On objective function value performance of the scenario approach under regularity conditions

Zheming Wang [*1,2] and Raphaël M. Jungers [†3]

[1]The Department of Automation, Department of Automation, Zhejiang University of Technology, Hangzhou 310023, China.
[2]The Zhejiang Provincial United Key Laboratory of Embedded Systems, Hangzhou 310023, China.
[3]The ICTEAM Institute, UCLouvain, Louvain-la-Neuve,1348, Belgium

Abstract

We study the objective function value performance of the scenario approach for robust convex optimization. A novel method is proposed to derive probabilistic bounds for the objective value from scenario programs with a finite number of samples. This method relies on a max-min reformulation and on the concept of complexity of robust optimization problems. With additional continuity and regularity conditions, via sensitivity analysis, we also provide explicit bounds which outperform the previously existing bounds. To illustrate our contribution, we also provide numerical examples. Finally, we apply our method to a planar antenna array synthesis problem, where we investigate the overfitting issue based on the derived probabilistic objective value bounds.

Keywords— Scenario approach, probabilistic guarantees, sensitivity analysis

1 Introduction

In real-world applications, optimization problems with uncertainties often arise due to the lack of knowledge on system parameters and measurement/estimation errors [1]. Classic approaches to tackle uncertain optimization problems include stochastic optimization in which uncertainties are treated as random variables [2, 3, 4] and robust optimization which seeks to obtain robust solutions against all realizations of uncertainties [5, 6, 7, 8]. An alternative approach for handling uncertainties is called the scenario approach or scenario optimization, which is somehow at the intersection of stochastic optimization and robust optimization as it is able to provide solutions to robust optimization and chance-constrained stochastic optimization problems in a probabilistic sense, see, e.g., some early works [9, 10]. We refer to [11, 12] for a comprehensive view on this subject.

The scenario approach is typically focused on constrained optimization problems with robustness constraints or chance constraints. By randomly sampling a finite set of constraints, a scenario program is formulated and a sample-based solution can be obtained. Chance-constrained theorems are then derived for this sample-based solution in terms of the measure of the violating subset of

*wangzheming@zjut.edu.cn

†raphael.jungers@uclouvain.be Raphaël M. Jungers is a FNRS honorary Research Associate. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under grant agreement No 864017 - L2C. Raphaël M. Jungers is also supported by the Walloon Region and the Innoviris Foundation.

the uncertain constraints for both convex [9, 10, 13, 14, 15, 16] and non-convex [17, 18] cases by using the concepts of support/essential constraints. In particular, the bounds in [13, 14] have been proved to be tight for a special class of uncertain convex programs called fully supported problems. While these theoretic results bring a lot of new perspectives into uncertain optimization problems, it takes further developments to apply the scenario approach to practical applications. For instance, many real-life physical properties of interest are directly related with the *objective function* of the underlying optimization problem, which means that the aforementioned chance-constrained theorems alone do not allow us to learn such properties. For this reason, a formal analysis of the scenario approach is needed with respect to objective function value performance. To the best of our knowledge, the first work providing such results is [19] under some continuity and regularity conditions on the basis of the chance-constrained theorems in [13, 14]. Though the results in [19] are built on an intermediate bound that is tight in terms of the measure of the violating subset of the uncertain constraints (as proved in [13, 14]), tightness is not provided in terms of the objective function. In this paper, we show that these results are indeed not necessarily tight in terms of the objective function by providing improved bounds under the same conditions. In addition, the chance-constrained theorems in [13, 14] require a non-degeneracy assumption. We show that such an assumption can be released when one is only concerned with the objective function.

This paper is motivated by recent progress on data-driven analysis of complicated systems, see, e.g., [20, 21, 22, 23, 24, 25, 26, 27, 28]. In particular, for switched linear systems [24, 25, 26, 27, 28], the scenario approach has been applied to the computation of the growth rate of the system using different types of Lyapunov functions after reformulating the stability analysis as a robust optimization problem in which the Lyapunov inequality is satisfied for all realizations of the trajectory. Different techniques are then proposed in these works to derive probabilistic guarantees on the computed growth rate. It appears that some of the results can be generalized from the techniques in [24, 25, 26, 27, 28] to the general scenario approach. In [27], it is shown that the stability analysis problem for switched systems reduces to a finite set of support constraints whose cardinality is bounded by the dimension of the decision variable. In other words, the notion of support constraints is generalized from scenario programs to robust optimization problems. This observation is then leveraged to derive probabilistic stability certificates using a sensitivity analysis argument on these support constraints. In this paper, we extend this sensitivity analysis method to general robust optimization. Analogously to [16], we formally call the minimal number of such support constraints the *complexity* of the robust optimization problem.

The contributions of this paper are two-fold: First, we present a new general method of deriving probabilistic guarantees on objective value performance of the scenario approach for robust convex programs. Second, we derive improvements of the bounds from [19] under the same regularity conditions.

This paper extends preliminary results presented in a conference paper [29]. Additionally to [29], the present work contains the detailed proofs of all the lemmas, propositions and theorems. We also apply our results to min-max optimization problems and a special class of non-convex problems, in which probabilistic guarantees are derived taking the underlying structure into account. Furthermore, we provide more numerical examples, among which is a planar antenna synthesis problem.

The rest of the paper is organized as follows. This section ends with the notation, followed by Section 2 on the review of preliminary results on the scenario approach. Section 3 presents a general method of deriving objective value performance guarantees for the scenario approach using a max-min reformulation and uniform level-set bounds defined implicitly. In Section 4, under additional regularity conditions, we show explicit objective value performance bounds via sensitivity analysis. In Section 5, we consider special cases and the extension to non-convex programs. Finally, we apply our results to two numerical examples in Section 6.

Notation. The non-negative real number set and the non-negative integer set are denoted by $\mathbb{R}^+$ and $\mathbb{Z}^+$ respectively. For any $p \geq 1$, the ℓ_p norm of a vector $x \in \mathbb{R}^n$ is $\|x\|_p$ ($\|x\|$ is the ℓ_2 norm by default). Given a metric space $(\mathcal{M}, \rho)$, consider any two subsets S and X, $\|S\|_X := \sup_{s\in S} \inf_{x\in X} \rho(s, x)$; $\mathrm{cl}(S)$ denotes the closure of S; $|S|$ denotes the cardinality of S for any $S \subset \mathcal{M}$; 2^S denotes the power set of S.

2 Preliminaries

We consider the following robust convex optimization problem (RCP):

$$J^\star := \inf_{x\in X} c^\top x \tag{1a}$$

$$\text{s.t. } f(x,\delta)\le 0, \forall \delta\in\Delta \tag{1b}$$

where $x\in\mathbb{R}^d$ is the decision variable, δ is the uncertain parameter contained in the uncertain set Δ, $X\subseteq\mathbb{R}^d$ is some static constraint set, Δ is a subset of some metric space $(\mathcal{M},\rho)$ with $\Delta\subset\mathcal{M}$ and $\rho:\mathcal{M}\times\mathcal{M}\to\mathbb{R}^+$ is a metric, and $f: X\times\mathcal{M}\to\mathbb{R}$ is the uncertain constraint function. Let us mention that a linear cost function is considered without loss of generality [9, 10]. We would also like to mention that, while there is only one uncertain constraint in (1), the results and discussions in the paper extend trivially to cases with multiple uncertain constraints. In this paper, we consider the case where (1b) contains infinitely many inequality constraints. Such a problem is also known as a semi-infinite program (SIP), see, e.g., [30]. The following assumptions are made.

Assumption 1 *$f: X\times\mathcal{M}\to\mathbb{R}$ is lower semi-continuous (l.s.c.) in $X\times\Delta$ and $X\subseteq\mathbb{R}^d$ and $\Delta\subset\mathcal{M}$ are compact.*

Assumption 2 *There exists $x\in X$ such that $f(x,\delta)\le 0$ for every $\delta\in\Delta$.*

Assumption 3 *$X\subseteq\mathbb{R}^d$ is convex and $f: X\times\mathcal{M}\to\mathbb{R}$ is convex in $x\in X$ for any $\delta\in\Delta$.*

These are mild and standard assumptions, which make the objective of performance approximation well-posed. In particular, Assumption 2 is a feasibility condition which prevents the optimal $J^\star$ from being ∞. Such an assumption is also made in [19] (see Assumption 3.3 in [19]).

Let $(\Delta,\mathcal{B}(\Delta),\mathbb{P})$ be a probability space where $\mathcal{B}(\Delta)$ is the Borel σ-algebra of Δ and $\mathbb{P}:\mathcal{B}(\Delta)\to[0,1]$ is a probability measure.

In the framework of the scenario approach (also called scenario optimization) [9, 10, 12], the solution of the robust optimization problem in (1) is approximated with a finite number of samples. More precisely, we randomly generate $N\in\mathbb{Z}^+$ samples, denoted by $(\delta_1,\delta_2,\cdots,\delta_N)\in\Delta^N$, from some probability measure with the support Δ, and formulate the following scenario convex program (SCP)

$$\mathcal{P}(\omega): \quad J(\bar\omega) := \min_{x\in X} c^\top x \tag{2a}$$

$$\text{s.t. } f(x,\delta)\le 0, \forall\delta\in\bar\omega. \tag{2b}$$

where $\omega := (\delta_1,\delta_2,\cdots,\delta_N)$, and $\bar\omega$ denote the unordered sample set, i.e., $\bar\omega := \{\delta_1,\delta_2,\cdots,\delta_N\}$, and $J: 2^\Delta\to\mathbb{R}$ with $J(\Delta)=J^\star$.

