This article has been accepted for publication in IEEE Control Systems Letters. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/LCSYS.2023.3348688

Feasibility Analysis and Regularity Characterization of Distributionally Robust Safe Stabilizing Controllers

Pol Mestres Kehan Long Nikolay Atanasov Jorge Cortés

***Abstract*— This paper studies the well-posedness and regularity of safe stabilizing optimization-based controllers for control-affine systems in the presence of model uncertainty. When the system dynamics contain unknown parameters, a finite set of samples can be used to formulate distributionally robust versions of control barrier function and control Lyapunov function constraints. Control synthesis with such distributionally robust constraints can be achieved by solving a (convex) second-order cone program (SOCP). We provide one necessary and two sufficient conditions to check the feasibility of such optimization problems, characterize their computational complexity and numerically show that they are significantly faster to check than direct use of SOCP solvers. Finally, we also analyze the regularity of the resulting control laws.**

***Index Terms*— Safety-critical control, control barrier functions, distributionally robust control synthesis.**

I. Introduction

RECENT years have seen increasing deployment of control systems and robots to aid transportation, warehouse management, and home automation. In these applications, it is crucial to implement controllers with provable safety and stability guarantees despite uncertainty in the system models and operational conditions. Recent work [1]–[6] tackles this when some prior information about the uncertainty is known. Instead, here we rely on a line of work initiated in [7] that circumvents the need for knowledge about the uncertainty distribution and uses only uncertainty samples to formulate distributionally robust constraints for control synthesis. This approach is robust to distributional shift at deployment time and enjoys provable out-of-sample performance. However, it also introduces several challenges, which we focus on here: the characterization of the quality and number of uncertainty samples needed to guarantee the feasibility of the safety and stability constraints, and the study of the regularity properties of the resulting controllers.

Literature Review: Control Lyapunov functions (CLFs) [8] are a well-established tool to design stabilizing controllers for nonlinear systems. More recently, control barrier functions (CBFs) [9] have gained popularity as a tool to render a desired subset of the system state space safe. If the system is control affine, CLF and CBF constraints are linear in the control input and can be incorporated in a quadratic program (QP) [10] that, if feasible, can be solved efficiently to obtain control inputs guaranteeing safety and stability. Recent work has explored alternative optimization formulations when the system model is uncertain. Under the assumption that the uncertainty follows a Gaussian Process (GP) or satisfies worst-case bounds, [1]–[3], [5], [11], [12] formulate second-order cone constraints that can be used to design controllers achieving safe stabilization of the true system. The paper [4] gives sufficient conditions for the feasibility of such second-order cone constraints. Our work here is closely related to [7], which leverages ideas from distributionally robust optimization (DRO) [13], [14] to model the uncertainty. The DRO framework constructs an *ambiguity set* of probability distributions that contains the true (unknown) one with high confidence. Such ambiguity sets are constructed with only finitely many samples and are used to formulate distributionally robust versions of the control design problem.

Statement of Contributions: We study the problem of safe stabilization of control-affine systems under uncertainty. We assume that the distribution of the uncertainty is unknown and formulate a second-order cone program (SOCP) using distributionally robust versions of the CLF and CBF constraints constructed on the basis of uncertainty samples. Our first contribution is the derivation of a necessary condition and two sufficient conditions for the feasibility of the optimization problem. We characterize the computational complexity of these conditions and show that, for a large number of samples, it is significantly smaller than solving the SOCP directly, which makes them useful to efficiently check whether the problem is feasible without having to solve it. Our first sufficient condition is dependent on the quality of the uncertainty samples but is limited to a single control objective. Our second sufficient condition is only dependent on the number of samples but can be used for any number of constraints. Our final contribution shows that the solution of this distributionally robust optimization problem is point-Lipschitz, and hence continuous, which means that solutions of the closed-loop system are guaranteed to exist and the controller obtained from it can be implemented without inducing chattering.

II. Preliminaries

We review distrib. robust chance-constrained programs and control Lyapunov and barrier functions under uncertainty.

A. Distributionally Robust Chance Constrained Programs

Given a random vector $\boldsymbol{\xi}$ following distribution $\mathbb{P}^*$ supported on set $\Xi \subseteq \mathbb{R}^k$ and a closed convex set $\mathcal{Z} \subset \mathbb{R}^n$, let $G : \mathcal{Z} \times \Xi \to \mathbb{R}$ define a probabilistic constraint $G(z, \boldsymbol{\xi}) \leq 0$. We are interested in satisfying this constraint with a prescribed confidence $1 - \epsilon$, with $\epsilon \in (0, 1)$, while minimizing a convex

The authors are with the Contextual Robotics Institute, UC San Diego (e-mails: {pomestre,k3long,natanasov,cortes}@ucsd.edu).

The authors gratefully acknowledge support from NSF under grants IIS-2007141 and CCF-2112665.

objective function $c : \mathcal{Z} \to \mathbb{R}$. To achieve this[1], define the chance-constrained program:

$$\begin{aligned} &\min_{z\in\mathcal{Z}} c(z) \\ &\text{s.t. } \mathbb{P}^*(G(z,\boldsymbol{\xi}) \le 0) \ge 1-\epsilon. \end{aligned} \tag{1}$$

The feasible set of (1) is not convex in general. Nemirosvski and Shapiro [15, Section 2] propose a convex approximation of the feasible set of (1) by replacing the chance constraint with a conditional value-at-risk (CVaR) constraint. CVaR of $G(z,\boldsymbol{\xi})$ can be formulated as the following convex program:

$$\mathrm{CVaR}^{\mathbb{P}^*}_{1-\epsilon}(G(z,\boldsymbol{\xi})) := \inf_{t\in\mathbb{R}} [\epsilon^{-1}\mathbb{E}_{\mathbb{P}^*}[(G(z,\boldsymbol{\xi})+t)_+]-t]. \tag{2}$$

The resulting problem

$$\begin{aligned} &\min_{z\in\mathcal{Z}} c(z) \\ &\text{s.t. } \mathrm{CVaR}^{\mathbb{P}^*}_{1-\epsilon}(G(z,\boldsymbol{\xi})) \le 0, \end{aligned} \tag{3}$$

is convex and its feasible set is contained in that of (1).

