MITSUBISHI ELECTRIC RESEARCH LABORATORIES
https://www.merl.com

Sample quantile-based programming for non-convex separable chance constraints

Vinod, Abraham P.; Di Cairano, Stefano

TR2023-062 June 01, 2023

Abstract

We propose a sampling-based approximation to non-convex, separable, chance constrained optimization prob-lems using sample quantiles. The proposed approach does not require any prior knowledge of the distribution or the moments of the uncertainty, and accommodates chance constraints that are non-convex in the decision variables. We prescribe the finite number of samples and the tightening necessary to produce a feasible solution to the original chance constrained optimization problem with bounded suboptimality. The proposed approximation has a low computational cost since the number of sample-based constraints does not grow with the number of samples, and the number of samples needed is independent of the number of decision variables. We show the effectiveness of the proposed solution in a stochastic motion planning problem with non-convex obstacle avoidance constraints.

American Control Conference (ACC) 2023

© 2023 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Mitsubishi Electric Research Laboratories, Inc.
201 Broadway, Cambridge, Massachusetts 02139

Sample quantile-based programming for non-convex separable chance constraints

Abraham P. Vinod* and Stefano Di Cairano

Abstract— **We propose a sampling-based approximation to non-convex, separable, chance constrained optimization problems using sample quantiles. The proposed approach does not require any prior knowledge of the distribution or the moments of the uncertainty, and accommodates chance constraints that are non-convex in the decision variables. We prescribe the finite number of samples and the tightening necessary to produce a feasible solution to the original chance constrained optimization problem with bounded suboptimality. The proposed approximation has a low computational cost since the number of sample-based constraints does not grow with the number of samples, and the number of samples needed is independent of the number of decision variables. We show the effectiveness of the proposed solution in a stochastic motion planning problem with non-convex obstacle avoidance constraints.**

I. Introduction

Chance constrained programming is a mathematical framework for decisions under *stochastic* uncertainty [1], [2]. In these optimization problems, the choice of the decision variables must ensure that the likelihood of violation of constrains remains below a (small) pre-specified threshold. In contrast to their robust counterpart (where the decisions are made based on the *worst-case* realization of the uncertainty), allowing a (small) likelihood of constraint violation typically yields significant improvements in the optimal value [1]. We focus on chance constrained programs with constraints that are non-convex in the decision variables, and no prior knowledge is available about the distribution or the moments of the uncertainty. We propose *a tractable algorithm based on sample quantile functions that guarantees feasibility and bounded suboptimality with a finite number of samples*.

Existing approaches for chance constrained programming can be split into two groups — sample-free approaches and sample-based approaches. Sample-free approaches assume that the information on the distribution or the moments of the uncertainty is available *a priori*, and use the information to inner-approximate the feasible set via non-stochastic constraints. Examples of such approaches include quantile-based reformulations [3], [4], moment-based approaches like the Chebyshev-Cantelli bound [5], and convexification using Bernstein approximation [6]. Sample-free approaches typically produce convex approximations for computational tractability. Consequently, they impose structural assumptions on the optimization problem and on the stochastic nature of the uncertainty [1], [2]. Sample-based approaches, on the other hand, replace the chance constraint with a collection of constraints that are evaluated at the available samples of the uncertainty. The samples of the uncertainty may be collected before or at runtime. Unlike sample-free approaches, sample-based approaches do not require any prior knowledge of the stochastic uncertainty and typically require less restrictive structural assumptions on the problem.

Sample-based approaches can be further classified based on the enforcement of the sample constraint into two groups — *all-sample* and *fraction-sample approximations*.

All-sample approximations require the satisfaction of the constraints at all realizations [1], [7]–[10], and control the probability of constraint satisfaction by determining the number of samples considered. All-sample approximations typically inherit the structural properties of the original chance constrained program including convexity and differentiability [1], [8]–[10]. In [8], finite lower bounds on the samples are provided for the case where the cost and the constraint functions in the chance constrained program are convex. However, these approaches typically return conservative, i.e. suboptimal, solutions and are sensitive to outliers. As an alternative, sample-and-discard approaches have been proposed that further lower the optimal value by trading-off its feasibility [8]–[10]. For the more general non-convex setting, bounds on the samples are obtained from statistical learning theory [7], though the resulting sample bounds are hard to compute in practice [9]. In [9], [10], the scenario optimization was generalized to handle non-convex costs and constraints. While the setting considered in [9], [10] is more general than the setting considered in this paper, our proposed approach has significantly lower computational burden — the number of constraints in our proposed approximation is independent of the number of samples, the minimum number of samples needed by the approximation is independent of the number of decision variables, and our approach does not require solving multiple sample approximations.

Fraction-sample approximations control the probability of constraint violation using the fraction of the sample constraints that are allowed to be violated [1], [11]–[14]. The computed optimal value is less susceptible to the outliers among the drawn samples of the uncertainty as compared to the all-sample approximations. A popular approach in fraction-sample approximation relies on mixed-integer programming formulations [1], [11], [12]. However, mixed-integer-based approaches become computationally prohibitive with a large number of samples. Recently, two fraction-sample approximations were proposed that do not require mixed-integer reformulations [13], [14] In [13],

* Corresponding author. Email: vinod@merl.com.
A. Vinod and S. Di Cairano are with Mitsubishi Electric Research Laboratories, Cambridge, MA 02139, USA.

the authors approximate the chance constraint using the empirical characteristic functions and impose convexity assumptions on the chance constraints for the sake of tractability. In [14], the authors use a smoothed sample quantile function to obtain a (possibly non-convex) nonlinear program for a chance constrained program with differentiable cost and constraints, and compute an approximate solution using trust-region methods. Compared to [13], [14], our approach does not impose any structural restrictions on the original program (specifically differentiability or convexity), guarantees feasibility, and provides a suboptimality bound.