Chance-constrained guarantees on the solution of the sampled problem are derived in [13, 14] for convex problems in terms of the violation probability, depending on the dimension of the decision variable and the number of samples, as shown below.

Theorem 1 ([13, Theorem 1],[14, Theorem 3.3]) *For any $N\in\mathbb{Z}^+$ with $N>d$, suppose ω are N independent and identically distributed (i.i.d.) samples drawn according to the probability measure $\mathbb{P}$. Consider $\mathcal{P}(\omega)$ as defined in (2), let us assume that $\mathcal{P}(\omega)$ admits a unique optimal solution. Then, with Assumptions 1–3 and additional conditions (see Remark 1), for any $\epsilon\in(0,1)$,*

$$\mathbb{P}^N\{\omega\in\Delta^N : \mathcal{V}(\bar\omega)>\epsilon\}\le\Phi_c(\epsilon;d,N) \tag{3}$$

where $\mathcal{V}(\bar\omega) := \mathbb{P}\{\delta : J(\bar\omega\cup\{\delta\})>J(\bar\omega)\}$, $\bar\omega$ is the unordered set of the elements in ω, $J(\bar\omega)$ is defined as in (2), $\mathbb{P}^N$ is the probability measure in the N-Cartesian product of Δ, and

$$\Phi_c(\epsilon;k,N) := \sum_{i=0}^{k-1}\binom{N}{i}\epsilon^i(1-\epsilon)^{N-i}, \forall k\ge 1. \tag{4}$$

In fact, the bound in (3) is proved to be tight for a special class of convex programs called fully supported programs, see [13, 14] for details. For non-convex optimization problems, similar results are derived in [17, 18] using an a posteriori estimate of the *complexity* of the SCP. While these probabilistic guarantees for both the convex and non-convex cases provide a fundamental understanding on the solution of the scenario program (2), they alone do not allow to bound the objective value of the original robust optimization.

Remark 1 *The bound in (3) is derived by [13] and [14] based on different conditions. More precisely, [13] requires that the feasible domain of $\mathcal{P}(\omega)$ has a nonempty interior for any ω; [14] assumes that $\mathcal{P}(\omega)$ is nondegenerate with probability one, see Definition 2.7 in [14] for the definition of degenerate problems.*

In [19], theoretical links between $\mathcal{P}(\omega)$ and the original robust optimization problem RCP in (1) are further established in terms of the objective function in a statistical sense under the assumptions in Theorem 1 and Slater's condition, i.e., there exists $x \in X$ such that $\sup_{\delta\in\Delta} f(x,\delta) < 0$. We briefly summarize the idea in [19] as follows. First, it is shown that the optimizer of SCP is a feasible solution to the following chance-constrained program (CCP) with some confidence level

$$\min_{x\in X} c^\top x \tag{5a}$$

$$\text{s.t. } \mathbb{P}\{f(x,\delta) \leq 0\} \geq 1-\epsilon \tag{5b}$$

where $\epsilon \in (0,1)$. It is then shown that the feasible set of the program CCP is contained in the feasible set of a relaxed RCP in the form of

$$\inf_{x\in X} c^\top x \tag{6a}$$

$$\text{s.t. } f(x,\delta) \leq \xi, \forall\delta \in \Delta \tag{6b}$$

where $\xi \geq 0$. Finally, with the relations above, probabilistic objective bounds for the scenario approach can be derived from the relaxed RCP under Lipschitz continuity and other regularity assumptions. We call this procedure a two-step method in the sense that it relies on the link from SCP to CCP (via the chance-constrained Theorem 1) to establish probabilistic objective function value bounds. As we will show in the sequel, the step from SCP to CCP brings conservatism into the resulting bound.

In this paper, we revisit the problem of deriving probabilistic objective function value bounds for the scenario approach from a sensitivity analysis perspective in the context of robust optimization. We develop a method that skips the step from SCP to CCP and enables us to directly establish probabilistic objective value bounds for SCP using a perturbation argument. Thus, our method does not require the conditions in Remark 1. We believe that this latter fact constitutes an important improvement because these conditions can be violated in some control problems, see [26] for such an example. The difference between our method and the two-step method [19] is illustrated in Figure 1.

3 Probabilistic objective value performance

This section presents a general method deriving probabilistic objective value bounds of the scenario approach for robust optimization problems. This method adopts a max-min reformulation of the original robust optimization problem (1).

3.1 Continuity properties

We first show some preliminary continuity results on parametric optimization, see, e.g., [31, 32, 33] for a comprehensive view. Given any $N \in \mathbb{Z}^+$, for any $\omega := (\delta_1, \delta_2, \cdots, \delta_N) \in \Delta^N$, for ease of

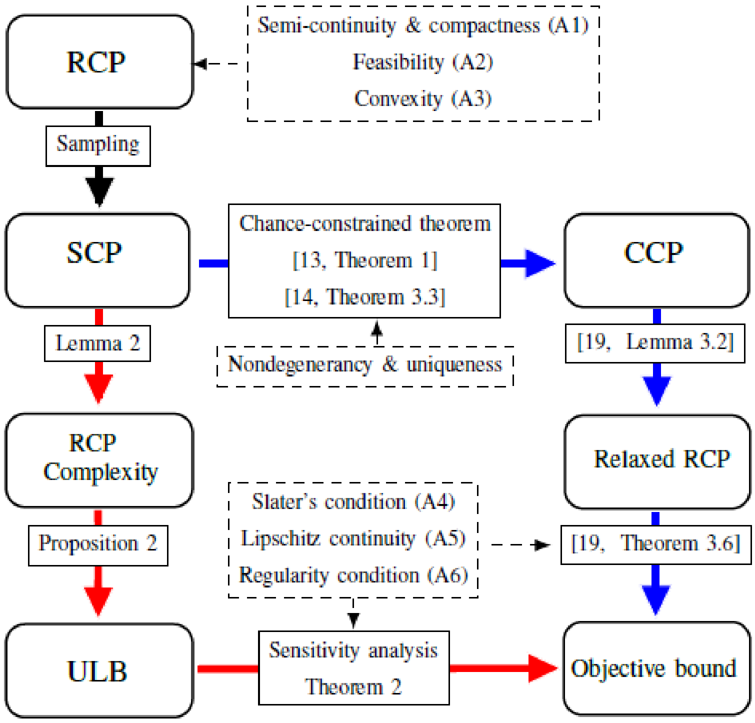


Figure 1: Flowchart of the proposed method and its comparison with the two-step method: The red arrows represent our method and the blue arrows represent the two-step method.

discussion, let us define

$$g_N(\omega) := J(\{\delta_1, \delta_2, \cdots, \delta_N\}), \tag{7}$$

$$F_N(\omega) := \{x \in X : f(x, \delta_i) \le 0, i = 1, 2, \cdots, N\}, \tag{8}$$

$$F_N^o(\omega) := \{x \in X : f(x, \delta_i) < 0, i = 1, 2, \cdots, N\}, \tag{9}$$

$$\phi_N(\omega) := \{x \in F_N(\omega) : c^\top x = g_N(\omega)\}, \tag{10}$$

Under Assumption 1, a semi-continuity property is stated below.

Lemma 1 *Consider the RCP in* (1) *and the SCP in* (2)*. Given any $N \in \mathbb{Z}^+$, for any $\omega \in \Delta^N$, let $g_N(\omega)$ be defined as in (7). Suppose Assumptions 1& 2 hold, then $g_N(\cdot)$ is l.s.c. in Δ^N.*

Proof: The proof is given in Appendix A. □

With Lemma 1, g_N can be even proved to be continuous under additional conditions, as stated in the following proposition.

Proposition 1 *Assume $f : X \times \mathcal{M} \to \mathbb{R}$ is continuous in $X \times \Delta$ and X is compact. Given any $N \in \mathbb{Z}^+$, for any $\omega \in \Delta^N$, let $g_N(\omega)$, $F_N(\omega)$, $F_N^o(\omega)$ and $\phi_N(\omega)$ be given as in (7), (8), (9) and (10) respectively. The following results hold:*
(i) For any $\omega \in \Delta^N$, if $F_N(\omega) \subseteq cl(F_N^o(\omega))$, then $g_N : \Delta^N \to \mathbb{R}$ is continuous at ω.
(ii) Suppose that for any $\omega \in \Delta^N$ there exists $x \in \phi_N(z)$ such that $x \in F_N^o(\omega)$, then $g_N : \Delta^N \to \mathbb{R}$ is continuous at ω.

Proof: The proof is given in Appendix B. □

3.2 Max-min reformulation

With the continuity property above, we then define, for any $N \in \mathbb{Z}^+$,

$$g_N^* := \sup_{\omega \in \Delta^N} g_N(\omega) = \sup_{\omega \in \Delta^N} \inf_{x \in F_N(\omega)} c^\top x \tag{11}$$

where $F_N(\omega)$ is defined as in (8). The next lemma states the equivalence between the RCP in (1) and the max-min formulation as defined in (11).

