell})_{\ell \in \mathbb{N}}$ from $(z^j)_{j \in \mathbb{N}}$ that converges to a point $\bar{z}$. When considering iteration $(j_\ell + 1)$, evaluating the constraints of Problem 5 at its solution $z^{j_\ell+1}$ gives

$$h(z^{j_\ell}) + \nabla h(z^{j_\ell}) \Delta z^{j_\ell+1} = 0 \tag{38}$$

$$g(z^{j_\ell}) + \nabla g(z^{j_\ell}) \Delta z^{j_\ell+1} \leq 0. \tag{39}$$

Thanks to Assumption 3, the quantities $\Delta z^{j_\ell+1}$ tend to 0 when ℓ tends to infinity. Under Assumption 1, passing to the limit in (38) gives that $\bar{z}$ is feasible for Problem 4.

Consider now Assumption 2. Then, Corollary 1 applies to the subsequence $(z^{j_\ell})_{\ell \in \mathbb{N}}$ guaranteeing the existence of multipliers $(\alpha^\ell, \lambda^\ell, \zeta^\ell)$ for Problems 5 associated to each z^{j_ℓ}, respectively for every $\ell \in \mathbb{N}$, and satisfying

$$\alpha^\ell \left(\nabla c(z^{j_\ell}) + 2 \frac{d\varphi}{dt} (\|\Delta z^{j_\ell+1}\|^2 - \Delta_{j_\ell}) z^{j_\ell+1} \right) \tag{40}$$

$$+ (\lambda^\ell)^\top \nabla h(z^{j_\ell}) + (\zeta^\ell)^\top \nabla g(z^{j_\ell}) = 0 \tag{41}$$

$$\zeta_i^\ell \left(g_i(z^{j_\ell}) + \nabla g_i(z^{j_\ell}) \Delta z^{j_\ell+1} \right) = 0, \tag{42}$$

$$\zeta_i^\ell \geq 0, \ i = 1, \ldots, s, \quad \|(\alpha^\ell, \lambda^\ell, \zeta^\ell)\| = 1, \ \alpha^\ell \geq 0. \tag{43}$$

Since the sequence $((\alpha^\ell, \lambda^\ell, \zeta^\ell))_{\ell \in \mathbb{N}}$ is bounded, up to some subsequence we can assume that it converges to a tuple (α, λ, ζ) satisfying $\|(\alpha, \lambda, \zeta)\| = 1$, $\zeta \geq 0$ and $\alpha \geq 0$. In particular, we have that (α, λ, ζ) is nontrivial. By passing to the limit in (40) we obtain that $\bar{z}$ satisfies (32) related to Problem 4 and the conclusion follows. $\square$

B. Full Formulation of (CC-LOCP)

In this section, we formulate the full linearized problem of (**CC-OCP**). At iteration $(j + 1)$, each nonconvex term is linearized around $z^j = (\mu^j, \nu^j)$. For instance, each nonconvex cost terms $\tilde{l}_f(\cdot), \tilde{l}_k(\cdot)$ and nonlinear dynamics $\tilde{f}(\cdot)$ are approximated using a first-order Taylor expansion as

$$l_k^j(\mu_k, \nu_k) := \tilde{l}(\mu_k^j, \nu_k^j) + \nabla_{\mu\nu} \tilde{l}(\mu_k^j, \nu_k^j) \begin{bmatrix} \mu_k - \mu_k^j \\ \nu_k - \nu_k^j \end{bmatrix} \tag{44}$$

$$l_f^j(\mu_N) := \tilde{l}_f(\mu_N^j) + \nabla_\mu \tilde{l}_f(\mu_N^j) \left(\mu_k - \mu_k^j \right) \tag{45}$$

$$\tilde{f}(\mu_k, \nu_k) \approx f_k^j + \nabla_\mu f_k^j \left(\mu_k - \mu_k^j \right) + \nabla_\nu f_k^j \left(\nu_k - \nu_k^j \right), \tag{46}$$

where $f_k^j := \tilde{f}(\mu_k^j, \nu_k^j)$. Note that if the original cost functions $\tilde{l}_f(\cdot), \tilde{l}_k(\cdot)$ are convex, linearizing these terms is not necessary.

