Randomized Solutions to Convex Programs with Multiple Chance Constraints*

Georg Schildbach† Lorenzo Fagiano‡ Manfred Morari†

May 16, 2017

Abstract

The scenario-based optimization approach ('scenario approach') provides an intuitive way of approximating the solution to chance-constrained optimization programs, based on finding the optimal solution under a finite number of sampled outcomes of the uncertainty ('scenarios'). A key merit of this approach is that it neither requires explicit knowledge of the uncertainty set, as in robust optimization, nor of its probability distribution, as in stochastic optimization. The scenario approach is also computationally efficient because it only requires the solution to a convex optimization program, even if the original chance-constrained problem is non-convex. Recent research has obtained a rigorous foundation for the scenario approach, by establishing a direct link between the number of scenarios and bounds on the constraint violation probability. These bounds are tight in the general case of an uncertain optimization problem with a single chance constraint.

This paper shows that the bounds can be improved in situations where the chance constraints have a limited 'support rank', meaning that they leave a linear subspace unconstrained. Moreover, it shows that also a combination of multiple chance constraints, each with individual probability level, is admissible. As a consequence of these results, the number of scenarios can be reduced from that prescribed by the existing theory for problems with the indicated structural property. This leads to an improvement in the objective value and a reduction in the computational complexity of the scenario approach. The proposed extensions have many practical applications, in particular high-dimensional problems such as multi-stage uncertain decision problems or design problems of large-scale systems.

Key words: Uncertain Optimization, Chance Constraints, Randomized Methods, Convex Optimization, Scenario Approach, Multi-Stage Decision Problems.

1 Introduction

Optimization is ubiquitous in modern problems found in engineering, logistics, and other sciences. A common pattern is that a decision or design variable $x \in \mathbb{R}^d$ has to be selected from a subset of $\mathbb{R}^d$, as described by constraints $f_i : \mathbb{R}^d \to \mathbb{R}$, and its quality is measured against some objective or cost function $f_0 : \mathbb{R}^d \to \mathbb{R}$:

$$\min_{x\in\mathbb{R}^d} \quad f_0(x) \ , \tag{1.1a}$$

$$\text{s.t.} \quad f_i(x) \leq 0 \qquad \forall\ i = 1, 2, \ldots, N \ . \tag{1.1b}$$

†Automatic Control Laboratory, Swiss Federal Institute of Technology, Zurich, Switzerland (schildbach|morari@control.ee.ethz.ch).

‡Dipartimento di Automatica e Informatica, Politecnico di Torino, Torino, Italy and Department of Mechanical Engineering, University of California at Santa Barbara, Santa Barbara (CA), United States (lorenzo.fagiano@polito.it).

1.1 Chance-Constrained Optimization

Unfortunately, in many practical applications the underlying problem data is uncertain. This uncertainty shall be represented with an abstract variable $\delta \in \Delta$, where Δ is an uncertainty set whose nature is not specified. The uncertainty may affect the objective function f_0 and/or the constraints f_i. Thus for a particular decision x it becomes uncertain what objective value is achieved and/or whether the constraints are indeed satisfied. The second situation represents a particular challenge, as good solutions are usually located on the boundary of the feasible set.

This gives rise to a trade-off problem between the (uncertain) objective value and the robustness of the chosen decision to a constraint violation. A large variety of approaches addressing this issue have been proposed in the areas of robust and stochastic optimization [3–5,14,15,17,19,21], with the preferred method of choice depending on the requirements of the application at hand.

In many practical applications, δ can be assumed to be of a stochastic nature. In this case, the formulation of *chance constraints*, where the decision variable x has to be feasible with a least probability $(1-\varepsilon)$ for $\varepsilon \in (0,1)$, has proven to be an appropriate concept for handling the uncertainty in the constraints. However, chance-constrained optimization problems are usually very difficult to solve. The *scenario approach*, as explained below, represents an attractive method for finding an 'approximate solution' to stochastic programs, since it is both intuitive and computationally efficient.

1.2 The Scenario Approach

Recent contributions [8–12] have revealed the theoretical links between the scenario approach and the solution to an optimization problem with a linear objective function and a single chance constraint (SCP):

$$\min_{x \in \mathbb{X}} \quad c^{\mathrm{T}} x \ , \tag{1.2a}$$

$$\text{s.t.} \quad \Pr\big[f(x,\delta) \leq 0\big] \geq (1-\varepsilon) \ . \tag{1.2b}$$

Here $\mathbb{X} \subset \mathbb{R}^d$ is a compact and convex set, c^{T} denotes the transpose of a vector $c \in \mathbb{R}^d$, $\Pr[\cdot]$ is the probability measure on the uncertainty set Δ, $f : \mathbb{R}^d \times \Delta \to \mathbb{R}$ is a convex function in its first argument $x \in \mathbb{R}^d$ for Pr-almost every uncertainty $\delta \in \Delta$, and ε is some value in the open real interval $(0,1)$.

The chance constraint (1.2b) is interpreted as follows. For any given $x \in \mathbb{R}^d$, the left-hand side represents the probability of the event that x indeed belongs to the feasible set. Written more properly,

$$\Pr\big[f(x,\delta) \leq 0\big] := \Pr\big\{\delta \in \Delta \,\big|\, f(x,\delta) \leq 0\big\} \ , \tag{1.3}$$

however the left-hand side notation is kept throughout for brevity. Note that x is considered to be a *feasible point* of the chance constraint (1.2b) if this probability is at least $(1-\varepsilon)$.

Remark 1.1 (Problem Formulation) *The formulation of the* SCP *encompasses a vast range of problems, namely any uncertain optimization problem that becomes convex if the value of δ were fixed. (a) Any uncertain convex objective function $f(\cdot,\delta)$ can be included by an epigraph reformulation, with the new objective being a scalar and hence linear [7, Sec. 3.1.7]. (b) Joint chance constraints, where x must satisfy multiple convex constraints simultaneously with probability $(1-\varepsilon)$, are covered since the intersection of convex sets is convex. (c) Additional deterministic, convex constraints can be included by intersection with the compact set $\mathbb{X}$.*

The characterization of the feasible set of a chance constraint requires exact knowledge of the probability distribution of δ. Moreover, the feasible set is non-convex and difficult to express explicitly, except for very special cases [5,14,19,21]. This makes the SCP, in full generality and especially in higher dimensions d, an extremely difficult problem to solve.

The scenario approach can be used to find an *approximate solution* to the SCP, which is considered to be any point in $\mathbb{X}$ that is feasible for the chance constraint with some given (very high) *confidence*

$(1-\theta) \in (0,1)$. This problem is usually not as hard, if an approximate solution is chosen in a low-violation region of the decision space (with high confidence). However, then the resulting objective value may be poor, in which case the approximate solution shall be called '*conservative*'. Clearly, it is of major interest to find approximate solutions that are the least conservative (i.e. with an objective value as low as possible), and this is the goal of the scenario approach.

The basic idea of the scenario approach is to draw a specific number $K \in \mathbb{N}$ of samples ('*scenarios*') from the uncertainty δ, and to take the optimal solution that is feasible under all of these scenarios ('*scenario solution*') as an approximate solution. Computing the scenario solution involves a deterministic optimization program ('*scenario program*'), which is obtained by replacing the chance constraint (1.2b) with the K sampled deterministic constraints.

By construction, the scenario program is a deterministic, convex optimization program that can be solved efficiently by standard algorithms [7, 16, 18]. Moreover, the scenario approach is distribution-free in the sense that it does not rely on a particular mathematical model for the distribution of δ, or even its support set Δ. In fact, both may be unknown; the only requirements are stated in the following assumption.

Assumption 1.2 (Uncertainty) *(a) The uncertainty δ is a random variable with (possibly unknown) probability measure* $\Pr$ *and support set Δ. (b) A sufficient number of independent random samples from δ can be obtained.*

Note that Assumption 1.2 is fairly general. It could even be argued that the scenario approach is at the heart of any robust and stochastic optimization method, because either the uncertainty set Δ or the probability distribution of δ are usually constructed based on some (necessarily finite) experience of the uncertainty.

Tight bounds for the proper choice of the sample size K are established by [9, 11], when linking it directly to the probability with which the scenario solution violates the chance constraint (1.2b). Moreover, [9, 12] show that the theory can be extended to the case where $R \leq K$ sampled constraints are discarded *a posteriori*, that is after observing the outcomes of the K samples. While this increases the complexity of the scenario approach (in terms of data requirement and computation), it can be used to improve the objective value achieved by the scenario solution. In fact, the scenario solution can be shown to converge to the exact solution of (1.2) when the number of discarded constraints are increased, given that some mild technical assumptions hold, cf. [12, Sec. 4.4]

1.3 Novel Contributions

From a practical point of view, the strongest appeal of the scenario approach is the facility of its application and the low computational complexity. It becomes particularly attractive for uncertain optimization problems in higher dimensions, as these occur frequently in fields such as engineering or logistics. In these cases, an uncertain constraint will often not involve all decision variables simultaneously, as allowed by the general case of (1.2b). Instead, multiple uncertain constraints may be present, each of them involving only a subset of the decision variables.

Example 1.3 (Multi-Stage Decision Problems) An important example are uncertain *multi-stage decision problems* [5, Cha. 7], [14, Cha. 8] [19, Cha. 13] [21, Cha. 3], which occur in many fields such as production planning, portfolio optimization, or control theory. The basic setting is that some *decision* (e.g. on production quantities, buy/sell orders, or control inputs) has to be taken repeatedly at a finite number of time steps. Each decision affects the *state* of the system (e.g. inventory level, portfolio, or state variable) at the subsequent time step. Besides the decision, the state is also subject to uncertain influences (e.g. customer demand, price fluctuations, or dynamic disturbances). If constraints on the state variables are present (e.g. service levels, value at risk, or safety regions), this adds multiple uncertain constraints (one for the state of each time step) to the overall decision problem. Further deterministic constraints may hold for the decision variables, for example. The special structure of such a problem is that a constraint on the state at some time step involves only the decisions made prior to this time step, while the decisions afterwards are not involved.

This paper extends the theory of the scenario approach for problems where a single (or multiple) chance constraint(s) are present that involve only a subset of the decision variables. More precisely, the chance constraint(s) may affect only a certain subspace of the decision space, whose dimension will be called its '*support rank*'. Other constraints, either deterministic or uncertain, cover the directions that are left unconstrained, so that the solution remains bounded.

The main result of this paper is that an uncertain constraint with a lower support rank can only supply a lower number of *support constraints* [9–11], and therefore its associated sample size can be reduced. This leads to a subtle shift from the idea of a '*problem dimension*' in the existing theory to that of a '*support dimension*' of a particular chance constraint. Moreover, it requires an extension of the existing theory to cope with multiple chance constraints in the uncertain optimization program. Finally, the approach of constraint removal *a posteriori* is carried over almost analogously to this extended setting.