The main contribution of this work is a fraction-sample approximation for non-convex separable chance constrained programs based on the sample quantile function. The proposed approximation does not require any prior knowledge of the distribution or the moments of the uncertainty, and does not impose any hard structural assumptions on the problem such as differentiability or convexity of the cost and the constraint functions. We use the Dvoretzky–Kiefer–Wolfowitz inequality to arrive at the finite sample lower bound that guarantees feasibility and bounded suboptimality of the approximate solution, in the considered general setting. The proposed reformulation is less susceptible to outliers because of its fraction-sample nature, preserves the structural properties of the original chance constrained program, and yields a suboptimality bound on the computed solution. Thus, our approach retains the advantages of both all-sample and fraction-sample approximations for non-convex chance constrained programs with separable constraints, without their major limitations.

II. Preliminaries & Problem statement

We employ the following notation throughout the paper: $\mathbb{N}_{[a,b]}$ denotes the interval of natural numbers between a, b. We denote random vectors in bold, the Euclidean norm with $\|\cdot\|$, the n-dimensional identity matrix as I_n, and 0_n and $0_{n\times m}$ as a vector and matrix of zeros respectively.

A. Sample distribution and quantile functions

Consider a random variable $\boldsymbol{x} \in \mathbb{R}$ with probability measure $\mathbb{P}_{\boldsymbol{x}}$. For any $x \in \mathbb{R}$ and $p \in [0,1]$, the distribution function $\Phi_{\boldsymbol{x}}$ and the quantile function $Q_{\boldsymbol{x}}$ of $\boldsymbol{x}$ are

$$\Phi_{\boldsymbol{x}}(x) = \mathbb{P}_{\boldsymbol{x}}\{\boldsymbol{x} \leq x\}, \tag{1a}$$
$$Q_{\boldsymbol{x}}(p) = \inf\{x \in \mathbb{R} : \Phi_{\boldsymbol{x}}(x) \geq p\}. \tag{1b}$$

Let $\left\{x_{\text{sample}}^{(j)}\right\}_{j=1}^{M}$ be $M \in \mathbb{N}$ samples of $\boldsymbol{x}$ drawn according to $\mathbb{P}_{\boldsymbol{x}}$. For any $x \in \mathbb{R}$ and $p \in [0,1]$, the corresponding *sample distribution function* and the *sample quantile function* of $\boldsymbol{x}$ are

$$\hat{\Phi}_{\boldsymbol{x}}^{M}(x) \triangleq \frac{1}{M}\sum\nolimits_{j=1}^{M} \mathbb{I}\{x_{\text{sample}}^{(j)} \leq x\}, \tag{2a}$$
$$\hat{Q}_{\boldsymbol{x}}^{M}(p) \triangleq \inf\{x \in \mathbb{R} : \hat{\Phi}_{\boldsymbol{x}}^{M}(x) \geq p\}. \tag{2b}$$

Here, for each $j \in \mathbb{N}_{[1,M]}$, the indicator function $\mathbb{I}\{x_{\text{sample}}^{(j)} \leq x\} = 1$ when $x_{\text{sample}}^{(j)} \leq x$, and is zero otherwise.

Lemma 1. *[15, Lem. A.1.1] For any $x \in \mathbb{R}$ and $p \in [0,1]$, $\Phi_{\boldsymbol{x}}(x) \geq p$ if and only if $x \geq Q_{\boldsymbol{x}}(p)$, and $x \geq \hat{Q}_{\boldsymbol{x}}^{M}(p)$ if and only if $\hat{\Phi}_{\boldsymbol{x}}^{M}(x) \geq p$.*

We will use Lemma 1 to reformulate constraints on distribution functions as constraints on quantile functions.

Lemma 2 (Dvoretzky–Kiefer–Wolfowitz [16, Sec. 2.3.2])**.** *Let $\boldsymbol{x}$ have a continuous distribution $\Phi_{\boldsymbol{x}}$. Then, for any $\varepsilon > 0$ and $\hat{\Phi}_{\boldsymbol{x}}^{M}$ defined using M i.i.d. samples, $\mathbb{P}\left\{\sup_{x\in\mathbb{R}} |\Phi_{\boldsymbol{x}}(x) - \hat{\Phi}_{\boldsymbol{x}}^{M}(x)| > \varepsilon\right\} \leq 2\exp(-2M\varepsilon^2)$.*

Lemma 2 characterizes an exponential decay (with respect to M) in the probability that the maximum pointwise difference between the sample and the true distribution exceeds some $\varepsilon > 0$.

B. Problem statement

Consider the following, possibly non-convex, chance constrained optimization problem,

$$\begin{aligned} &\text{minimize} && f_0(z) && \text{(3a)}\\ &\text{subject to} && z \in \mathcal{Z}, && \text{(3b)}\\ &\forall i \in \mathbb{N}_{[1,N]}, && \mathbb{P}_{\boldsymbol{w}}\{f_i(z) + g_i(\boldsymbol{w}) \leq 0\} \geq \delta_i. && \text{(3c)} \end{aligned}$$

Here, $z \in \mathbb{R}^{n_z}$ is the decision variable, $\mathcal{Z} \subset \mathbb{R}^{n_z}$ is a set of deterministic constraints on z, $\boldsymbol{w} \in \mathbb{R}^{n_w}$ is the random vector with an unknown probability measure $\mathbb{P}_{\boldsymbol{w}}$, $f_i : \mathbb{R}^{n_z} \to \mathbb{R}$, $\forall i \in \mathbb{N}_{[0,N]}$ and $g_i : \mathbb{R}^{n_w} \to \mathbb{R}$, $\forall i \in \mathbb{N}_{[1,N]}$ are known functions, and $\delta_i \in (0,1)$ are known risk thresholds. In the literature, (3c) is commonly known as a *separable probabilistic constraint* [1, Sec. 4.3].