Then, combining all convexified terms and using the problem formulation from Sections IV and V, the convexified problem (Problem 3) is given as

Chance-Constrained Linearized OCP (CC-LOCP)

$$\min_{\mu,\nu} \ l_f^j(\mu_N) + \sum_{k=0}^{N-1} l_k^j(\mu_k, \nu_k) + \tag{47a}$$

$$\omega_j \left(\sum_{k=1}^{N} \varphi(\|\mu_k - \mu_k^j\|^2 - \Delta_j) + \sum_{k=0}^{N-1} \varphi(\|\nu_k - \nu_k^j\|^2 - \Delta_j) \right)$$

$$\text{s.t.} \quad \mu_{k+1} = f_k^j + \nabla_\mu f_k^j \left(\mu_k - \mu_k^j \right) + \nabla_\nu f_k^j \left(\nu_k - \nu_k^j \right) \tag{47b}$$
$$k = 0, \ldots, N-1$$

$$\mathbf{H}_i \cdot \mu_k + \Phi_{\delta_x^l}^{-1} \left(\|\mathbf{H}_{i\cdot}^T\|_{\Sigma_k^j} + \nabla \|\mathbf{H}_{i\cdot}^T\|_{\Sigma_k^j} \cdot (z - z^j) \right) \leq h_i,$$
$$i = 1, \ldots, l, \ k = 1, \ldots, N \tag{47c}$$

$$d_i(\mu_k^j) + \nabla d_i(\mu_k^j)(\mu_k - \mu_k^j)$$
$$- \Phi_{\delta_x}^{-1} \left(\|\mathbf{n}_{ik}^j\|_{\Sigma_k^j} + \nabla \|\mathbf{n}_{ik}^j\|_{\Sigma_k^j} \cdot (z - z^j) \right) \geq 0, \tag{47d}$$
$$i = 1, \ldots, M, \ k = 1, \ldots, N-1$$

$$\mathbf{P}_i \cdot \mu_N + \Phi_{\delta_N}^{-1} \left(\|\mathbf{P}_{i\cdot}\|_{\Sigma_N^j} + \nabla \|\mathbf{P}_{i\cdot}\|_{\Sigma_N^j} (z - z^j) \right) \leq p_i, \tag{47e}$$
$$i = 1, \ldots, q$$

$$\mathbf{G}_i^u \cdot \nu_k + \Phi_{\delta_u}^{-1} \left(\|\mathbf{G}_{i\cdot}^u \mathbf{K}\|_{\Sigma_k^j} + \nabla \|\mathbf{G}_{i\cdot}^u \mathbf{K}\|_{\Sigma_k^j} \cdot (z - z^j) \right) \leq b_i,$$
$$i = 1, \ldots, r, \ k = 0, \ldots, N-1 \tag{47f}$$

$$\mu_0 = x(0), \quad \Sigma_0 = 0_{n^2}, \tag{47g}$$

where $\Sigma_k^j = \Sigma_k(\mu_k^j)$. Note that the chance constraints in (**CC-OCP**) are also linearized. Denoting $\mathbf{a}(z)$: $\text{dom}(z) \to \mathbb{R}^n$, $\Sigma(z)$: $\text{dom}(z) \to \mathbb{R}^{n \times n}$, with $\Sigma(z) \geq 0 \ \forall z$, the gradients in (47c)-(47f) can be computed as

$$\nabla_z (\|\mathbf{a}\|_\Sigma) = \frac{1}{2\sqrt{\mathbf{a}^T \Sigma \mathbf{a}}} \left(2\mathbf{a}^T \Sigma \nabla_z \mathbf{a} + \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j \nabla_z \Sigma_{ij} \right),$$

where $\mathbf{a} := \mathbf{a}(z)$ and $\Sigma := \Sigma(z)$ for conciseness. As discussed in Section VI-C, the gradients $\nabla \Sigma_{ij}$ and $\nabla \mathbf{a}$ are C^1 for the classes of functions considered in this work. In practice, auto-differentiation packages can be used to efficiently compute the gradient of the constraints directly.