From a practical point of view, these extensions improve on the merits of the scenario approach for problems that have a structure described above. In particular, the lower sample sizes reduce the computational complexity of the scenario approach and simultaneously improve the objective value of the scenario solution. At the same time, the feasibility guarantees for the scenario solution remain as strong as before. Hence the extensions of this paper, when applicable, offer only advantages over the existing results on the scenario approach.

1.4 Organization of the Paper

Section 2 contains the problem statement. Section 3 introduces some background on its properties, and states the rigorous definitions for the '*support dimension*' and the '*support rank*' of a chance constraint. Section 4 contains the main results of this paper, which give the improved sample bounds in the presence of a single (or multiple) chance constraint(s) of limited support rank. Section 5 extends this theory to the sampling-and-discarding procedure, which can be used to improve the objective value of the scenario solution, at the price of larger data requirements and an increased computational complexity. Section 6 presents a brief numerical example that demonstrates the application of the presented theory, as well as its potential benefits when compared to existing results.

2 Problem Formulation

This section introduces the generalized problem formulation with multiple chance constraints, the corresponding scenario program, and some basic terminology.

2.1 Stochastic Program with Multiple Chance Constraints

Consider the following extension of the SCP to an optimization problem with linear objective function and multiple chance constraints (MCP):

$$\min_{x\in\mathbb{X}} \quad c^{\mathrm{T}}x \ , \tag{2.1a}$$

$$\text{s.t.} \quad \Pr\bigl[f_i(x,\delta)\leq 0\bigr] \geq (1-\varepsilon_i) \qquad \forall\ i\in\mathbb{N}_1^N \ , \tag{2.1b}$$

where i is the chance constraint index in $\mathbb{N}_1^N := \{1,2,...,N\}$. The remarks for the SCP in Section 1.2 apply analogously; in particular the following key assumption is made.

Assumption 2.1 (Convexity) *The constraint functions $f_i : \mathbb{R}^d \times \Delta \to \mathbb{R}$ of all chance constraints $i \in \mathbb{N}_1^N := \{1,...,N\}$ are convex in their first argument $x \in \mathbb{R}^d$ for $\Pr$-almost every $\delta \in \Delta$.*

Other than Assumption 2.1, the dependence of the functions $f_i(x, \delta)$ on the uncertainty δ is completely generic.

The use of '$\min$' instead of '$\inf$' in (2.1a) is justified by the fact that the feasible set of a single chance constraint is closed under fairly general assumptions [14, Thm. 2.1]. This implies that the feasible set of the MCP is compact, due to the presence of $\mathbb{X}$, and the infimum is indeed attained.

It remains a standing assumption that the σ-algebra of $\Pr$-measurable sets in Δ is large enough to contain all sets whose probability is measured in this paper, like the ones in (2.1b), cf. [11, p. 4].

In order to avoid technical issues, which are of little relevance for most practical applications, the following is assumed, cf. [11, Ass. 1].

Assumption 2.2 (Existence and Uniqueness) *(a) Problem* (2.1) *admits at least one feasible point. By the compactness of* $\mathbb{X}$*, this implies that there exists at least one optimal point of* (2.1)*. (b) If there are multiple optimal points of* (2.1)*, a unique one is selected by the help of a* tie-break rule *(e.g. the lexicographic order on* $\mathbb{R}^d$*).*

In principle, an approximate solution to the MCP can be obtained by the classic scenario approach. Namely, a SCP can be setup with the same objective function (1.2a) as the MCP, and a chance constraint (1.2b) defined by

$$f(x, \delta) := \max\bigl\{f_1(x, \delta), \ldots, f_N(x, \delta)\bigr\} \qquad \text{and} \qquad \varepsilon := \min\bigl\{\varepsilon_1, \varepsilon_2, ..., \varepsilon_N\bigr\} \ . \tag{2.2}$$

Note that $f(x, \delta)$ is convex in x for almost every δ, since the pointwise maximum of convex functions is convex. Any feasible point of this SCP is also a feasible point of the MCP, and hence an approximate solution to the SCP with confidence $(1 - \theta)$ is also an approximate solution to the MCP with confidence $(1 - \theta)$.

However, this procedure introduces a considerable amount of conservatism, because it requires the scenario solution to simultaneously satisfy *all* constraints $i = 1, ..., N$ with the *highest* of all probabilities $(1 - \varepsilon_i)$. Clearly, this conservatism becomes more severe if the number of chance constraints N is large and there is a great variation in the values of ε_i.

2.2 The Extended Scenario Approach

The extended scenario approach of this paper can be used to compute an approximate solution of the MCP, which is a feasible point of every chance constraint $i = 1, ..., N$ with a given confidence probability of $(1 - \theta_i)$. The key difference from the classic scenario approach is that each chance constraint $i \in \mathbb{N}_1^N$ is sampled separately, and with an individual sample size $K_i \in \mathbb{N}$.

Let the *random samples* pertaining to constraint i be denoted $\delta^{(i,\kappa_i)}$, where $\kappa_i \in \{1, ..., K_i\}$, and for brevity also as the collective *multi-sample* $\omega^{(i)} := \{\delta^{(i,1)}, ..., \delta^{(i,K_i)}\}$. The collection of all samples is combined in an overall multi-sample $\omega := \{\omega^{(1)}, ..., \omega^{(N)}\}$, with the total number of samples given by $K := \sum_{i=1}^N K_i$. All of these samples can be considered 'identical copies' of the random uncertainty δ, in the sense that they are themselves random variables and satisfy the following key assumption.

Assumption 2.3 (Independence and Identical Distribution) *The sampling procedure is designed such that the set of all random samples, together with the actual random uncertainty,*

$$\bigcup_{i\in\mathbb{N}_1^N}\big\{\delta^{(i,1)},...,\delta^{(i,K_i)}\big\}\cup\big\{\delta\big\}$$

form a set of independent and identically distributed (i.i.d.) *random variables.*

The multi-sample ω is an element of Δ^K, the K-th product of the uncertainty set Δ, and it is distributed according to Pr^K, the K-th product of the measure Pr. The scenario program for multiple chance constraints ($\mathrm{MSP}[\omega^{(1)},...,\omega^{(N)}]$) is constructed as follows:

$$\min_{x\in\mathbb{X}} \quad c^{\mathrm{T}}x \ , \tag{2.3a}$$

$$\text{s.t.} \quad f_i\big(x,\delta^{(i,\kappa_i)}\big)\leq 0 \qquad \forall\ \kappa_i\in\mathbb{N}_1^{K_i},\ \ \forall\ i\in\mathbb{N}_1^N\ \ . \tag{2.3b}$$

In problem (2.3), the objective function of the MCP is minimized, while forcing x to lie inside the constrained sets for all samples $\delta^{(i,\kappa_i)}$ substituted into the corresponding constraint $i\in\mathbb{N}_1^N$. Clearly, the solution to problem (2.3) is itself a random variable, as it depends on the random multi-sample ω. For this reason, the scenario approach is a *randomized method* for finding an approximate solution to the MCP.

Of course, the MSP is actually solved for the observations of the random samples, leading to its deterministic instance ($\overline{\mathrm{MSP}}[\bar{\omega}^{(1)},...,\bar{\omega}^{(N)}]$):

$$\min_{x\in\mathbb{X}} \quad c^{\mathrm{T}}x \ , \tag{2.4a}$$

$$\text{s.t.} \quad f_i\big(x,\bar{\delta}^{(i,\kappa_i)}\big)\leq 0 \qquad \forall\ \kappa_i\in\mathbb{N}_1^{K_i},\ \ \forall\ i\in\mathbb{N}_1^N\ \ . \tag{2.4b}$$

Note that (2.4) arises from (2.3) by replacing the *(random) samples* $\delta^{(i,\kappa_i)}$, $\omega^{(i)}$, ω with their *(deterministic) outcomes* $\bar{\delta}^{(i,\kappa_i)}$, $\bar{\omega}^{(i)}$, $\bar{\omega}$. Throughout the paper, these outcomes are indicated by a bar, to distinguish them from the corresponding random variables. By Assumption (2.1), $\overline{\mathrm{MSP}}$ constitutes a convex program that can be solved efficiently by a suitable algorithm for convex optimization, cf. [7, 16, 18].

Note that (2.3) remains important for analyzing the (probabilistic) properties of the (random) scenario solution. In fact, the subsequent theory is mainly concerned with showing that, with a very high confidence, the scenario solution is a feasible point of the chance constraints (2.1b), provided that the sample sizes $K_1,...,K_N$ are appropriately selected.

2.3 Randomized Solution and Violation Probability

In order to avoid unnecessary complications, the following technical assumption ensures that there always exists a feasible solution to the MSP, cf. [11, p. 3].

Assumption 2.4 (Feasibility) *(a) For any number of samples $K_1,...,K_N$, the* MSP *admits a feasible solution almost surely. (b) For the sake of notational simplicity, any* Pr*-null set for which (a) may not hold is assumed to be removed from* Δ.

Assumption 2.4 can be taken for granted in the majority of practical problems. When it does not hold in a particular case, a generalization of the presented theory accounting for the infeasible case can be developed along the lines of [9].

Hence the existence of a solution to $\overline{\mathrm{MSP}}$ is ensured, and uniqueness holds by Assumption 2.1 and by carry-over of the tie-break rule of Assumption 2.2(b), see [20, Thm. 10.1, 7.1]. Therefore the *solution map*

$$\bar{x}^\star:\Delta^K\to\mathbb{X} \tag{2.5}$$

is well-defined, returning the unique optimal point $\bar{x}^\star(\bar{\omega}^{(1)},...,\bar{\omega}^{(N)})$ of the $\overline{\mathrm{MSP}}$ for a given outcome of the multi-samples $\{\bar{\omega}^{(1)},...,\bar{\omega}^{(N)}\}\in\Delta^K$. The solution map can also be applied to the MSP, for which it is denoted by $x^\star : \Delta^K \to \mathbb{X}$. Now $x^\star(\omega^{(1)}, ..., \omega^{(N)})$ represents a random vector of unknown probability distribution, which is also referred to as the *scenario solution*. In fact, its distribution is a complicated function of the geometry and the parameters of the problem.

Note that there are two levels of randomness present in the analysis. The first is introduced by the random samples in ω, which affect the choice of the scenario solution. The second is the actual random uncertainty δ, which determines whether or not the scenario solution is feasible with respect to the chance constraints (2.3b). For this reason, the scenario approach presented here is also called a *double-level-of-probability approach* [8, Rem. 2.3].

