ISSN 0005-1179 (print), ISSN 1608-3032 (online), Automation and Remote Control, 2025, Vol. 86, No. 8, pp. 740–755.

SPECIAL ISSUE

Tight Approximations of Chance Constrained Sets Through Pack-Based Probabilistic Scaling

V. Mirasierra[*,a], **M. Mammarella**[**,b], **F. Dabbene**[**,c], **and T. Alamo**[*,d]

[*] *Departamento de Ingeniería de Sistemas y Automática, Universidad de Sevilla, Escuela Superior de Ingenieros, Camino de los Descubrimientos s/n, 41092 Sevilla, Spain*
[**] *Cnr-Istituto di Elettronica e di Ingegneria dell'Informazione e delle Telecomunicazioni (CNR-IEIIT), c/o Politecnico di Torino, Corso Duca Degli Abruzzi, 10129, Italy*
e-mail: [a] *vmirasierra@us.es,* [b] *martina.mammarella@cnr.it,* [c] *fabrizio.dabbene@cnr.it,* [d] *talamo@us.es*

Received March 3, 2025
Revised May 20, 2025
Accepted June 27, 2025

Abstract—The computation of probabilistic safe regions remains an evergreen problem in stochastic settings. Although the exact computation of safe regions may be possible for some specific problems, the results are generally overly complex (e.g., nonconvex, nonconnected) making them impractical for real-time applications. In this work, we present a sample-based procedure to obtain tight inner approximations of the safe region. The proposed approach does not require any assumption on the underlying probability distribution and the computation of the inner approximation set can be done offline. Unlike similar approaches, the proposed pack-based probabilistic scaling includes a tightening constraint, which tunes the level of conservativeness of the resulting approximation.

Keywords: randomized algorithms, probabilistic robustness, uncertain systems, statistical learning theory

DOI: 10.31857/S0005117925080049

1. INTRODUCTION

Real-world systems are often not deterministic and subject to uncertainty, necessitating the development of robust and stochastic control strategies. In robust control [1–3], the uncertainty is assumed to be unknown, but confined in a compact region, and the controller is designed to guarantee constraint satisfaction for all admissible values of the uncertainty. In contrast, stochastic control [4–6], incorporates probabilistic considerations introducing the concept of chance constraints [7]. Unlike hard constraints, chance constraints can be occasionally violated, provided that the probability of satisfaction remains above a specified threshold.

Relaxing the constraints and taking probabilities into account make stochastic schemes less conservative than their robust counterpart. Moreover, they make it possible to deal with infinite support uncertainties. In return, the resulting design process is much more intricate for two main reasons: First, it is highly difficult to check whether solutions of chance-constrained problems are feasible, and second, chance constraints usually involve nonconvexity (see, e.g., [8, Fig. 1; 9, Fig.1].

In the last decade, sampling-based schemes have emerged as a valid tool to deal with stochastic problems. Notably, Prof. Boris Polyak played a pivotal role in this field, being among the first scholars to recognize the potential of randomized methods in tackling optimization problems under stochastic uncertainty; for instance, see the works [10–12]. These works paved the way for subsequent results combining sampling and optimization, as the scenario approach proposed in [13]. For an overview of these techniques, the reader is referred to [14, 15].

The probabilistic safe region or *chance-constrained set* (CCS) is defined as the region that contains all the points satisfying the chance constraints. In a general setting, the exact computation of the CCS is cumbersome and requires the uncertainty to follow a certain distribution [16, 17]. Besides, the complexity of its geometry can make it ill-suited for real-time applications [18]. Because of these limitations, it is pertinent to address the problem of approximating the safe regions using sets of manageable complexity.

For the stochastic control problem, several relaxations have been proposed, which rely on computationally efficient approximations of the chance-constrained set. These relaxations can be either based on some concentration inequalities, e.g. exploiting previous knowledge about the structure of the uncertainty [19], or they can be constructed using random sampling methods [20, 21].

The present work stems from the results in [9, 22], where a sample-based methodology to inner approximate the CCS named *probabilistic scaling* is presented. This approach computes first a simple approximating set, which is then scaled to meet the required probabilistic guarantees. These operations are all performed *offline* and the trade-off between the number of samples required and the tightening of the approximation can be adjusted by the user.

In this paper, we discuss and extend the pack-based probabilistic scaling approach presented in the preliminary conference publication [22], by defining a novel measure of the tightening of the approximating set. Then, we show how to design the approximating set to meet the required probabilistic guarantees while incorporating the specified tightening constraint. In this way, the user is given the capability to control at the same time the complexity and the fitting of the resulting approximating set, balancing the trade-off between the required number of samples and the computational complexity of the approximation problem (which is computed offline).

The paper is structured as follows. In Section 2 we introduce the problem of approximating the chance constrained set and the numerical example used to compare the different approaches. In Section 3 we go through statistical learning theory solutions to the problem, first introducing the classical probabilistic scaling approach (Section 3.1) and later describing the extension to the pack-based framework (Section 3.2). Then, Section 4 is dedicated to the tight immersed pack-based probabilistic scaling, which is the main contribution of this work. Last, Section 5 includes the comparative analysis of the different approaches in terms of conservativeness.

Notation: $\mathbb{N}_{\geqslant 0}$ is the set of natural numbers including 0. The notation $\oplus$ refers to the Minkowski sum of sets. Given a set of N scalars $\{x_1, x_2, \dots, x_N\}$, we denote $x_{1:N}$ the smallest one, $x_{2:N}$ the second smallest one, and so on and so forth until $x_{N:N}$, which is the largest. By the definition of $x_{1+r:N}$, for a given $r \geqslant 0$, no more than r elements of $\{x_1, \dots, x_N\}$ are strictly smaller than $x_{1+r:N}$. We refer to the binomial distribution as

$$B(s; N, \varepsilon) = \sum_{i=0}^{s} \binom{N}{i} \varepsilon^i (1-\varepsilon)^{N-i}.$$

2. APPROXIMATING CHANCE-CONSTRAINED SETS

Let us consider a robustness problem, where the controller parameters and the auxiliary variables are parameterized by means of a decision variable vector $\theta \in \Theta \subseteq \mathbb{R}^{n_\theta}$, which is denoted as *design parameter*. The uncertainty vector w represents one of the admissible uncertainty realizations of a random vector with given probability distribution $\Pr_{\mathcal{W}}$ with (possibly unbounded) support $\mathcal{W}$. Then, the generic uncertain constraint can be defined as

$$g(\theta, w) \leqslant 0, \tag{1}$$

where the function $g : \mathbb{R}^{\Theta \times \mathcal{W}} \to \mathbb{R}$ captures the requirement for θ given w. In particular, in a robust setting, one requires that the constraint (1) holds for all possible values of w. Clearly, there might

be situations where dealing with this kind of constraint in a fully robust manner is senseless, e.g., when the support of w is unbounded [23]. In that case, one may accept that the constraint (1) is violated by a fraction of the elements of $\mathcal{W}$. This concept is rigorously formalized in the definition of chance constraints.

Definition 1 [set of probability ε-CCS [9]]. Consider a probability measure $\Pr_{\mathcal{W}}$ over $\mathcal{W}$. Given the violation level $\varepsilon \in (0,1)$, we define the chance-constrained set of probability ε (ε-CCS) as follows

$$\Omega(\varepsilon) = \{\theta \in \Theta \mid \Pr_{\mathcal{W}}\{g(\theta, w) > 0\} \leqslant \varepsilon\}.$$

Recently, several approaches have been proposed to construct a probabilistically guaranteed approximation of the chance-constrained set. These approaches are based on sample-based results (see e.g., [21, 24, 25]). Given $\mathcal{W}$, consider a collection of N independent identically distributed (i.i.d.) samples $\mathbf{z} = \{w_1, \ldots, w_N\}$ drawn from $\mathcal{W}$. In this case, we say that $\mathbf{z}$ belongs to the Cartesian product $\mathcal{W}^N \doteq \mathcal{W} \times \cdots \times \mathcal{W}$ (N times) and, correspondingly, we say that $\mathbf{z}$ is drawn according to the product probability measure $\Pr_{\mathcal{W}^N}$. Let us introduce the concept of an indicator function, later used to redefine the chance-constrained set.