C. Approximate Uncertainty Propagation

In this section, we derive the expressions in (4) for the first moments of the state distribution $x_k \sim \mathcal{N}(\mu_k, \Sigma_k)$. The dynamics are given as $x_{k+1} = f(x_k, u_k(x_k), w_k)$, the control policy is $u_k(x_k) = \nu_k + \mathbf{K}(x_k)(x_k - \mu_k)$, $\mathbb{E}(w_k) = \bar{w}$ and $\text{Var}(w_k) = \Sigma_w$. We assume that x_k, w_k are independent for all k.

Denoting $(x, u, w) = (x_k, u_k, w_k)$, and using a first-order Taylor approximation at $(\mu, \nu, \bar{w})$, we obtain

$$x_{k+1} \approx f(\mu, \nu, \bar{w}) + (\nabla_x f(x, u(x), w))|_{(\mu,\nu,\bar{w})} \cdot (x - \mu)$$
$$+ (\nabla_w f(x, u(x), w))|_{(\mu,\nu,\bar{w})} \cdot (w - \bar{w}),$$

where

$$(\nabla_x f(x, u(x), w))|_{(\mu,\nu,\bar{w})} =$$
$$(\nabla_x f(x, u, w) + \nabla_u f(x, u, w) \nabla_x(u(x)))|_{(\mu,\nu,\bar{w})}.$$

We note that $\nabla_x(u(x)) = \nabla_x \mathbf{K}(x)(x - \mu) + \mathbf{K}(x)$. Thus, after evaluating the above at $(\mu, \nu, \bar{w})$, we obtain

$$\nabla_x f(\mu, \nu, \bar{w}) = \nabla_\mu f(\mu, \nu, \bar{w}) + \nabla_\nu f(\mu, \nu, \bar{w}) \mathbf{K}(\mu).$$

From the linearized expression computed above, we obtain

$$\boldsymbol{\mu}_{k+1} = \mathbb{E}(\boldsymbol{x}_{k+1}) = \boldsymbol{f}(\boldsymbol{\mu}_k, \boldsymbol{\nu}_k, \mathbb{E}(\boldsymbol{w})),$$
$$\boldsymbol{\Sigma}_{k+1} = \text{Var}(\boldsymbol{x}_{k+1}) = \mathbb{E}((\boldsymbol{x}_{k+1} - \boldsymbol{\mu}_{k+1})(\boldsymbol{x}_{k+1} - \boldsymbol{\mu}_{k+1})^T)$$
$$= \nabla_x \boldsymbol{f}_{\text{cl}} \cdot \boldsymbol{\Sigma}_k \cdot \nabla_x \boldsymbol{f}_{\text{cl}}^T + \nabla_w \boldsymbol{f} \cdot \boldsymbol{\Sigma}_w \cdot \nabla_w \boldsymbol{f}^T,$$

where $\nabla_x \boldsymbol{f}_{\text{cl}} = (\nabla_x \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}(\boldsymbol{x}), \boldsymbol{w}))|_{(\boldsymbol{\mu}, \boldsymbol{\nu}, \bar{\boldsymbol{w}})}$, and the last result is obtained after using the fact that $\mathbb{E}((\boldsymbol{x} - \boldsymbol{\mu})(\boldsymbol{w} - \bar{\boldsymbol{w}})^T) = \boldsymbol{0}$ since $\boldsymbol{x}$ and $\boldsymbol{w}$ are uncorrelated.

Similarly, to obtain the approximation $\boldsymbol{u} \sim \mathcal{N}(\boldsymbol{\nu}, \mathbf{K}\boldsymbol{\Sigma}\mathbf{K})$ used in (15), we compute

$$\boldsymbol{u}(\boldsymbol{x}) \approx \boldsymbol{\nu} + (\nabla_x(\boldsymbol{u}(\boldsymbol{x})))|_{(\boldsymbol{\mu}, \boldsymbol{\nu}, \bar{\boldsymbol{w}})}(\boldsymbol{x} - \boldsymbol{\mu})$$
$$= \boldsymbol{\nu} + \mathbf{K}(\boldsymbol{\mu})(\boldsymbol{x} - \boldsymbol{\mu}),$$

from which we derive $\text{Var}(\boldsymbol{u}_k) = \mathbf{K}(\boldsymbol{\mu})\boldsymbol{\Sigma}_k(\boldsymbol{\mu}, \boldsymbol{\nu}, \bar{\boldsymbol{w}})\mathbf{K}(\boldsymbol{\mu})$.