Lemma 2 *Consider Problems (1) and (11) for any $N \in \mathbb{Z}^+$ with the optimal values $J^\star$ and g_N^* respectively. Suppose Assumptions 1 – 3 hold, then it holds that $J^\star = g_N^*$ for any $N \geq d$;*

Proof: From the definitions in (1) and (11), it holds that $g_N(\omega) \leq J^\star$ for any $\omega \in \Delta^N$, which implies that $g_N^* \leq J^\star$ for any $N \in \mathbb{Z}^+$. As $\{g_N^*\}_{N\in\mathbb{Z}^+}$ is a monotonically increasing sequence, bounded from above by $J^\star$, we only need to show that $g_d^* = J^\star$. First, we show that $g_{d+1}^* = J^\star$. The optimality of g_{d+1}^* in (11) implies that the set $\mathcal{X}_{(\delta_1,\delta_2,\cdots,\delta_{d+1})} := \{x \in X : c^\top x \leq g_d^*, f(x,\delta_i) \leq 0, i = 1,2,\cdots,d+1\}$ is nonempty for any $(\delta_1,\delta_2,\cdots,\delta_{d+1}) \in \Delta^{d+1}$. Since $f(\cdot,\cdot)$ is l.s.c. in $X \times \Delta$ and X is a compact convex set, $\mathcal{X}_{(\delta_1,\delta_2,\cdots,\delta_{d+1})}$ is also a compact convex set. Hence, from Helly's theorem (see, e.g., Theorem 21.6 in [34]), $\{x \in X : c^\top x \leq g_{d+1}^*, f(x,\delta) \leq 0, \delta \in \Delta\}$ is nonempty, which means that $J^\star \leq g_{d+1}^*$ and thus $g_{d+1}^* = J^\star$. Now, we show that $g_d^* = g_{d+1}^*$. For any $\omega := (\delta_1,\delta_2,\cdots,\delta_{d+1}) \in \Delta^{d+1}$, from Theorem 2 in [9] (or Theorem 3 in [10]), there exists i such that $J(\bar{\omega}) = J(\bar{\omega} \setminus \{\delta_i\})$, where $\bar{\omega} := \{\delta_1,\delta_2,\cdots,\delta_{d+1}\}$. As a result, $g_d^* = g_{d+1}^*$, which implies that $g_d^* = J^\star$. $\square$

In view of the result in Lemma 2, we define the concept of *complexity* for robust optimization problems with infinitely many constraints, following the definition of *complexity* for scenario programs with a finite number of constraints in [16].

Definition 1 *Consider the RCP in the form of (1), the* complexity *of RCP is the smallest integer $k \in \mathbb{Z}^+$ such that $g_k^* = J^\star$.*

Similar to the *complexity* of SCP in [16], the *complexity* of RCP is intuitively the minimal number of samples needed to obtain the true objective value $J^\star$. With this formal definition, Lemma 2 basically means that the *complexity* of RCP is bounded by the number of decision variables under Assumptions 1 – 3. When RCP is not feasible, i.e., Assumption 2 does not hold (which means that $J^\star = +\infty$), from Helly's theorem, it still holds that $J^\star = g_{d+1}^*$ but not $J^\star = g_d^*$ in Lemma 2. Thus, the *complexity* of RCP is bounded by $d+1$ in any case.

3.3 Uniform level-set bounds

To establish probabilistic objective function value performance guarantees, inspired from [35, 19], we define the following tail probability for any $N \in \mathbb{Z}^+$

$$p_N(\alpha) := \mathbb{P}^N\{\omega \in \Delta^N : g_N^* - \alpha \leq g_N(\omega)\}, \forall \alpha \geq 0. \tag{12}$$

From the semi-continuity property of g_N in Lemma 1, $p_N(\alpha)$ is well defined and monotonically increasing. In particular, when $N \geq d$, from Lemma 2, p_N defined as in (12) becomes

$$p_N(\alpha) = \mathbb{P}^N\{\omega \in \Delta^N : J^\star - \alpha \leq g_N(\omega)\}, \forall \alpha \geq 0. \tag{13}$$

The continuity property of this tail probability function is stated in the following lemma.

Lemma 3 *Consider the RCP* (1) *and the SCP* (2)*. Suppose Assumptions 1 - 3 hold. For any $N \in \mathbb{Z}^+$, let the tail probability function $p_N : \mathbb{R}^+ \to [0,1]$ be defined as in (12). Then, the following property holds: i) There exists a finite $C \in \mathbb{R}^+$ such that $p_N(\alpha) = 1$ for any $\alpha \geq C$. ii) $p_N(\cdot)$ is upper semi-continuous (u.s.c.) at any $\alpha \in \mathbb{R}^+$.*

Proof: i) Let $\tilde{J} := \min_{x\in X} c^\top x$. As a result, $g_N(\omega) \geq \tilde{J}$ for any $z \in \Delta$. Hence, for any $\alpha \geq g_N^* - \tilde{J}$, $p_N(\alpha) = 1$.
ii) It is easy to verify that $p_N(\cdot)$ is monotonically nondecreasing. Now, we show that for any $\bar{\alpha} \in \mathbb{R}^+$, it holds that $p_N(\bar{\alpha}) \geq \limsup_{\alpha\to\bar{\alpha}} p_N(\alpha)$, i.e., $\lim_{i\to\infty} \alpha_i = \bar{\alpha}$ implies $p_N(\bar{\alpha}) \geq \limsup_{i\to\infty} p_N(\alpha_i)$. From the monotonicity of $p_N(\alpha)$, we only need to consider the case that $\alpha_{i+1} \leq \alpha_i$ for all $i \in \mathbb{Z}^+$. We then define: $S_i = \{\omega \in \Delta^N : J^\star - \alpha_i \leq g_N(\omega))\}$ for $i \in \mathbb{Z}^+$. Thus, $\limsup_{i\to\infty} p_N(\alpha_i) =$

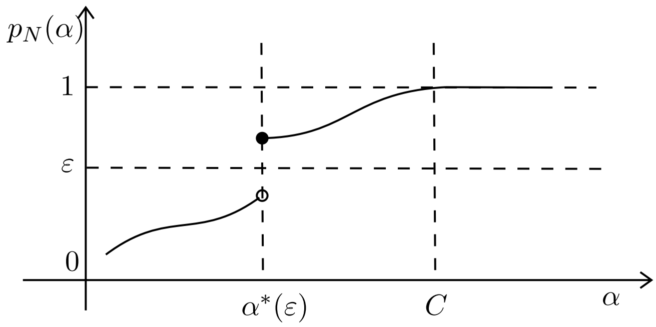


Figure 2: Illustration of the tail probability function and the optimal ULB: Given $\varepsilon \in (0, 1)$, $\alpha^*(\varepsilon)$ is the supremum of α satisfying $p_N(\alpha) \le \varepsilon$.

$\lim_{i\to\infty} \mathbb{P}^N\{S_i\} = \mathbb{P}^N\{\lim_{i\to\infty} S_i\} = p_N(\bar{\alpha})$ since $S_i \supseteq S_{i+1}$ and $\mathbb{P}^N\{S_i\}$ is bounded for any $i \in \mathbb{Z}^+$. This proves upper semi-continuity of p_N. □

We then call $\tilde{\alpha} : (0, 1) \to \mathbb{R}$ a *uniform level-set bound (ULB)* if it satisfies

$$p_N(\tilde{\alpha}(\varepsilon)) \ge \varepsilon, \forall \varepsilon \in (0, 1). \tag{14}$$

The *optimal ULB* is defined as follows:

$$\alpha^*(\varepsilon) \coloneqq \sup_{\alpha \ge 0} \alpha : p_N(\alpha) \le \varepsilon, \forall \varepsilon \in (0, 1). \tag{15}$$

Let us mention that an important difference with [35, 19] is that the definition of ULBs in this paper is directly based on the objective function instead of the worst-case violation of the constraint.

With the discussions above, we arrive at the main result of this section.

Proposition 2 *Consider the RCP in* (1) *and the SCP in* (2)*. Suppose Assumptions 1 – 3 hold. For any* $N \in \mathbb{Z}^+$ *with* $N \ge d$ *and* N *i.i.d. samples, denoted by* ω*, drawn according to the probability measure* $\mathbb{P}$*, consider Problem (2) with the optimal value* $g_N(\omega)$ *as defined in* (7)*. For any* $\varepsilon \in (0, 1)$*, let* $\alpha^*(\varepsilon)$ *be defined as in (15). Then, it holds that*

$$\mathbb{P}^N\{\omega \in \Delta^N : J^\star \le g_N(\omega) + \alpha^*(\varepsilon)\} \ge \varepsilon \tag{16}$$

where $J^\star$ *is given in (1).*

Proof: With Lemma 2 at hand, we only need to show $p_N(\alpha^*(\epsilon)) \ge \varepsilon$ for any $\varepsilon \in (0, 1)$ as $J^\star = g_N^*$. We consider two cases: i) the point at the optimal is continuous; ii) the point at the optimal is not continuous. In the first case, from the monotonicity of $p_N(\cdot)$, $p_N(\alpha) \le \varepsilon$ is active, meaning that $p_N(\alpha^*(\varepsilon)) = \varepsilon$. In the second case, from the upper semi-continuity shown in Lemma 3, $p_N(\alpha^*(\varepsilon)) \ge \varepsilon$. Thus, we conclude that $p_N(\alpha^*(\varepsilon)) \ge \varepsilon$. □

While this is an elementary result, following from the bound on the *complexity* of RCP and the u.s.c. of $p_N(\cdot)$, as illustrated in Figure 2, it provides a general way of bounding the original objective value $J^\star$. In practice, as computing the optimal ULB is not always possible, we turn to upper bounds of $\alpha^*(\varepsilon)$.