Definition 2 [indicator function of g]. Given $\theta \in \Theta$ and $w \in \mathcal{W}$, then the indicator function $I^g : \Theta \times \mathcal{W} \to \{0,1\}$ of constraint (1) is defined as

$$I^g(\theta, w) \doteq \begin{cases} 0 & \text{if } g(\theta, w) \leq 0 \\ 1 & \text{otherwise.} \end{cases}$$

In the context of statistical learning theory, we can compute approximations of the ε-CCS by means of a constraint on the *empirical mean* defined as

$$\frac{1}{N}\sum_{i=1}^{N} I^g(\theta, w_i).$$

That is, given $\mathbf{z} = \{w_1, \ldots, w_N\} \in \mathcal{W}^N$ and a discarding parameter $r \geqslant 0$, then the parameter $\rho = \frac{r}{N}$ bounds the empirical mean so that the set

$$\Phi_{\rho_N}(\mathbf{z}) \doteq \left\{\theta \in \Theta : \frac{1}{N}\sum_{i=1}^{N} I^g(\theta, w_i) \leqslant \rho\right\} \tag{2}$$

constitutes an approximation of $\Omega(\varepsilon)$. Note that the expression $\frac{1}{N}\sum_{i=1}^{N} I^g(\theta, w_i) \leqslant \frac{r}{N}$ means that the constraint $g(\theta, w_i) \leq 0$ is violated by no more than r elements of $\mathbf{z}$.

Remark 1. We note that, given ε, $\Omega(\varepsilon)$ is a fixed set. On the other hand, when the ε-CCS is approximated by means of sampling techniques (see e.g., [9, 26]), then the corresponding approximated set has a random nature, being generated from the random samples $\mathbf{z} \in \mathcal{W}^{\mathbf{N}}$.

Assuming that the indicator function I^g has finite Vapnik–Chervonenkis (VC) dimension [27] and that $\rho < \varepsilon$, then the probability of $\Phi_{\rho_N}(\mathbf{z})$ being an inner approximation of $\Omega(\varepsilon)$, i.e.,

$$\Pr_{\mathcal{W}^N}\{\Phi_{\rho_N}(\mathbf{z}) \subseteq \Omega(\varepsilon)\}$$

converges to 1 as the number of samples N converges to infinity. In [28], the sample complexity bounds for N are explicitly computed, which guarantee that $\Phi_{\rho_N}(\mathbf{z})$ is included in $\Omega(\varepsilon)$ with a given confidence $\delta \in (0,1)$, i.e., $\Pr_{\mathcal{W}^N}\{\Phi_{\rho_N}(\mathbf{z}) \subseteq \Omega(\varepsilon)\} \geqslant 1 - \delta$.

The resulting sample complexity grows linearly with the VC dimension of I^g multiplied by a factor larger than $\frac{1}{\varepsilon}$. However, as shown in [9], this approximation may be very conservative. Also, when the function g is not convex, the resulting approximation is generally non-convex and is often non-connected. This may hinder its practical application and makes it generally unsuitable for real-time problems.

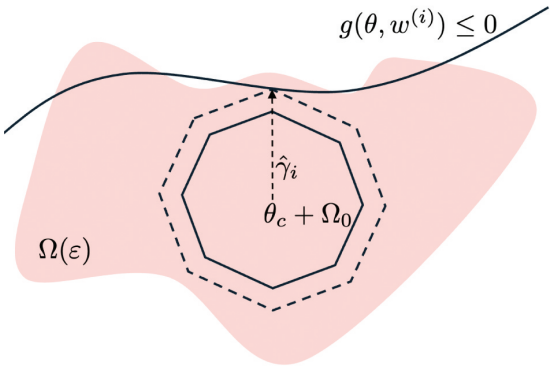

Fig. 1. Scheme of the probabilistic scaling approach.

Building on these notions, [9, 22] introduced the probabilistic scaling idea. At the basis of this approach is the introduction of an initial *simple approximating set* (SAS) $\theta_c \oplus \Omega_0$, which has to possess two main characteristics: i) be able to capture sufficiently well the "shape" of the probabilistic set $\Omega(\varepsilon)$, while at the same time being ii) sufficiently simple. This initial SAS does not need to offer any guarantee of probabilistic nature, but it should be able to capture the shape of the ε-CCS.

In [9] it was shown how to *scale* this set around its center θ_c to obtain a scalable SAS

$$\Omega(\gamma) = \theta_c \oplus \gamma\Omega_0,$$

and a sample-based procedure was introduced to construct a probabilistically meaningful approximation of the ε-CCS. Specifically, given a shape Ω_0 and a scaling center θ_c, the goal of probabilistic scaling is to find the largest scaling factor $\bar{\gamma}$ such that

$$\Pr_{\mathcal{W}}\{\theta_c \oplus \bar{\gamma}\Omega_0 \subseteq \Omega(\varepsilon)\} \geqslant 1 - \delta, \tag{3}$$

and therefore, also the chance constraint

$$\Pr_{\mathcal{W}}\{g(\theta, w) \leqslant 0\} \geq 1 - \varepsilon \tag{4}$$

is satisfied with a probability not lower than $1 - \delta$.

The procedure for constructing such an approximation is discussed in detail in [9], and recalled formally in Section 3.1.

In Fig. 1, we give a simple illustration of the approach, where we assume that the red area represents the ε-CCS, which as observed can be in general nonconvex. Then:

(1) Select "candidate" approximating set $\theta_c + \Omega_0$ (black polygon);
(2) To design the optimal scaling $\bar{\gamma}$, extract N samples $\mathbf{z} = \{w_1, \ldots, w_N\} \in \mathcal{W}^N$;
(3) For each random sample w_i, compute the maximum scaling γ_i so that the scaled set (dashed polygon) does not violate the constraint corresponding to w_i;
(4) Select the optimal scaling as $\bar{\gamma} = \gamma_{1+r:N}$, i.e., as the r smallest value of γ_i.

Then, (3) holds for $B(r; N, \varepsilon) \leqslant \delta$.

Despite the undeniable benefits of exploiting *probabilistic scaling*, especially in the extended version where the computational complexity is further reduced by employing the so-called simple-approximating sets (SAS) [9], the scaling solution may result to be conservative. This issue is illustrated by means of the following example from [22].

Example 1. We consider a problem involving individual chance constraints, where every constraint is tangent to the unit circle of a given dimension at a random point, drawn from a uniform distribution. In this case, clearly, the unit ball is the safe region with probability 1, whereas the

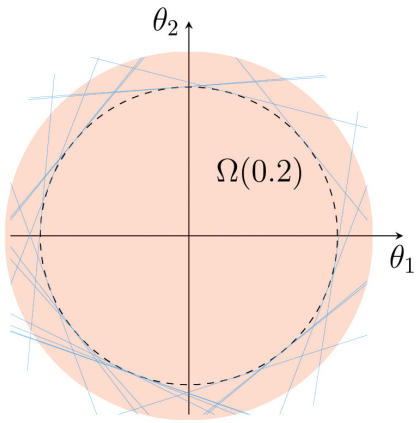

Fig. 2. The red circle represents $\Omega(0.2)$, the dashed black circle is the SAS (unit circle), and the cyan lines are the sampled constraints.