D. Reformulation of Obstacle Avoidance Constraints

Proposition 1. (first statement)

Define $\boldsymbol{x}, \boldsymbol{x}^j \in \mathbb{R}^n$, $\mathcal{X}_{\text{free}} = \mathbb{R}^n \backslash \mathcal{X}_{\text{obs}}$, $d_{\min} \geq 0$ a safety margin, $d_i^j = d_i(\boldsymbol{x}^j)$, and $\mathbf{n}_i^j = (\boldsymbol{x}^j - \boldsymbol{x}_{\text{obs}})/d_i^j$.

If $\mathcal{X}_{\text{obs}} = \bigcup_{i=1}^M \mathcal{O}_i$, $\mathcal{O}_i \subset \mathbb{R}^n$ are closed convex sets, and

$$d_i^j + \mathbf{n}_i^{jT}(\boldsymbol{x} - \boldsymbol{x}^j) \geq d_{\min} \quad \forall i = 1, \dots, M, \qquad (48)$$

then $\boldsymbol{x} \in \mathcal{X}_{\text{free}}$.

Proof. This proof is inspired from the proof of the Hyperplane Separating Theorem [33] in [34]. First, we show by contradiction that $\forall \boldsymbol{x} \in \mathcal{O}_i$, the linearized constraint $g_i(\boldsymbol{x}) := d_i^j + \mathbf{n}_i^{jT}(\boldsymbol{x} - \boldsymbol{x}^j) - d_{\min} \geq 0$ is not satisfied. Hence, satisfaction of all constraints implies that $\boldsymbol{x} \notin \mathcal{X}_{\text{obs}}$. Finally, we conclude that $\boldsymbol{x}$ must lie in the safe set $\mathcal{X}_{safe}$.

Without loss of generality, assume that $\boldsymbol{x}^j \notin \mathcal{O}_i$[4], such that $d_i^j = \|\boldsymbol{x}^j - \boldsymbol{x}_{\text{obs}}\|$ and $\mathbf{n}_i^j = \frac{(\boldsymbol{x}^j - \boldsymbol{x}_{\text{obs}})}{\|\boldsymbol{x}^j - \boldsymbol{x}_{\text{obs}}\|}$. Then, assume, for the sake of contradiction, that $\exists \bar{\boldsymbol{x}} \in \mathcal{O}_i$ s.t. $g_i(\bar{\boldsymbol{x}}) > 0$, i.e.

$$g_i(\bar{\boldsymbol{x}}) = d_i^j + \mathbf{n}_i^{jT}(\bar{\boldsymbol{x}} - \boldsymbol{x}^j) - d_{\min}$$
$$= \|\boldsymbol{x}^j - \boldsymbol{x}_{\text{obs}}\| + \frac{(\boldsymbol{x}^j - \boldsymbol{x}_{\text{obs}})^T}{\|\boldsymbol{x}^j - \boldsymbol{x}_{\text{obs}}\|}(\bar{\boldsymbol{x}} - \boldsymbol{x}^j) - d_{\min} > 0.$$

Define $d^2(\boldsymbol{x}) = \|\boldsymbol{x} - \boldsymbol{x}^j\|^2$. We claim that $(\bar{\boldsymbol{x}} - \boldsymbol{x}_{\text{obs}})$ is a descent direction for $d^2(\boldsymbol{x})$ at $\boldsymbol{x}_{\text{obs}}$. Indeed, using the inequality above and $d_{\min} \geq 0$,