Remark 2 *From the proof of Lemma 3, when there is no measure concentration on the level sets of* g_N*, i.e.,* $\mathbb{P}^N(\omega \in \Delta^N : g_N(\omega) = u) = 0$ *for any* $u \in \mathbb{R}$*, it can be shown that* $p_N(\cdot)$ *is continuous. Thus, the optimal ULB can be defined by* $p_N(\alpha^*(\varepsilon)) = \varepsilon$*.*

Remark 3 *Let us also highlight that the result in Proposition 2 is still valid without optimality or uniqueness. That is, for any algorithm that generates a feasible solution to* $\mathcal{P}(\omega)$ *with the corresponding objective value function* $\tilde{g}_N(\omega)$*, it holds that* $\mathbb{P}^N\{\omega \in \Delta^N : J^\star \le \tilde{g}_N(\omega) + \alpha^*(\varepsilon)\} \ge \varepsilon$ *for any* $\varepsilon \in (0, 1)$ *under the same conditions as stated in Proposition 2.*

4 Explicit probabilistic guarantees via sensitivity analysis

We now explicitly establish probabilistic objective function value bounds for the scenario approach via sensitivity analysis under regularity conditions. With this, we are also able to compute explicit ULBs as defined in (14).

4.1 Lipschitz continuity

We first derive some continuity results on the optimal objective function of the scenario program (2). This relies on Slater's condition of the RCP as given below.

Assumption 4 *(Slater's condition) There exists $x_0 \in X$ such that $\sup_{\delta\in\Delta} f(x_0,\delta) < 0$.*

In fact, there can exist multiple solutions x_0 that satisfy Slater's condition. For convenience, we choose any x_0 according to Assumption 4, and define the constant below following [19]:

$$L_{SP} := \frac{\min_{x\in X} c^\top x - c^\top x_0}{\sup_{\delta\in\Delta} f(x_0,\delta)} \tag{17}$$

We then assume Lipschitz continuity of $f(x,\delta)$ with respect to δ inside the closure of Δ.

Assumption 5 *For any $x \in X$, $f(x,\cdot)$ is Lipschitz continuous with constant L_δ in Δ, i.e., $\forall \delta_1, \delta_2 \in \Delta$,*

$$\|f(x,\delta_1) - f(x,\delta_2)\| \leq L_\delta \|\delta_1 - \delta_2\|. \tag{18}$$

Based on these two additional assumptions, we present a Lipschitz continuity result in the next lemma.

Lemma 4 *Given any $N \in \mathbb{Z}^+$, let $g_N(\omega)$ be defined as in (7) for any $\omega \in \Delta^N$. Suppose Assumptions 1 – 5 hold, then, for any $\omega := (\delta_1, \delta_2, \cdots, \delta_N) \in \Delta^N$ and $\omega' := (\delta'_1, \delta'_2, \cdots, \delta'_N) \in \Delta^N$, it holds that*

$$\|g_N(\omega) - g_N(\omega')\| \leq L_{SP} L_\delta \max_{i\in\{1,2,\cdots,N\}} \|\delta_i - \delta'_i\|. \tag{19}$$

where L_{SP} and L_δ are given in (17) and (18) respectively.

Proof: For any two points $\omega := (\delta_1, \delta_2, \cdots, \delta_N), \omega' := (\delta'_1, \delta'_2, \cdots, \delta'_N) \in \Delta^N$, we want to show (19). First, we defined the perturbed problem for ω with the perturbation $L_\delta \max_j \|\delta_j - \delta'_j\|$:

$$\tilde{g}_N(\omega) := \min_{x\in X} c^\top x \tag{20a}$$

$$\text{s.t.} \quad f(x,\delta_i) \leq L_\delta \max_j \|\delta_j - \delta'_j\|, \forall i = 1,2,\cdots,N \tag{20b}$$

Similarly, the perturbed problem for ω' is also defined:

$$\tilde{g}_N(\omega') := \min_{x\in X} c^\top x \tag{21a}$$

$$\text{s.t.} \quad f(x,\delta'_i) \leq L_\delta \max_j \|\delta_j - \delta'_j\|, \forall i = 1,2,\cdots,N \tag{21b}$$

From strong duality and perturbation analysis, see, e.g., Section 5.6 in [36] and Lemma 3.4 in [19], the optimal cost function of the perturbed problem is Lipschitz continuous with respect to the perturbation with constant L_{SP}. Thus, $0 \leq g_N(\omega) - \tilde{g}_N(\omega) \leq L_{SP} L_\delta \max_j \|\delta_j - \delta'_j\|$ and $0 \leq g_N(\omega') - \tilde{g}_N(\omega') \leq L_{SP} L_\delta \max_j \|\delta_j - \delta'_j\|$. From the Lipschitz continuity of $f(x,\delta)$ in δ, $|f(x,\delta_i) - f(x,\delta'_i)| \leq L_\delta \|\delta_i - \delta'_i\| \leq L_\delta \max_j \|\delta_j - \delta'_j\|$ for all $i = 1,2,\cdots,N$, which implies that $\{x : f(x,\delta_i) \leq 0, i = 1,2,\cdots,N\} \subseteq \{x : f(x,\delta'_i) \leq L_\delta \max_j \|\delta_j - \delta'_j\|, i = 1,2,\cdots,N\}$ and $\{x : f(x,\delta'_i) \leq 0, i = 1,2,\cdots,N\} \subseteq \{x : f(x,\delta_i) \leq L_\delta \max_j \|\delta_j - \delta'_j\|, i = 1,2,\cdots,N\}$. Hence, $\tilde{g}_N(\omega) \leq g_N(\omega')$ and $\tilde{g}_N(\omega') \leq g_N(\omega)$. Finally, we conclude that $|g_N(\omega') - g_N(\omega)| \leq L_{SP} L_\delta \max_j \|\delta_j - \delta'_j\|$. $\square$

With the continuity result above, a deterministic relation between the robust optimization problem RCP and the scenario program SCP is established in the following proposition.

Proposition 3 *Suppose Assumptions 1 – 5 hold. For any $\bar{\omega} \subset \Delta$, let $J(\bar{\omega})$ be defined as in (2). Then, the following results hold: (i) There exists a set $\bar{\omega}^* \subset \Delta$ with $|\bar{\omega}^*| = d$ such that $J(\bar{\omega}^*) = J^\star$. (ii) Given any d-element set $\bar{\omega}^*$ satisfying $J(\bar{\omega}^*) = J^\star$, for any finite subset $\bar{\omega} \subseteq \Delta$ with $|\bar{\omega}| \geq d$, $J(\bar{\omega}) \geq J^\star - L_{SP} L_\delta \|\bar{\omega}^*\|_{\bar{\omega}}$, where L_{SP} and L_δ are given in (17) and (18) respectively.*

Proof: (i) From Lemma 2, we know that $J^\star = g_d^* := \sup_{\omega \in \Delta^d} g_d(\omega)$. With the result in Lemma 4, $g_d(\cdot)$ is Lipschitz continuous in Δ^d, which means that, according to Weierstrass's theorem, there exist $\omega \in \Delta^d$ such that $g_d^* = g(\omega)$.
(ii) Given an extreme set $\bar{\omega}^* := \{\delta_1^*, \delta_2^*, \cdots, \delta_d^*\}$, let us define the set $\bar{\omega}_{proj} := \{\delta_1, \delta_2, \cdots, \delta_d\} \subseteq \bar{\omega}$ as follows: $\delta_i := \arg\min_{\delta \in \bar{\omega}} \|\delta_i^* - \delta\|$ for any $i = 1, 2, \cdots, d$. Thus, from (19),

$$\begin{aligned} J(\bar{\omega}^*) - J(\bar{\omega}) &\leq J(\bar{\omega}^*) - J(\bar{\omega}_{proj}) = g_d(\omega^*) - g_d(\omega_{proj}) \\ &\leq L_{SP} L_\delta \max_{i \in \{1,2,\cdots,d\}} \|\delta_i^* - \delta_i\| \\ &= L_{SP} L_\delta \|\bar{\omega}^*\|_{\bar{\omega}}, \end{aligned}$$

where $\omega^* := (\delta_1^*, \delta_2^*, \cdots, \delta_d^*)$ and $\omega_{proj} := (\delta_1, \delta_2, \cdots, \delta_d)$. $\square$

4.2 Explicit probabilistic bounds

In the rest of the section, we provide explicit probabilistic bounds for the scenario approach using the continuity results above. To do so, we also need a regularity condition on Δ with respect to the measure $\mathbb{P}$. Such a condition is often needed for domain or boundary estimation problems [37]. For comparison, we consider the same condition as in [19].

Assumption 6 *For the probability space $(\Delta, \mathcal{B}(\Delta), \mathbb{P})$, there exists a strictly increasing function $\varphi : \mathbb{R}^+ \to [0, 1]$ such that*

$$\forall \delta \in \Delta, \forall r \in \mathbb{R}^+, \mathbb{P}\{\mathbb{B}_r(\delta) \cap \Delta\} \geq \varphi(r), \tag{22}$$

where $\mathbb{B}_r(\delta)$ is an open ball centered at δ with radius r in the metric space $(\mathcal{M}, \rho)$.

Following [37], we can also consider other regularity conditions[1]. The following lemma shows some elementary probability results.