To highlight the two probability levels more clearly, suppose first that the multi-sample $\bar{\omega}$ has already been observed, so that the scenario solution $\bar{x}^\star(\bar{\omega}^{(1)}, ..., \bar{\omega}^{(N)})$ is fixed. Then for each chance constraint $i = 1, ..., N$ in (2.1b), the *a posteriori violation probability* $\bar{V}_i(\bar{\omega}^{(1)}, ..., \bar{\omega}^{(N)})$ is given by

$$\bar{V}_i\big(\bar{\omega}^{(1)}, ..., \bar{\omega}^{(N)}\big) := \Pr\big[f_i\big(\bar{x}^\star(\bar{\omega}^{(1)}, ..., \bar{\omega}^{(N)}), \delta\big) > 0\big] \ . \tag{2.6}$$

In particular, each $\bar{V}_i$ has a deterministic, yet generally unknown, value in $[0, 1]$. If the multi-sample ω has not yet been observed, the scenario solution $x^\star(\omega^{(1)}, ..., \omega^{(N)})$ is a random vector and so the *a priori violation probability*

$$V_i\big(\omega^{(1)}, ..., \omega^{(N)}\big) := \Pr\big[f_i\big(x^\star(\omega^{(1)}, ..., \omega^{(N)}), \delta\big) > 0\big] \tag{2.7}$$

becomes itself a random variable on $(\Delta^K, \Pr^K)$, with support $[0, 1]$. Hence the goal is to choose appropriate sample sizes $K_1, ..., K_N$ which ensure that $V_i(\omega^{(1)}, ..., \omega^{(N)}) \leq \varepsilon_i$ for all $i = 1, ..., N$, with a sufficiently high confidence $(1 - \theta_i)$. Before these results are derived however, some structural properties of scenario programs and technical lemmas ought to be discussed.

3 Structural Properties of the Constraints

In this section, a structural property of a chance constraint is introduced which yields a reduction in the number of samples below the levels given by the existing theory [9–11]. This property relates to the new concept of the *support dimension* or, in a form that is more easily checked for many practical instances, the *support rank*.

3.1 Support Constraints

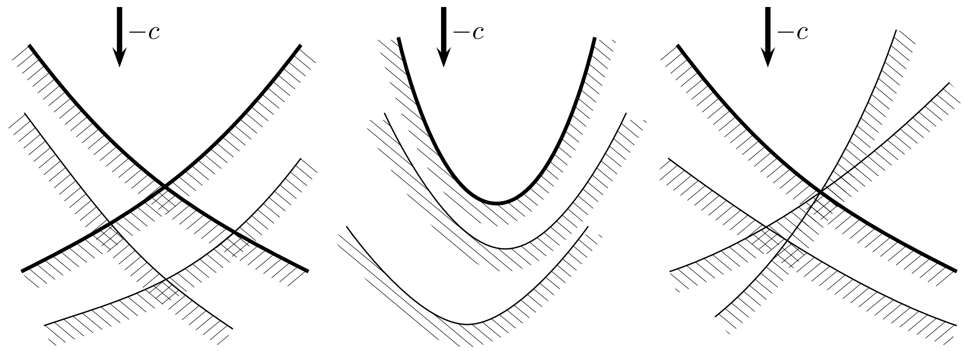
The concept of a *support constraint* carries over from the SCP case, cf. [10, Def. 4]. An illustration is given in Figure 3.1.

Definition 3.1 (Support Constraint) *Consider the $\overline{\mathrm{MSP}}$ for some outcome of the multi-sample $\bar{\omega}$. (a) For some $i \in \mathbb{N}_1^N$ and $\kappa_i \in \mathbb{N}_1^{K_i}$, constraint $f_i(x, \bar{\delta}^{(i,\kappa_i)}) \leq 0$ is a* support constraint *of* (2.4) *if its removal from the problem entails a change in the optimal solution:*

$$\bar{x}^\star\big(\bar{\omega}^{(1)}, ..., \bar{\omega}^{(N)}\big) \neq \bar{x}^\star\big(\bar{\omega}^{(1)}, ..., \bar{\omega}^{(i-1)}, \bar{\omega}^{(i)} \setminus \{\bar{\delta}^{(i,\kappa_i)}\}, \bar{\omega}^{(i+1)}, ..., \bar{\omega}^{(N)}\big) \ .$$

In this case the sample $\bar{\delta}^{(i,\kappa_i)}$ is also said 'to generate *this support constraint.' (b) For each $i \in \mathbb{N}_1^N$, the indices κ_i of all samples that generate a support constraint of the $\overline{\mathrm{MSP}}$ are included in the set $\overline{\mathrm{Sc}}_i$. Moreover, the tuples (i, κ_i) of all support constraints of the $\overline{\mathrm{MSP}}$ are collected in the* support (constraint) set $\overline{\mathrm{Sc}}$. *With some abuse of this notation, $\overline{\mathrm{Sc}} = \bigcup_{i=1}^N \overline{\mathrm{Sc}}_i$.*

Definition 3.1(a) can be stated equivalently in terms of the objective function: a sampled constraint is a support constraint if and only if the optimal objective function value (or its preference by the tie-break rule) is strictly larger than when the constraint were removed. To be more precise, Definition 3.1(b), $\overline{\mathrm{Sc}}$ may also



(a) Two Support Constraints. (b) One Support Constraint. (c) One Support Constraint.

Figure 3.1: Illustration of Definition 3.1 in $\mathbb{R}^2$. The arrow indicates the optimization direction, the bold lines are the *support constraints* of the respective configuration.

account for the set $\mathbb{X}$ as an additional support constraint. This minor subtlety is tacitly understood in the sequel.

In the stochastic setting of the $\mathrm{MSP}[\omega^{(1)},...,\omega^{(N)}]$, whether or not a particular random sample $\delta^{(i,\kappa_i)}$ generates a support constraint becomes a random event, which can be associated with a certain probability. Similarly, the support constraint set Sc, and its subsets $\mathrm{Sc}_1,...,\mathrm{Sc}_N$ contributed by the various chance constraints, are naturally random sets.

3.2 Support Dimension

The link between the sample sizes $K_1,...,K_N$ and the corresponding violation probability of the scenario solution depends decisively on the 'dimensions' of the problem. The following lower bounds represent a mild technical condition, cf. [9, Thm. 3.3] and [11, Def. 2.3].

Assumption 3.2 *The sample sizes satisfy $K_1,...,K_N \geq d$.*

In the existing literature, the dimension of the SCP has been characterized by *Helly's dimension*, cf. [9, Def. 3.1]. In this paper, there is a subtle shift from the problem dimension to the dimension of chance constraint i in the MCP, embodied by its *support dimension*.

Definition 3.3 (Support Dimension) *(a) Denote by $|\,\mathrm{Sc}\,|$ the (random) cardinality of the set* Sc*.* Helly's dimension *is the smallest integer ζ that satisfies*

$$\operatorname*{ess\,sup}_{\omega\in\Delta^K} |\,\mathrm{Sc}\,| \leq \zeta \ .$$

(b) The support dimension *of a chance constraint $i \in \mathbb{N}_1^N$ in the* MSP *is the smallest integer ζ_i that satisfies*

$$\operatorname*{ess\,sup}_{\omega\in\Delta^K} |\,\mathrm{Sc}_i\,| \leq \zeta_i \ .$$

From a basic argument using Helly's Theorem, the number of support constraints $|\,\mathrm{Sc}\,|$ of any (feasible) convex optimization problem in $\mathbb{R}^d$ is upper bounded by the dimension of the decision space d, cf. [10, Thm. 2]. This result implies that finite integers ζ and $\zeta_1,...,\zeta_N$ matching Definition 3.3 always exist, so that the concepts of 'Helly's dimension' and 'support dimension' are indeed well-defined. Moreover, the result provides immediate upper bounds on the support dimension of each chance constraint $i \in \mathbb{N}_1^N$ in (2.3), namely $\zeta_i \leq \zeta \leq d$.

It turns out that the support dimension ζ_i directly relates to the minimum sample size K_i that is required for a given violation level ε_i and residual probability θ_i. The basic mechanism shall be illustrated by the proposition below, for the simpler case of a *single-level of probability* problem, cf. [10, Thm. 1].

Proposition 3.4 (Probability Bound) *Consider a particular constraint* $i \in \mathbb{N}_1^N$ *in the* $\mathrm{MSP}[\omega^{(1)}, ..., \omega^{(N)}]$ *with some fixed sample size* K_i*, and let* $\hat{\zeta}_i$ *be an upper bound for its support dimension* ζ_i*. Then the following holds:*

$$\mathrm{Pr}^{K+1}\bigl[f_i\bigl(x^\star(\omega^{(1)}, ..., \omega^{(N)}), \delta\bigr) > 0\bigr] \leq \tfrac{\hat{\zeta}_i}{K_i+1} \quad . \tag{3.1}$$

Proof. Consider $\mathrm{MSP}' := \mathrm{MSP}[\omega^{(1)}, ..., \omega^{(i-1)}, \omega^{(i)} \cup \{\delta\}, \omega^{(i+1)}, ..., \omega^{(N)}]$ and let $\mathrm{Sc}_i' \subset \{1, ..., K_i, K_i+1\}$ denote the set of support constraints generated by samples from $\omega^{(i)} \cup \{\delta\}$, where $(K_i+1) \in \mathrm{Sc}_i'$ stands for δ generating a support constraint. Note that the event where $f_i\bigl(x^\star(\omega^{(1)}, ..., \omega^{(N)}), \delta\bigr) > 0$ can be equivalently expressed as δ generating a support constraint of MSP'. Hence condition (3.1) can be reformulated as

$$\mathrm{Pr}^{K+1}\bigl[(K_i+1) \in \mathrm{Sc}_i'\bigr] \leq \tfrac{\hat{\zeta}_i}{K_i+1} \quad . \tag{3.2}$$

To analyze the event $(K_i+1) \in \mathrm{Sc}_i'$, observe that by Assumption 2.3 all samples in $\omega^{(i)} \cup \{\delta\}$ are i.i.d., whence all sampled instances of constraint i in (2.3b) along with '$f_i(\,\cdot\,, \delta) \leq 0$' are probabilistically identical. In particular, they are all equally likely to become a support constraint of MSP'. Hence if the number of support constraints $|\,\mathrm{Sc}_i'\,|$ were known, then

$$\mathrm{Pr}^{K+1}\bigl[(K_i+1) \in \mathrm{Sc}_i'\bigr] = \tfrac{|\,\mathrm{Sc}_i'\,|}{K_i+1} \quad .$$

Even though $|\,\mathrm{Sc}_i'\,|$ is a random variable, by Definition 3.3(b) $|\,\mathrm{Sc}_i'\,| \leq \zeta_i$ almost surely, and by assumption $\zeta_i \leq \hat{\zeta}_i$. This immediately yields (3.1). □

3.3 The Support Rank

In many practical cases, the support dimension ζ_i of a chance constraint $i \in \mathbb{N}_1^N$ in the MSP is not known exactly. Then it has to be replaced by some upper bound. As argued above, the existing upper bound is given by the dimension d of the decision space. However, this bound may not be tight in the case where the constraints satisfy a certain structural property, namely when they have a limited *support rank*.

Intuitively speaking, the support rank is the dimension d of the decision space less the maximal dimension of an (almost surely) *unconstrained subspace*. The latter is understood as a linear subspace of $\mathbb{R}^d$ that cannot be constrained by the sampled instances of constraint i, for almost every value of the multi-sample $\omega^{(i)}$.

Before the support rank is introduced in a rigorous manner, three examples of constraint classes with bounded support rank are described, in order to equip the reader with the necessary intuition behind this concept. They also show that very common constraint classes possess this property, and that in practical problems it can often be spotted easily.