$$\nabla d^2(\boldsymbol{x}_{\text{obs}})(\bar{\boldsymbol{x}} - \boldsymbol{x}_{\text{obs}}) = 2(\boldsymbol{x}_{\text{obs}} - \boldsymbol{x}^j)^T(\bar{\boldsymbol{x}} - \boldsymbol{x}_{\text{obs}})$$
$$= 2\left((\boldsymbol{x}_{\text{obs}} - \boldsymbol{x}^j)^T(\bar{\boldsymbol{x}} - \boldsymbol{x}^j) + (\boldsymbol{x}_{\text{obs}} - \boldsymbol{x}^j)^T(\boldsymbol{x}^j - \boldsymbol{x}_{\text{obs}})\right)$$
$$< 2\left(\|\boldsymbol{x}^j - \boldsymbol{x}_{\text{obs}}\|^2 - d_{\min}\|\boldsymbol{x}^j - \boldsymbol{x}_{\text{obs}}\| - \|\boldsymbol{x}^j - \boldsymbol{x}_{\text{obs}}\|^2\right)$$
$$= 2\left(-d_{\min}\|\boldsymbol{x}^j - \boldsymbol{x}_{\text{obs}}\|\right) < 0.$$

Hence, $\exists \lambda^+ > 0$ such that $\forall \lambda \in (0, \lambda^+)$, $\tilde{\boldsymbol{x}} = \boldsymbol{x}_{\text{obs}} + \lambda(\bar{\boldsymbol{x}} - \boldsymbol{x}_{\text{obs}})$ satisfies $\tilde{\boldsymbol{x}} \in \mathcal{O}_i$ (by convexity of $\mathcal{O}_i$) and

$$d^2(\tilde{\boldsymbol{x}}) < d^2(\boldsymbol{x}_{\text{obs}}).$$

---

[4]If $\boldsymbol{x}^j \in \mathcal{O}_i$, define $\tilde{\boldsymbol{x}}^j = \boldsymbol{x}^j + 2d_i^j \mathbf{n}_i^j$, the squared distance function $\tilde{d}^2(\boldsymbol{x}) = \|\boldsymbol{x} - \tilde{\boldsymbol{x}}^j\|^2$ and follow the same reasoning to reach the same conclusion.

Hence, $\boldsymbol{x}_{\text{obs}}$ is not the closest point in $\mathcal{O}_i$ to $\boldsymbol{x}^j$. This is a contradiction. Therefore, $g_i(\boldsymbol{x}) < 0, \forall \boldsymbol{x} \in \mathcal{O}_i$.

By negation,

$$d_i^j + \mathbf{n}_i^{jT}(\boldsymbol{x} - \boldsymbol{x}^j) \geq d_{\min} \implies \boldsymbol{x} \notin \mathcal{O}_i.$$

Then, since $\mathcal{X}_{\text{obs}} = \bigcup_{i=1}^M \mathcal{O}_i$, we have

$$d_i^j + \mathbf{n}_i^{jT}(\boldsymbol{x} - \boldsymbol{x}^j) \geq d_{\min} \ \forall i = 1, \dots, M \implies \boldsymbol{x} \notin \mathcal{X}_{\text{obs}}.$$

Finally, since $\mathcal{X}_{safe} = \mathbb{R}^n \setminus \mathcal{X}_{\text{obs}}$, we have that $\boldsymbol{x} \in \mathcal{X}_{safe} \iff \boldsymbol{x} \notin \mathcal{X}_{\text{obs}}$. Therefore,

$$d_i^j + \mathbf{n}_i^{jT}(\boldsymbol{x} - \boldsymbol{x}^j) \geq d_{\min} \ \forall i = 1, \dots, M \implies \boldsymbol{x} \in \mathcal{X}_{safe}.$$

which concludes the proof. $\square$

Note that as shown on Figure 5, this result does not hold in the case of non-convex obstacles. A solution consists of decomposing each non-convex obstacle $\mathcal{O}_i$ into convex sets which union includes $\mathcal{O}_i$, or to use a convex outer approximation.
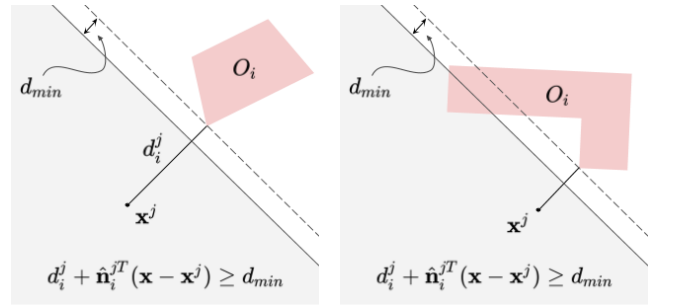
Fig. 5: Obstacle constraints convex reformulation using the signed distance function (21). On the left, this is conservative since the obstacle is convex. On the right, it is not.