Lemma 5 *Given the probability space $(\Delta, \mathcal{B}(\Delta), \mathbb{P})$, for any $k \in \mathbb{Z}^+ (k \geq 1)$, consider k subsets $\{B_1, B_2, \cdots, B_k\}$ of Δ with $\mathbb{P}\{B_1\} = \mathbb{P}\{B_2\} = \cdots = \mathbb{P}\{B_k\} = \epsilon$ for some $\epsilon \in [0, 1]$. For any $N \in \mathbb{Z}^+$, let ω be N independent and identically distributed (i.i.d.) samples drawn according to the probability measure $\mathbb{P}$. Then, for any $\epsilon \in [0, 1]$,*

$$\begin{aligned} \mathbb{P}^N\{\omega \in \Delta^N : B_i \cap \bar{\omega} \neq \emptyset, \forall i\} &\geq 1 - \Psi(\epsilon; k, N) \qquad (23) \\ &\geq 1 - \Phi_a(\epsilon; k, N) \qquad (24) \end{aligned}$$

where $\bar{\omega}$ is the unordered set of elements in ω,

$$\Phi_a(\epsilon; k, N) := k(1 - \epsilon)^N, \text{ and} \tag{25}$$

$$\Psi(\epsilon; k, N) := \lfloor \frac{k}{k_\epsilon} \rfloor \Phi(\epsilon; k_\epsilon, N) + \Phi(\epsilon; k (\text{mod } k_\epsilon), N) \tag{26}$$

with $k_\epsilon := \lfloor \frac{1}{\epsilon} \rfloor$, $\Phi(\epsilon; 0, N) := 0$ and

$$\Phi(\epsilon; k, N) := \sum_{i=1}^{k} (-1)^{i-1} \binom{k}{i} (1 - i\epsilon)^N. \tag{27}$$

[1] As an example, another regularity condition is that there exist $c > 0$ and $\bar{r} > 0$ such that

$$\mathbb{P}\{\mathbb{B}_r(\delta) \cap \Delta\} \geq c\mu(\mathbb{B}_r(\delta)), \forall \delta \in \Delta, 0 < r \leq \bar{r}$$

where $\mu(\cdot)$ denotes the Lebesque measure.

Proof: The proof is given in Appendix C. □

Remark 4 *From the arguments in the proof of Lemma 5, one observation is that, for any positive integer* $k' \le k_\epsilon$, *it holds that* $\mathbb{P}^N\{\omega \in \Delta^N : B_i \cap \bar{\omega} \neq \emptyset, \forall i\} \ge 1 - \lfloor \frac{k}{k'} \rfloor \Phi(\epsilon; k', N) - \Phi(\epsilon; k(\text{mod } k'), N)$. *In particular, when* $k' = 1$, $\lfloor \frac{k}{k'} \rfloor \Phi(\epsilon; k', N) + \Phi(\epsilon; k(\text{mod } k'), N) = k(1-\epsilon)^N$, *which show the connection between the two bounds in Lemma 5.*

To illustrate the difference among the bounds in (4), (25), and (26), a numerical comparison is made in Figure 3 with $k = 8$ and $N = 500$. Note that, in this figure, $\Psi(\epsilon; k, N) = \Phi(\epsilon; k, N)$ for all the values of ϵ as it remains in the interval $[0, 1/8]$. It can be seen that $\Phi(\epsilon; k, N)$ is the best among these three bounds. As ϵ increases, $\Phi_a(\epsilon; k, N)$ becomes close to $\Phi(\epsilon; k, N)$ and outperforms $\Phi_c(\epsilon; k, N)$ eventually.

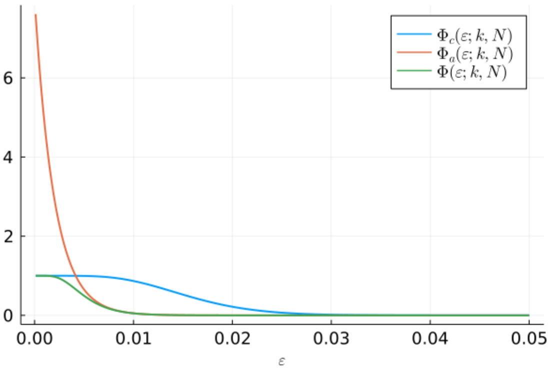


Figure 3: The curves of $\Phi_c(\epsilon; k, N)$, $\Phi_a(\epsilon; k, N)$ and $\Phi(\epsilon; k, N)$ with $k = 8$ and $N = 500$: $\Psi(\epsilon; k, N) = \Phi(\epsilon; k, N)$ for any $\epsilon \in [0, 1/8]$.

Putting the results above together, we then present the main result of this section.

Theorem 2 *Consider Problem (1) with the optimal objective value* $J^\star$, *suppose Assumptions 1 – 6 hold. For any* $N \in \mathbb{Z}^+$ *with* $N \ge d$, *let* $g_N(\omega)$ *be defined as in (7) for any* $\omega \in \Delta^N$. *Let* ω *be* N *i.i.d. samples drawn according to the probability measure* $\mathbb{P}$. *Then, for any* $\epsilon \in (0, 1)$,

$$\begin{aligned} &\mathbb{P}^N\{\omega \in \Delta^N : J^\star \le g_N(\omega) + L_{SP} L_\delta \varphi^{-1}(\epsilon)\} \\ &\quad \ge 1 - \Psi(\epsilon; d, N) \end{aligned} \tag{28}$$

where $\Psi(\epsilon; d, N)$ *is given in* (26), *and* L_{SP} *and* L_δ *are given in (17) and (18) respectively.*

Proof: From property (i) of Proposition 3, there exists a set $\bar{\omega}^* := \{\delta_1^*, \delta_2^*, \cdots, \delta_d^*\} \subset \Delta$ such that $J(\bar{\omega}_d^*) = J^\star$. For any $\epsilon \in (0, 1)$, let r_i be such that $\mathbb{P}\{\mathbb{B}_{r_i}(\delta_i^*) \cap \Delta\} = \epsilon$ for any $i = 1, 2, \cdots, d$. Under Assumption 6, we can see that $r_i \le \varphi^{-1}(\epsilon)$ for any $i = 1, 2, \cdots, d$. For convenience, we define $B_i = \mathbb{B}_{r_i}(\delta_i^*) \cap \Delta$ for any $i = 1, 2, \cdots, d$. With Lemma 5, we know that $\mathbb{P}^N\{\omega \in \Delta^N : B_i \cap \bar{\omega} \neq \emptyset, \forall i = 1, 2, \cdots, d\} \ge 1 - \Psi(\epsilon; d, N)$, where $\bar{\omega}$ is the unordered set of the samples ω. Since $B_i \cap \bar{\omega}$ implies that $|\bar{\delta}_i^*|_{\bar{\omega}} \le r_i \le \varphi^{-1}(\epsilon)$ for any $i = 1, 2, \cdots, d$, from property (ii) of Proposition 3, we get (28). □

Remark 5 *From the arguments in the proof of Theorem 2, Assumption 6 can be further relaxed. For Theorem 2 to be valid, we only need the regularity condition* (22) *to hold for any set* $\bar{\omega}^* = \{\delta_1^*, \delta_2^*, \cdots, \delta_d^*\} \subset \Delta$ *such that* $J(\bar{\omega}^*) = J^\star$, *i.e.,* $\forall \delta \in \bar{\omega}_d^*, \forall r \in \mathbb{R}^+, \mathbb{P}\{\mathbb{B}_r(\delta) \cap \Delta\} \ge \varphi(r)$. *In addition, let us mention that Theorem 2 is also valid without the closedness of* Δ *by imposing the regularity condition* (22) *for any* $\delta \in cl(\Delta)$, *while we implicitly assume that* Δ *is closed in Assumption 1.*

Theorem 2 then allows to derive explicit ULBs. More precisely, the following corollary provides a general explicit ULB.

Corollary 1 *Suppose the conditions in Theorem 2 hold. For any $\beta \in (0,1)$, it holds that*

$$\begin{aligned}\mathbb{P}^N\{\omega \in \Delta^N : & J^\star - g_N(\omega) \\ & \leq L_{SP} L_\delta \varphi^{-1}(\Psi^{-1}(\beta; d, N))\} \geq 1-\beta.\end{aligned} \tag{29}$$

where $\Psi^{-1}(\cdot; d, N)$ denote the inverse function of $\Psi(\cdot; d, N)$ with fixed d and N.

Proof: This is a direct consequence of Theorem 2. □

Using the bound in (24), an explicit and closed-form ULB can be derived in the following corollary.

Corollary 2 *Suppose the conditions in Theorem 2 hold, a ULB satisfying (14) can be given as follow:*

$$\tilde{\alpha}(\varepsilon) = L_{SP} L_\delta \varphi^{-1}(1 - \sqrt[N]{\frac{1-\varepsilon}{d}}), \forall \varepsilon \in (0,1). \tag{30}$$

Proof: As shown in Lemma 5, for any $\epsilon' \in (0,1)$ and $k \in \mathbb{Z}^+$, $\Psi(\epsilon'; k, N) \leq \Phi_a(\epsilon'; k, N) = k(1-\epsilon')^N$, which together with Theorem 2 implies that

$$\begin{aligned}&\mathbb{P}^N\{\omega \in \Delta^N : J^\star \leq g_N(\omega) + L_{SP} L_\delta \varphi^{-1}(\epsilon')\} \\ &\geq 1 - d(1-\epsilon')^N.\end{aligned} \tag{31}$$

Letting $\varepsilon = 1 - d(1-\epsilon')^N$, (30) is a direct consequence of (31). □

With the discussions above, the overall procedure of deriving objective value performance is summarized in Figure 1.