We make the following assumption on f_i, g_i for the well-posedness of (3), see [1, Sec. 5.1] for more details, and for enabling the application of Lemma 2.

Assumption 1. *For each constraint $i \in \mathbb{N}_{[1,N]}$, f_i is continuous and $g_i(\boldsymbol{w})$ has a continuous distribution function.*

The all-sample approximation of (3) solves

$$\begin{aligned} &\text{minimize} && f_0(z) && \text{(4a)}\\ &\text{subject to} && z \in \mathcal{Z}, && \text{(4b)}\\ &\forall i \in \mathbb{N}_{[1,N]} && f_i(z) + \max_{j\in\mathbb{N}_{[1,M]}} g_i\left(w_{\text{sample}}^{(j)}\right) \leq 0. && \text{(4c)} \end{aligned}$$

By (2a), (4c) is a reformulation of the constraint $\{z \in \mathbb{R}^{n_z} : \hat{\Phi}_{f_i(z)+g_i(\boldsymbol{w})}^{M}(0) = 1\}$. The main advantage of (4) is that it inherits the structural properties of f_i. However, (4) is highly sensitive to outliers in the drawn samples of $\boldsymbol{w}$, and typically return a conservative, i.e., suboptimal, solution to (3). To mitigate the conservativeness, sample-and-discard approaches lower the optimal value by reducing δ_i, thereby trading-off its feasibility [8], [9].

Instead, the fraction-sample approximation of (3) is

$$\begin{aligned} &\text{minimize} && f_0(z) && \text{(5a)}\\ &\text{subject to} && z \in \mathcal{Z}, && \text{(5b)}\\ &\forall i \in \mathbb{N}_{[1,N]} && \hat{\Phi}_{f_i(z)+g_i(\boldsymbol{w})}^{M}(0) \geq \Delta_i. && \text{(5c)} \end{aligned}$$

We refer to $\Delta_i \in [0,1]$ as the *sample risk thresholds*. When $\Delta_i = \delta_i$ for each $i \in \mathbb{N}_{[1,N]}$, the optimal solution of (5) approaches the optimal solution of (3) as $M \to \infty$ under Assumption 1 [1, Sec. 5.1]. By construction, (5c) requires $f_i(z) + g_i\left(w_{\text{sample}}^{(j)}\right) \leq 0$ only for a subset of samples in $\mathbb{N}_{[1,M]}$ [11]–[14]. Thus, (5) does not suffer from the drawbacks of (4), since the solution of (5) has a lower sensitivity to the sample outliers with higher M.

A popular approach to solve (5) is via a mixed-integer reformulation of (5c) based on (2a). Existing works [11], [12] use a binary variable for each sample to encode the satisfaction of the constraint $f_i(z) + g_i\left(w_{\text{sample}}^{(j)}\right) \leq 0$, and constraint the sum of the resulting binary variables be no smaller than $\lceil \Delta_i M \rceil$. These approaches approximate (5c) as

$$\frac{1}{M}\sum_{j=1}^{M} \mathbb{I}\{f_i(z) + g_i\left(w_{\text{sample}}^{(j)}\right) \leq 0\} \geq \Delta_i. \quad (6)$$

However, the resulting mixed-integer programs are computationally expensive to solve, and impractical for large M.

The main goal of this paper is to solve (3) by combining the strengths of the formulations (4) and (5), while overcoming some of their respective limitations. Sampling-based approaches must contend with the possibility that we can obtain a very poor approximation of (3) due to the drawn set of samples [8], [12]. Similarly to [8], the proposed approximation must guarantee that the probability of such an (unlucky) event is below a user-specified *probability of unreliability* $\beta \in (0,1)$.

Problem 1. *Given a user-specified probability of unreliability $\beta \in (0,1)$, prescribe a sufficient number of samples M and appropriate sample risk thresholds Δ_i to obtain a well-defined sample quantile function-based approximation of* (3). *Additionally, every feasible solution of the approximation must be feasible for* (3) *with a probability* $1-\beta$.

Apart from Problem 1, we will also compute a bound on the suboptimality of a solution that solves the approximation. In what follows, we will use $\boldsymbol{g}_i$ to denote $g_i(\boldsymbol{w})$ for brevity.

Remark 1. *Despite a relatively strong assumption of separability in the chance constraints, we emphasize that f_i and g_i can be non-convex in z and $\boldsymbol{w}$ respectively, unlike some of the existing approaches [8], [11]. This enables the use of our approach in tackling non-convex motion planning problems under uncertainty.*

III. Sample quantile-based programming

To address Problem 1, we first propose a sample quantile-based approximation of (3). Next, we show that the feasible set defined by the chance constraints can be inner- and outer-approximated using the level sets of the sample quantile functions with a user-specified probability of reliability. Finally, we use the set approximations to compute a feasible solution of (3), and conclude with a discussion of various features of the proposed approach.

A. Sample quantile-based approximation of (3)

Consider the following approximation to (3),

$$\begin{aligned} &\text{minimize} && f_0(z) && (7a)\\ &\text{subject to} && z \in \mathcal{Z}, && (7b)\\ &\forall i \in \mathbb{N}_{[1,N]}, && f_i(z) + \hat{Q}_{\boldsymbol{g}_i}^M(\Delta_i) \leq 0. && (7c)\end{aligned}$$

Given Δ_i, the constraints (7c) are deterministic as (7c) depend only on the samples of $\boldsymbol{w}$. (7) can be solved by any off-the-shelf nonlinear optimization solver. While our approach does not require it, (7) also permits exploiting additional properties of f_i such as differentiability and convexity.