E. Differentiability of the Signed Distance Function

Theorem 3. *Let $D \subseteq \mathbb{R}^3$ be either a plane or a disc, whose center is denoted by $\boldsymbol{c}$. Consider the signed distance function*

$$d_D^s : \mathcal{A} \to \mathbb{R} : \boldsymbol{x} \mapsto \begin{cases} -d_{\partial D}(\boldsymbol{x}) & \text{if } \boldsymbol{x} \in D \\ d_{\partial D}(\boldsymbol{x}) & \text{otherwise}, \end{cases}$$

where $d_S(\boldsymbol{x}) = \inf_{\boldsymbol{y} \in S} \|\boldsymbol{x} - \boldsymbol{y}\|$, and $\mathcal{A} = \mathbb{R}^3$ in the case that D is a plane whereas $\mathcal{A} = \mathbb{R}^3 \setminus \boldsymbol{c}$ in the case that D is a disc. The mapping d_D^s is of class C^∞.

Proof. For sake of conciseness and clarity of notation, we provide a proof in the case $D = D_1(0)\backslash 0$, i.e, the unitary disc centered at 0, for which $\partial D = S^2 := \{\boldsymbol{y} \in \mathbb{R}^3 \mid \|\boldsymbol{y}\| = 1\}$. The proof of the other cases goes similarly.

For every $\boldsymbol{x} \in \mathbb{R}^3 \setminus 0$, there exists a unique point $\boldsymbol{y}(\boldsymbol{x}) \in S^2$ such that $\|\boldsymbol{x} - \boldsymbol{y}(\boldsymbol{x})\| = d_{S^2}(\boldsymbol{x})$. From this, we have

$$\boldsymbol{x} = (1 + d_D^s(\boldsymbol{x}))\boldsymbol{y}(\boldsymbol{x}) \qquad (49)$$

which is satisfied for every $\boldsymbol{x} \in \mathbb{R}^3 \setminus 0$. We use relation (49) to show that d_D^s is a smooth function of $\boldsymbol{x}$ by leveraging the inverse function theorem locally around any $\boldsymbol{x} \in \mathbb{R}^3 \setminus 0$.

Fix $\boldsymbol{x}_0 \in \mathbb{R}^3 \setminus 0$ and let $\boldsymbol{y}_0 = \boldsymbol{y}(\boldsymbol{x}_0) \in S^2$ such that $\|\boldsymbol{x}_0 - \boldsymbol{y}_0\| = d_{S^2}(\boldsymbol{x}_0)$. We may assume that $\boldsymbol{y}_0 \neq \boldsymbol{N} := (0,0,1)$ (otherwise, we would have $\boldsymbol{y}_0 \neq \boldsymbol{S} := (0,0,-1)$ and the proof goes similarly). Consider the stereographic projection

$$\boldsymbol{\varphi} : S^2 \setminus \boldsymbol{N} \to \mathbb{R}^2 : (y_1, y_2, y_3) \mapsto \left(\frac{y_1}{1-y_3}, \frac{y_2}{1-y_3} \right)$$

which has C^∞ inverse

$$\boldsymbol{\varphi}^{-1} : \mathbb{R}^2 \to S^2 \setminus \boldsymbol{N}$$
$$(z_1, z_2) \mapsto \left(\frac{2z_1}{z_1^2 + z_2^2 + 1}, \frac{2z_2}{z_1^2 + z_2^2 + 1}, \frac{z_1^2 + z_2^2 - 1}{z_1^2 + z_2^2 + 1} \right),$$