Remark 6 *As a comparison, we also show the bound from [19, Theorem 3.6] below, for any $\beta \in (0,1)$,*

$$\begin{aligned}\mathbb{P}^N\{\omega \in \Delta^N : & J^\star - g_N(\omega) \\ & \leq L_{SP} L_\delta \varphi^{-1}(\Phi_c^{-1}(\beta; d, N))\} \geq 1-\beta.\end{aligned} \tag{32}$$

The difference between these bounds lies in the numerical value of the functions $\Phi_a(\cdot; d, N)$, $\Phi(\cdot; d, N)$ (or $\Psi(\cdot; d, N)$) and $\Phi_c(\cdot; d, N)$, which essentially provide the confidence level of the estimates.

5 Min-max optimization and extension to non-convex programs

In this section, we show that the proposed method is applicable to min-max optimization problems and a special class of non-convex programs. More importantly, we find that, in these problems, Slater's condition is not necessarily needed in the derivation of explicit probabilistic objective function value bounds.

5.1 Min-max optimization

Let us first consider min-max optimization problems in the form of

$$\inf_{x \in X} \sup_{\delta \in \Delta} f(x, \delta). \tag{33}$$

To put (33) into the form of (1), we use the epigraph form below

$$J^\star_{epi} := \inf_{x \in X, \gamma} \gamma \tag{34a}$$

$$\text{s.t. } f(x, \delta) \leq \gamma, \forall \delta \in \Delta. \tag{34b}$$

For any $N \in \mathbb{Z}^+$ and any $\omega := (\delta_1, \delta_2, \cdots, \delta_N) \in \Delta^N$, we define the sampled problem of (34) below:

$$\gamma_N(\omega) := \min_{x \in X, \gamma} \gamma \tag{35a}$$

$$\text{s.t. } f(x, \delta_i) \leq \gamma, \forall i \in \{1, 2, \cdots, N\}. \tag{35b}$$

With these definitions, objective function value performance of the scenario approach for min-max optimization is analyzed in the following theorem, which can be viewed as a variant of Theorem 2.

Theorem 3 *Consider Problem* (34) *with the optimal value* $J^\star_{epi}$. *Suppose Assumptions 1, 3 & 5–6 hold. For any* $N \in \mathbb{Z}^+$ *with* $N \geq d$, *let* $\gamma_N(\omega)$ *be defined as in (35) and* $\phi_N(\omega) := \{x \in X : \max_i f(x, \delta_i) = \gamma_N(\omega)\}$ *for any* $\omega \in \Delta^N$. *Let* ω *be* N *i.i.d. samples drawn according to the probability measure* $\mathbb{P}$. *Then, the following results hold:*
(i) For any $\epsilon \in (0, 1)$,

$$\begin{aligned} &\mathbb{P}^N\{\omega \in \Delta^N : J^\star_{epi} \leq \gamma_N(\omega) + L_\delta \varphi^{-1}(\epsilon)\} \\ &\qquad \geq 1 - \Psi(\epsilon; d+1, N) \end{aligned} \tag{36}$$

where $\Psi(\epsilon; d+1, N)$ *is given in* (26) *and* L_δ *satisfies* (18).
(ii) Suppose there exists a continuous function $\mathcal{L}_\delta : X \to \mathbb{R}^+$ *such that,* $\forall \delta_1, \delta_2 \in \Delta$ *and* $\forall x \in X$

$$\|f(x, \delta_1) - f(x, \delta_2)\| \leq \mathcal{L}_\delta(x) \|\delta_1 - \delta_2\|. \tag{37}$$

For any $\epsilon \in (0, 1)$,

$$\begin{aligned} &\mathbb{P}^N\{\omega \in \Delta^N : J^\star_{epi} \leq \gamma_N(\omega) + \min_{x \in \phi_N(\omega)} \mathcal{L}_\delta(x) \varphi^{-1}(\epsilon)\} \\ &\qquad \geq 1 - \Psi(\epsilon; d+1, N). \end{aligned} \tag{38}$$

Proof: (i) Given any $N \in \mathbb{Z}^+$, let $\gamma^*_N := \sup_{\omega \in \Delta^N} \gamma_N(\omega)$. It follows from the arguments of the proof of Lemma 2 that $\gamma^*_{d+1} = J^\star_{epi}$. For any $\omega := (\delta_1, \delta_2, \cdots, \delta_{d+1}) \in \Delta^{d+1}$ and $\omega' := (\delta'_1, \delta'_2, \cdots, \delta'_{d+1}) \in \Delta^{d+1}$, there exist $x \in \phi_N(\omega)$ and $x' \in \phi_N(\omega)$ such that $\gamma_{d+1}(\omega) = \max_i f(x, \delta_i)$ and $\gamma_{d+1}(\omega') = \max_i f(x', \delta'_i)$. From the Lipschitz condition (18) in Assumption 5, $\max_i f(x, \delta_i) - \max_i f(x', \delta'_i) \geq \max_i f(x, \delta_i) - \max_i f(x, \delta'_i) \geq -L_\delta \max_i \|\delta_i - \delta'_i\|$. Similarly, it holds that $\max_i f(x, \delta_i) - \max_i f(x', \delta'_i) \leq L_\delta \max_i \|\delta_i - \delta'_i\|$, which means that $|\gamma_{d+1}(\omega) - \gamma_{d+1}(\omega')| \leq L_\delta \max_{i \in \{1,2,\cdots,d+1\}} \|\delta_i - \delta'_i\|$. This implies Lipschitz continuity of $\gamma_{d+1}(\cdot)$ in Δ^{d+1}. Thus, according to Weierstrass's theorem, there exists $\omega^* := (\delta^*_1, \delta^*_2, \cdots, \delta^*_{d+1})$ such that $\gamma_{d+1}(\omega^*) = J^\star_{epi}$. The rest follows from the arguments in the proof of Theorem 2.
(ii) From (i), there exists $x^* \in \phi_N(\omega^*)$ such that $J^\star_{epi} = \max_{i=1,2,\cdots,d+1} f(x^*, \delta^*_i)$. Note that, for any $\omega := (\delta_1, \delta_2, \cdots, \delta_N) \in \Delta^N$, the set $\phi_N(\omega)$ is compact as $f : X \times \Delta \to \mathbb{R}$ is l.s.c in $X \times \Delta$. Given any $x \in \phi_N(\omega)$, we know that $J^\star_{epi} = \max_{i=1,2,\cdots,d+1} f(x^*, \delta^*_i) \leq \max_{i=1,2,\cdots,d+1} f(x, \delta^*_i) \leq \gamma_N(\omega) + \mathcal{L}_\delta(x) \max_{i=1,2,\cdots,d+1} \|\delta^*_i\|_{\bar{\omega}}$ where the first inequality is from optimality, the second inequality is due to the condition (37), and $\bar{\omega} := \{\delta_1, \delta_2, \cdots, \delta_N\}$. As x is arbitrary, $J^\star_{epi} \leq \gamma_N(\omega) + \min_{z \in \phi_N(\omega)} \mathcal{L}_\delta(x) \max_{i=1,2,\cdots,d+1} \|\delta^*_i\|_{\bar{\omega}}$. Again, the rest follows from the arguments in the proof of Theorem 2. □

Remark 7 *The two-step method (as summarized in Figure 1) can be also applied to min-max optimization. In general, under the same conditions as in Theorem 3, the corresponding bound can be given as,* $\forall \epsilon \in (0, 1)$,

$$\begin{aligned} &\mathbb{P}^N\{\omega \in \Delta^N : J^\star_{epi} \leq \gamma_N(\omega) + L_\delta \varphi^{-1}(\epsilon)\} \\ &\qquad \geq 1 - \Phi_c(\epsilon; d+1, N). \end{aligned} \tag{39}$$

5.2 Extension to non-convex programs

In this section, we show that our results can be extended to a special class of non-convex programs. In fact, the proposed method is applicable to non-convex programs in which the complexity (see Definition 1) or its upper bound is known a priori with proper continuity and regularity conditions. More precisely, consider the scenario program (2) and g_N^* as defined in (11) for any $N \in \mathbb{Z}^+$, suppose there exists $\bar{k}$ such that $g_{\bar{k}}^* = J^\star$, with proper conditions on the constraint and uncertain set in (1) such that $\alpha^*(\varepsilon)$ in (15) is well defined, Proposition 2 is still applicable for any $N \geq \bar{k}$.

In particular, we consider a special type of non-convex problems called quasi-linear programs as defined in [26]. Such problems are motivated by data-driven stability analysis of switched linear systems [24, 27].

Given two proper functions $v : \Delta \to \mathbb{R}^d$ and $h : \Delta \to \mathbb{R}^d$, we define the following robust optimization problem

$$J_{quasi}^\star := \inf_{x \in X, \lambda \geq 0} \lambda \tag{40a}$$

$$\text{s.t. } v(\delta)^\top x \leq \lambda h(\delta)^\top x, \forall \delta \in \Delta. \tag{40b}$$

For any $N \in \mathbb{Z}^+$ and any $\omega := (\delta_1, \delta_2, \cdots, \delta_N) \in \Delta^N$, we define the following problem with a finite number of constraints

$$\lambda_N(\omega) := \min_{x \in X, \lambda \geq 0} \lambda \tag{41a}$$

$$\text{s.t. } v(\delta_i)^\top x \leq \lambda h(\delta_i)^\top x, \forall i \in \{1, 2, \cdots, N\}. \tag{41b}$$

Let $\phi_N(\omega)$ denote the optimal set of the variable x for the given $\omega \in \Delta^N$. Similar to (11), we also define the following max-min formulation for (40)

$$\lambda_N^* := \sup_{z \in \Delta^N} \lambda_N(z), \forall N \in \mathbb{Z}^+. \tag{42}$$

An a-priori bound on the *complexity* of quasi-linear programs is provided in the following lemma.