The motivation for (7) comes from the observation that, using Lemma 1 and (1), the chance constraint (3c) is equivalent to the constraint,

$$\forall i \in \mathbb{N}_{[1,N]}, \quad f_i(z) + Q_{\boldsymbol{g}_i}(\delta_i) \leq 0. \quad (8)$$

Next, we formalize the relationship between (5) and (7) in Proposition 1 (see Appendix for a proof sketch). Two optimization problems are *equivalent*, if the solution of one solves the other, and vice versa [17].

Proposition 1. *Problems* (5) *and* (7) *are equivalent.*

We highlight two advantages of (7) over the all-sample approximation (4) and the mixed-integer reformulation of (5). First, unlike (4), the sensitivity of the solution of (7) to the outliers decreases with increasing M. Second, unlike the mixed-integer reformulation of (5) using (6), (7) does not incur any additional computation cost for dealing with the fractional-sample constraint (5c) since $\hat{Q}_{\boldsymbol{g}_i}^M(\Delta_i)$ can be pre-computed before solving (7).

B. Tuning sample-based chance constraint approximations

The sample quantile-based approximation (7c) has two parameters, M and Δ_i. We now propose a tuning strategy for these parameters to obtain inner- and outer-approximations of the chance constraint (3c). Later, we construct a *safe approximation* of (3) using the inner-approximation, and compute the suboptimality bound using the outer-approximation.

Consider the feasible set defined by N chance constraints (3c) with $\delta_i \in (0,1)$ for all $i \in \mathbb{N}_{[1,N]}$,

$$\mathscr{C} = \left\{z \in \mathbb{R}^{n_z} : \mathbb{P}_{\boldsymbol{w}}\{f_i(z) + g_i(\boldsymbol{w}) \leq 0\} \geq \delta_i,\ \forall i \in \mathbb{N}_{[1,N]}\right\}.$$

Proposition 2. *Given a probability of unreliability $\beta \in (0,1)$. Then, with probability greater than $1-\beta$,*

$$\mathscr{C} \supseteq \bigcap_{i \in \mathbb{N}_{[1,N]}} \left\{z \in \mathbb{R}^{n_z} : f_i(z) + \hat{Q}_{\boldsymbol{g}_i}^M(\Delta_i^{\text{inner}}) \leq 0\right\}, \quad (9a)$$

$$\mathscr{C} \subseteq \bigcap_{i \in \mathbb{N}_{[1,N]}} \left\{z \in \mathbb{R}^{n_z} : f_i(z) + \hat{Q}_{\boldsymbol{g}_i}^M(\Delta_i^{\text{outer}}) \leq 0\right\}, \quad (9b)$$

where

$$\Delta_i^{\text{inner}} = \delta_i + \sqrt{\frac{\ln(N) + \ln(1/\beta)}{2M}} \in (\delta_i, 1], \quad (10a)$$

$$\Delta_i^{\text{outer}} = \delta_i - \sqrt{\frac{\ln(N) + \ln(1/\beta)}{2M}} \in [0, \delta_i), \quad (10b)$$

provided the number of samples available M satisfies $M \geq M_{\min} \triangleq (\ln(N) + \ln(1/\beta)) / (2\Delta_{min}^2)$*, where* $\Delta_{min} = \min\left(\min_{i\in\mathbb{N}_{[1,N]}} (1-\delta_i), \min_{i\in\mathbb{N}_{[1,N]}} \delta_i\right) \in (0,1)$.

Proposition 2 (see Appendix for a proof sketch) provides a sufficient bound on the number of samples $M_{\min}$ as a function of the number of constraints N, probability of unreliability β, and risk thresholds δ_i. It also prescribes sample risk thresholds Δ^{inner} and Δ^{outer} that can inner-approximate (9a) and outer-approximate (9b) the set of decision variables that satisfy the constraint set $\mathscr{C}$.

Remark 2. (SIMPLIFIED LOWER BOUND ON M) *For $\delta_i \in [0.5, 1)$ for all $i \in \mathbb{N}_{[1,N]}$,*

$$M_{\min} = \frac{\ln(N) + \ln(1/\beta)}{2\big(1 - \max_{i\in\mathbb{N}_{[1,N]}} \delta_i\big)^2}. \tag{11}$$

Recall that the finite sample lower bound in [8] for solving (3) with *convex* f_i is $M_{\min}^{\text{convex}} = \frac{2}{1-\delta}(\ln(1/\beta) + n_z)$. Observe that $M_{\min}^{\text{convex}}$ has a linear dependence on the number of decision variables n_z. In contrast, despite the possibility of non-convex f_i in (3), $M_{\min}$ is independent of n_z. Both $M_{\min}$ and $M_{\min}^{\text{convex}}$ have a logarithmic dependence on $1/\beta$. The dependence of $M_{\min}$ on δ degrades to a squared dependence on $1/(1-\delta)$ from a linear dependence. Thus, as δ approaches one, the rate of increase in $M_{\min}^{\text{convex}}$ is smaller than $M_{\min}$, which may be attributed to the structural assumption of convex f_i used to derive $M_{\min}^{\text{convex}}$.

We conclude by observing that a sample-based, inner-approximation of $\mathscr{C}$ is non-empty only if $\delta \in (0, \delta_{\text{max}}]$, where $\delta_{\text{max}} \triangleq 1 - \sqrt{\frac{\ln(N)+\ln(1/\beta)}{2M}}$, for given N, M and β by (10a).