Both (1) and (3) assume that $\mathbb{P}^*$ is known. Instead, suppose that it is unknown and we only have access to samples $\{\boldsymbol{\xi}_i\}_{i\in[N]}$ from $\mathbb{P}^*$. We describe a way of constructing a set of distributions that could have generated the samples. Let $\mathcal{P}_p(\Xi)$ be the set of probability measures with finite p-th moment supported on Ξ. Let $\hat{\mathbb{P}}_N := \frac{1}{N}\sum_{i=1}^N \delta_{\boldsymbol{\xi}_i}$ be the empirical distribution constructed from the samples $\{\boldsymbol{\xi}_i\}_{i=1}^N$. Let W_p be the p-Wasserstein distance [14, Definition 3.1] between two probability measures in $\mathcal{P}_p(\Xi)$ and let $\mathcal{M}_N^r := \{\mu \in \mathcal{P}_p(\Xi) : W_p(\mu, \hat{\mathbb{P}}_N) \le r\}$ be the ball of radius r centered at $\hat{\mathbb{P}}_N$. We define a distributionally robust chance-constrained program:

$$\begin{aligned} &\min_{z\in\mathcal{Z}} c(z) \\ &\text{s.t. } \inf_{\mathbb{P}\in\mathcal{M}_N^r} \mathbb{P}(G(z,\boldsymbol{\xi}) \le 0) \ge 1-\epsilon. \end{aligned} \tag{4}$$

We can use CVaR to obtain a convex conservative approximation of (4):

$$\begin{aligned} &\min_{z\in\mathcal{Z}} c(z) \\ &\text{s.t. } \sup_{\mathbb{P}\in\mathcal{M}_N^r} \mathrm{CVaR}^{\mathbb{P}}_{1-\epsilon}(G(z,\boldsymbol{\xi})) \le 0. \end{aligned} \tag{5}$$

If (5) is feasible, then (4) is also feasible [15, Section 2].

We say that a distribution $\mathbb{P}$ is light-tailed if there exists $a > 0$ such that $A := \mathbb{E}_{\mathbb{P}}[\exp \|\boldsymbol{\xi}\|^a] = \int_\Xi \exp \|\boldsymbol{\xi}\|^a \mathbb{P}(d\boldsymbol{\xi}) < \infty$. If $\mathbb{P}^*$ is light-tailed, the following observation specifies how the radius of $\mathcal{M}_N^r$ should be selected so that the true distribution lies in the ball with high confidence.

[1]We denote by $\mathbb{Z}_{>0}$, $\mathbb{R}$ and $\mathbb{R}_{\ge 0}$ the set of positive integers, real, and nonnegative real numbers, resp. We denote by $\mathbf{0}_n$ the n-dimensional zero vector. We write $\partial\mathcal{S}$ for the boundary of the set $\mathcal{S}$. Given $N \in \mathbb{Z}_{>0}$, we denote $[N] = \{1,\dots,N\}$. Given $x \in \mathbb{R}^n$, $\|x\|$ denotes the Euclidean norm of x. For $x \in \mathbb{R}$, we define $(x)_+ = \max(x, 0)$. A function $\beta : \mathbb{R}_{\ge 0} \to \mathbb{R}$ is of class $\mathcal{K}_\infty$ if $\beta(0) = 0$, β is strictly increasing and $\lim_{t\to\infty} \beta(t) = \infty$. A function $V : \mathbb{R}^n \to \mathbb{R}$ is positive definite if $V(0) = 0$ and $V(x) > 0$ for all $x \neq 0$, and proper in a set Γ if $\{x \in \Gamma : V(x) \le c\}$ is compact for any $c \ge 0$. Given an $m \times n$ matrix A and two integers i, j such that $1 \le i < j \le m$, $A_{i:j}$ denotes the $(j-i+1)\times n$ matrix obtained by selecting the rows from i to j of A. A function $f : \mathbb{R}^n \to \mathbb{R}^q$ is point-Lipschitz at a point $x_0 \in \mathbb{R}^n$ if there exists a neighborhood U of x_0 and a constant $L_{x_0} > 0$ such that $\|f(x) - f(x_0)\| \le L_{x_0}\|x - x_0\|$ for all $x \in U$.

Remark 2.1: (Choice of Wasserstein ball radius): If the true distribution $\mathbb{P}^*$ is light-tailed, the choice of $r = r_N(\bar{\epsilon})$ given in [14, Theorem 3.5],

$$r_N(\bar{\epsilon}) = \begin{cases} \left(\frac{\log(c_1\bar{\epsilon}^{-1})}{c_2 N}\right)^{\frac{1}{\max\{k,2\}}} & \text{if } N \ge \frac{\log(c_1\bar{\epsilon}^{-1})}{c_2}, \\ \left(\frac{\log(c_1\bar{\epsilon}^{-1})}{c_2 N}\right)^{\frac{1}{a}} & \text{else,} \end{cases} \tag{6}$$

where c_1, c_2 and a are positive constants that only depend on a, A and k (cf. [14, Theorem 3.4]), ensures that the ball $\mathcal{M}_N^{r_N(\bar{\epsilon})}$ contains $\mathbb{P}^*$ with probability at least $1-\bar{\epsilon}$. Then, a solution z^* of (5) satisfies the constraint $\mathrm{CVaR}^{\mathbb{P}^*}_{1-\epsilon}(G(z^*,\boldsymbol{\xi})) \le 0$ with probability at least $1-\bar{\epsilon}$. Note that c_1, c_2 and a can be computed by knowing the class of distributions to which $\mathbb{P}^*$ belongs to, without having actual knowledge of $\mathbb{P}^*$. If exact values are not known, but upper and lower bounds are, these can be used instead to compute an upper bound of $r_N(\bar{\epsilon})$. •

Remark 2.2: (Choice of ϵ): The parameter ϵ determines the confidence level $1-\epsilon$ for constraint satisfaction. Throughout the paper, we assume $\epsilon \le \frac{1}{N}$, albeit results are valid generally, with explicit expressions becoming more involved. •

B. Distributionally Robust Safety and Stability

The notions of CLF [8] and CBF [9] can be used to design controllers in uncertainty-free systems that enforce stability and safety, resp. Here we extend these notions for systems with uncertainty in the dynamics. Consider a nominal model F and a linear combination of k perturbations,

$$\dot{x} = (F(x) + \sum_{j=1}^k W_j(x)\xi_j)\underline{u}, \tag{7}$$

where for $1 \le j \le k$, $W_j(x) \in \mathbb{R}^{n\times(m+1)}$ denotes known model perturbations, and $\xi_j \in \mathbb{R}$ denotes the corresponding unknown weight, and $\underline{u} = [1; u] \in \underline{\mathcal{U}} := \{1\} \times \mathbb{R}^m$. We let $\boldsymbol{\xi} = [\xi_1, \xi_2, \dots, \xi_k]^T \in \Xi \subseteq \mathbb{R}^k$. We assume that $\boldsymbol{\xi}$ follows an unknown distribution $\mathbb{P}^*$ but a set of samples $\{\boldsymbol{\xi}_i\}_{i=1}^N$ is available. We are interested in extending the notions of CLF [8] and CBF [9] for systems of the form (7). To do so, note that as shown in [7, Section IV], the CBF condition for a system of the form (7) and a function $h : \mathbb{R}^n \to \mathbb{R}$ reads as $\mathrm{CBC}(x, \underline{u}, \boldsymbol{\xi}) := \underline{u}^T q_h(x) + \underline{u}^T R_h(x)\boldsymbol{\xi} \ge 0$, where the exact forms of q_h and R_h are given in [7, Section IV] and depend on h and its gradient. Now, since $\boldsymbol{\xi}$ follows a distribution $\mathbb{P}^*$, we extend the definition of CBF by requiring that for all x in the safe set, there exist $\underline{u} \in \underline{\mathcal{U}}$ such that

$$\mathbb{P}^*(\mathrm{CBC}(x, \underline{u}, \boldsymbol{\xi}) \ge 0) \ge 1-\epsilon. \tag{8}$$