Lemma 6 *Consider the problem* (40) *with the optimal objective value* $J_{quasi}^\star$ *and assume* X *is convex and compact. Suppose that* $v : \Delta \to \mathbb{R}^d$ *and* $h : \Delta \to \mathbb{R}^d$ *are continuous in* Δ *and such that* $h(\delta)^\top x > 0$ *for any* $x \in X$ *and* $\delta \in \Delta$*. Let* λ_N^* *be defined as in (42) for any* $N \in \mathbb{Z}^+$*. Then,* $\lambda_N^* = J_{quasi}^\star$ *for any* $N \geq d$*.*

Proof: The proof is given in Appendix D. □

Let us point out that, while the result above is similar to Theorem 2 in [26], there are two fundamental differences. First, Theorem 2 in [26] is only proved for scenario programs with a finite number of constraints while Lemma 6 is valid for the robust optimization problem in (40) with an infinite number of constraints. Second, Theorem 2 in [26] require the uniqueness of the solution while this condition is not needed in Lemma 6.

With the knowledge of the Lipschitz constant of $\lambda_d(\cdot)$ and some regularity condition on Δ (as mentioned in Assumption 6), we are able to compute explicit probabilistic bounds on $J_{quasi}^\star$ based on the scenario program (41) following the reasoning in Section 4. In the rest of this section, we provide the details of the derivation.

The following Lipschitz continuity condition is needed.

Assumption 7 *The functions* $v : \Delta \to \mathbb{R}^d$ *and* $h : \Delta \to \mathbb{R}^d$ *are Lipschitz continuous in* Δ *with constant* L_v *and* L_h *respectively i.e.,* $\forall \delta_1, \delta_2 \in \Delta$,

$$\|v(\delta_1) - v(\delta_2)\| \leq L_v \|\delta_1 - \delta_2\|, \tag{43}$$

$$\|h(\delta_1) - h(\delta_2)\| \leq L_h \|\delta_1 - \delta_2\|. \tag{44}$$

In addition, there exists a constant $\sigma > 0$ *such that,*

$$\forall (x, \delta) \in X \times \Delta, h(\delta)^\top x \geq \sigma. \tag{45}$$

With these assumptions, a continuity property of the optimal value function (41) is provided in the following lemma.

Lemma 7 *For any $N \in \mathbb{Z}^+$, let $\lambda_N(\omega)$ be defined as in (41) for any $\omega \in \Delta^N$. Suppose Assumption 7 hold and $X \subset \mathbb{R}^d$ is convex and compact. Then, for any $\omega := (\delta_1, \delta_2, \cdots, \delta_N) \in \Delta^N$ and $\omega' := (\delta'_1, \delta'_2, \cdots, \delta'_N) \in \Delta^N$,*

$$\begin{aligned}\lambda_N(\omega') \leq & \lambda_N(\omega) \\ & + \frac{L_v + \lambda_N(\omega) L_h}{\sigma} \|x\| \max_{i \in \{1,2,\cdots,N\}} \|\delta_i - \delta'_i\|\end{aligned} \tag{46}$$

where x is any element in the optimal set $\phi_N(\omega) := \{x \in X : v(\delta_i)^\top x \leq \lambda_N(\omega) h(\delta_i)^\top x, i = 1, 2, \cdots, N\}$ and σ is given in Assumption 7.

Proof: For any $x \in \phi_N(\omega)$, we have that, $\forall i, v(\delta_i)^\top x \leq \lambda_N(\omega) h(\delta_i)^\top x$, which implies that $v(\delta'_i)^\top x - cL_v\|x\| \leq \lambda_N(\omega) h(\delta'_i)^\top x + \lambda_N(\omega) cL_h\|x\|$, where $c = \max_j \|\delta_j - \delta'_j\|$. From (45), the inequality above implies that, $\forall i$,

$$v(\delta'_i)^\top x \leq \left(\lambda_N(\omega) + \frac{cL_v\|x\| + \lambda_N(\omega) cL_h\|x\|}{\sigma}\right) h(\delta'_i)^\top x.$$

Thus, by optimality, we arrive at (46). $\square$

Using the continuity property in Lemma 7, we derive a probabilistic objective value function bound for quasi-linear programs as defined in (40).

Theorem 4 *Consider Problem (40) with the optimal value $J^\star_{quasi}$. Suppose X is convex and compact, Δ is compact and Assumptions 6 & 7 hold. For any $N \in \mathbb{Z}^+$ with $N \geq d$, let $\lambda_N(\omega)$ be defined as in (41) for any $\omega \in \Delta^N$. Let ω be N i.i.d. samples drawn according to the probability measure $\mathbb{P}$. Then, for any $\epsilon \in (0, 1)$,*

$$\begin{aligned}\mathbb{P}^N\{\omega \in \Delta^N : & J^\star_{quasi} \leq \lambda_N(\omega) \\ & + \frac{L_v + \lambda_N(\omega) L_h}{\sigma} \min_{x \in \phi_N(\omega)} \|x\| \varphi^{-1}(\epsilon)\} \\ \geq & \ 1 - \Psi(\epsilon; d, N)\end{aligned} \tag{47}$$

where $\Psi(\epsilon; d, N)$ is given in (26) and $\phi_N(\omega) := \{x \in X : v(\delta_i)^\top x \leq \lambda_N(\omega) h(\delta_i)^\top x, i = 1, 2, \cdots, N\}$.

Proof: We first show that, for the case when $N = d$, there exists L_λ such that, for any $\omega := (\delta_1, \delta_2, \cdots, \delta_d) \in \Delta^d$ and $\omega' := (\delta'_1, \delta'_2, \cdots, \delta'_d) \in \Delta^d$,

$$|\lambda_d(\omega) - \lambda_d(\omega')| \leq L_\lambda \max_{i \in \{1,2,\cdots,d\}} \|\delta_i - \delta'_i\|. \tag{48}$$

Since X is compact, there exists r such that $X \subseteq r\mathbb{B}_d$. Under Assumption 7, $v(\delta)^\top x$ is always bounded for any $(x, \delta) \in X \times \Delta$. Thus, there exists $\bar{\lambda}$ such that $\lambda_d(\omega) \leq \bar{\lambda}$ for any $\omega \in \Delta^d$. Thus, from Lemma 7, (48) is satisifed with $L_\lambda = \frac{L_v + \bar{\lambda} L_h}{\sigma} r$. This implies Lipschitz continuity of $\lambda_d(\cdot)$ in Δ^d. Hence, according to Weierstrass's theorem, there exists some $\omega^* := (\delta^*_1, \delta^*_2, \cdots, \delta^*_d)$ in Δ^d such that $\lambda^*_d = \lambda_d(\omega^*)$. From Lemmas 6 & 7, we then conclude that, for any $\omega \in \Delta^N$ and any $x \in \mathcal{X}(\omega)$,

$$J^\star_{quasi} \leq \lambda_N(\omega) + \frac{L_v + \lambda_N(\omega) L_h}{\sigma} \|x\| \|\bar{\omega}^*\|_{\bar{\omega}}. \tag{49}$$

where $\bar{\omega}^* = \{\delta^*_1, \delta^*_2, \cdots, \delta^*_N\}$ and $\bar{\omega} = \{\delta_1, \delta_2, \cdots, \delta_N\}$ represent the unordered sets of ω^* and ω respectively. Note that $\phi_N(\omega)$ is closed. It also holds that, for any $\omega := (\delta_1, \delta_2, \cdots, \delta_N) \in \Delta^N$,

$$J^\star_{quasi} \leq \lambda_N(\omega) + \frac{L_v + \lambda_N(\omega) L_h}{\sigma} \min_{x \in \phi_N(\omega)} \|x\| \|\bar{\omega}^*\|_{\bar{\omega}}. \tag{50}$$

Following the same reasoning as in the proof of Theorem 2, we then obtain the result above. $\square$

Theorem 4 provides an a-posteriori objective gap bound because the deviation to the actual optimal value $J^\star_{quasi}$ depends on the sample ω. From the compactness of X and Δ, we can immediately derive an a-priori objective gap bound as follows.

Corollary 3 *Suppose that the same conditions in Theorem 4 hold. Let D_x and D_v be such that $\|x\| \leq D_x$ for any $x \in X$ and $\|v(\delta)\| \leq D_v$ respectively. Then, for any $\epsilon \in (0,1)$,*

$$\begin{aligned}\mathbb{P}^N\{\omega \in \Delta^N : J^\star_{quasi} \leq & \lambda_N(\omega) \\ & + \frac{L_v\sigma + D_v D_x L_h}{\sigma} D_x \phi^{-1}(\epsilon)\} \\ \geq & 1 - \Psi(\epsilon; d, N) \end{aligned} \tag{51}$$

where $\Psi(\epsilon; d, N)$ is given in (26)*.*

Proof: With the bounds D_x and D_v, it can be seen that $\lambda_N(\omega) \leq \frac{D_v D_x}{\sigma}$ for any $\omega \in \Delta^N$. This implies that $\frac{L_v + \lambda_N(\omega) L_h}{\sigma} \min_{x \in \mathcal{X}(\omega)} \|x\| \leq \frac{L_v\sigma + D_v D_x L_h}{\sigma} D_x$ for any $\omega \in \Delta^N$. With this, the result above is a direct consequence of Theorem 4. □

6 Numerical examples

In this section, we apply our method to two numerical examples: One is an illustrative example where the exact objective function value bound is available; the other is a canonical robust optimization problem in planar antenna synthesis.