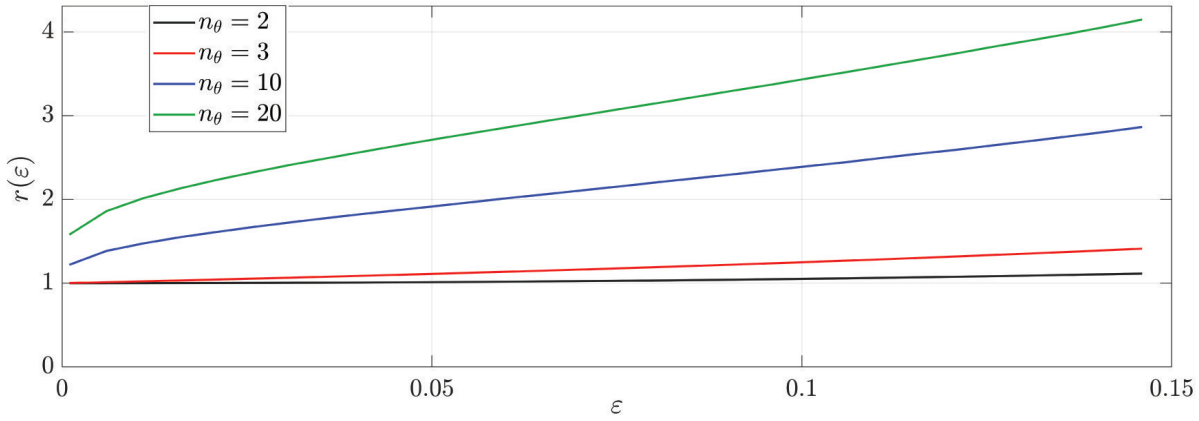

Fig. 3. Radius r of $\Omega(\varepsilon)$ as a function of ε for different problem dimensions n_θ.

ε-CCS is always a slightly larger scaled version of the unit ball as ε increases. In particular, it can be easily shown that the exact radius corresponding to the chance constrained region $\Omega(\varepsilon)$ can be computed using some transcendental functions. Figure 2 illustrates this example in $\mathbb{R}^2$: where the dashed line is the unit circle in $\mathbb{R}^2$ and the outer red circle represents the chance constrained set $\Omega(\varepsilon)$ for the specific value $\varepsilon = 0.20$.

Assume that we want to approximate the ε-CCS using the empirical mean approximation $\Phi_{\rho_N}(\mathbf{z})$ introduced in (2). To this end, we generate N random linear constraints tangent to points drawn from a uniform probability distribution on the surface of the unit hypersphere and construct the approximation as the intersection of them (possibly discarding the "worst" ones). It is clear that such an approximation will fail to capture the red circle.

Additionally, assume we want to use a probabilistic scaling approach, and we choose the unit ball as the initial approximation $\theta_c \oplus \Omega_0$ of the chance-constrained set $\Omega(\varepsilon)$. Then, applying the previously described procedure, it would be possible to scale this initial geometry around its center (the origin) to obtain an inner approximation of $\Omega(\varepsilon)$ with a given confidence level $\delta \in (0, 1)$. However, it is evident that the scaling scheme will always yield the unit hypersphere as a final result, as each sampled constraint is tangent to it, implying that all computed scaling factors will be equal to one. Hence, simple sampling-based procedures will fail to capture the radius of the true set $\Omega(\varepsilon)$. Note that this radius may be significantly larger than one, especially when the n_θ increases, as shown in Fig. 3.

On the other hand, for the given example, one may notice that larger scale factors can be obtained if one scales the unit-circle taking into consideration only the regions in which more than a given number of constraints are violated. In this paper, we resort to the *pack-based strategy*, successfully employed in the context of statistical learning theory [28] and convex scenario [29], to obtain less conservative sample complexities and to guarantee that the obtained scaled set is included into the chance constrained set with a given confidence level. Specifically, the goal is to extend the pack-based strategy first proposed in [22] to obtain sample-based approximations of $\Omega(\varepsilon)$ with tunable complexity, which do not require any previous knowledge of the problem, e.g., symmetry. The ability to reduce the conservativeness of the proposed approach will be later demonstrated against the illustrative Example 1.

3. PRELIMINARY NOTIONS

In this section, we first recall some notions from the pack-based strategy, which are propaedeutical to the main results of this paper. Then, in the next section we present the pack-based probabilistic scaling (PBPS) approach discussed in [22], which will be later extended to further reduce the conservativeness of the approximating set. First, we introduce the definition of *pack of samples*.

Definition 3 [pack of L samples]. Given an integer L, a collection of L samples $\mathbf{z} = \{w_1, \dots, w_L\} \in \mathcal{W}^L$ is said to be a pack of dimension L.

Then, we extend the definition of indicator function in Definition 2 to the pack-based framework.

Definition 4 [pack indicator function]. Given integers s and L such that $0 \leqslant s < L$ and a pack $\mathbf{z} \in \mathcal{W}^L$ of dimension L, the *pack indicator function* $I_s^g : \Theta \times \mathcal{W}^L \to \{0, 1\}$ is defined as

$$I_s^g(\theta, \mathbf{z}) \doteq \begin{cases} 0 & \text{if } \sum_{\ell=1}^{L} I^g(\theta, w_\ell) \leqslant s \\ 1 & \text{otherwise,} \end{cases} \tag{5}$$

where $I_s^g(\theta, \mathbf{z})$ indicates whether the point θ violates more than s of the constraints associated with the uncertainty realizations of the pack $\mathbf{z}$.

Definition 5 [pack safe region]. The pack safe region $\Phi_s(\mathbf{z})$ is defined as the set of points which violate no more than s of the constraints associated with the uncertainty realizations of $\mathbf{z}$, and can be expressed as

$$\Phi_s(\mathbf{z}) \doteq \{\theta \in \Theta \,|\, I_s^g(\theta, \mathbf{z}) = 0\}.$$

In the next section, we present a generalization of the results on probabilistic scaling applied in the framework of pack-based strategy. In detail, we show how to scale the set $\theta_c \oplus \Omega_0$ around its center θ_c to guarantee with confidence level $\delta \in (0, 1)$, the inclusion of the scaled, pack-based set into $\Omega(\varepsilon)$.

3.1. Generalized Probabilistic Scaling

First, we introduce the definition of scaling factor in the pack-based framework.

Definition 6 [pack scaling factor]. Given a scalable SAS $\Omega(\gamma)$ defined by a scaling center $\theta_c \in \Theta$ and a shape Ω_0, and a pack $\mathbf{z} \in \mathcal{W}^L$, we define the pack scaling factor of $\Omega(\gamma)$ relative to the random constraints $g(\theta, w_i) \leqslant 0, \forall w_i \in \mathbf{z}$ as

$$\gamma^s(\theta_c, \Omega_0, \mathbf{z}) \doteq \begin{cases} 0 & \text{if } \theta_c \notin \Phi_s(\mathbf{z}) \\ \max\limits_{\theta_c \oplus \gamma\Omega_0 \subseteq \Phi_s(\mathbf{z})} \gamma & \text{otherwise.} \end{cases} \tag{6}$$

Now, we formalize the *generalized probabilistic scaling problem*, considering M i.i.d. packs $\mathbf{z}_i$, each one of dimension L. Note that the problem generalizes the probabilistic scaling introduced in [9], which can be obtained by letting $M = N$ and $L = 1$ (i.e., considering N packs of dimension 1).