The CLF condition for (7) takes a similar form and is written as $\mathrm{CLC}(x, \underline{u}, \boldsymbol{\xi}) \le 0$ (cf. [7, Section IV]). As shown in Section II-A, CVaR can be used as a convex approximation of (8) and its analogue with CLC. We use

$$\mathrm{CVaR}^{\mathbb{P}^*}_{1-\epsilon}(\mathrm{CBC}(x, \underline{u}, \boldsymbol{\xi})) \ge 0, \tag{9a}$$

$$\mathrm{CVaR}^{\mathbb{P}^*}_{1-\epsilon}(\mathrm{CLC}(x, \underline{u}, \boldsymbol{\xi})) \ge 0, \tag{9b}$$

as the distributionally robust analogues of the CLF and CBF conditions from [8] and [9], resp. The existence of a controller satisfying (9) implies the existence of a controller that makes the CLC (resp. the CBC) condition hold at every point with

probability at least $1-\epsilon$, paving the way for the design of controllers that make the system stable (resp. safe) with arbitrarily high probability.

III. Problem Statement

Consider the system model in (7) with distributional uncertainty, meaning that the true distribution $\mathbb{P}^*$ of the parameter $\boldsymbol{\xi}$ is unknown. We assume that the system admits a CLF and a CBF, which allow us to formulate the constraints (9). Given a nominal controller specified by a smooth function $\underline{k}:\mathbb{R}^n\to\underline{\mathcal{U}}$, we would like to synthesize a controller closest to it that respects safety and stability constraints. Using (2), this problem can be written in general form as

$$\min_{\underline{u}\in\underline{\mathcal{U}}} \|\underline{u}-\underline{k}(x)\|^2 \tag{10}$$
$$\text{s.t.} \sup_{\mathbb{P}\in\mathcal{M}_N^r} \inf_{t\in\mathbb{R}} [\epsilon^{-1}\mathbb{E}_{\mathbb{P}}[(G_l(x,\underline{u},\boldsymbol{\xi})+t)_+]-t]\le 0,\ \forall l\in[M],$$

where $M\in\mathbb{Z}_{>0}$ and each $G_l:\mathbb{R}^n\times\underline{\mathcal{U}}\times\Xi\to\mathbb{R}$ is an affine function in $\underline{u}$ and $\boldsymbol{\xi}$, $G_l(x,\underline{u},\boldsymbol{\xi})=\underline{u}^Tq_l(x)+\underline{u}^TR_l(x)\boldsymbol{\xi}$, for smooth functions $q_l:\mathbb{R}^n\to\mathbb{R}^{m+1}$ and $R_l:\mathbb{R}^n\to\mathbb{R}^{(m+1)\times k}$. With $M=2$ and constraints corresponding to CBC and CLC, this corresponds to a stable and safe control synthesis problem. The case $M=1$ with the constraint CBC corresponds to a distributionally robust version of a safety filter of $\underline{k}$.

Although the constraints in (10) are convex, the program is intractable due to the search of suprema over the Wasserstein set. Fortunately, [7, Proposition IV.1] shows that when $\Xi=\mathbb{R}^k$ and $p=1$, the following SOCP is equivalent to (10):

$$\min_{\underline{u}\in\underline{\mathcal{U}},y\in\mathbb{R},t\in\mathbb{R},s_i\in\mathbb{R}} y \tag{11a}$$
$$\text{s.t.}\quad r\left\|R_l^T(x)\underline{u}\right\|+\frac{1}{N}\sum_{i=1}^N s_i-t\epsilon\le 0,\ \forall l\in[M], \tag{11b}$$
$$s_i\ge G_l(x,\underline{u},\boldsymbol{\xi}_i)+t,\ \forall i\in[N],\ \forall l\in[M], \tag{11c}$$
$$s_i\ge 0,\ \forall i\in[N], \tag{11d}$$
$$y+1\ge\sqrt{\|2(\underline{u}-\underline{k}(x))\|^2+(y-1)^2}\ . \tag{11e}$$

We refer to (11) as the DRO-SOCP and take $\Xi=\mathbb{R}^k$ and $p=1$ Wasserstein distance throughout the paper.

A critical observation about problem (11) is that, in general, it might be infeasible, leading to controllers that are undefined. Furthermore, even if the problem is feasible, the controller obtained from it might not be continuous, hence resulting in implementation problems (it might induce chattering behavior when implemented on physical systems) and theoretical problems (lack of existence of solutions of the closed-loop system). Hence, our goal in this paper is twofold. First, we derive conditions to ensure the feasibility of (11). Given the complexity of obtaining characterizations for the feasibility of such problems, we focus on identifying conditions that are easy to evaluate computationally as opposed to directly attempting to solve the optimization problem: either sufficient conditions, to quickly ensure feasibility, or necessary, to quickly discard it. Second, assuming that the problem (11) is feasible, we characterize the regularity properties of the resulting controller.

IV. Feasibility Analysis

In this section, we study the feasibility properties of (11). We start by giving a necessary condition for its feasibility.

Proposition 4.1: (Necessary condition for feasibility of DRO-SOCP): Let $\epsilon\in(0,\frac{1}{N}]$ and $r>0$. For $x\in\mathbb{R}^n$, let

$$\bar{Q}_l(x)=rR_l(x)_{2:(m+1)}\in\mathbb{R}^{m\times k},\ \bar{r}_l(x)=rR_l(x)_1\in\mathbb{R}^{1\times k},$$
$$\bar{w}_{l,i}(x)=(-\epsilon q_l(x)-\epsilon R_l(x)\boldsymbol{\xi}_i)_{2:(m+1)}\in\mathbb{R}^m,$$
$$\bar{v}_{l,i}(x)=(-\epsilon q_l(x)-\epsilon R_l(x)\boldsymbol{\xi}_i)_1\in\mathbb{R},$$
$$\bar{F}_{l,i}(x)=\bar{Q}_l(x)\bar{Q}_l(x)^T-\bar{w}_{l,i}(x)\bar{w}_{l,i}(x)^T\in\mathbb{R}^{m\times m},$$
$$\bar{J}_{l,i}(x)=\bar{r}_l(x)\bar{Q}_l(x)^T-\bar{v}_{l,i}(x)\bar{w}_{l,i}^T\in\mathbb{R}^{1\times m},$$
$$\bar{H}_{l,i}(x)=\begin{pmatrix}(\bar{r}_l\bar{r}_l^T-\bar{v}_{l,i}^2)(x) & \bar{J}_{l,i}(x)\\ \bar{J}_{l,i}^T(x) & \bar{F}_{l,i}(x)\end{pmatrix}\in\mathbb{R}^{(m+1)\times(m+1)}$$

for $l\in[M]$ and $i\in[N]$. Let $\bar{\lambda}_{l,i}(x)$ be the minimum eigenvalue of $\bar{F}_{l,i}(x)$ and suppose $\bar{Q}_l(x)\bar{Q}_l(x)^T$ is invertible for all $l\in[M]$. If (11) is feasible, then for each $l\in[M]$, there exists $i\in[N]$ such that $\bar{H}_{l,i}(x)$ is not positive definite and one of the following holds:

(i) $\bar{\lambda}_{l,i}(x)<0$,
(ii) $\bar{\lambda}_{l,i}(x)>0$ and $\left(\bar{v}_{l,i}-\bar{w}_{l,i}^T\bar{F}_{l,i}^{-1}(\bar{Q}_l\bar{r}_l^T-\bar{w}_{l,i}\bar{v}_{l,i})\right)(x)\ge 0$,
(iii) $\bar{\lambda}_{l,i}(x)=0$, and $\left(\bar{v}_{l,i}-\bar{w}_{l,i}^T(\bar{Q}_l\bar{Q}_l^T)^{-1}\bar{Q}_l\bar{r}_l^T\right)(x)>0$.

Proof: Note that (10) (and hence (11)) is equivalent to

$$\min_{\underline{u}\in\underline{\mathcal{U}}} \|\underline{u}-\underline{k}(x)\|^2 \tag{12}$$
$$\text{s.t. } r\left\|R_l^T(x)\underline{u}\right\|+\inf_{t\in\mathbb{R}}\left[\frac{1}{N}\sum_{i=1}^N(G_l(x,\underline{u},\boldsymbol{\xi}_i)+t)_+-t\epsilon\right]\le 0,$$

for $l\in\{1,\dots,M\}$, cf. [7, Proposition IV.1]. For $(x,\underline{u})\in\mathbb{R}^n\times\underline{\mathcal{U}}$, the function $A^l_{x,\underline{u}}(t)=\frac{1}{N}\sum_{i=1}^N(G_l(x,\underline{u},\boldsymbol{\xi}_i)+t)_+-t\epsilon$ is a piecewise linear function in t. Since $\epsilon\le\frac{1}{N}$, it is decreasing for $t<t_l^*(x,\underline{u}):=\min_{i\in[N]}-G_l(x,\underline{u},\boldsymbol{\xi}_i)$ and increasing for $t>t_l^*(x,\underline{u})$. Hence, it achieves its minimum at $t_l^*(x,\underline{u})$. Thus, (11) is feasible if and only if for all $l\in[M]$ the following inequalities are simultaneously feasible:

$$r\left\|R_l^T(x)\underline{u}\right\|+\epsilon\underline{u}^Tq_l(x)+\epsilon\max_{i\in[N]}\underline{u}^TR_l(x)\boldsymbol{\xi}_i\le 0. \tag{13}$$

Note that, if for some $l\in[M]$, the constraint $r\left\|R_l^T(x)\underline{u}\right\|+\epsilon\underline{u}^Tq_l(x)+\epsilon\underline{u}^TR_l(x)\boldsymbol{\xi}_i\le 0$ is infeasible for all $i\in[N]$, then (11) is infeasible. Note that this is only a sufficient, but not necessary, condition for infeasibility (or equivalently, a necessary, but not sufficient, condition for feasibility). The result follows from [2, Theorem 2], which characterizes the feasibility of a single second-order cone constraint. ■

Next, we state a sufficient condition for the feasibility of (11) in the case $M=1$.

Proposition 4.2: (Sufficient condition for feasibility of DRO-SOCP with one constraint): Let $r>0$, $M=1$, and $0<\epsilon\le\frac{1}{N}$. Given $x\in\mathbb{R}^n$, define

$$\hat{Q}(x)=(r+\epsilon\max_{i\in[N]}\|\boldsymbol{\xi}_i\|)R_1(x)_{2:(m+1)}\in\mathbb{R}^{m\times k},$$
$$\hat{r}(x)=(r+\epsilon\max_{i\in[N]}\|\boldsymbol{\xi}_i\|)R_1(x)_1\in\mathbb{R}^{1\times k},$$
$$\hat{w}(x)=-\epsilon q_1(x)_{2:(m+1)}\in\mathbb{R}^m,\ \hat{v}(x)=-\epsilon q_1(x)_1\in\mathbb{R},$$
$$\hat{F}(x)=Q(x)Q(x)^T-w(x)w(x)^T\in\mathbb{R}^{m\times m},$$

This article has been accepted for publication in IEEE Control Systems Letters. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/LCSYS.2023.3348688

$$\hat{J}(x) = \hat{r}(x)\hat{Q}(x)^T - \hat{v}(x)\hat{w}(x)^T \in \mathbb{R}^{1\times m},$$
$$\hat{H}(x) = \begin{pmatrix} (\hat{r}\hat{r}^T - \hat{v}^2)(x) & \hat{J}(x) \\ \hat{J}(x)^T & \hat{F}(x) \end{pmatrix} \in \mathbb{R}^{(m+1)\times(m+1)}.$$

Let $\hat{\lambda}(x)$ be the minimum eigenvalue of $\hat{F}(x)$. Suppose that $Q(x)Q(x)^T$ is invertible, $\hat{H}(x)$ is not positive definite and one of the following holds:

(i) $\hat{\lambda}(x) < 0$,
(ii) $\hat{\lambda}(x) > 0$ and $\big(\hat{v} - \hat{w}^T\hat{F}^{-1}(\hat{Q}\hat{r}^T - \hat{w}\hat{v})\big)(x) \geq 0$,
(iii) $\hat{\lambda}(x) = 0$ and $\big(\hat{v} - \hat{w}^T(\hat{Q}\hat{Q}^T)^{-1}\hat{Q}\hat{r}^T\big)(x) > 0$.

Then, (11) is feasible at x.