C. Feasible solutions of (3) with suboptimality bounds

Motivated by Proposition 2, we consider the following optimization problems to address Problem 1,

$$\text{(Safe-Approx)} : \begin{array}{ll} \text{min.} & f_0(z) \\ \text{s. t.} & z \in \mathcal{Z}, \\ & f_i(z) + \hat{Q}^M_{\boldsymbol{g}_i}(\Delta_i^{\text{inner}}) \leq 0 \end{array}$$

$$\text{(Unsafe-Approx)} : \begin{array}{ll} \text{min.} & f_0(z) \\ \text{s. t.} & z \in \mathcal{Z}, \\ & f_i(z) + \hat{Q}^M_{\boldsymbol{g}_i}(\Delta_i^{\text{outer}}) \leq 0 \end{array}$$

where the sample risk thresholds $\Delta_i^{\text{inner}}, \Delta_i^{\text{outer}}$ are selected based on (10), and $M \geq M_{\min}$.

We will use the solution of (Safe-Approx) to obtain a suboptimal solution of (3), and use the solution of (Unsafe-Approx) to provide a bound on the suboptimality. We provide the justification for these steps in Theorem 1.

Theorem 1. *For any user-specified probability of unreliability $\beta \in (0,1)$, the following statements are true with probability $1-\beta$:*
a) Any feasible solution of (Safe-Approx) *is feasible for* (3).
b) Every feasible solution $z^\dagger$ of (Safe-Approx) *satisfies the sub-optimality bound,*

$$0 \leq f_0(z^\dagger) - f_0(z^*) \leq f_0(z^\dagger) - \kappa_M, \tag{12}$$

where z^* is the unknown global optimum of (3)*, and κ_M is the finite global optimal value of* (Unsafe-Approx).

Algorithm 1: Sample quantile-based programming

Input: Problem (3) with $\delta_i \in [0.5, 1)$, $\nu \geq 1$, probability of unreliability $\beta \in (0,1)$, sample generator for $\boldsymbol{w}$
Output: $z^\dagger$, a suboptimal solution of (3)
1: Compute $M_{\min}$ using (11) based on β, δ_i, N
2: Obtain $M = \lfloor \nu M_{\min} \rfloor$ samples of $\boldsymbol{w}$
3: **for** $i \in \mathbb{N}_{[1,N]}$
4: Compute Δ_i^{inner} using (10a)
5: Compute $\hat{Q}^M_{\boldsymbol{g}_i}(\Delta_i^{\text{inner}})$ using (2b)
6: Solve (Safe-Approx) to local/global optimality for $z^\dagger$
7: *Optional*: Return suboptimality bound $f_0(z^\dagger) - \kappa_M$, κ_M is the global opt. value of (Unsafe-Approx)

We summarize the proposed approach in Algorithm 1, that computes a suboptimal solution of (3) using Theorem 1. The sample scaling ν in Step 2 helps the user trade-off the availability of samples of $\boldsymbol{w}$ with the desire to reduce $\Delta_i^{\text{inner}} - \delta_i$ that indirectly reduces the conservativeness of $z^\dagger$. Note that $\hat{Q}^M_{\boldsymbol{g}_i}$ in Step 5 can be evaluated efficiently using a `max-heap` data structure [18, Sec. 6.1]. By design, the computation of $z^\dagger$ in Step 6 is unaffected by the size of M.

D. Discussion

1) What happens if (Unsafe-Approx) *does not produce a finite global optimum?:* Theorem 1 excludes two cases when solving (Unsafe-Approx) — infeasiblity ($\kappa_M = \infty$) and unboundedness ($\kappa_M = -\infty$).

When (Unsafe-Approx) is infeasible, (3) is infeasible with a probability of $1-\beta$. On the other hand, (12) yields a trivial suboptimality bound of ∞ when (Unsafe-Approx) is unbounded. (Unsafe-Approx) is guaranteed to have a bounded (global) optimal value when (3) satisfies additional requirements (e.g., $\mathcal{Z}$ is compact and f_0 is continuous).

2) Comparison with moment-based constraint tightening: A popular approach to enforce chance constraints is via moments. For example, the Chebyshev-Cantelli bound prescribes a constraint tightening for the chance constraint relying only on the first and second moments of the random variable $g(\boldsymbol{w})$.

Lemma 3 (CHEBYSHEV-CANTELLI BOUND [5])**.** *Let the random variable $g(\boldsymbol{w})$ have a finite mean $\mu_{\boldsymbol{g}}$ and standard deviation $\sigma_{\boldsymbol{g}}$. Then, for any $z \in \mathbb{R}^{n_z}$ and threshold $\delta \in (0,1)$,*

$$f(z) + \mu_{\boldsymbol{g}} + \sigma_{\boldsymbol{g}}\sqrt{\delta/(1-\delta)} \leq 0 \;\Rightarrow\; \mathbb{P}_{\boldsymbol{w}}\{f(z) + g(\boldsymbol{w}) \leq 0\} \geq \delta.$$

Consider (3c) in the form of (8) with $N = 1$, $\boldsymbol{w} \sim \mathcal{N}(0,1)$ where $n_w = 1$, an arbitrary f, g as an identity map, and varying threshold δ. By (1b), (3c) is equivalent to $\{z : f(z) + Q_{\boldsymbol{w}}(\delta) \leq 0\}$ for each δ.