Example 3.5 For each of the following cases, a visual illustration can be found in Figure 3.2.

(a) Single Linear Constraint. Suppose some chance constraint $i \in \mathbb{N}_1^N$ of (2.1b) takes the linear form

$$f_i(x, \delta) \equiv a^{\mathrm{T}} x - b(\delta) \quad , \tag{3.3}$$

where $a \in \mathbb{R}^d$, and $b : \Delta \to \mathbb{R}$ is a scalar depending on the uncertainty in a generic way. Note that these constraints in the MSP are unable to constrain any direction in the subspace orthogonal to the span of a, $\mathrm{span}\{a\}^\perp$, regardless of the outcome of the multi-sample $\omega^{(i)}$. Hence the support rank α of the chance constraint (3.3) is equal to 1.

(b) Multiple Linear Constraints. As a generalization of case (a), suppose that some chance constraint $i \in \mathbb{N}_1^N$ of (2.1b) is given by

$$f_i(x,\delta) \equiv A(\delta)x - b(\delta) \ , \tag{3.4}$$

where $A : \Delta \to \mathbb{R}^{r\times d}$ and $b : \Delta \to \mathbb{R}^r$ represent a matrix and a vector that depend on the uncertainty δ. Moreover, suppose that the uncertainty enters the matrix $A(\delta)$ in such a way that the dimension of the linear span of its rows $A_{j,\cdot}(\delta)$, for $j = 1, ..., r$, satisfies

$$\dim \operatorname{span}\bigl\{A_{j,\cdot}(\delta) \bigm| j \in \mathbb{N}_1^r,\ \delta \in \Delta\bigr\} \leq \beta < d \ .$$

Note that these constraints in the MSP are unable to constrain any direction in $\operatorname{span}\bigl\{A_{j,\cdot}(\delta) \bigm| j \in \mathbb{N}_1^r,\ \delta \in \Delta\bigr\}^{\perp}$, regardless of the outcome of the multi-sample $\omega^{(i)}$. Hence the support rank of the chance constraint (3.4) is equal to β.

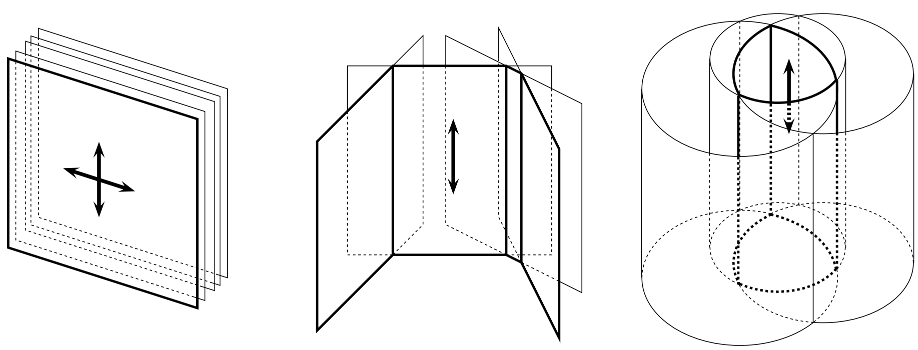
(c) Quadratic Constraint. For a nonlinear example, consider the case where some chance constraint $i \in \mathbb{N}_1^N$ of (2.1b) is given by

$$f_i(x,\delta) \equiv \bigl(x - x_c(\delta)\bigr)^{\mathrm{T}} Q \bigl(x - x_c(\delta)\bigr) - r(\delta) \ , \tag{3.5}$$

where $Q \in \mathbb{R}^{d\times d}$ is positive semi-definite with $\operatorname{rank} Q = \gamma < d$, and $x_c : \Delta \to \mathbb{R}^d$, $r : \Delta \to \mathbb{R}_+$ represent a vector and scalar that depend on the uncertainty. Note that these constraints in the MSP are unable to constrain any direction in the null space of the matrix Q, regardless of the outcome of the multi-sample $\omega^{(i)}$. Since this null space has dimension $d - \gamma$, the support rank of the chance constraint (3.5) is equal to γ.

To introduce the support rank in a rigorous manner, pick a chance constraint $i \in \mathbb{N}_1^N$ of the MCP. For each point $x \in \mathbb{X}$ and each uncertainty $\delta \in \Delta$, denote the corresponding level set of $f_i : \mathbb{R}^d \times \Delta \to \mathbb{R}$ by

$$F_i(x,\delta) := \bigl\{\xi \in \mathbb{R}^d \bigm| f_i(x+\xi,\delta) = f_i(x,\delta)\bigr\} \ . \tag{3.6}$$



(a) Single Linear Constraint. (b) Multiple Linear Constraints. (c) Quadratic Constraint.

Figure 3.2: Illustration of Example 3.5 in $\mathbb{R}^3$. The arrows indicate the dimension of the *unconstrained subspace*, equal to 3 minus the respective *support rank* α, β, or γ.

Let $\mathcal{L}$ be the collection of all linear subspaces in $\mathbb{R}^d$. In order to be unconstrained, select only those subspaces that are contained in almost all level sets $F_i(x,\delta)$:

$$\mathcal{L}_i := \bigcap_{\delta\in\Delta} \bigcap_{x\in\mathbb{R}^d} \bigl\{L \in \mathcal{L} \bigm| L \subset F_i(x,\delta)\bigr\} \ . \tag{3.7}$$

Introduce '$\preceq$' as the partial order on $\mathcal{L}_i$ defined by set inclusion; i.e. for any two subspaces $L, L' \in \mathcal{L}_i$, $L \preceq L'$ if and only if $L \subseteq L'$. Then the following concepts are well-defined, as shown in Proposition 3.7 below.

Definition 3.6 (Unconstrained Subspace, Support Rank) *(a) The* unconstrained subspace L_i *of chance constraint* $i \in \mathbb{N}_1^N$ *is the unique maximal element in* $\mathcal{L}_i$*, in the sense that* $L \preceq L_i$ *for all* $L \in \mathcal{L}_i$*. (b) The* support rank $\rho_i \in \mathbb{N}_0^d$ *of chance constraint* $i \in \mathbb{N}_1^N$ *equals to* d *minus the dimension of* L_i*,*

$$\rho_i := d - \dim L_i \quad .$$

It is a minor technicality in Definition 3.6 that any $\Pr$-null set that adversely influences the dimension of the unconstrained subspace can be removed from Δ; this is tacitly understood.

Observe that if $\mathcal{L}_i$ contains only the trivial subspace, then the support rank is actually equal to Helly's dimension d. On the other hand, if $\mathcal{L}_i$ contains more than the trivial subspace, then the support rank becomes strictly less than d.

Proposition 3.7 (Well-Definedness of Unconstrained Subspace) *The collection* $\mathcal{L}_i$ *contains a unique maximal element* L_i *in the set-inclusion sense, i.e.* L_i *contains all other elements of* $\mathcal{L}_i$ *as subsets.*

Proof. First, note that $\mathcal{L}_i$ is always non-empty, because for every $x \in \mathbb{X}$ and every $\delta \in \Delta$ the level set $F_i(x, \delta)$ includes the origin by its definition in (3.6). Therefore $\mathcal{L}_i$ contains (at least) the trivial subspace $\{0\}$.

Second, since every chain in $\mathcal{L}_i$ has an upper bound (namely $\mathbb{R}^d$), *Zorn's Lemma* (or the *Axiom of Choice*, cf. [6, p. 50]) implies that $\mathcal{L}_i$ has at least one maximal element in the '$\preceq$'-sense.

Third, in order to prove that the maximal element is unique, suppose that $L_i^{(1)}, L_i^{(2)}$ are two maximal elements of $\mathcal{L}_i$. It will be shown that their direct sum $L_i^{(1)} \oplus L_i^{(2)} \in \mathcal{L}_i$, so that $L_i^{(1)} \neq L_i^{(2)}$ would contradict their maximality. According to (3.7), it must be shown that $L_i^{(1)} \oplus L_i^{(2)} \subset F_i(x, \delta)$ for any fixed values $x \in \mathbb{X}$ and $\delta \in \Delta$. To see this, pick

$$\xi \in L_i^{(1)} \oplus L_i^{(2)} \quad \Longrightarrow \quad \xi = \xi^{(1)} + \xi^{(2)} \quad \text{for } \xi^{(1)} \in L_i^{(1)},\ \xi^{(2)} \in L_i^{(2)} \quad .$$

Then apply (3.6) twice to obtain

$$f_i(x + \xi^{(1)} + \xi^{(2)}, \delta) = f_i(x + \xi^{(1)}, \delta) = f_i(x, \delta) \quad ,$$

because $\xi^{(2)} \in L_i^{(2)}$ and $\xi^{(1)} \in L_i^{(1)}$. $\square$

3.4 The Support Rank Lemma

The following lemma provides the link between the support rank of a chance constraint and its support dimension.

Lemma 3.8 (Support Rank) *Suppose that a chance constraint* $i \in \mathbb{N}_1^N$ *has the support rank* $\rho_i \in \mathbb{N}_1^d$*. Then its support dimension in the* MSP *is bounded by* $\zeta_i \leq \rho_i$*.*

Proof. Without loss of generality, the proof is given for the first chance constraint $i = 1$. Pick any random multi-sample $\bar{\omega} \in \Delta^K$ (less any $\Pr^K$-null set for which the support rank condition may not hold).

By the assumption, there exists a linear subspace $L_1 \subset \mathbb{R}^d$ of dimension $d - \rho_1$ for which

$$f_1(x + \xi) = f_1(x) \qquad \forall\ x \in \mathbb{X},\ \ \forall\ \xi \in L_1 \quad .$$

The orthogonal complement of L_1, $L_1^\perp$, is also a linear subspace of $\mathbb{R}^d$ with dimension ρ_1, and every vector in $\mathbb{R}^d$ can be uniquely written as the orthogonal sum of vectors in L_1 and $L_1^\perp$, cf. [6, p. 135].

For the sake of a contradiction, suppose that $i = 1$ contributes more than ρ_1 support constraints to the resulting $\overline{\text{MSP}}$, i.e. $|\,\overline{\text{Sc}}_1\,| \geq \rho_1 + 1$. For any $\kappa_1 \in \overline{\text{Sc}}_1$, let

$$\bar{x}^\star_{\kappa_1} := \bar{x}^\star\big(\bar{\omega}^{(1)} \setminus \{\bar{\delta}^{(1,\kappa_1)}\}, \bar{\omega}^{(2)}, ..., \bar{\omega}^{(N)}\big)$$

be the solution obtained if this support constraint is omitted. By Definition 3.1, if a support constraint is omitted from $\overline{\text{MSP}}$, its solution moves away from $\bar{x}^\star_0$, i.e. $\bar{x}^\star_0 \neq \bar{x}^\star_{\kappa_1}$ for all $\kappa_1 \in \overline{\text{Sc}}_1$. Denote the collection of all solutions by

$$X := \big\{\bar{x}^\star_{\kappa_1} \,\big|\, \kappa_1 \in \overline{\text{Sc}}_1\big\} \cup \{\bar{x}^\star_0\} \ ,$$

so that $|X| \geq \rho_1 + 2$. Observe that each $\bar{x}^\star_{\kappa_1}$ is feasible with respect to all constraints of the $\overline{\text{MSP}}$, except for the one generated by $\delta^{(1,\kappa_1)}$, which is necessarily violated according to Definition 3.1.