Proof: By repeating an argument similar to the one in the proof of Proposition 4.1, (11) is feasible in the case $M = 1$ if and only if the following inequality is feasible:

$$r\left\|R(x)^T\underline{u}\right\| + \epsilon\underline{u}^T q(x) + \epsilon \max_{i\in[N]} \underline{u}^T R(x)\boldsymbol{\xi}_i \leq 0. \qquad (14)$$

Using the Cauchy-Schwartz inequality, the following inequality being feasible implies that (14) is feasible,

$$(r + \epsilon \max_{i\in[N]} \|\boldsymbol{\xi}_i\|)\left\|R(x)^T\underline{u}\right\| + \epsilon\underline{u}^T q(x) \leq 0. \qquad (15)$$

If (15) is feasible, there exists $\hat{\underline{u}}$ such that $r\left\|\hat{\underline{u}}^T R(x)\right\| + \epsilon\hat{\underline{u}}^T q(x) + \epsilon\hat{\underline{u}}^T R(x)\boldsymbol{\xi}_i \leq 0$ for all $i \in [N]$, and thus $\hat{\underline{u}}$ satisfies (14). The result follows by [2, Thm. 2]. ∎

Remark 4.3: (More data leads to better feasibility guarantees): For a fixed r, the addition of new data points (larger N) implies that there are more chances that either of (i)-(iii) in Proposition 4.1 are satisfied for each $l \in \{1,\dots,M\}$. Moreover, if $\mathbb{P}^*$ is light-tailed, $r_N(\bar{\epsilon})$ decreases with N. The choice $r = r_N(\bar{\epsilon})$ means that for each fixed $i \in [N]$ and $l \in [M]$, the feasible set of the inequality $r\left\|R_l(x)^T\underline{u}\right\| + \epsilon\underline{u}^T q_l(x) + \epsilon\underline{u}^T R_l(x)\boldsymbol{\xi}_i \leq 0$ increases, which from the proof of Proposition 4.1, also means that there are more chances that either of (i)-(iii) are met. Similarly, under the assumption that the norm of additional samples is upper bounded by $\max_{i\in[N]}\|\boldsymbol{\xi}_i\|$, the choice $r = r_N(\bar{\epsilon})$ also leads to a larger feasible set of (15) and thus the sufficient condition in Proposition 4.2 has more chances of being satisfied. •

We next give a sufficient condition for the feasibility of (11) with high probability for an arbitrary number of constraints.

Proposition 4.4: (Sufficient condition for feasibility of DRO-SOCP): Let $r > 0$, $\epsilon \in (0,1)$ and $\bar{\epsilon} \in (0,1)$. Suppose that there exists a controller $\hat{k} : \mathbb{R}^n \to \mathcal{U}$ and non-negative functions $S_l : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ for $l \in [M]$ satisfying

$$\mathrm{CVaR}^{\mathbb{P}^*}_{1-\epsilon}(G_l(x,\hat{k}(x),\boldsymbol{\xi})) \leq -S_l(x), \quad \forall l \in [M]. \qquad (16)$$

Moreover, suppose that $\mathbb{P}^*$ is light-tailed and let $r_N(\bar{\epsilon})$ be defined as in (6). Let $x \in \mathbb{R}^n$ be such that $\|R_l(x)\| \neq 0$ for all $l \in [M]$, and let $B : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ be an upper bound on the norm of $\hat{k}$. Then, if

$$r_N(\bar{\epsilon}) < \min_{l\in[M]} \frac{\epsilon S_l(x)}{2\,\|R_l(x)\|\,B(x)}, \qquad (17)$$

(11) is strictly feasible at x with probability at least $1-\bar{\epsilon}$ for any $r \leq r_N(\bar{\epsilon})$.

Proof: Note that by definition, the first component of $\hat{k}(x)$ is 1 for all $x \in \mathbb{R}^n$. Hence, $B(x) \geq \left\|\hat{k}(x)\right\| \geq 1$ for all $x \in \mathbb{R}^n$ so (17) is well-defined. Let $t_1^* \in \mathbb{R}$ be such that

$$\mathrm{CVaR}^{\mathbb{P}^*}_{1-\epsilon}(G_1(x,\underline{u},\boldsymbol{\xi})) = \frac{1}{\epsilon}\mathbb{E}_{\mathbb{P}^*}[(G_1(x,\underline{u},\boldsymbol{\xi}) + t_1^*)_+] - t_1^*,$$

and define $\hat{G}(x,\boldsymbol{\xi}) = \frac{1}{\epsilon}(G_1(x,\hat{k}(x),\boldsymbol{\xi}) + t_1^*)_+ - t_1^*$. Note that for any $\boldsymbol{\xi},\boldsymbol{\xi}' \in \mathbb{R}^k$,

$$|\hat{G}(x,\boldsymbol{\xi}) - \hat{G}(x,\boldsymbol{\xi}')| \leq \frac{1}{\epsilon}\|R_1(x)\|\cdot\left\|\hat{k}(x)\right\|\cdot\left\|\boldsymbol{\xi} - \boldsymbol{\xi}'\right\|, \qquad (18)$$

where we have used the fact that the operator $(\cdot)_+$ is Lipschitz with constant 1. Using (18) in [14, Theorem 3.2], we conclude that for any $\hat{\mathbb{P}} \in \mathcal{P}_p(\Xi)$, $|\mathbb{E}_{\mathbb{P}^*}(\hat{G}(x,\boldsymbol{\xi})) - \mathbb{E}_{\hat{\mathbb{P}}}(\hat{G}(x,\boldsymbol{\xi}))| \leq \frac{1}{\epsilon}\|R_1(x)\|\cdot\left\|\hat{k}(x)\right\|\cdot W_1(\mathbb{P}^*,\hat{\mathbb{P}})$.

From (17), together with the fact that $\mathcal{M}_N^{r_N(\bar{\epsilon})}$ contains $\mathbb{P}^*$ with probability at least $1-\bar{\epsilon}$, cf. Remark 2.1, and since the maximum Wasserstein distance between two distributions in $\mathcal{M}_N^{r_N(\bar{\epsilon})}$ is $2r_N(\bar{\epsilon})$, with probability at least $1-\bar{\epsilon}$,

$$|\mathrm{CVaR}^{\mathbb{P}^*}_{1-\epsilon}(G_1(x,\hat{k}(x),\boldsymbol{\xi})) - \mathbb{E}_{\hat{\mathbb{P}}}(\hat{G}(x,\boldsymbol{\xi}))| < S_1(x). \qquad (19)$$

for any $\hat{\mathbb{P}} \in \mathcal{M}_N^{r_N(\bar{\epsilon})}$. By definition of CVaR, cf. (2), for any $\hat{\mathbb{P}} \in \mathcal{P}_p(\Xi)$, $\mathrm{CVaR}^{\hat{\mathbb{P}}}_{1-\epsilon}(G_1(x,\hat{k}(x),\boldsymbol{\xi})) \leq \mathbb{E}_{\hat{\mathbb{P}}}(\hat{G}(x,\boldsymbol{\xi}))$. Combining this with (19) and (16), we get that with probability at least $1-\bar{\epsilon}$, $\mathrm{CVaR}^{\hat{\mathbb{P}}}_{1-\epsilon}(G_1(x,\hat{k}(x),\boldsymbol{\xi})) < 0$ for all $\hat{\mathbb{P}} \in \mathcal{M}_N^{r_N(\bar{\epsilon})}$. This argument holds for $l \in \{2,\dots,N\}$, implying that $\hat{k}(x)$ is strictly feasible for (10) (and hence, (11)) with probability at least $1-\bar{\epsilon}$ for any $r \leq r_N(\bar{\epsilon})$. ∎