Figure 1 compares the exact constraint tightening $Q_{\boldsymbol{w}}(\delta)$ with the alternatives discussed in this paper — the proposed constraint tightening approach ($\hat{Q}^M_{\boldsymbol{g}}(\Delta^{\text{inner}})$ in Proposition 2) and the moment-based constraint tightening (Lemma 3). We

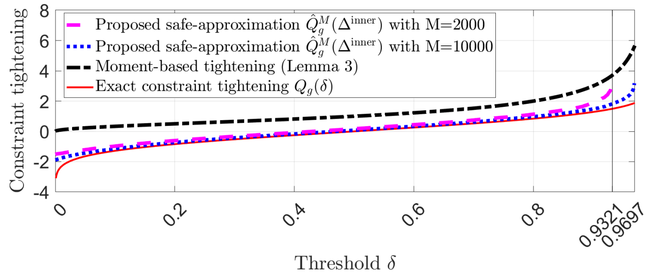

Fig. 1. Proposed constraint tightening $\hat{Q}_{\boldsymbol{g}}^M(\Delta^{\text{inner}})$ for $M \in \{2000, 10^5\}$ is less conservative than the moment-based tightening (Lemma 3).

TABLE I

PROPOSED SOLUTION (Safe-Approx) COMPUTES A SAFE, NEAR-OPTIMAL MOTION PLAN USING JUST THE SAMPLES OF $\boldsymbol{r}_\ell$.

	Proposed solution		Optimal solution	Moments-based
Formulation	Safe-Approx	Unsafe-Approx	(13) with (15b)	(13) with (15a)
Min. over all (t, ℓ) collision-avoidance probability (13d)	0.977	0.923	0.950	0.987
Cost	122	115	117	129
Suboptimality	5	–	–	12

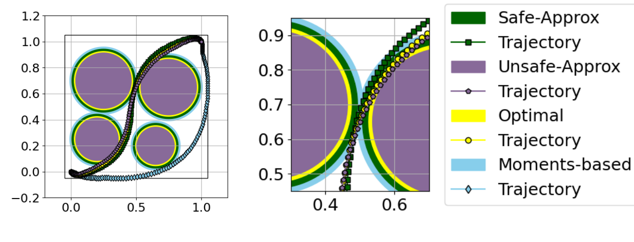

Fig. 2. (Safe-Approx) (our proposed approach) computes a safe near-optimal motion plan, while the moment-based approach produces a conservative motion plan. (Left) The computed motion plans are safe. (Right) Zoomed-in view between the obstacles illustrate the clearances.

performed the comparison over varying δ for two different values of $M \in \{2000, 10000\}$, and β set to 10^{-8}. We used the true mean and the standard deviation of $\boldsymbol{w}$ in the moment-based tightening.

Figure 1 shows that the proposed tightening $\hat{Q}_{\boldsymbol{g}}^M(\Delta^{\text{inner}})$ provides a more precise approximation (smaller upper bound of the true quantile-based tightening $Q_{\boldsymbol{g}}$) for all $\delta \in (0, \delta_{\max}]$, when compared to the moment-based tightening. Additionally, as M increases, the proposed safe-approximation yields a smaller constraint tightening and has a higher $\delta_{\max}$.

We emphasize that the proposed tightening only needs samples of $\boldsymbol{w}$ to compute $\hat{Q}_{\boldsymbol{g}}^M(\Delta^{\text{inner}})$, unlike the moment-based tightening and the exact constraint tightening. The latter approaches require prior knowledge of the first two moments of $\boldsymbol{w}$ and the (true) quantile function of $\boldsymbol{w}$ respectively, which if estimated from samples can invalidate the approximation guarantees. Additionally, the dependence on the finite (and fixed in Figure 1) number of samples also explains the (relatively) faster increase of the proposed tightening $\hat{Q}_{\boldsymbol{g}}^M(\Delta^{\text{inner}})$ for δ closer to one, compared to the latter approaches.

IV. APPLICATION: MOTION PLANNING WITH STOCHASTIC OBSTACLES

We consider the problem of a robot navigating in a constrained environment with $L \in \mathbb{N}$ obstacles that have non-deterministic geometries. Such problems arise when accurate descriptions of the geometry are unavailable, possibly due to sensing limitations.

The robot has double-integrator dynamics in x and y coordinates, $x_{t+1} = Ax_t + Bu_t$, for each $t \in \mathbb{N}_{[0,T-1]}$, with state $x_t \in \mathbb{R}^4$ denoting its position and velocity in each dimension, input $u_t \in \mathcal{U}$ where the convex and compact set $\mathcal{U} \subset \mathbb{R}^2$ encodes the actuation constraints, and $T \in \mathbb{N}$ is the planning horizon.

We denote the desired target state as $x_{\text{target}} \in \mathbb{R}^4$. We also require that the robot stays within a convex and compact *keep-in* set $\mathcal{K} \subset \mathbb{R}^4$. We assume that each of the L obstacles are balls with known centers $c_\ell \in \mathbb{R}^2$ but a stochastic radius $\boldsymbol{r}_\ell$ for all $\ell \in \mathbb{N}_{[1,L]}$. We do not assume the knowledge of the probability measure or the distribution function associated with $\boldsymbol{r}_\ell$. Instead, we assume the ability to generate samples of $\boldsymbol{r}_\ell$ for any $M \in \mathbb{N}$ (e.g., noisy readings of the obstacle geometry from sensors).