Property 1 [generalized probabilistic scaling]. Given the accuracy parameter $\varepsilon \in (0, 1)$ and the confidence level $\delta \in (0, 1)$, consider the discarding integer parameter $r \geqslant 0$ and suppose that M is chosen such that

$$B(r; M, \varepsilon) \leqslant \delta. \tag{7}$$

Draw M i.i.d. L-dimensional packs, $\mathbf{z}_i \in \mathcal{W}^L$, $i = 1, \dots M$. For each pack $\mathbf{z}$, compute the corresponding pack scaling factor γ_i as

$$\gamma_i \doteq \gamma^s(\theta_c, \Omega_0, \mathbf{z})$$

according to (6) and define $\bar{\gamma} \doteq \gamma_{1+r:M} > 0$. Then, with probability no smaller than $1 - \delta$,

$$\Pr_{\mathcal{W}^L}\{\theta_c \oplus \bar{\gamma}\Omega_0 \not\subseteq \Phi_s(\mathbf{z})\} \leqslant \varepsilon.$$

Proof. This property can be demonstrated by particularizing the results of convex scenario [21, 26] to the case of a scalar decision variable. Another possibility is to derive the results using the properties of the generalized max function [30, Property 3]. Consider the following optimization problem:

$$\begin{aligned} &\max_{\gamma} \quad \gamma \\ &s.t. \quad \theta_c \oplus \gamma\Omega_0 \subseteq \Phi_s(\mathbf{z}), \quad i = 1, \dots, M. \end{aligned} \tag{8}$$

If this problem has a feasible solution, then we can rewrite it using the definition of $\gamma^s(\cdot)$ as

$$\begin{aligned} &\max_{\gamma} \quad \gamma \\ &s.t. \quad \gamma \leqslant \gamma^s(\theta_c, \Omega_0, \mathbf{z}), \quad i = 1, \dots, M. \end{aligned} \tag{9}$$

It has been proved in [21, 26] that if one discards no more than r constraints on a convex problem with M random constraints, then the probability of violating the constraints with the solution obtained from the random convex problem is no larger than ε, with probability no smaller than $1 - \delta$, where

$$\delta = \begin{pmatrix} d + r - 1 \\ d - 1 \end{pmatrix} B(d + r - 1; M, \varepsilon),$$

and d is the number of decision variables. We first notice that (9) is convex and has a unique scalar decision variable γ, i.e., $d = 1$. Also, the assumptions required in the application of the results of [21, 26] can be easily checked. In particular, non-degeneracy is implied by the fact that the problem is scalar, while uniqueness can be enforced by introducing a tie-break rule. Hence, if we allow r violations in the above minimization problem, then with probability no smaller than $1 - \delta$, with $\delta = B(r; M, \varepsilon)$, the optimal solution $\bar{\gamma}$ of problem (9) with no more than r constraint removed satisfies

$$\Pr_{\mathcal{W}^L}\{\bar{\gamma} > \gamma^s(\theta_c, \Omega_0, \mathbf{z})\} \leqslant \varepsilon.$$

Hence, we can conclude that with probability no smaller than $1 - \delta$

$$\Pr_{\mathcal{W}^L}\{\theta_c \oplus \bar{\gamma}\Omega_0 \not\subseteq \Phi_s(\mathbf{z})\} \leqslant \varepsilon.$$

Note that problem (9) with constraint removal can be solved directly by ordering the values $\gamma_i = \gamma^s(\theta_c, \Omega_0, \mathbf{z}_i)$. It is clear that if $r \geqslant 0$ violations are allowed, then the optimal value for γ is $\bar{\gamma} = \gamma_{1+r:N}$. Smaller values of γ would meet the inclusion of constraint but will not be optimal, while larger values of γ would no longer meet the inclusion constraint □.

As discussed before, the result in [22, Property 1] can be particularized from Property 2 by setting $M = N$ and $L = 1$. This is summarized in the next corollary.

Corollary 1 [classic probabilistic scaling]. *Suppose that N is chosen such that*

$$B(r; N, \varepsilon) \leqslant \delta.$$

Let $\mathbf{z} \in \mathcal{W}^N$. For each constraint $i = 1, \ldots, N$, define $\gamma_i = \gamma^r(\theta_c, \Omega_0, \mathbf{z}_i)$. Suppose that $\bar{\gamma} = \gamma_{1+r:N} > 0$. Then, with probability no smaller than $1 - \delta$,

$$\Pr_{\mathcal{W}}\{\theta_c \oplus \bar{\gamma}\Omega_0 \not\subseteq \Omega(\varepsilon)\} \leqslant \varepsilon.$$

The proof is straightforward and follows directly from Definition 1. The above corollary shows that the probabilistic scaling approach in [9] can be viewed as a special case of a more general pack-based scheme.

Calculating approximations of the ε-CCS using classical probabilistic scaling is generally easy to compute, does not require any assumption on the underlying probabilities (such as finite VC dimension), provides probabilistic guarantees to the scaled region, and its effectiveness has been proven [9]. Despite all its advantages, classical probabilistic scaling may still lead to very conservative solutions, as shown in Example 1. In that case, having that $\gamma_i = 1$, for all $i = 1, \ldots, N$ and all the constraints are taken into account independently, the act of discarding some of them has no effect on the resulting scaled approximating set.

In the next section, we outline the so-called pack-based probabilistic scaling, first proposed in [22]. For the same initial SAS, this variant of the classical probabilistic scaling applied in the framework of pack-based strategy may lead to less conservative results at the expense of (possibly) more demanding computational cost.

It is important to highlight that the discarding parameter r is set by the user and it should be selected taking into account that large values of r make the resulting set less sensitive to extreme values, at the expense of a larger sample complexity N. On the other hand, the convexity of the approximating scaled set is independent of the discarding parameter r and only depends on the choice of the SAS geometry.

Remark 2. Property 1 can also be particularized for the case $r = 0$. Suppose that M is such that $(1 - \varepsilon)^M \leqslant \delta$. Draw M i.i.d. L-dimensional packs $\mathbf{z}_i \in \mathcal{W}^L$ and define $\gamma_i \doteq \gamma^s(\theta_c, \Omega_0, \mathbf{z}_i)$. Suppose that $\bar{\gamma} = \gamma_{1:M} > 0$. Then, with probability no smaller than $1 - \delta$, $\Pr_{\mathcal{W}^L}\{\theta_c \oplus \bar{\gamma}\Omega_0 \not\subseteq \Phi_s(\mathbf{z})\} \leqslant \varepsilon$.

3.2. Pack-Based Probabilistic Scaling

The main underlying idea of pack-based probabilistic scaling is to divide the uncertainty samples into packs and to allow some constraint violations inside each pack. As opposed to regular probabilistic scaling, where the scaling factor associated with each constraint is computed *independently*, in the pack-based approach the constraints inside each pack are taken into account *together*. Ultimately, this can lead to tighter approximations of the ε-CCS and reduced sample complexity.

Let the N sampled constraints be divided into M packs of L constraints each, i.e., $\mathbf{z} = \{\mathbf{z}_1, \ldots \mathbf{z}_M\} = \{w_1, \ldots, w_N\}$, with $\mathbf{z} \in \mathcal{W}^N$ and $\mathbf{z}_i \in \mathcal{W}^L$ for $i = 1, \ldots, M$. The following theorem shows how to determine the scaling factor using a pack-based approach so that the scaled SAS is fully contained in the ε-CCS with given confidence δ.