Remark 4.5: (Dependency of sufficient condition on slack terms): Condition (16) on the controller $\hat{k}$ guarantees the satisfaction of the constraints in (10) with a slack term $S_l(x)$ on the righthand side. Larger values of these slack terms mean that fewer samples are needed to satisfy (17). Moreover, for the constraints in (9), [4, Remark 2.3] shows how to obtain such functions S_l, even without knowledge of $\hat{k}$. •

Remark 4.6: (Applicability of the sufficient condition): Checking condition (17) does not require precise knowledge of $\hat{k}$, just an upper bound of its norm. In particular, if bounds on the control norm are included as constraints in (11), those can be used to construct B. Moreover, unlike Proposition 4.2, condition (17) is agnostic to the samples $\{\boldsymbol{\xi}_1,\dots,\boldsymbol{\xi}_N\}$ and instead solely depends on its number N. Note that for each $x \in \mathbb{R}^n$ with $\|R_l(x)\| \neq 0$ for all $l \in [M]$, if $S_l(x) > 0$ for all $l \in [M]$, there exists $\hat{N}_x$ such that condition (17) holds for all $N \geq \hat{N}_x$. This is because $r_N(\bar{\epsilon})$ is decreasing in N and $\lim_{N\to\infty} r_N(\bar{\epsilon}) = 0$. The value $\hat{N}_x$ is state-dependent, larger for smaller values of ϵ, $S_l(x)$ and larger values of $B(x)$. •

Remark 4.7: (Checking for (in)feasibility efficiently): A commonly used algorithm for solving SOCPs is the method in [16]. For an SOCP with r_S constraints and optimization variable of dimension n_S, it requires solving $\sqrt{r_S}$ linear systems of dimension n_S, and hence has complexity $\mathcal{O}(\sqrt{r_S}n_S^3)$, cf. [17]. Therefore, (11) has complexity $\mathcal{O}(\sqrt{MN}(m+N)^3)$. Instead, since checking the positive definiteness of a symmetric matrix of dimension n_P can be done by checking if its Cholesky factorization exists (which has complexity $\mathcal{O}(n_P^3)$), the complexity of checking the condition in Proposition 4.1 is

TABLE I: Complexity of SOCP solver versus the results in this section.

Method	Necessary/Sufficient	Complexity	M
Prop. 4.1	Necessary	$\mathcal{O}(NMm^3)$	any
Prop. 4.2	Sufficient	$\mathcal{O}(N+m^3)$	1
Prop. 4.4	Sufficient	$\mathcal{O}(M)$	any
SOCP solver	Necessary and sufficient	$\mathcal{O}(\sqrt{NM}(m+N)^3)$	any

$\mathcal{O}(NMm^3)$. Hence, for large N, it is much more efficient than solving the SOCP (11) directly. We also note that the scaling in M for the complexity of the SOCP solver is more favorable than that of checking the necessary condition. On the other hand, the complexity of checking the sufficient condition in Proposition 4.2 reduces to finding a maximum of N numbers (which has complexity linear in N) and checking the positive definiteness of two symmetric matrices of dimension $m+1$ and m, resp. Hence, its complexity is $\mathcal{O}(N+m^3)$, which is also more efficient than solving the SOCP. Finally, note that the complexity of checking the conditions in Proposition 4.4 is constant in N and m, and is linear in M due to the minimum in (17). Table I summarizes this complexity analysis. •

Proposition 4.1 provides necessary conditions for feasibility. If the conditions are not met, it is reasonable to gather more data for verifying feasibility without having to directly solve the program. Moreover, if the conditions in Propositions 4.2 and 4.4 are not met (which does not mean that (11) is not feasible), this might be an indication that more data is needed to certify feasibility, cf. Remarks 4.3 and 4.6.

V. Regularity Analysis

In this section, we show that the controller obtained by solving (11) is point-Lipschitz.

Proposition 5.1: (Point-Lipschitzness of SOCP DRO): Let $r > 0$, $0 < \epsilon \leq \frac{1}{N}$ and suppose R_l and q_l are twice continuously differentiable for all $l \in [M]$. Let $\underline{u}^* : \mathbb{R}^n \to \mathbb{R}^m$ be the function mapping $x \in \mathbb{R}^n$ to the solution of (11) in $\underline{u}$ at x. If (10) is strictly feasible at $x_0 \in \mathbb{R}^n$ (i.e., there exists a solution satisfying all the constraints strictly), then $\underline{u}^*$ is point-Lipschitz at x_0.

Proof: We first show the result for $M = 1$. Let $\mathcal{I} := \arg\max_{i\in[N]} G_1(x_0, \underline{u}^*(x_0), \boldsymbol{\xi}_i)$, note that the set $\mathcal{I}$ is dependent on x_0, but we omit this dependency to simplify the notation. Note also that since $G_1(x, \underline{u}, \boldsymbol{\xi}_i)$ is continuous in x and $\underline{u}$ for all $i \in [N]$, there exists a neighborhood $\mathcal{N} = \mathcal{N}_x \times \mathcal{N}_{\underline{u}} \subset \mathbb{R}^n \times \underline{\mathcal{U}}$ of $(x_0, \underline{u}^*(x_0))$ such that for all $(\hat{x}, \hat{\underline{u}}) \in \mathcal{N}$, there exists $i_{\hat{x},\hat{\underline{u}}} \in \mathcal{I}$ such that $i_{\hat{x},\hat{\underline{u}}} \in \arg\max_{i\in[N]} G_1(\hat{x}, \hat{\underline{u}}, \boldsymbol{\xi}_i)$. Recall from the proof of Proposition 4.1 that, for any $x, \underline{u} \in \mathbb{R}^n \times \underline{\mathcal{U}}$, the function $A_{x,\underline{u}}(t) := \frac{1}{N}\sum_{i=1}^{N}(G_1(x, \underline{u}, \boldsymbol{\xi}_i)+t)_+ - t\epsilon$ attains its minimum at $t^*(x, \underline{u}) := \max_{i\in[N]} G_1(x, \underline{u}, \boldsymbol{\xi}_i)$. Therefore, for $(\hat{x}, \hat{\underline{u}}) \in \mathcal{N}$, $t^*(\hat{x}, \hat{\underline{u}}) = G_1(\hat{x}, \hat{\underline{u}}, \boldsymbol{\xi}_{i_{\hat{x},\hat{\underline{u}}}})$.