Given the initial state of the robot $x_0 \in \mathcal{K}$, we wish to solve the following (non-convex) motion planning problem,

$$\begin{aligned}
&\text{minimize} && \textstyle\sum_{t=1}^{T} b_t && \text{(13a)}\\
&\text{subject to} && x_{t+1} = Ax_t + Bu_t,\ \forall t \in \mathbb{N}_{[0,T-1]}, && \text{(13b)}\\
&\forall t \in \mathbb{N}_{[1,T]} && u_{t-1} \in \mathcal{U},\ b_t \in \{0,1\},\ x_t \in \mathcal{K}, && \text{(13c)}\\
&{\scriptstyle \forall t \in \mathbb{N}_{[1,T]}, \atop \ell \in \mathbb{N}_{[1,L]}}, && \mathbb{P}\{\|Cx_t - c_\ell\| \geq \boldsymbol{r}_\ell\} \geq \delta, && \text{(13d)}\\
&\forall t \in \mathbb{N}_{[1,T]}, && \|x_t - x_{\text{target}}\| \leq b_t D_{\mathcal{K}}. && \text{(13e)}
\end{aligned}$$

with $C = [I_2\ 0_{2\times 2}] \in \mathbb{R}^{2\times 4}$ and $D_{\mathcal{K}} = \sup_{x,y\in\mathcal{K}} \|x - y\|$ as the diameter of the set $\mathcal{K}$. Problem (13) is a mixed-integer nonlinear program with continuous decision variables $x_t \in \mathbb{R}^4$ for $t \in \mathbb{N}_{[1,T]}$ and $u_t \in \mathbb{R}^2$ for $t \in \mathbb{N}_{[0,T-1]}$, and binary decision variables $b_t \in \{0,1\}$ for $t \in \mathbb{N}_{[1,T]}$. The constraints in (13c) ensure that the motion plan satisfies actuation constraints and stays within the keep-in set. The constraint (13d) requires that the risk of collision at each time step $t \in \mathbb{N}_{[1,T]}$ and for each obstacle $j \in \mathbb{N}_{[1,L]}$ is below $1 - \delta$, for a user-specified $\delta \in (0,1)$. Using the "big-M" formulation [11], (13a) and (13e) enforce that the optimal solution of (13) has the robot taking the least number of steps to reach the target with $b_t = 1$ whenever $\|x_t - x_{\text{target}}\| > D_{\mathcal{K}}$.

We now discuss various reformulations of (13d) that is non-convex in the decision variables. First, we obtain (Safe-Approx) and (Unsafe-Approx) used in Theorem 1 by replacing (13d) in (13) with the following constraints,

$$\|Cx_t - c_\ell\| \geq \hat{Q}_{\boldsymbol{r}_\ell}^M(\Delta^{\text{inner}}), \tag{14a}$$
$$\|Cx_t - c_\ell\| \geq \hat{Q}_{\boldsymbol{r}_\ell}^M(\Delta^{\text{outer}}), \tag{14b}$$

for every $t \in \mathbb{N}_{[1,T]}$ and $\ell \in \mathbb{N}_{[1,L]}$ respectively. Here, the sample risk thresholds $\Delta^{\text{inner}}, \Delta^{\text{outer}}$ in (14) are computed via (10) using M samples of $\boldsymbol{r}_\ell$, and a user-specified probability of unreliability $\beta \in (0,1)$. We also compare the proposed sample-based approximations with existing moment-based approach (Lemma 3) and exact quantile-based approach (8),

$$\|Cx_t - c_\ell\| \geq \mu_{\boldsymbol{r}_\ell} + \sigma_{\boldsymbol{r}_\ell}\sqrt{\delta/(1-\delta)}, \tag{15a}$$

$$\|Cx_t - c_\ell\| \geq Q_{\boldsymbol{r}_\ell}(\delta). \tag{15b}$$

Unlike the proposed approach, the constraint reformulations in (15) require prior knowledge of the distribution or the moments of $\boldsymbol{r}_\ell$. The traditional scenario-based approaches [8] cannot be used to solve (13) due to the non-convexity (constraints (13d) and (13e) and the binary decision variables).

The reformulations of (13) using either (14a), (14b), (15a), or (15b) result in a mixed-integer program with non-convex quadratic constraints. We use GUROBI [19] to solve the resulting optimization problems to global optimality. However, we emphasize that the results in Theorem 1 hold even if a locally optimal motion plan was obtained for (Safe-Approx). For example, under some feasible assignment of binary variables, (13) simplifies to a non-convex optimization problem with only continuous variables, which can be solved to a local optimum via a local nonlinear optimization solver.

Figure 2 shows the various motion plans for $\delta = 0.95$, $\beta = 10^{-8}$, $T = 150$, $L = 4$ (thus, $N = 600$ and $M_{\min} = 4043$), $M = 10^4$ with $\nu \approx 2.5$, $x_0 = [0, 0, 0, 0]^\top$, $x_{\text{target}} = [1, 1, 0, 0]^\top$, and $\mathcal{K} = [-0.1, 1.1]^2 \times [-5, 5]^2$. We considered four obstacles with their radii given by an exponential of mean 0.025. The motion plan of the proposed sample quantile-based safe-approximation ((13) with (14a)) is similar to the optimal motion plan ((13) with (15b)). However, due to the inherent conservativeness of the moments-based bound, the computed motion plan (using (13) with (15a)) is unnecessarily conservative, resulting in longer path.

Table I summarizes the results along with a Monte-Carlo simulation-based validation as well as the (sub)optimality of each approach. We used 10^7 samples in the Monte-Carlo simulations to compute the satisfaction of the constraint $\|Cx_t - c_\ell\| \geq \boldsymbol{r}_\ell$ for every $t \in \mathbb{N}_{[1,T]}$ and $\ell \in \mathbb{N}_{[1,L]}$, and reported the lowest collision-avoidance probability (13d) over all obstacles and time steps. As expected, the collision-avoidance probability of (Safe-Approx), optimal solution, and the moment-based solution are above the specified threshold of $\delta = 0.95$, while the collision-avoidance probability of (Unsafe-Approx) is below $\delta = 0.95$. Moreover, the collision-avoidance probability of (Safe-Approx) is closer to δ than the collision-avoidance probability of the moment-based method highlighting again the conservativeness of the moment-based approach. Using the solution from (Unsafe-Approx), we have a suboptimality bound $f_0(z^\dagger) - \kappa_M = 7$ on the solution of (Safe-Approx).