Since $\mathbb{R}^d$ is the orthogonal direct sum of L_1 and $L_1^\perp$, for each point in X there is a unique orthogonal decomposition of

$$\bar{x}^\star_{\kappa_1} = v_{\kappa_1} + w_{\kappa_1} \ , \qquad \text{where} \ \ v_{\kappa_1} \in L_1, \ \ w_{\kappa_1} \in L_1^\perp \ ,$$

where $\kappa_1 \in \overline{\text{Sc}}_1 \cup \{0\}$. Consider the set

$$W := \big\{w_{\kappa_1} \,\big|\, \kappa_1 \in \overline{\text{Sc}}_1 \cup \{0\}\big\} \ .$$

By the hypothesis, W contains at least $\rho_1 + 2$ distinct points in the ρ_1-dimensional subspace $L_1^\perp$. According to Radon's Theorem [23, p. 151], W can be split into two disjoint subsets, W_A and W_B, such that there exists a point $\tilde{w}$ in the intersection of their convex hulls:

$$\tilde{w} \in \operatorname{conv}\big\{W_A\big\} \cap \operatorname{conv}\big\{W_B\big\} \ . \tag{3.8}$$

Split the indices in $\overline{\text{Sc}}_1 \cup \{0\}$ correspondingly into I_A and I_B, and observe that every $w_A \in W_A$ satisfies the constraints in I_B:

$$f_1\big(w_A, \bar{\delta}^{(1,\kappa_1)}\big) \leq 0 \quad \forall\ \kappa_1 \in I_B \qquad \Longrightarrow \qquad f_1\big(\tilde{w}, \bar{\delta}^{(1,\kappa_1)}\big) \leq 0 \quad \forall\ \kappa_1 \in I_B \ .$$

The last implication follows because $\tilde{w} \in \operatorname{conv}\{W_A\}$ and $f_1(\,\cdot\,, \bar{\delta}^{(1,\kappa_1)})$ is convex. Similarly, every point $w_B \in W_B$ satisfies the constraints in I_A:

$$f_1\big(w_B, \bar{\delta}^{(1,\kappa_1)}\big) \leq 0 \quad \forall\ \kappa_1 \in I_A \qquad \Longrightarrow \qquad f_1\big(\tilde{w}, \bar{\delta}^{(1,\kappa_1)}\big) \leq 0 \quad \forall\ \kappa_1 \in I_A \ .$$

Combining both statements thus yields

$$f_1(\tilde{w}, \bar{\delta}^{(1,\kappa_1)}) \leq 0 \qquad \forall\ \kappa_1 \in \overline{\text{Sc}}_1 \ . \tag{3.9}$$

According to (3.8), $\tilde{w}$ can be expressed as a convex combination of elements in W_A or W_B. Splitting the points in X into X_A and X_B correspondingly and applying the same convex combination yields some

$$\tilde{x} \in \operatorname{conv}\big\{X_A\big\} \cap \operatorname{conv}\big\{X_B\big\} \ , \tag{3.10}$$

and thereby also some $\tilde{v} \in L_1$ with $\tilde{x} = \tilde{v} + \tilde{w}$.

To establish the contradiction two things remain to be verified: first that $\tilde{x}$ is feasible with respect to all constraints, and second that it has a lower cost (or a better tie-break value) than $\bar{x}^\star_0$. For the first, $\tilde{x} \in \mathbb{X}$ because all points of X lie in $\mathbb{X}$ and $\tilde{x} \in \operatorname{conv}\{X\}$. Moreover, thanks to (3.9),

$$f_1\big(\tilde{x}, \bar{\delta}^{(1,\kappa_1)}\big) = f_1\big(\tilde{w}, \bar{\delta}^{(1,\kappa_1)}\big) \leq 0 \qquad \forall\ \kappa_1 \in \overline{\text{Sc}}_1 \ .$$

For the second, pick the set from X_A and X_B that does not contain $\bar{x}^\star_0$; without loss of generality, say this is X_A. By construction, all elements of X_A have a strictly lower objective function value (or at least a better tie-break value) than $\bar{x}^\star_0$. By linearity this also holds for all points in $\operatorname{conv}\{X_A\}$, where $\tilde{x}$ lies according to (3.10). $\square$

Remark 3.9 (Support Rank versus Support Dimension) *While the support rank ρ_i is a property of chance constraint i alone, the support dimension ζ_i may depend on the overall setup of the* MSP. *The support dimension ζ_i constitutes the relevant basis for selecting the sample size K_i. However, it may be difficult to determine for practical problems, as it may depend on the interactions of multiple chance constraints (see Example 3.10 below). The support rank ρ_i provides an easier-to-handle upper bound to ζ_i, which can be used in place of ζ_i for selecting K_i.*

Example 3.10 (Upper Bounding of Support Dimension) To illustrate the statements in Remark 3.9, consider a small example of (2.1) in dimension $d = 3$. Let $\mathbb{X} = [-1, 1]^3$ be the unit cube, $c^{\mathrm{T}} = [0\,1\,1]$ with a lexicographic tie-break rule, and two chance constraints $i = 1, 2$. Both constraints affect only the first and second coordinates x_1 and x_2, leaving the choice of $x_3 = -1$ for the third coordinate. For $i = 1$, the constraints are parallel hyperplanes constraining x_1 from below, where the lower bound is given by the first uncertainty δ_1:

$$f_1(x, \delta) = -x_1 + \delta_1 \ .$$

For $i = 2$, the constraints are V-shaped, with the vertex located at $x_1 = -\delta_2$ and $x_2 = -1$:

$$f_2(x, \delta) = \big|x_1 + \delta_2\big| - x_2 - 1 \ .$$

Both uncertainties $\delta := \{\delta_1, \delta_2\}$ are uniformly distributed on the interval $[0, 1]$. The setup is illustrated in Figure 3.3.

In this case, the support dimensions are $\zeta_1 = 1$, $\zeta_2 = 1$ and the support ranks are $\rho_1 = 1$, $\rho_2 = 2$ for the constraints $i = 1, 2$. Notice that for $i = 2$ the support rank is strictly greater that its support dimension, due to the presence of constraint 1. Hence there is some conservatism in the upper bound, although both bounds are better than the existing upper bound by the dimension of the decision space $d = 3$ [10, Thm. 2].

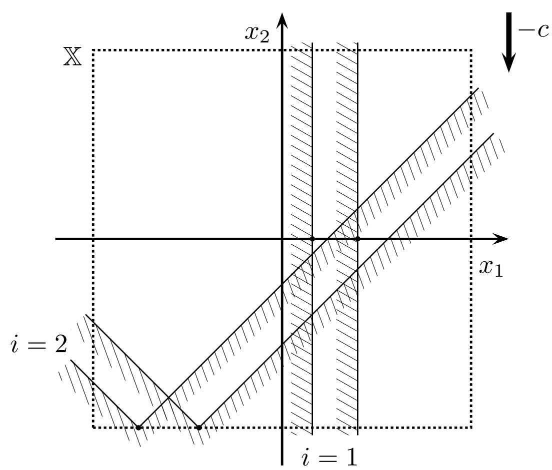


Figure 3.3: Illustration of Example 3.10. The plot shows a projection on the x_1, x_2-plane for $x_3 = -1$. The unit box $\mathbb{X}$ is depicted by a dotted line. Two (possible) samples are shown for the linear constraint $i = 1$ ($x_1 \geq \delta_1$) and for the V-shaped constraint $i = 2$ ($x_2 \geq \big|x_1 + \delta_2\big| - 1$).

4 Feasibility of the Scenario Solution

In the first part of this section, it is shown that for a proper choice of the sample sizes $K_1, ..., K_N$ the scenario solution $x^\star\big(\omega^{(1)}, ..., \omega^{(N)}\big)$ is an approximate solution of the MCP (i.e. it is a feasible point of each chance constraint $i = 1, ..., N$ in (2.1b) with a high confidence $(1 - \theta_i)$). In the second part of this section, an explicit formula for computing the sample sizes $K_1, ..., K_N$ for given residual probabilities θ_i is provided.

4.1 The Sampling Theorem

Denote by $\mathrm{B}(\cdot\,;\cdot,\cdot)$ the beta distribution function, cf. [1, p. 26.5.3, 26.5.7]:

$$\mathrm{B}(\varepsilon; n, K) := \sum_{j=0}^{n} \binom{K}{j} \varepsilon^j (1-\varepsilon)^{K-j} \ . \tag{4.1}$$

Theorem 4.1 (Sampling Theorem) *Consider problem* (2.3) *under Assumptions 2.1, 2.2, 2.3, 2.4, 3.2. Then*

$$\mathrm{Pr}^K\big[V_i(\omega^{(1)}, ..., \omega^{(N)}) > \varepsilon_i\big] \leq \mathrm{B}(\varepsilon_i; \rho_i - 1, K_i) \ , \tag{4.2}$$

for each chance constraint $i \in \mathbb{N}_1^N$, whose support rank is ρ_i.

Proof. The result is an extension of [11, Thm. 2.4] for the classic scenario approach, which is also used as a basis for this proof.[1]

Without loss of generality, consider the first chance constraint $i = 1$; the result for the other chance constraints $i = 2, ..., N$ follows analogously. Consider the conditional probability

$$\mathrm{Pr}^K\big[V_1(\omega^{(1)}, ..., \omega^{(N)}) > \varepsilon_1 \,\big|\, \omega^{(2)}, ..., \omega^{(N)}\big] \ , \tag{4.3}$$

i.e. the probability of drawing $\omega^{(1)}$ such that $x^\star(\omega^{(1)}, ..., \omega^{(N)})$ has a probability of violating '$f_1(\,\cdot\,, \delta) \leq 0$' that is higher than ε_1, given fixed values for the other samples $\omega^{(2)}, ..., \omega^{(N)}$.