Theorem 1 [pack-based probabilistic scaling]. *Consider a shape Ω_0, a scaling center θ_c, accuracy parameter $\varepsilon \in (0, 1)$, confidence level $\delta \in (0, 1)$, and nonnegative integers M, L, s with $L > s$ so that*

$$B(s; L, \varepsilon)^M \leqslant \delta. \tag{10}$$

For each pack of constraints $i = 1, \ldots, M$, let $\mathbf{z}_i \in \mathcal{W}^L$ and define $\gamma_i \doteq \gamma^s(\theta_c, \Omega_0, \mathbf{z}^{(i)})$ as in (6). Suppose that $\bar{\gamma} = \gamma_{1:M} > 0$. Then, with probability no smaller than $1 - \delta$,

$$\theta_c \oplus \bar{\gamma}\Omega_0 \subseteq \Omega(\varepsilon).$$

Proof. Let $p = 1 - B(s; L, \varepsilon)$. From Remark 2, we know that if M is chosen such that $B(s; L, \varepsilon)^M \leqslant \delta$ and $\bar{\gamma} = \gamma_{1:M} > 0$, then with probability no smaller than $1 - \delta$ we have

$$\Pr_{\mathcal{W}^L}\{\theta_c \oplus \bar{\gamma}\Omega_0 \not\subseteq \Phi_s(\mathbf{z})\} \leqslant p.$$

Equivalently, $\Pr_{\mathcal{W}^L}\{I_s^g(\theta, \mathbf{z}) = 1, \ \ \forall \theta \in \theta_c \oplus \bar{\gamma}\Omega_0\} \leqslant p$. Moreover, from Property 3 in Appendix A we have that

$$\Pr_{\mathcal{W}^L}\{I_s^g(\theta, \mathbf{z}) = 1\} \leqslant p \iff \Pr_{\mathcal{W}}\{I^g(\theta, w) = 1\} \leqslant \varepsilon. \tag{11}$$

Thus, we conclude that $\Pr_{\mathcal{W}}^L\{I^g(\theta, w) = 1\} \leqslant \varepsilon$, $\forall \theta \in \theta_c \oplus \bar{\gamma}\Omega_0$, equivalent to $\theta_c \oplus \bar{\gamma}\Omega_0 \subseteq \Omega(\varepsilon)$. □

Unlike regular probabilistic scaling, the sample complexity in PBPS is given by two parameters, namely the number of packs M and the size of each pack L. The sample complexity is calculated as $N = ML$. Consequently, condition (10) is defined by three tunable parameters: M, L and s. Similar to the discarding parameter r of regular probabilistic scaling, large values of the discarding parameter of each pack s make the approximating set more insensitive to extreme values. As for M and L, one could choose them according to any criterion, e.g., minimize the sample complexity N. Further details can be found in [22].

In the next section, we extend the PBPS approach introducing a constraint tightening scheme, namely the *tight immersion*, to obtain a tighter approximation of the ε-CCS. Moreover, this extension will provide a clear way to select the tuning parameters, as discussed in Section 4.1.

4. TIGHT IMMERSION

First, we introduce the notion of tight immersion.

Definition 7. τ-tight immersed. The set $\mathcal{S}$ is τ-tight immersed in the ε-CCS $\Omega(\varepsilon)$ if

$$\mathcal{S} \subseteq \Omega(\varepsilon), \tag{12a}$$
$$\mathcal{S} \not\subseteq \Omega(\tau\varepsilon), \tag{12b}$$

where $\tau \in [0, 1)$ is a measure of tightening.

Remark 3. If the ε-CCS $\Omega(\varepsilon)$ is strictly increasing with respect to ε, i.e., $\forall \tau_1, \tau_2 \in [0, 1)$ with $\tau_1 < \tau_2$, it follows that $\Omega(\tau_1\varepsilon) \subset \Omega(\tau_2\varepsilon)$. Hence, the larger τ is, the larger $\Omega(\tau\varepsilon)$ will be.

Tight immersion guarantees not only that the approximation set is *inside* the ε-CCS (12a), but also that it will be *not inside* a conservative set characterized by τ (12b). Therefore, it imposes a more restrictive condition than the regular *inner approximation*. However, tight immersion should never be used to compare the goodness of two different geometries. Indeed, as illustrated in Fig. 4,

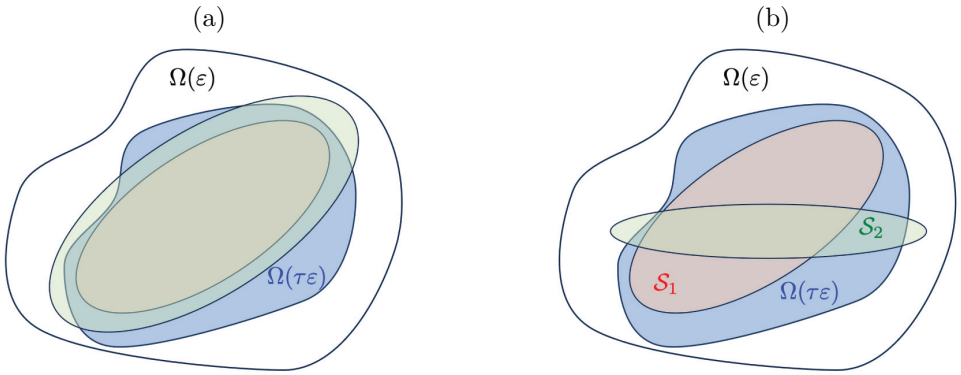

Fig. 4. Illustration of the concept of tight immersion.

we note that for the same geometry, the set with the largest value of τ fits the ε-CCS better (Fig. 4a). Instead, from Fig. 4b we note that for different geometries tight immersion by itself does not imply good approximation.

The following property is complementary to the Definition 7 of tight immersion.

Property 2. If the approximating set $\underline{\Omega}(\varepsilon)$ is τ-tight immersed in the set $\Omega(\varepsilon)$, then it is also $\tilde{\tau}$-tight immersed in it, with $\tilde{\tau} \in [0, \tau)$.

Proof. From Definition 7, we know that being $\underline{\Omega}(\varepsilon)$ τ-tight immersed in the set $\Omega(\varepsilon)$ implies $\underline{\Omega}(\varepsilon) \not\subseteq \Omega(\tau\varepsilon)$. Then, for any $\tilde{\tau} \in [0, \tau)$, we have $\Omega(\bar{\tau}\varepsilon) \subseteq \Omega(\tau\varepsilon)$. Consequently, the condition $\underline{\Omega}(\varepsilon) \not\subseteq \Omega(\bar{\tau}\varepsilon)$ (12b) holds for any $\tilde{\tau} \in [0, \tau)$ and this concludes the proof. □

Next, we finally demonstrate how to determine the pack parameters (M, L, s) so that, upon pack-based probabilistic scaling, the condition (12b) is met with confidence $1 - \bar{\delta}$, with $\bar{\delta} \in (0, 1)$. Hence, given a SAS Ω_0 centered in θ_c, we aim to determine the optimal scaling factor $\bar{\gamma}$ so that the scaled set $\mathcal{S} = \theta_c \oplus \bar{\gamma}\Omega_0$ is tight-immersed in $\Omega(\varepsilon)$.

Theorem 2 [tight-immersed pack-based probabilistic scaling]. *Consider the SAS with shape Ω_0 and scaling center θ_c, accuracy parameter $\varepsilon \in (0, 1)$, confidence level $\bar{\delta} \in (0, 1)$, tightening parameter $\tau \in [0, 1)$, and non negative integers M, L, s, with $L > s$ and such that the following condition holds*

$$B(s; L, \tau\varepsilon)^M \geqslant 1 - \bar{\delta}. \tag{13}$$

Draw M i.i.d. multisamples $\mathbf{z}_i \in \mathcal{W}^L$, with $i = 1, \ldots, M$, and define the pack scaling factor related to each i pack of random constraints as in (6), i.e., $\gamma_i = \gamma^s(\theta_c, \Omega_0, \mathbf{z}_i)$. Suppose that $\bar{\gamma} = \gamma_{1:M} > 0$. Then, with probability no smaller than $1 - \bar{\delta}$,

$$\theta_c \oplus \bar{\gamma}\Omega_0 \subseteq \Omega(\varepsilon), \quad \theta_c \oplus \bar{\gamma}\Omega_0 \not\subseteq \Omega(\tau\varepsilon).$$