For each $i \in \mathcal{I}$, let $\underline{u}_i^* : \mathbb{R}^n \to \mathbb{R}^m$ be defined as:

$$\underline{u}_i(x) := \min_{\underline{u}\in\underline{\mathcal{U}}} \|\underline{u} - \underline{k}(x)\|^2 \quad \text{s.t. } r\left\|R_1(x)^T\underline{u}\right\| + \epsilon G_1(x, \underline{u}, \boldsymbol{\xi}_i) \leq 0. \tag{20}$$

Note that since (10) is strictly feasible at x_0, there exists $\tilde{\underline{u}} \in \underline{\mathcal{U}}$ such that $r\left\|R_1(x_0)^T\tilde{\underline{u}}\right\| + \max_{i\in[N]} \epsilon G_1(x_0, \tilde{\underline{u}}, \boldsymbol{\xi}_i) < 0$. By continuity of R_1 and G_1 in x, there exists a neighborhood $\tilde{\mathcal{N}}_x \subset \mathcal{N}_x$ of x_0 such that $r\left\|R_1(x)^T\tilde{\underline{u}}\right\| + \epsilon G_1(x, \tilde{\underline{u}}, \boldsymbol{\xi}_i) < 0$ for all $x \in \tilde{\mathcal{N}}_x$ and $i \in \mathcal{I}$. This implies that (20) is strictly feasible for any $x \in \tilde{\mathcal{N}}_x$. Hence, by [4, Proposition 5.4], $\underline{u}_i^*$ is point-Lipschitz at x_0 for each $i \in \mathcal{I}$. Now, since for all $y \in \mathcal{N}_x$ there exists $i \in \mathcal{I}$ such that $\underline{u}^*(y) = \underline{u}_i^*(y)$, and $\tilde{\mathcal{N}}_x \subset \mathcal{N}_x$, it follows $\|\underline{u}^*(y) - \underline{u}^*(x_0)\| = \|\underline{u}_i^*(y) - \underline{u}_i^*(x_0)\| \leq \gamma_i \|y - x_0\|$ for some $\gamma_i > 0$. Now, by taking $\gamma := \max_{i\in\mathcal{I}} \gamma_i$, it follows that $\|\underline{u}^*(y) - \underline{u}^*(x_0)\| \leq \gamma \|y - x_0\|$ for all $y \in \mathcal{N}_x$ and hence $\underline{u}^*$ is point-Lipschitz at x_0. The argument if $M > 1$ is analogous, defining a set $\mathcal{I}_l$ similar to $\mathcal{I}$ for each $l \in [M]$. ■

Proposition 5.1 implies in particular that u^* is continuous at x_0. Note also that the strict feasibility assumption in Proposition 5.1 is satisfied with a prescribed probability if the hypothesis of Proposition 4.4 is satisfied.

VI. Simulations

In this section, we evaluate our results in a ground-robot navigation example. We model the robot motion using unicycle kinematics and take a small distance $a = 0.05$ off the wheel axis, cf. [18] to obtain a relative-degree-one model:

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{\theta}\end{bmatrix} = \left(\begin{bmatrix}0 & \cos(\theta) & -a\sin(\theta)\\ 0 & \sin(\theta) & a\cos(\theta)\\ 0 & 0 & 1\end{bmatrix} + \sum_{j=1}^{3} W_j(x)\xi_j\right)\begin{bmatrix}1\\ v\\ \omega\end{bmatrix},$$

where v, ω are the linear and angular velocity, and

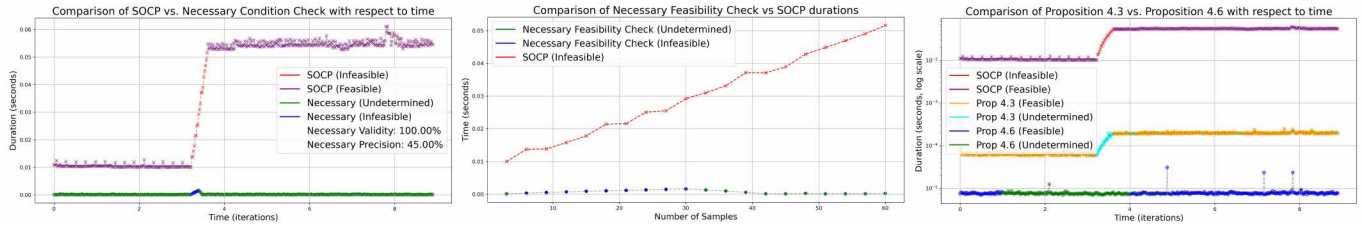
$$W_1(x) = \begin{bmatrix}0.02 & 0 & 0\\ 0.02 & 0 & 0\\ 0.01 & 0 & 0\end{bmatrix},\ W_2(x) = \begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & -0.02\end{bmatrix},$$
$$W_3(x) = \begin{bmatrix}0 & 0.02\cos(\theta) & -0.02a\sin(\theta)\\ 0 & 0.02\sin(\theta) & 0.02a\cos(\theta)\\ 0 & 0 & 0\end{bmatrix},$$

represent the model perturbations in the drift, angular velocity, and orientation. We consider uncertainty samples: $\xi_1 \sim \mathcal{N}(0.5, 1)$, $\xi_2 \sim \mathcal{U}(-1, 1)$, and $\xi_3 \sim \mathcal{B}(2, 0.2)$, where $\mathcal{N}$, $\mathcal{U}$, $\mathcal{B}$ denote normal, uniform, and beta distributions, resp. The optimization programs are solved using the Embedded Conic Solver in CVXPY [19] with an Intel i7 9700K CPU.

We first consider the problem of stabilizing the uncertain unicycle system to a goal position $[x_1^*, x_2^*] = [7, 7]$ with initial state $[0, 0, 0]$, so we take $M = 1$ in (10). At the initial state, the robot is assumed to have 3 samples $\{\boldsymbol{\xi}_i\}_{i=1}^3$ and initial Wasserstein radius $r = 0.5$ with risk tolerance $\epsilon = 0.01$. As the robot moves, each unsuccessful solver attempt prompts the collection of additional samples, and a corresponding reduction in the ambiguity radius as prescribed by (6). In all the figures presented, the x-axis represents the simulation timestep, where each timestep is equivalent to 0.02 seconds, and the y-axis denotes the time spent for carrying out the necessary and sufficient condition checks, as well as for running the solver at each timestep.

The time complexity, validity, and precision of Proposition 4.1 are explored in Fig. 1a and Fig. 1b. Fig. 1a compares the time complexity of checking the necessary condition in Proposition 4.1 and of solving the corresponding SOCP along the whole robot trajectory. Notably, the SOCP becomes infeasible at around $t = 3$ s and more uncertainty samples are given until feasibility is regained. As expected, when Proposition 4.1 predicts the program is infeasible, such inference is consistently mirrored by the solver. Fig. 1b specifically emphasizes the time complexity during data collection stages. As the

This article has been accepted for publication in IEEE Control Systems Letters. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/LCSYS.2023.3348688



(a) (b) (c)

Fig. 1: (a): Time complexity comparison between necessary condition verification (cf. Proposition 4.1) and SOCP solver along the robot trajectory. The label "undetermined" means that the necessary condition is met, from which we may not know if the problem is feasible or not. The label "precision" represents the ratio of instances where the necessary condition indicated infeasibility against the total number of instances where the SOCP was actually infeasible. (b): Time complexity of necessary condition verification and SOCP solver with increasing uncertainty samples (constraints). (c): Log-scaled time complexity comparison of two sufficient conditions (cf. Proposition 4.2 and Proposition 4.4) with the SOCP solver along the robot trajectory.