and denote $(z_1^0, z_2^0) = \boldsymbol{\varphi}(\boldsymbol{y}_0)$, $d_0 = d_{S^2}(\boldsymbol{x}_0) = \|\boldsymbol{x}_0 - \boldsymbol{y}_0\|$. Then, define the C^∞ vector function

$$\boldsymbol{F} : \mathcal{U} \to \mathbb{R}^3 : (z_1, z_2, d) \mapsto (1+d)\boldsymbol{\varphi}^{-1}(z_1, z_2)$$

where $\mathcal{U}$ is any neighborhood of (z_1^0, z_2^0, d_0) not containing $(0, 0, -1)$. Straightforward computations provide that

$$\det \nabla \boldsymbol{F}(z_1, z_2, d) = -\frac{4(1+d)^2}{(z_1^2 + z_2^2 + 1)^2} \neq 0$$

for every $(z_1, z_2, d) \in \mathcal{U}$. The inverse function theorem applied to $\boldsymbol{F}$ at (z_1^0, z_2^0, d_0) provides a neighborhood $\mathcal{V}_1$ of (z_1^0, z_2^0, d_0) and a neighborhood $\mathcal{V}_2$ of $\boldsymbol{x}_0$ such that the restriction $\boldsymbol{F}|_{\mathcal{V}_1} : \mathcal{V}_1 \to \mathcal{V}_2$ is a C^∞ diffeomorphism. We have that $d_D^s(\boldsymbol{x}) = \pi_3((\boldsymbol{F}|_{\mathcal{V}_1})^{-1}(\boldsymbol{x}))$ for every $\boldsymbol{x} \in \mathcal{V}_2$, where π_3 is the projection on the third component. Indeed, for $\boldsymbol{x} \in \mathcal{U}$, relation (49) gives $\boldsymbol{F}(\boldsymbol{\varphi}(\boldsymbol{y}(\boldsymbol{x})), d_D^s(\boldsymbol{x})) = \boldsymbol{x}$.

The conclusion follows. $\qquad\square$

F. Limit number of obstacles

As stated in Remark 2, using ellipsoidal confidence sets to reformulate chance constraints is usually more conservative than using equivalent reformulations for radial probability distributions. However, due to the risk allocation with Boole's inequality, this remark does not hold for a large number of constraints, such as an environment with multiple obstacles. For Gaussian random variables, it is possible to directly compare the conservatism of both approaches as in Figures 6, where the quantiles used to reformulate chance constraints used in (5) and (8) are compared. Values of the minimum number of linear chance constraints from which it is less conservative to use (8) than (5) are also reported.
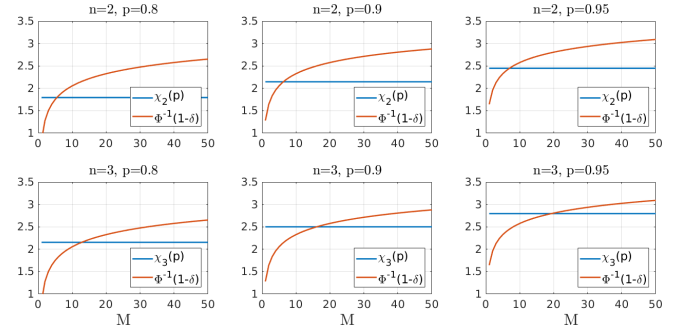

Fig. 6: *Comparison of the quantiles $\chi_n(p)$ and $\Phi^{-1}(1-\delta)$, where $\delta = \frac{1-p}{M}$, depending on M, the number of constraints.*

		Probability p			
		0.800	0.900	0.950	0.975
Dim. n	2	6	7	7	8
	3	13	17	20	23

TABLE I: *From Figure 6, it is possible to determine the number of linear chance constraints from which using ellipsoidal confidence sets with (8) is less conservative than leveraging Boole's inequality and the exact reformulation of linear chance constraints for Gaussian-distributed variables (5). For instance, for more than $M = 20$ obstacles for a 3-dimensional statespace with $p_x = 95\%$, using ellipsoidal confidence sets is less conservative.*