Proof. Let $p = B(s; L, \tau\varepsilon)$. According to Property 5 in Appendix C, if we select the parameters (M, L, s) such that (13) holds, then we have that the optimal scaling factor is $\bar{\gamma} = \gamma_{1:M} > 0$ satisfies, with probability no smaller than $1 - \delta$, $\Pr_{\mathcal{W}^L}\{\theta_c \oplus \bar{\gamma}\Omega_0 \subseteq \Phi_s(\mathbf{z})\} \leqslant p$, equivalently rewritten as

$$\Pr_{\mathcal{W}^L}\{I_s^g(\theta, \mathbf{z}) = 0, \ \ \forall \theta \in \theta_c \oplus \bar{\gamma}\Omega_0\} \leqslant p.$$

Then, from Property 4 in Appendix B, we know that

$$\Pr_{\mathcal{W}^L}\{I_s^g(\theta, \mathbf{z}) = 0\} \leqslant p \iff \Pr_{\mathcal{W}}\{I^g(\theta, w) = 0\} \leqslant 1 - \tau\varepsilon. \tag{14}$$

Therefore, we can conclude that

$$\Pr_{\mathcal{W}}\{I^g(\theta, w) = 0\} \leqslant 1 - \tau\varepsilon, \quad \forall \theta \in \theta_c \oplus \bar{\gamma}\Omega_0,$$

i.e., $\theta_c \oplus \bar{\gamma}\Omega_0 \not\subseteq \Omega(\tau\varepsilon)$ with probability no smaller than $1 - \bar{\delta}$. □

Remark 4. We note that in Theorem 2 we use the tightening confidence $1 - \bar{\delta}$ instead of the original confidence $1 - \delta$. This tightening confidence is *user-defined* and can be set lower than the original confidence to limit the sample complexity.

4.1. Design of the Pack Parameters

In this section, we show how to design the parameters (M, L, s) of the pack-based approach to meet tight immersion with confidences δ and $\bar{\delta}$, respectively. From Property 1 and Theorem 2, we know that conditions (12a) and (12b) hold if the pack parameters (M, L, s) are selected such that

$$M \ln B(s; L, \varepsilon) \leqslant \ln \delta, \tag{15a}$$

$$M \ln B(s; L, \tau\varepsilon) \geqslant \ln(1 - \bar{\delta}). \tag{15b}$$

We note that (15a) embeds the probabilistic guarantees whereas (15b) is only used to tighten the solution. Since $B(s;L,\varepsilon)$ is a negative quantity, we can divide (15a) by $\ln B(s;L,\varepsilon)$ to obtain

$$M \geqslant \frac{\ln \delta}{\ln B(s;L,\varepsilon)}.$$

Hence, to satisfy (15a), it suffices to select M such that

$$M = \left\lceil \frac{\ln \delta}{\ln\, B(s;L,\varepsilon)} \right\rceil. \tag{16}$$

Analogously, for (15b) we have that M shall be selected so that the following condition holds

$$M \leqslant \frac{\ln(1-\bar{\delta})}{\ln B(s;L,\tau\varepsilon)}. \tag{17}$$

For a given set of probabilistic and tightening parameters $(\varepsilon, \delta, \tau, \bar{\delta})$, there exist multiple combinations of (M, L, s) that meet (16) and (17). In this paper, we propose two different criteria ζ: (i) minimize the number of possible combinations of $s+1$ constraints, i.e., $\zeta = M\binom{L}{s+1}$, or (ii) minimize the total sample complexity, i.e., $\zeta = ML$. Then, the pack parameters (M, L, s) are the solution of the following optimization problem

$$\begin{aligned} (M^o, L^o, s^o) = \underset{M,L,s\in\mathbb{N}_{\geqslant 0}}{\operatorname{argmin}} \quad & \zeta \\ \text{s.t.} \quad & M \leqslant \frac{\ln(1-\bar{\delta})}{\ln B(s;L,\tau\varepsilon)} \\ & M = \left\lceil \frac{\ln \delta}{\ln B(s;L,\varepsilon)} \right\rceil, \\ & L \geqslant s+1. \end{aligned} \tag{18}$$

To solve Problem (18), we exploit the *exhaustive search* approach [31] to find a proper combination of the pack parameters (M, L, s), as shown in the following example.

Example 2. Given $\varepsilon = 0.05$, $\delta = 0.001$, $\bar{\delta} = 0.1$, for each $s = [1, 30]$, we set M according to (16). Then, we test the values $L = [s+1, \cdots, s+300]$ and check if the pairs (L, s) satisfy (17). Last, among all the pairs that satisfy (17), we select the one that minimizes the ζ criterion (either $\zeta = M\binom{L}{s+1}$ or $\zeta = ML$). In Table 1, we report the pack parameters (M, L, s) obtained by solving Problem (18) with the proposed approach using both criteria ζ. Table 1 shows that, for either criterion, increasing the tightening parameter τ results in an increase in both the number of samples required (N) and the combinatorial complexity $(M\binom{L}{s+1})$. When the number of available samples

Table 1. Pack parameters, sample complexity and number of possible combinations of TI-PBPS for different values of $\tau = [0.2, 0.3, 0.4, 0.5]$ minimizing the two different criteria ζ

Criterion: minimize $M\binom{L}{s+1}$						Criterion: Minimize N				
τ	M	L	s	N	$M\binom{L}{s+1}$	M	L	s	N	$M\binom{L}{s+1}$
0.2	43	27	2	1.16e+03	1.25e+05	2	195	4	3.90e+02	4.46e+09
0.3	155	27	3	4.19e+03	2.72e+06	2	303	8	6.06e+02	1.05e+17
0.4	2681	20	4	5.36e+04	4.16e+07	4	278	10	1.11e+03	6.28e+19
0.5	15033	29	6	4.36e+05	2.35e+10	8	309	14	2.47e+03	9.68e+25

is limited, it is possible to achieve tight immersion with as little as 390 samples at the cost of a high combinatorial complexity. Moreover, it is possible to add any limitation on the number of samples $N = ML$ as a constraint in the optimization problem (18) and obtain the pack parameters that minimize the combinatorial complexity while satisfying the constraint on sample complexity. Notice that tight immersion is usually computed offline; therefore, the complexity of calculating the approximation of the CCS does not interfere with online control loops.

5. RESULTS

In this section, we use Example 1 to evaluate the approximations of the 0.05-CCS set by employing both regular PS and TI-PBPS for various problem dimensions n_θ. Hence, we choose the unit ball centered in the origin as the initial SAS approximation $\theta_c \oplus \Omega_0$. Therefore, the resulting approximating sets are balls centered in the origin and with radius α. Moreover, by means of a Monte Carlo simulation, we draw 10^7 random constraints from a uniform distribution of the constraints tangent to the unit circle of each studied dimension. Taking advantage of the symmetry of the problem, we calculate the points where the random constraints intersect a fixed axis and use them to compute the exact value of the radius for 0.05-CCS. Then, we compute the radii of the approximating sets obtained by employing regular probabilistic scaling and the novel TI-PBPS for different levels of tightening τ. To reduce variability, the TI-PBPS radii correspond to the median radius of three separated experiments, each containing different realizations of the constraints.

In Fig. 5 we can observe how, for this particular problem, the TI-PBPS is able to substantially improve the result from regular PS (dashed black line), providing approximating radii more similar to the real one (dashed red line). Moreover, as expected, the tightening of TI-PBPS improves as τ increases.