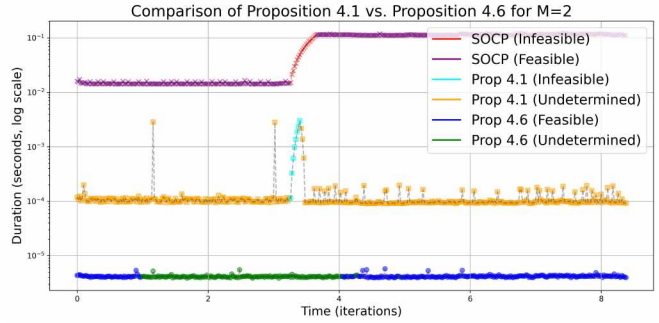


Fig. 2: Time complexity comparison of necessary (cf. Proposition 4.1) and sufficient (cf. Proposition 4.4) conditions.

number of samples increases, the SOCP's time complexity escalates at a much faster rate than the necessary condition verification, in agreement with Remark 4.7.

Fig. 1c compares the time complexity of solving the SOCP and of checking the sufficient conditions in Propositions 4.2 and 4.4. As expected, feasibility validation by either result ensures the actual feasibility of the program by the solver. Checking Proposition 4.2 is more time-consuming than checking Proposition 4.4, cf. Remark 4.7, but has greater accuracy in validating feasibility. Notably, both checks are significantly more efficient than solving the SOCP problem.

We also consider the safe stabilizing problem of the unicycle system. The stabilization goal is $[x_1^*, x_2^*] = [5, 5]$ while the safety goal is to avoid a circular obstacle centered at $[3, 2]$ with radius 1. Fig. 2 compares the time complexity and conservativeness of Proposition 4.1 and Proposition 4.4 for the case $M = 2$ in (10). Proposition 4.1 is valid and requires significantly less time than solving the SOCP, while Proposition 4.4 is also valid and even more efficient.

VII. Conclusions

We studied the feasibility of SOCP problems whose solution provide safe stabilizing controllers for uncertain systems with no prior knowledge of the uncertainty distribution and only a finite number of samples available. We provided a necessary condition and two sufficient conditions to check feasibility, characterized their computational complexity, and showed through simulations their usefulness in practical scenarios. We also showed that the controller obtained by solving the SOCP is point-Lipschitz under fairly general conditions. Future work will consider leveraging the identified feasibility conditions to guide online data-gathering policies that aim to reduce uncertainty about the system dynamics.

References

[1] K. Long, V. Dhiman, M. Leok, J. Cortés, and N. Atanasov, "Safe control synthesis with uncertain dynamics and constraints," *IEEE Robotics and Automation Letters*, vol. 7, no. 3, pp. 7295–7302, 2022.

[2] F. Castañeda, J. J. Choi, B. Zhang, C. J. Tomlin, and K. Sreenath, "Pointwise feasibility of Gaussian process-based safety-critical control under model uncertainty," in *IEEE Conference on Decision and Control*, Austin, Texas, USA, 2021, pp. 6762–6769.

[3] A. J. Taylor, V. D. Dorobantu, S. Dean, B. Recht, Y. Yue, and A. D. Ames, "Towards robust data driven control synthesis for nonlinear systems with actuation uncertainty," in *IEEE Conf. on Decision and Control*, Austin, Texas, USA, 2021, pp. 6469–6476.

[4] P. Mestres and J. Cortés, "Feasibility and regularity analysis of safe stabilizing controllers under uncertainty," *Automatica*, 2023, submitted.

[5] A. Lederer, A. Begzadic, N. Das, and S. Hirche, "Safe learning-based control of elastic joint robots via control barrier functions," arXiv 2212.00478, 2023.

[6] P. Seiler, M. Jankovic, and E. Hellstrom, "Control barrier functions with unmodeled input dynamics using integral quadratic constraints," *IEEE Control Systems Letters*, vol. 6, pp. 1664–1669, 2022.

[7] K. Long, Y. Yi, J. Cortés, and N. Atanasov, "Safe and stable control synthesis for uncertain system models via distributionally robust optimization," in *American Control Conference*, Jun. 2023, pp. 4651–4658.

[8] E. D. Sontag, *Mathematical Control Theory: Deterministic Finite Dimensional Systems*, 2nd ed., ser. TAM. Springer, 1998, vol. 6.

[9] A. D. Ames, S. Coogan, M. Egerstedt, G. Notomista, K. Sreenath, and P. Tabuada, "Control barrier functions: theory and applications," in *European Control Conference*, Naples, Italy, Jun. 2019, pp. 3420–3431.

[10] A. D. Ames, X. Xu, J. W. Grizzle, and P. Tabuada, "Control barrier function based quadratic programs for safety critical systems," *IEEE Transactions on Automatic Control*, vol. 62, no. 8, pp. 3861–3876, 2017.

[11] F. Castañeda, J. J. Choi, B. Zhang, C. J. Tomlin, and K. Sreenath, "Gaussian process-based min-norm stabilizing controller for control-affine systems with uncertain input effects and dynamics," in *American Control Conference*, New Orleans, LA, May 2021, pp. 3683–3690.

[12] K. Long, C. Qian, J. Cortés, and N. Atanasov, "Learning barrier functions with memory for robust safe navigation," *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 4931–4938, 2021.

[13] A. Ben-Tal, L. E. Ghaoui, and A. Nemirovski, *Robust Optimization*, ser. Applied Mathematics Series. Princeton University Press, 2009.

[14] P. M. Esfahani and D. Kuhn, "Data-driven distributionally robust optimization using the Wasserstein metric: performance guarantees and tractable reformulations," *Mathematical Programming*, vol. 171, no. 1-2, pp. 115–166, 2018.

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

[16] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Levret, "Applications of second-order cone programming," *Linear algebra and its applications*, vol. 284, pp. 193–228, 1997.

[17] A.-L. Cholesky, "Sur la résolution numerique des systèmes d'équations linéaires," *Bulletin de la société des amis de la bibliothèque de l'École polytechnique*, vol. 39, 2005.

[18] J. Cortés and M. Egerstedt, "Coordinated control of multi-robot systems: A survey," *SICE Journal of Control, Measurement, and System Integration*, vol. 10, no. 6, pp. 495–503, 2017.

[19] S. Diamond and S. Boyd, "CVXPY: A Python-embedded modeling language for convex optimization," *Journal of Machine Learning Research*, 2016.