Clearly, the quantity in (4.3) generally depends on the multi-samples $\omega^{(2)}, ..., \omega^{(N)}$. However, for $\mathrm{Pr}^{K_2+...+K_N}$-almost every value of these multi-samples (4.3) can be bounded by

$$\mathrm{Pr}^K\big[V_1(\omega^{(1)}, ..., \omega^{(N)}) > \varepsilon_1 \,\big|\, \omega^{(2)}, ..., \omega^{(N)}\big] \leq \mathrm{B}(\varepsilon_1; \rho_1 - 1, K_1) \ . \tag{4.4}$$

Indeed, by Assumption 2.1, for $\mathrm{Pr}^{K_2+...+K_N}$-almost every $\omega^{(2)}, ..., \omega^{(N)}$ the function $\tilde{f} : \mathbb{R}^d \to \mathbb{R}$ defined by

$$\tilde{f}(x) \equiv \max_{i \in \mathbb{N}_2^N} \max_{\kappa_i \in \mathbb{N}_1^{K_i}} f_i\big(x, \delta^{(i,\kappa_i)}\big)$$

is convex, as it is the point-wise maximum of convex functions. Then all sampled constraints of $i = 2, ..., N$ can be expressed as the deterministic convex constraint '$\tilde{f}(x) \leq 0$', which can be considered as part of the convex set $\mathbb{X}$. Thus for $\mathrm{Pr}^{K_2+...+K_N}$-almost every $\omega^{(2)}, ..., \omega^{(N)}$ the problem takes the form of a classic SCP, to which the results of [11] apply. In particular, [11, Thm. 2.4] yields (4.4) for $\mathrm{Pr}^{K_2+...+K_N}$-almost every $\omega^{(2)}, ..., \omega^{(N)}$.

The difference from using the support rank ρ_1 in place of the optimization dimension d in [11, Thm. 2.4] is minor. The key fact is that ρ_1 provides an upper bound for the number of support constraints contributed by constraint 1, according to Lemma 3.8, and hence it can replace d in [11, Prop. 2.2] and all subsequent results.

The final result is obtained by deconditioning the probability in (4.3):

$$\begin{aligned}
&\mathrm{Pr}^K\big[V_1(\omega^{(1)}, ..., \omega^{(N)}) > \varepsilon_1\big] = \\
&\quad = \int_{\omega^{(2)},...\omega^{(N)}} \mathrm{Pr}^K\big[V_1(\omega^{(1)}, ..., \omega^{(N)}) > \varepsilon_1 \,\big|\, \omega^{(2)}, ..., \omega^{(N)}\big] \mathrm{Pr}^{K_2}\big[\mathrm{d}\,\omega^{(2)}\big] ... \mathrm{Pr}^{K_N}\big[\mathrm{d}\,\omega^{(N)}\big] \\
&\quad \leq \int_{\omega^{(2)},...\omega^{(N)}} \Phi(\varepsilon_1; \rho_1 - 1, K_1) \mathrm{Pr}^{K_2}\big[\mathrm{d}\,\omega^{(2)}\big] ... \mathrm{Pr}^{K_N}\big[\mathrm{d}\,\omega^{(N)}\big] \\
&\quad = \Phi(\varepsilon_1; \rho_1 - 1, K_1) \ ,
\end{aligned}$$

[1]The authors thank an anonymous reviewer for his/her helpful suggestions on simplifying the proof.

based on [22, pp. 183,222], where the third line uses (4.4). □

4.2 Explicit Bounds on the Sample Sizes

Formula (4.2) in Theorem 4.1 ensures that with a *confidence level* of $1 - \mathrm{B}(\varepsilon_i; \rho_i - 1, K_i)$, the violation probability $V_i(\omega^{(1)}, ..., \omega^{(N)}) \leq \varepsilon_i$. However, in practical applications a given confidence level $(1 - \theta_i) \in (0, 1)$ is often imposed, while an appropriate sample size K_i has to be identified.

The most accurate way of finding this sample size is by observing that $\mathrm{B}(\varepsilon_i; \rho_i - 1, K_i)$ is a monotonically decreasing function in K_i and applying a numerical procedure (e.g. regula falsi) for computing the smallest sample size that ensures $\mathrm{B}(\varepsilon_i; \rho_i - 1, K_i) \leq \theta_i$. The resulting K_i shall be referred to as the *implicit bound* on the sample size.

For a qualitative analysis of the behavior of this implicit bound as ε_i and θ_i vary (and also for a good initialization of the regula falsi procedure), it is useful to derive an *explicit bound* on the sample size K_i. Since formula (4.2) cannot be readily inverted, the beta distribution function must first be controlled by some upper bound, which is then inverted.

A straightforward approach is to use a Chernoff bound [13], as shown in [8, Rem. 2.3] and [9, Sec. 5]. This provides a simple explicit formula for K_i:

$$K_i \geq \frac{2}{\varepsilon_i} \left[\log\Big(\frac{1}{\theta_i}\Big) + \rho_i - 1 \right] \ , \tag{4.6}$$

where $\log(\cdot)$ denotes the natural logarithm. As shown in [2, Cor. 1], this can be further improved to a better, albeit more complicated bound for K_i:

$$K_i \geq \frac{1}{\varepsilon_i} \left[\log\Big(\frac{1}{\theta_i}\Big) + \sqrt{2(\rho_i - 1)\log\Big(\frac{1}{\theta_i}\Big)} + \rho_i - 1 \right] \ . \tag{4.7}$$

5 The Sampling-and-Discarding Approach

The sampling-and-discarding approach has previously been proposed for the classic scenario approach [9, 12]; this section describes its extension to problems with multiple chance constraints.

The fundamental goal is to reduce the objective value of the scenario solution, while maintaining the same confidence levels for feasibility with respect to the chance constraints (see Section 1.2). To this end, the sample sizes K_i are deliberately increased above the bounds derived in Section 4, in exchange for allowing a certain number of R_i sampled constraints to be discarded *a posteriori*, i.e. after the outcomes of the samples have been observed.

In this section, first the possible procedures for discarding constraints are recalled. Second, the main result on the sampling-and-discarding approach for the MCP is stated. It provides an implicit formula for the selection of appropriate sample-and-discarding pairs (K_i, R_i), which may again vary for different chance constraints $i = 1, ..., N$. Third, explicit bounds for the choice of pairs (K_i, R_i) are provided.

5.1 Constraint Discarding Procedure

For each chance constraint of the MCP, if $R_i \geq 0$ sampled constraints are to be discarded a posteriori, the discarding procedure is performed by a pre-defined *(sample) removal algorithm*.

Definition 5.1 (Removal Algorithm) *For each chance constraint $i = 1, ..., N$, the* (sample) removal algorithm $\mathcal{A}_i^{(K_i,R_i)} : \Delta^K \to \Delta^{K_i - R_i}$ *is a deterministic function on the overall multi-sample $\omega \in \Delta^K$. It returns a subset of samples $\tilde{\omega}^{(i)} \in \Delta^{K_i - R_i}$, in which R_i out of the K_i samples in $\omega^{(i)} \in \Delta^{K_i}$ have been removed.*

Obviously, the algorithm should aim at improving the objective value from $\mathrm{MSP}[\omega^{(1)},...,\omega^{(N)}]$ to $\mathrm{MSP}[\tilde{\omega}^{(1)},...,\tilde{\omega}^{(N)}]$ as much as possible. Various possible removal algorithms are described in [9, Sec. 5.1], and further references are found in [12, Sec. 2]. Brief descriptions of the most important removal algorithms are listed below.

Example 5.2 *(a) Optimal Constraint Removal.* The best improvement of the objective function value is achieved by solving the reduced problem for all possible ways of removing R_i of the K_i samples. However, a major drawback of this removal algorithm is its combinatorial complexity. Therefore the algorithm becomes computationally intractable for larger values of R_i, in particular when samples have to be removed for multiple constraints.

(b) Greedy Constraint Removal. Starting by solving the $\mathrm{MSP}[\omega^{(1)},...,\omega^{(N)}]$ for all K_i samples, the R_i samples are removed in R_i sequentially steps. In each step, a single sample is removed by the optimal constraint removal procedure. Between multiple constraints i, the removal algorithm can either proceed in a fixed order or again greedy-based. For most practical problems this algorithm can be expected to work almost as good as (a), while carrying a much lower computational burden.

(c) Marginal Constraint Removal. The R_i samples are removed in R_i sequential steps, where the removed sample in each step is selected according to the highest Lagrange multiplier. Compared to the greedy constraint removal, the decision is thus based on the highest marginal cost improvement [7, Cha. 5]), instead of the highest total cost improvement. In the case of multiple constraints i, the removal algorithm can either handle them all together, or proceed sequentially.

The existing theory for the SCP [9, Sec. 4.1.1] and [12, Ass. 2.2] assumes that all of the removed constraints are violated by the relaxed scenario solution.

Assumption 5.3 (Violation of Discarded Constraints) *Every chance constraint $i \in \mathbb{N}_1^N$ with $R_i > 0$ satisfies the following condition: for almost every $\omega \in \Delta^K$, each of the constraints discarded by the removal algorithm $\mathcal{A}_i^{(K_i,R_i)}(\omega)$ is violated by the solution of the reduced problem, i.e.*

$$f_i\big(x^\star(\tilde{\omega}^{(1)},...,\tilde{\omega}^{(N)}),\delta^{(i,\kappa_i)}\big) > 0 \qquad \forall\; \delta^{(i,\kappa_i)} \in \big(\omega \setminus \tilde{\omega}\big)\;. \tag{5.1}$$

While Assumption 5.3 is sufficient for the MCP as well, it may turn out to be too restrictive for some problem instances. In fact, due to the interplay of multiple chance constraints, it may not be possible to find R_i constraints that are violated by the relaxed scenario solution (this situation may also occur for a single chance constraint, in the presence of a deterministic constraint set $\mathbb{X}$). In this case, the *monotonicity property*, as introduced below, provides a possible alternative.

Definition 5.4 (Monotonicity Property) *A chance constraint $i \in \mathbb{N}_1^N$ is called* monotonic *if for all $K_i \in \mathbb{N}$ and almost every $\omega^{(i)} \in \Delta^{K_i}$ the following condition holds: Every point in the feasible set of sampled instances of chance constraint i,*

$$\mathbb{X}_i(\omega^{(i)}) := \big\{\xi \in \overline{\mathbb{R}}^d \;\big|\; f_i(\xi,\delta^{(i,\kappa_i)}) \leq 0 \;\;\forall\; \kappa_i \in \mathbb{N}_1^{K_i}\big\}\;, \tag{5.2}$$

where $\overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$, is violated by a new sampled constraint only if also the optimal point in $\mathbb{X}_i(\omega^{(i)})$,

$$x_i^\star(\omega^{(i)}) := \arg\min\big\{c^T\xi \;\big|\; \xi \in \mathbb{X}_i(\omega^{(i)})\big\} \tag{5.3}$$

is violated. In other words, for every $\xi \in \mathbb{X}_i(\omega^{(i)})$ and almost every $\delta \in \Delta$,

$$f_i\big(\xi,\delta\big) > 0 \qquad \Longrightarrow \qquad f_i\big(x_i^\star(\omega^{(i)}),\delta\big) > 0\;. \tag{5.4}$$

Assumption 5.5 (Monotonicity of Chance Constraints) *Every chance constraint $i \in \mathbb{N}_1^N$ enjoys the* monotonicity property.

Definition (5.4) is easy to check for most practical problems, without involving any calculations. The following example illustrates the intuition behind this concept.