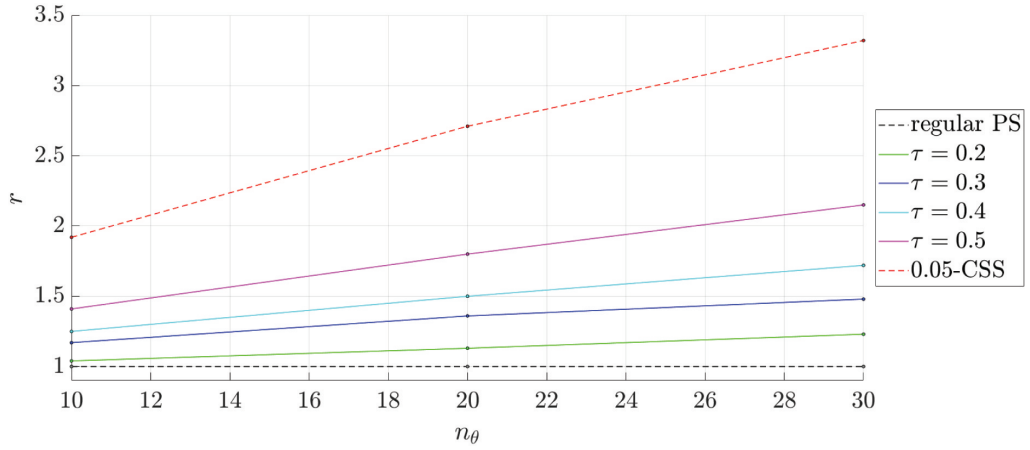

Fig. 5. Comparison of the radius (α) of the approximation set for different problem dimensions n_θ obtained by applying regular PS and TI-PBPS for $\varepsilon = 0.05$, $\delta = 0.001$, and $\bar{\delta} = 0.1$.

6. CONCLUSIONS

In this paper we have presented the probabilistic scaling approach to compute sample-based approximations of a chance constrained set. The proposed approach allows the user to first choose any set and then apply a linear transformation to approximate the safe region with the desired probabilistic guarantees. As a result, the complexity of the approximation is tuned a priori. A pack-based variant of probabilistic scaling with a tight-immersed approach is proposed, which prevents the solution from being conservative. The trade-off between the number of samples, problem complexity,

and the level of conservativeness of this approach can be tuned by the user. Future research directions point towards improving the proposed solution, e.g. applying importance sampling schemes, inspired by [32].

APPENDIX A

Property 3. Consider the integer parameters $L > s \geqslant 0$, the pack $\mathbf{z} \in \mathcal{W}^L$, and the probability parameter $\varepsilon \in (0,1)$. Then, for $w \in \mathcal{W}$, it holds

$$\Pr_{\mathcal{W}^L}\{I_s^g(\theta,\mathbf{z})=1\} \leqslant 1 - B(s;L,\varepsilon) \iff \Pr_{\mathcal{W}}\{I^g(\theta,w)=1\} \leqslant \varepsilon. \tag{A.1}$$

Proof. Define $E(\theta) = \Pr_{\mathcal{W}}\{I^g(\theta,w)=1\}$. Then, we have

$$\Pr_{\mathcal{W}^L}\{I_s^g(\theta,\mathbf{z})=0\} = \sum_{i=0}^{s}\binom{L}{i}E(\theta)^i(1-E(\theta))^{L-i} = B(s;L,E(\theta)). \tag{A.2}$$

Denote $p = 1-B(s;L,\varepsilon)$. Since $B(s;L,\varepsilon)$ is a strictly decreasing function of ε (see [23, Property 4]), we have

$$B(s;L,E(\theta)) \geqslant B(s;L,\varepsilon) = 1-p \iff E(\theta) \leqslant \varepsilon. \tag{A.3}$$

Therefore, we have

$$\Pr_{\mathcal{W}^L}\{I_s^g(\theta,\mathbf{z})=1\} \leqslant p \iff \Pr_{\mathcal{W}}\{I^g(\theta,w)=1\} \leqslant \varepsilon, \tag{A.4}$$

which concludes the proof. □

APPENDIX B

Property 4. Consider the integer parameters $L > s \geqslant 0$, the pack $\mathbf{z} \in \mathcal{W}^L$, the sample $w \in \mathcal{W}$, and probability parameter $\varepsilon \in (0,1)$. Then, we have

$$\Pr_{\mathcal{W}^L}\{I_s^g(\theta,\mathbf{z})=0\} \leqslant B(s;L,\tau\varepsilon) \iff \Pr_{\mathcal{W}}\{I^g(\theta,w)=0\} \leqslant 1-\tau\varepsilon.$$

Proof. Recalling the definition of $E(\theta) = \Pr_{\mathcal{W}}\{I^g(\theta,w)=1\}$, we have

$$\Pr_{\mathcal{W}^L}\{I_s^g(\theta,\mathbf{z})=0\} = \sum_{i=0}^{s}\binom{L}{i}E(\theta)^i(1-E(\theta))^{L-i} = B(s;L,E(\theta)).$$

Since $B(s;L,\tau\varepsilon)$ is strictly decreasing with respect of $\tau\varepsilon$ (Property 4 of [23]), we obtain

$$\Pr_{\mathcal{W}^L}\{I_s^g(\theta,\mathbf{z})=0\} = B(s;L,E(\theta)) \leqslant B(s;L,\tau\varepsilon) \iff \Pr_{\mathcal{W}}\{I^g(\theta,w)=1\} = E(\theta) \geqslant \tau\varepsilon.$$

Therefore,

$$\Pr_{\mathcal{W}^L}\{I_s^g(\theta,\mathbf{z})=0\} \leqslant B(s;L,\tau\varepsilon) \iff \Pr_{\mathcal{W}}\{I^g(\theta,w)=0\} \leqslant 1-\tau\varepsilon. \qquad \square$$

APPENDIX C

Property 5. Given the accuracy parameter $p \in (0,1)$ and the confidence level $\bar{\delta} \in (0,1)$, suppose that the number of packs M is chosen such that the following condition holds

$$1 - p^M \leqslant \bar{\delta},$$

Then, for each pack of constraints $i = 1, \dots, M$, draw the M i.i.d. multisamples $\mathbf{z} \sim \Pr_{\mathcal{W}^L}$ and define $\gamma_i = \gamma^s(c, \Omega_0, \mathbf{z}_i)$. Suppose that $\bar{\gamma} = \gamma_{1:M} > 0$. Then, with probability no smaller than $1 - \bar{\delta}$,

$$\Pr^M_{\mathcal{W}^L}\{c \oplus \bar{\gamma}\Omega_0 \subseteq \Phi^g_s(\mathbf{z})\} \leqslant p.$$

Proof. The proof follows the one of Property 1. Consider the following optimization problem

$$\begin{aligned} \min_{\gamma} \quad & \gamma \\ s.t. \quad & c \oplus \gamma\Omega_0 \not\subseteq \Phi^g_s(\mathbf{z}_)), \quad i = 1, \dots, M. \end{aligned} \tag{C.1}$$

If problem (C.1) has a feasible solution, according to (6) we can rewrite (C.1) as

$$\begin{aligned} \min_{\gamma} \quad & \gamma \\ s.t. \quad & \gamma > \gamma^s(c, \Omega_0, \mathbf{z}_)), \quad i = 1, \dots, M. \end{aligned} \tag{C.2}$$

According to the sampling-and-discard approach [21, 26], if one discards no more than $M - 1$ constraints and the number of decision variables d is $d = 1$, then the probability of violating the scaled approximating constraint set is no larger than $p \in (0, 1)$, with probability no smaller than $1 - \bar{\delta}$, where the confidence level $\bar{\delta}$ is defined as follows:

$$\begin{aligned} \bar{\delta} &= \begin{pmatrix} M-1 \\ M-1 \end{pmatrix} B(M-1; M, p) = \sum_{i=0}^{M-1} \begin{pmatrix} M \\ i \end{pmatrix} p^i (1-p)^{M-i} \\ &= 1 - \sum_{i=M}^{M} \begin{pmatrix} M \\ i \end{pmatrix} p^i (1-p)^{M-i} = 1 - p^M \end{aligned}$$