V. Conclusion and Future Work

This paper proposes a sample quantile-based approach to solve non-convex separable chance constrained programs without any prior knowledge of the distribution or the moments of the uncertainty. We characterize tractable approximations of the chance constrained programs using the sample quantile function. We also determine the minimum number of samples and the appropriate sample risk thresholds needed to compute a feasible solution with suboptimality bounds. Unlike existing scenario-based approaches, we impose no requirements on the convexity of the constraint functions. However, a limitation of our current approach is the separability requirement on the chance constraints. Our future work will investigate more general classes of chance constraints that can be reformulated using the sample quantile function.

References

[1] A. Shapiro, D. Dentcheva, and A. Ruszczynski, *Lectures on stochastic programming: modeling and theory*. SIAM, 2021.

[2] A. Prékopa, *Stochastic programming*. Springer, 2013, vol. 324.

[3] F. Oldewurtel, C. Jones, A. Parisio, and M. Morari, "Stochastic model predictive control for building climate control," *IEEE Trans. Ctrl. Syst. Tech.*, vol. 22, no. 3, pp. 1198–1205, 2013.

[4] S. Priore, C. Petersen, and M. Oishi, "Approximate quantiles for stochastic optimal control of LTI systems with arbitrary disturbances," in *Amer. Ctrl. Conf.*, 2022, pp. 1814–1821.

[5] J. Paulson, E. Buehler, R. Braatz, and A. Mesbah, "Stochastic model predictive control with joint chance constraints," *Int'l J. Ctrl.*, 2020.

[6] A. Nemirovski and A. Shapiro, "Convex approximations of chance constrained programs," *J. Opt*, vol. 17, pp. 969–996, 2007.

[7] M. Vidyasagar, *Learning and generalisation: with applications to neural networks*. Springer Science & Business Media, 2013.

[8] M. Campi and S. Garatti, "A sampling-and-discarding approach to chance-constrained optimization: feasibility and optimality," *J. Opt. Th. App.*, vol. 148, no. 2, pp. 257–280, 2011.

[9] S. Grammatico, X. Zhang, K. Margellos, P. Goulart, and J. Lygeros, "A scenario approach for non-convex control design," *IEEE Trans. Auto. Ctrl.*, vol. 61, no. 2, pp. 334–345, 2015.

[10] M. C. Campi, S. Garatti, and F. A. Ramponi, "A general scenario theory for nonconvex optimization and decision making," *IEEE Trans. Auto. Ctrl.*, vol. 63, no. 12, pp. 4067–4078, 2018.

[11] L. Blackmore, M. Ono, A. Bektassov, and B. Williams, "A probabilistic particle-control approximation of chance-constrained stochastic predictive control," *IEEE Trans. Robot.*, vol. 26, pp. 502–517, 2010.

[12] J. Luedtke and S. Ahmed, "A sample approximation approach for optimization with probabilistic constraints," *J. Opt.*, vol. 19, 2008.

[13] V. Sivaramakrishnan and M. Oishi, "Fast, convexified stochastic optimal open-loop control for linear systems using empirical characteristic functions," *IEEE Ctrl. Syst. Lett.*, vol. 4, pp. 1048–1053, 2020.

[14] A. Peña-Ordieres, J. Luedtke, and A. Wachter, "Solving chance-constrained problems via a smooth sample-based nonlinear approximation," *J. Opt.*, vol. 30, pp. 2221–2250, 2020.

[15] R. Reiss, *Approximate Distributions of Order Statistics*, ser. Springer Series in Statistics. New York, NY: Springer New York, 1989.

[16] R. J. Serfling, *Approximation theorems of mathematical statistics*. John Wiley & Sons, 2009, vol. 162.

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

[18] T. Cormen, C. Leiserson, R. Rivest, and C. Stein, *Introduction to algorithms*. MIT press, 2022.

[19] Gurobi Optimization, LLC, "Gurobi Optimizer Reference Manual," 2022. [Online]. Available: https://www.gurobi.com

Appendix

1) Proof sketch of Proposition 1: In (5) and (7), the objective f_0 and the constraint $z \in \mathcal{Z}$ are identical. To prove that the constraint (7c) is an exact reformulation of (5c), we use (2b) to show that $\hat{\Phi}^M_{f_i(z)+g_i(\boldsymbol{w})}(0) = \hat{\Phi}^M_{\boldsymbol{g}_i}(-f_i(z))$ for all $i \in \mathbb{N}_{[1,N]}$. Then, we complete the proof with Lemma 1.

2) Proof sketch of Proposition 2: Define $\varepsilon = \sqrt{\frac{\ln(N)+\ln(1/\beta)}{2M}}$, then, $2\exp(-2M\varepsilon^2) = \frac{\beta}{N}$. The proof for (9) follows from Boole's inequality, Proposition 1, Lemma 2, and the sample risk thresholds Δ_i^{inner} and Δ_i^{outer} defined in (10) for all $\delta_i \in (0, 1)$.

The lower bound on M follows from the restrictions on M to ensure that $\Delta_i^{\text{inner}} \leq 1$ and $\Delta_i^{\text{outer}} \geq 0$ for every $i \in \mathbb{N}_{[1,N]}$, given δ_i, β, and N.