Example 5.6 (Monotonic Chance Constraints) Consider an MSP in $d = 2$ dimensions, where $\mathbb{X} = [-100, 100]^2 \subset \mathbb{R}^2$ and $c = [\,0\;\;1\,]^{\mathrm{T}}$, $\delta = [\delta_1\;\delta_2\;\delta_3]$ belongs to $\Delta = \{-1, 1\} \times [-1, 1] \times [-1, 1]$, and there are $N = 2$ chance constraints.

(a) Monotonic Chance Constraint. Let the first chance constraint $i = 1$ be of the linear form

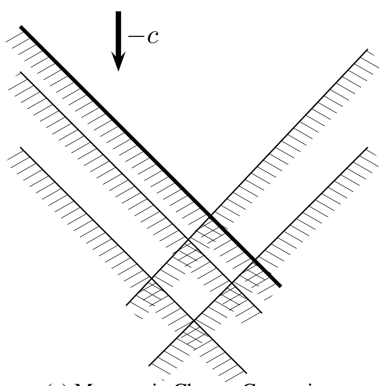
$$\begin{bmatrix}\delta_1^{(1,\kappa_1)} & 1\end{bmatrix} x - \delta_2^{(1,\kappa_1)} \leq 0 \qquad \forall\; \kappa_1 = 1, ..., K_1 \;\; .$$

Observe that for any number $K_1 \in \mathbb{N}$ and every possible sample values $\omega^{(1)}$, an additional sample δ either cuts off no point from $\mathbb{X}_1(\omega^{(1)})$, or the the point $x_1^\star(\omega^{(1)})$ becomes infeasible. This fact is illustrated in Figure 5.1(a). Therefore chance constraint $i = 1$ enjoys the monotonicity property.

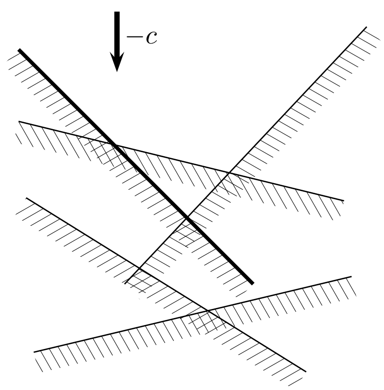
(b) Non-Monotonic Chance Constraint. Let the second chance constraint $i = 2$ be of the linear form

$$\begin{bmatrix}\delta_2^{(2,\kappa_2)} & 1\end{bmatrix} x - \delta_3^{(2,\kappa_2)} \leq 0 \qquad \forall\; \kappa_2 = 1, ..., K_2 \;\; ,$$

Observe that for any number K_2 there exist sample values $\omega^{(2)}$ that make it possible for a new sample δ to cut off some previously feasible point from $\mathbb{X}_2(\omega^{(2)})$, without rendering the point $x_2^\star(\omega^{(2)})$ infeasible. A possible configuration of this type is depicted in Figure 5.1(b). Therefore chance constraint $i = 2$ does not enjoy the monotonicity property.



(a) Monotonic Chance Constraint.



(b) Non-Monotonic Chance Constraint.

Figure 5.1: Illustration of Example 5.6. Non-bold constraints are generated by the multi-sample $\omega^{(i)} \in \Delta^{K_i}$ of chance constraint $i = 1, 2$; bold constraints are generated by the uncertainty $\delta \in \Delta$. In (b) a feasible point is made infeasible without affecting the optimum, which is not possible in the case of (a).

The usefulness of the monotonicity property is based on the following result, whose proof is an straightforward consequence of Definition 5.4 and therefore omitted.

Lemma 5.7 *Let $K_i \in \mathbb{N}$ and $R_i \leq K_i$. Suppose chance constraint $i \in \mathbb{N}_1^N$ of* MCP *is monotonic and the removal algorithm $\mathcal{A}_i^{(K_i,R_i)}$ is sequential. Then for almost every $\omega^{(i)} \in \Delta^{K_i}$ the following holds:*
(a) With probability one every point ξ in the set $\mathbb{X}_i(\omega^{(i)})$ has a violation probability less than or equal to that of the cost-minimal point $x_i^\star(\omega^{(i)})$:

$$\Pr\big[f_i(\xi, \delta) > 0\big] \leq \Pr\big[f_i(x_i^\star(\omega^{(i)}), \delta) > 0\big] \qquad \forall\; \xi \in \mathbb{X}_i(\omega^{(i)}) \;\; . \tag{5.5}$$

(b) The final solution $x_i^\star(\tilde{\omega}^{(i)})$, where $\tilde{\omega}^{(i)} = \mathcal{A}_i^{(K_i,R_i)}(\omega_i)$, violates all R_i removed constraints.

5.2 The Discarding Theorem

For the sampling-and-discarding approach, the following result holds for the MCP.

Theorem 5.8 (Discarding Theorem) *Consider the problem* (2.1) *under Assumptions 2.1, 2.2, 2.3, 2.4, 3.2, and either 5.3 or 5.5. Let $\mathcal{A}_i^{(K_i,R_i)}$ be sample removal algorithms for each of its chance constraints $i = 1, ..., N$, some of which may be trivial (i.e. $R_i = 0$). Then it holds that*

$$\Pr^K\big[V_i(\tilde{\omega}^{(1)}, ..., \tilde{\omega}^{(N)}) > \varepsilon_i\big] \leq \binom{R_i + \rho_i - 1}{R_i} \mathrm{B}(\varepsilon_i; R_i + \rho_i - 1, K_i) \ , \tag{5.6}$$

where ρ_i denotes the support rank of chance constraint i and $\mathrm{B}(\cdot;\cdot,\cdot)$ the beta distribution (A.1).

Proof. Here the MCP case is reduced to the SCP case, for which a detailed proof is available in [12, Sec. 5.1].

First, suppose that Assumption 5.3 holds. The proof in [12, Sec. 5.1] works analogously for an arbitrary chance constraint $i \in \mathbb{N}_1^N$, given that an upper bound of the violation distribution is readily available from Theorem 4.1.

Second, suppose that Assumption 5.5 holds. In this case the proof in [12, Sec. 5.1] can be applied directly to the SCP which arises from the MCP if all chance constraints other than a particular $i \in \mathbb{N}_1^N$ are omitted (and also $\mathbb{X}$ is omitted). In particular, (5.6) holds for the scenario solution of this SCP, using Lemma 5.7(b). Given that the chance constraint is monotonic and by virtue of Lemma 5.7(a), (5.6) also holds for any point in $\mathbb{X}_i(\omega^{(i)})$, in particular for the scenario solution of the MCP. $\square$

The work of [12] already provides an excellent account of the merits of the sampling-and-discarding approach, which does not require a restatement here. However, it should be emphasized that the scenario solution converges to the true solution of the MCP as the number of discarded constraints increases, provided that the constraints are removed by the optimal procedure of Example 5.2(a).

5.3 Explicit Bounds on the Sample-and-Discarding Pairs

Similar to Section 4, explicit bounds on the sample size K_i can also be derived for the sampling-and-discarding approach, assuming the number of discarded constraints R_i to be fixed. The technical details, using Chernoff bounds [13], are worked out in [9, Sec. 5]. The resulting explicit bound is indicated here for the sake of completeness,

$$K_i \geq \frac{2}{\varepsilon_i} \log\Big(\frac{1}{\theta_i}\Big) + \frac{4}{\varepsilon_i}\big(R_i + \rho_i - 1\big) \ , \tag{5.7}$$

where $\log(\cdot)$ denotes the natural logarithm.

Similarly, explicit bounds on the number of discarded constraints R_i can be obtained, assuming the sample size K_i to be fixed:

$$R_i \leq \varepsilon_i K_i - \rho_i + 1 - \sqrt{2\varepsilon_i K_i \log\Big(\frac{(\varepsilon_i K_i)^{\rho_i - 1}}{\theta_i}\Big)} \ . \tag{5.8}$$

The technical details of this are found in [12, Sec. 4.3].

6 Example: Minimal Diameter Cuboid

The following academic example has been selected to highlight the strengths of the extensions to the scenario approach presented in this paper.

6.1 Problem Statement

Let δ be a random point in $\Delta \subset \mathbb{R}^n$, whose distribution and support set are unknown, but sampled values can be obtained. The objective in this example is to construct the Cartesian product C of closed intervals in $\mathbb{R}^n$ ('n-cuboid') of minimal n-diameter W, which is large enough to contain the point δ in its i-th coordinate with probability $(1-\varepsilon_i)$. The setting is illustrated in Figure 6.1.

Let $z \in \mathbb{R}^n$ denote the center point of the cuboid and $t \in \mathbb{R}^n_+$ the interval widths in each dimension, so that

$$C = \left\{\xi \in \mathbb{R}^n \,\middle|\, |\xi_i - z_i| \leq t_i/2\right\} \ . \tag{6.1}$$

Then the corresponding stochastic program reads as follows:

$$\min_{z\in\mathbb{R}^n, t\in\mathbb{R}^n_+} \quad \|t\|_2 \ , \tag{6.2a}$$

$$\text{s.t.} \quad \Pr\big[z_i - t_i/2 \leq \delta_i \leq z_i + t_i/2\big] \geq (1-\varepsilon_i) \qquad \forall\, i \in \mathbb{N}_1^n \ . \tag{6.2b}$$

Since the objective function is not linear, (6.2) has to be reformulated (see Remark 1.1(a)) as

$$\min_{z\in\mathbb{R}^n, t\in\mathbb{R}^n_+, T\in\mathbb{R}} \quad T \ , \tag{6.3a}$$

$$\text{s.t.} \quad \|t\|_2 \leq T \ , \tag{6.3b}$$

$$\Pr\Big[\max\big\{z_i - t_i/2 - \delta_i, -z_i - t_i/2 + \delta_i\big\} \leq 0\Big] \geq (1-\varepsilon_i) \quad \forall\, i \in \mathbb{N}_1^n \, . \tag{6.3c}$$

Note that (6.3) takes the form of a MCP, for a $d = 2n+1$ dimensional search space and $N = n$ chance constraints: the objective function (6.3a) is linear; constraint (6.3b) is deterministic and convex; and each of the chance constraints in (6.3c) is convex in z, t for any fixed value of the uncertainty $\delta \in \Delta$.

Here each of the chance constraints $i = 1, ..., n$ depends on exactly two decision variables z_i and t_i, which is a special case of involving $[z; t; T] \in \mathbb{R}^{2n+1}$ (see Remark 1.1(c)). The convex and compact set $\mathbb{X}$ is constructed from the positivity constraints on t, the deterministic and convex constraint (6.3b), and some artificial bounds assumed on all variables. Existence of a feasible solution, and hence Assumption 2.2, holds automatically from the problem setup.

6.2 Solution via Scenario Approach

By inspection, each of the chance constraints $i = 1, ..., n$ has support rank $\rho_i = 2$, because it only involves the two variables z_i and t_i. For a fixed confidence level, e.g. $\theta = 10^{-6}$, the implicit sample sizes $K_1, ..., K_n$ in (4.2) can be computed for given values of n and $\varepsilon_1, ..., \varepsilon_n \in (0,1)$ by a bisection-based algorithm (see Section 4.2). For simplicity, all $\varepsilon_1 = ... = \varepsilon_n$ are selected as equal, and since $\rho_1 = ... = \rho_N = 2$, the implicit sample sizes $K_1 = ... = K_n$ are also identical.