If we remove no more than $M - 1$ constraints, the optimal solution to Problem (C.2) is given by $\bar{\gamma} = \gamma_{1:M}$, with $\gamma_i = \gamma^s(c, \Omega_0, \mathbf{z})$. Correspondingly, we have $\Pr^M_{\mathcal{W}^L}\{\bar{\gamma} \leqslant \gamma(c, \Omega_0, \mathbf{z})\} \leqslant p$, from which we can conclude that, with probability no smaller than $1 - \bar{\delta}$,

$$\Pr^M_{\mathcal{W}^L}\{\theta_c \oplus \bar{\gamma}\Omega_0 \subseteq \Phi^g_s(\mathbf{z})\} \leqslant p. \qquad \square$$

REFERENCES

1. Bemporad, A. and Morari, M., Robust Model Predictive Control: A Survey, in *Robustness in Identification and Control*, Garulli, A. and Tesi, A., Eds., *Lect. Notes Control Inf. Sci.*, vol. 245, London: Springer, 1999, pp. 1–25. https://doi.org/10.1007/BFb0109870
2. Mayne, D., Seron, M. and Raković, S., Robust Model Predictive Control of Constrained Linear Systems with Bounded Disturbances, *Automatica*, 2005, vol. 41, no. 2, pp. 219–224.
3. Mayne, D., Raković, S., Findeisen, R., and Allgöwer, F., Robust Output Feedback Model Predictive Control of Constrained Linear Systems, *Automatica*, 2006, vol. 42, no. 7, pp. 1217–1222.
4. Lorenzen, M., Dabbene, F., Tempo, R., and Allgöwer, F., Constraint-Tightening and Stability in Stochastic Model Predictive Control, *IEEE Trans. Autom. Control*, 2016, vol. 62, no. 7, pp. 3165–3177.
5. Farina, M., Giulioni, L., and Scattolini, R., Stochastic Linear Model Predictive Control with Chance Constraints – a Review, *J. Process Control*, 2016, vol. 44, pp. 53–67.
6. Mesbah, A., Stochastic Model Predictive Control: An Overview and Perspectives for Future Research, *IEEE Control Syst. Mag.*, 2016, vol. 36, no. 6, pp. 30–44.
7. Charnes, A. and Cooper, W., Chance Constraints and Normal Deviates, *J. Amer. Statist. Assoc.*, 1962, vol. 57, no. 297, pp. 134–148.

8. Kücükyavuz, S. and Jiang, R., Chance-Constrained Optimization under Limited Distributional Information: A Review of Reformulations Based on Sampling and Distributional Robustness, *EURO J. Comput. Optim.*, 2022, vol. 10, art. no. 100030.

9. Mammarella, M., Mirasierra, V., Lorenzen, M., Alamo, T., and Dabbene, F., Chance-Constrained Sets Approximation: A Probabilistic Scaling Approach, *Automatica*, 2022, vol. 137, art. no. 110108.

10. Polyak, B. and Tempo, R., Probabilistic Robust Design with Linear Quadratic Regulators, *Syst. Control Lett.*, 2001, vol. 43, no. 5, pp. 343–353.

11. Calafiore, G. and Polyak, B., Stochastic Algorithms for Exact and Approximate Feasibility of Robust LMIs, *IEEE Trans. Autom. Control*, 2001, vol. 46, no. 11, pp. 1755–1759.

12. Dabbene, F., Gay, P., and Polyak, B., Recursive Algorithms for Inner Ellipsoidal Approximation of Convex Polytopes, *Automatica*, 2003, vol. 39, pp. 1773–1781.

13. Calafiore, G. and Campi, M., The Scenario Approach to Robust Control Design, *IEEE Trans. Autom. Control*, 2006, vol. 51, no. 5, pp. 742–753.

14. Tempo, R., Calafiore, G., and Dabbene, F., *Randomized Algorithms for Analysis and Control of Uncertain Systems, with Applications*, London: Springer, 2013.

15. Polyak, B. and Shcherbakov, P., Randomization in Robustness, Estimation, and Optimization, in *Uncertainty in Complex Networked Systems: In Honor of Roberto Tempo*, Basar, T., Ed., Cham: Springer, 2018, pp. 181–208.

16. Kataoka, S., A stochastic Programming Model, *Econometrica: J. Econom. Soc.*, 1963, pp. 181–196.

17. Prékopa, A., Logarithmic Concave Measures with Application to Stochastic Programming, *Acta Scientiarum Mathematicarum*, 1971, vol. 32, pp. 301–316.

18. Geng, X. and Xie, L., Data-Driven Decision Making in Power Systems with Probabilistic Guarantees: Theory and Applications of Chance-Constrained Optimization, *Annu. Rev. Control*, 2019, vol. 47, pp. 341–363.

19. Soudjani, S., and Majumdar, R., Concentration of Measure for Chance-Constrained Optimization, *IFAC-PapersOnLine*, 2018, vol. 51, no. 16, pp. 277–282.

20. Lejeune, M. and Prékopa, A., Relaxations for Probabilistically Constrained Stochastic Programming Problems: Review and Extensions, *Annals of Operations Research*, 2018, pp. 1–22.

21. Campi, M. and Garatti, S., A Sampling-and-Discarding Approach to Chance-Constrained Optimization: Feasibility and Optimality, *J. Optim. Theory Appl.*, 2011, vol. 148, no. 2, pp. 257–280.

22. Alamo, T., Mirasierra, V., Dabbene, F., and Lorenzen, M., Safe Approximations of Chance Constrained Sets by Probabilistic Scaling, *Proc. 18th Eur. Control Conf. (ECC)*, IEEE, 2019, pp. 1380–1385.

23. Alamo, T., Tempo, R., Luque, A., and Ramirez, D., Randomized Methods for Design of Uncertain Systems: Sample Complexity and Sequential Algorithms, *Automatica*, 2015, vol. 52, pp. 160–172.

24. Lorenzen, M., Dabbene, F., Tempo, R., and Allgöwer, F., Stochastic MPC with Offline Uncertainty Sampling, *Automatica*, 2017, vol. 81, pp. 176–183.

25. Mammarella, M., Lorenzen, M., Capello, E., Park, H., Dabbene, F., Guglieri, G., Romano, M., and Allgöwer, F., An Offline-Sampling SMPC Framework with Application to Autonomous Space Maneuvers, *IEEE Trans. Control Syst. Technol.*, 2018, vol. 28, no. 2, pp. 388–402.

26. Calafiore, G., Random Convex Programs, *SIAM J. Optim.*, 2010, vol. 20, no. 6, pp. 3427–3464.

27. Vapnik, V., *The Nature of Statistical Learning Theory*, New York: Springer Science & Business Media, 1999.

28. Alamo, T., Tempo, R., and Camacho, E., Randomized Strategies for Probabilistic Solutions of Uncertain Feasibility and Optimization Problems, *IEEE Trans. Autom. Control*, 2009, vol. 54, no. 11, pp. 2545–2559.

29. Alamo, T., Tempo, R., and Camacho, E., Improved Sample Size Bounds for Probabilistic Robust Control Design: A Pack-Based Strategy, *Proc. 46th IEEE Conf. Decis. Control*, IEEE, 2007, pp. 6178–6183.

30. Alamo, T., Manzano, J., and Camacho, E., Robust Design through Probabilistic Maximization, in *Uncertainty in Complex Networked Systems*, Springer, 2018, pp. 247–274.

31. Nievergelt, J., Exhaustive Search, Combinatorial Optimization and Enumeration: Exploring the Potential of Raw Computing Power, *Proc. Int. Conf. Curr. Trends Theory Pract. Comput. Sci.*, Springer, 2000, pp. 18–35.

32. Lukashevich, A., Gorchakov, V., Vorobev, P., Deka, D., and Maximov, Y., Importance Sampling Approach to Chance-Constrained DC Optimal Power Flow, *IEEE Trans. Control Netw. Syst.*, 2023, vol. 11, no. 2, pp. 928–937.

This paper was recommended for publication by P.S. Shcherbakov, a member of the Editorial Board