Given the outcomes of all multi-samples, the $\overline{\text{MSP}}$ is easily solved by the smallest n-cuboid that contains all sampled points; see also Figure 6.1. In other words, here the $\overline{\text{MSP}}$ has an analytic solution.

Table 6.1(a) summarizes the implicit sample sizes required for guaranteeing various chance constraint levels ε_i in various dimensions n (all with $\theta = 10^{-6}$). These sample sizes are also compared to those from the classic scenario approach, based on a reformulation of (6.3) as an SCP according to the procedure outlined in Section 2.1.

Observe from Table 6.1 that the SCP-based sample sizes are always larger than those using the extensions of the MCP theory. This effect increases, in particular, as the dimension n of the optimization space grows larger. The reason is that the support dimension of each chance constraint remains constant for all n, whereas Helly's dimension grows as it equals to n. The marginal growth of the sample size of the MCP, despite the support rank $\rho_i = 2$ being constant, is the result of adjusting the confidence level θ to be (evenly) distributed among the chance constraints, i.e. $\theta_i = \theta/n$ for all $i = 1, ..., n$.

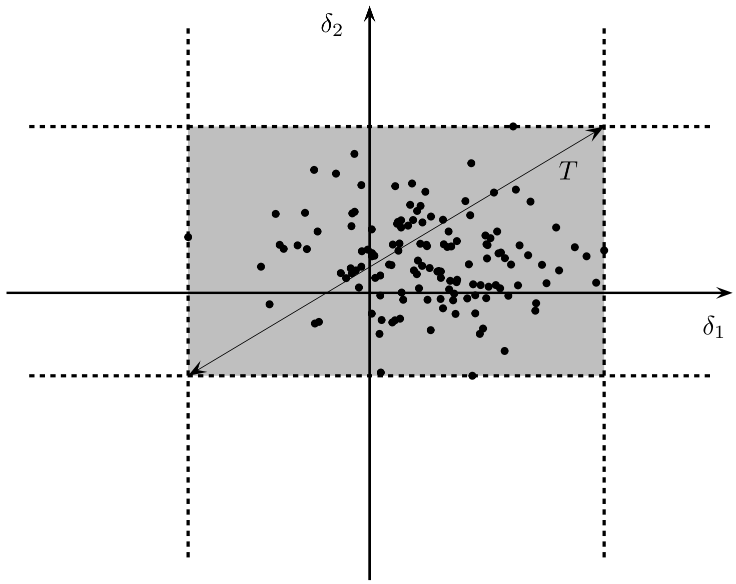


Figure 6.1: Illustration of the numerical example for $n = 2$. The point $\delta \in \Delta$ appears at random in $\mathbb{R}^2$, according to some unknown distribution; the points drawn here are 166 i.i.d. samples of δ. The objective is to construct the smallest product of two closed intervals ('2-cuboid'), drawn here as the shaded rectangle, such that the probability of failing to contain the realization of δ is smaller than ε_1 and ε_2 in dimension 1 and 2, respectively.

sample size K_i		cuboid dimension $n =$ 2	3	5	10	50	100	500
$\varepsilon_i =$	1%	1,734	1,777	1,831	1,903	2,072	2,144	2,311
	5%	341	349	360	374	407	421	454
	10%	166	170	176	182	199	205	221
	25%	62	63	65	67	73	76	82

(a) MCP-based Scenario Approach.

sample size K_i		cuboid dimension $n =$ 2	3	5	10	50	100	500
$\varepsilon_i =$	1%	2,334	2,722	3,431	5,020	15,588	27,535	115,786
	5%	459	536	677	992	3,095	5,477	23,093
	10%	225	263	332	488	1,533	2,719	11,506
	25%	84	99	125	186	595	1,063	4,550

(b) SCP-based Scenario Approach.

Table 6.1: Implicit sample sizes $K_1 = ... = K_n$ for the MCP-based and the SCP-based scenario approach, assuming a confidence level of $\theta = 10^{-6}$, for varying problem dimension n and chance constraint levels $\varepsilon_1 = ... = \varepsilon_n$.

The larger sample size of the SCP-based approach, as compared to the MCP-based approach, implies higher data requirements and higher computational efforts, but it also increases the conservatism of the scenario solution. The latter effect is quantified in Table 6.2, showing the relative excess of the (average) objective function values of the SCP-based solutions over those of the MCP-based solutions. Note that the objective values achieved by the SCP-based approach are always higher than those achieved by the MCP-based approach, with the effect becoming increasingly significant as the dimension n of the decision space grows larger.

relative obj. value		cuboid dimension $n =$ 2	3	5	10	50	100	500
$\varepsilon_i =$	1%	2.4%	3.4%	5.0%	7.5%	14.8%	18.4%	26.9%
	5%	3.3%	4.6%	6.6%	9.8%	19.3%	23.8%	34.4%
	10%	3.9%	5.4%	7.6%	11.5%	22.2%	27.4%	39.3%
	25%	5.0%	7.2%	10.1%	15.1%	28.5%	34.7%	49.1%

Table 6.2: Objective function value of SCP-based scenario solution as a percentage increase over the MCP-based scenario solution, based on the sample sizes in Table 6.1 and a multi-variate standard normal distribution for δ. Each of the indicated values represents an average over one million simulation runs.

Acknowledgement

The research of L. Fagiano has received funding from the European Union Seventh Framework Programme FP7/2007-2013 under grant agreement number PIOF-GA-2009-252284, Marie Curie project 'Innovative Control, Identification and Estimation Methodologies for Sustainable Energy Technologies'.

A Probability Distributions

Several basic probability-related functions are used throughout this paper. The *Binomial Distribution Function* [1, p. 26.1.20]

$$\Phi(x; K, \varepsilon) := \sum_{j=0}^{x} \binom{K}{j} \varepsilon^j (1-\varepsilon)^{K-j} \tag{A.1}$$

expresses the probability of seeing at most $x \in \mathbb{N}_0^K$ successes in $K \in \mathbb{N}$ independent Bernoulli trails, where the probability of success is $\varepsilon \in (0, 1)$ per trial. The (real) *Beta Function* [1, p. 6.2.1]

$$\mathrm{B}(a, b) := \int_0^1 \xi^{a-1}(1-\xi)^{b-1} \,\mathrm{d}\,\varepsilon \tag{A.2}$$

is defined for any parameters $a, b \in \mathbb{R}_+$, and $\xi \in (0, 1)$; it also satisfies the identity [1, p. 6.2.2]

$$\mathrm{B}(a, b) = \mathrm{B}(b, a) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} \ , \tag{A.3}$$

where $\Gamma : \mathbb{R}_+ \to \mathbb{R}_+$ denotes the (real) *Gamma Function* with $\Gamma(n+1) = n!$ for any $n \in \mathbb{N}_0^\infty$ [1, p. 6.1.5]. The corresponding *Incomplete Beta Function* [1, p. 6.6.1] is then given by

$$\mathrm{B}(\varepsilon; a, b) := \int_0^\varepsilon \xi^{a-1}(1-\xi)^{b-1} \,\mathrm{d}\,\xi = \int_{1-\varepsilon}^1 \xi^{b-1}(1-\xi)^{a-1} \,\mathrm{d}\,\xi \ , \tag{A.4}$$

where the last equality follows by a simple substitution. An important identity is obtained from [1, pp. 3.1.1, 6.6.2, 26.5.7],

$$\mathrm{B}(\varepsilon; a, b) = \mathrm{B}(a, b) \sum_{j=a}^{a+b-1} \binom{a+b-1}{j} \varepsilon^j (1-\varepsilon)^{a+b-1-j} \ , \tag{A.5}$$

which can written more compactly by use of the binomial distribution (A.1), see for instance [9, p. 3437]:

$$\mathrm{B}(\varepsilon; a, b) = \frac{1}{b} \binom{a+b-1}{b}^{-1} \Phi(b-1; a+b-1, 1-\varepsilon) \ . \tag{A.6}$$

References

[1] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*, Dover Publications, New York, 9th ed., 1970.

[2] T. Alamo, R. Tempo, and A. Luque, *On the sample complexity of probabilistic analysis and design methods*, in Perspectives in Mathematical System Theory, Control, and Signal Processing, J. W. et al., ed., Springer, Berlin et al., 2010, pp. 39–50.

[3] D. Bai, T. Carpenter, and J. Mulvey, *Making a case for robust optimization models*, Management Science, 43(7) (1997), pp. 895–907.

[4] A. Ben-Tal and A. Nemirovski, *Robust convex optimization*, Mathematics of Operations Research, 23(4) (1998), pp. 769–805.

[5] J. Birge and F. Louveaux, *Introduction to Stochastic Programming*, Springer, New York, 1997.

[6] B. Bollobás, *Linear Analysis*, Cambridge University Press, Cambridge et al., 2nd ed., 1999.

[7] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, 2004.

[8] G. Calafiore, *On the expected probability of constraint violation in sampled convex programs*, Journal of Optimization Theory and Applications, 143 (2009), pp. 405–412.

[9] ———, *Random convex programs*, SIAM Journal of Optimization, 20(6) (2010), pp. 3427–3464.

[10] G. Calafiore and M. Campi, *Uncertain convex programs: Randomized solutions and confidence levels*, Mathematical Programming, Series A, 102-1 (2005), pp. 25–46.

[11] M. Campi and S. Garatti, *The exact feasibility of randomized solutions of uncertain convex programs*, SIAM Journal of Optimization, 19 (2008), pp. 1211–1230.

[12] ———, *A sampling and discarding approach to chance-constrained optimization: Feasibility and optimality*, Journal of Optimization Theory and Applications, 148 (2011), pp. 257–280.

[13] H. Chernoff, *A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations*, The Annals of Mathematical Statistics, 23(4) (1952), pp. 493–507.

[14] P. Kall and J. Mayer, *Stochastic Linear Programming*, Springer, New York et al., 2nd ed., 2011.

[15] P. Kouvelis and G. Yu, *Robust Discrete Optimization and Its Applications*, Kluwer Academic Publishers, Dordrecht, 1997.

[16] D. Luenberger and Y. Ye, *Linear and Nonlinear Programming*, Springer, Berlin et al., 3rd ed., 2008.

[17] J. Mulvey and R. Vanderbei, *Robust optimization of large-scale systems*, Operations Research, 43(2) (1995), pp. 264–281.

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

[19] A. Prékopa, *Stochastic Programming*, Kluwer, Dordrecht et al., 1995.

[20] R. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970.

[21] A. Shapiro, D. Dentcheva, and A. Ruszczyński, *Lectures on Stochastic Programming, Modeling and Theory*, SIAM, Philadelphia, 2009.

[22] A. Shiryaev, *Probability*, Springer, New York et al., 2nd ed., 1996.

[23] G. Ziegler, *Lectures on Polytopes*, Springer, New York et al., 1st ed., 2007.