6.1 An illustrative example

We first present an example to illustrate the conservatism induced by the chance-constrained theorems in [13, 14]. Consider the following robust optimization problem

$$\begin{aligned} & \min_{x_1, x_2} x_2 \\ \text{s.t.} \quad & |x_1 + D(\delta)| \leq x_2, \forall \delta \in \Delta = \{\delta \in \mathbb{R} : |\delta| \leq 1\}, \\ & |x_1| \leq 1, \end{aligned}$$

where

$$D(\delta) := \begin{cases} \delta, & -0.5 \leq \delta \leq 0.5 \\ 1 - \delta, & 0.5 < \delta \leq 1 \\ -1 - \delta, & -1 \leq \delta < -0.5. \end{cases}$$

The optimum is $J^\star = 0.5$ with the optimizer being $x_1^\star = 0$ and $x_2^\star = 0.5$. Given N i.i.d. samples $\omega := (\delta_1, \delta_2, \cdots, \delta_N)$ for some $N \in \mathbb{Z}^+$, drawn from the uniform distribution $\mathbb{P}$ on Δ, let

$$\underline{\eta}(\omega) := \min_{i=1,2,\cdots,N} D(\delta_i), \quad \overline{\eta}(\omega) := \max_{i=1,2,\cdots,N} D(\delta_i).$$

With these definitions, the optimizer of the sampled problem SCP is

$$x_1^*(\omega) = -\frac{\overline{\eta}(\omega) + \underline{\eta}(\omega)}{2}, x_2^*(\omega) = \frac{\overline{\eta}(\omega) - \underline{\eta}(\omega)}{2}.$$

The tail probability $p_N(\cdot)$ defined as in (12) becomes

$$p_N(\alpha) = \mathbb{P}^N\{\omega \in \Delta^N : J^\star - \alpha \leq x_2^*(\omega)\}$$

for any $\alpha \in [0, 0.5]$. It can be verified that, for any $\alpha \in [0, 0.5]$,

$$p_N(\alpha) = \mathcal{T}(\alpha) := 1 - (1 - 2\alpha)^N - 2\alpha N (1 - 2\alpha)^{N-1}.$$

The details can be found in the appendix. Thus, the optimal ULB of p_N is given as

$$\alpha^*(\varepsilon) := \sup_{\alpha \geq 0}\{\alpha : p_N(\alpha) \leq \varepsilon\} = \mathcal{T}^{-1}(\varepsilon), \forall \varepsilon \in (0,1).$$

Note that this example is exactly in the form of (34) and that Assumption 5 is satisfied with $L_\delta = 1$. Thus, from Theorem 3, a probabilistic objective value function bound is given as, for any $\epsilon \in [0, 0.5]$,

$$\mathbb{P}^N\{\omega \in \Delta^N : J^\star \leq x_2^*(\omega) + \epsilon\} \geq 1 - \Phi(\epsilon; 2, N).$$

Finally, we conclude that, given the confidence level $\beta \in (0, 1)$

$$\begin{aligned}&\mathbb{P}^N\{\omega \in \Delta^N : J^\star - x_2^*(\omega) \leq \Phi^{-1}(\beta; 2, N)\}\\ \geq &\mathbb{P}^N\{\omega \in \Delta^N : J^\star - x_2^*(\omega) \leq \mathcal{T}^{-1}(1-\beta)\} \geq 1 - \beta.\end{aligned}$$

We now consider the two-step method as shown in Figure 1. From (39), we obtain the following bound: $\forall \epsilon \in (0, 1)$,

$$\mathbb{P}^N\{\omega \in \Delta^N : J^\star \leq x_2^*(\omega) + \epsilon\} \geq 1 - \Phi_c(\epsilon; d, N).$$

Hence, for any confidence level $\beta \in (0, 1)$,

$$\mathbb{P}^N\{\omega \in \Delta^N : J^\star - x_2^*(\omega) \leq \Phi_c^{-1}(\beta; 2, N)\} \geq 1 - \beta,$$

The comparison of the three objective function value bounds is conducted under two different settings. First, we consider the case where N is fixed at 100 and compute the objective function value bounds for different values of β. We then fix β at 0.01 and compute the objective function value bounds for different values of N. The comparison results are given in Figures 4 & 5. From these figures, we can see that our method provides considerable improvement, in particular when β is small or N is small.

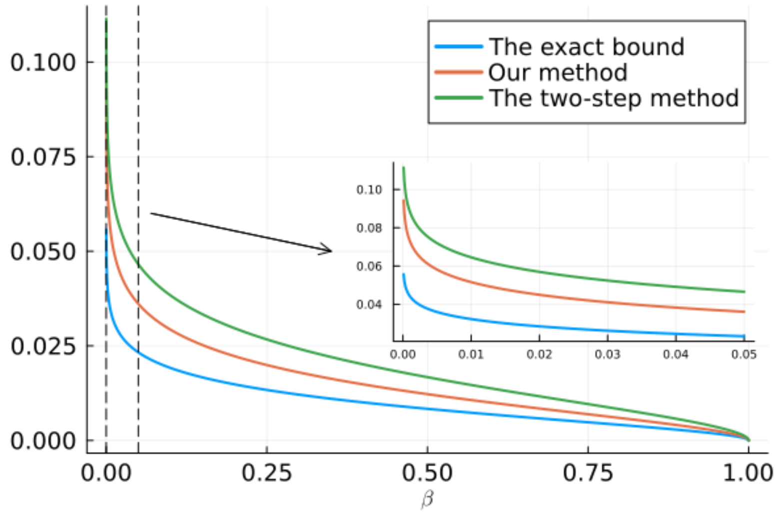


Figure 4: Comparison between our method and the two-step method from [19] when N is fixed at 100.

6.2 Planar antenna synthesis

We now apply the proposed analysis approach to a canonical robust optimization problem as described in Section 1.1 of [38], which is concerned with a monochromatic transmitting antenna placed at the origin. In particular, we consider a planar antenna comprised of a central circle and a number of concentric rings of the same areas as the circle in the XY-plane (which is considered as the Earth's surface). Let the wavelength be $\lambda = 0.5\,\mathrm{m}$, and the outer radius of the outer ring be 1 m.

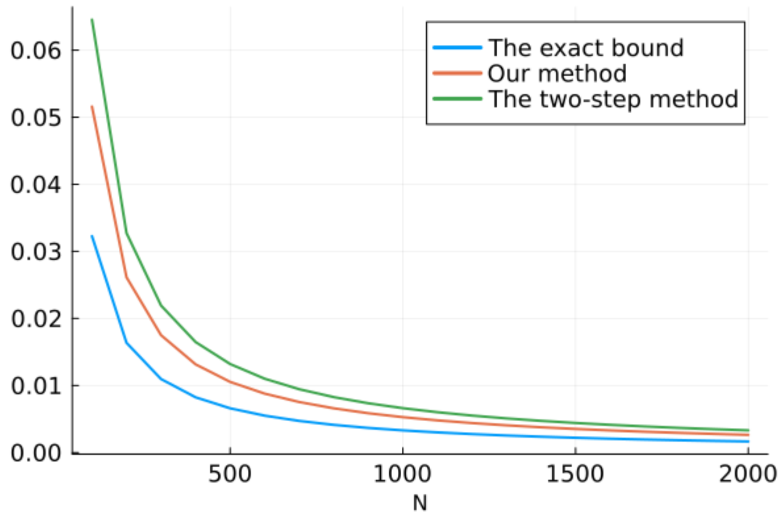


Figure 5: Comparison between our method and the two-step method from [19] when β is fixed at 0.01.

The directional distribution of energy sent by the antenna can be described in terms of the antenna's diagram which is a real-valued function $D(\theta)$ of the altitude (the angle between the direction of the XY-plane). Suppose that there are M rings. Let the diagram of the central circle be $D_1(\theta)$ and the diagrams of the rings be $D_2(\theta), D_3(\theta), \cdots, D_{M+1}(\theta)$. The overall diagram of the antenna is just the sum of the diagrams of all the elements. In this antenna design problem, the goal is to select the weights $\{x_i\}$ of the diagrams $\{D_i(\theta)\}$ (which in reality corresponds to modifying the output powers and shifting the phases of the elements) so that the resulting sum $\sum_{i=1}^{M+1} x_i D_i(\theta)$ is as close to the target diagram $D_*(\theta)$ as possible. As mentioned in Section 1.1 of [38], for any $i = 1, 2, \cdots, M+1$, the analytic expression of the diagram

$$D_i(\theta) = \frac{1}{2}\int_{r_{i-1}}^{r_i}\int_0^{2\pi} r\cos\left(2\pi r\lambda^{-1}\cos(\theta)\cos(\zeta)\right)d\zeta dr$$

where $\{r_i\}_{i=1}^M$ are the radii of the circles (see Figure 6 below for an example) with $r_0 = 0$ and $r_M = 1$. Let the target diagram be axial symmetric with respect to the Z-axis (which is the vertical direction of the XY-plane) and be "concentrated" in the cone $[\frac{5}{12}\pi, \frac{1}{2}\pi]$. Its expression is given as

$$D_*(\theta) = \begin{cases} 0, & \theta \in [0, \frac{5}{12}\pi) \\ \sin(6\theta - \frac{5}{2}\pi), & \theta \in [\frac{5}{12}\pi, \frac{1}{2}\pi] \end{cases}.$$

We formulate the following robust optimization problem to select the weights $\{x_i\}$:

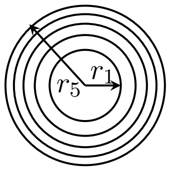
$$\begin{aligned} &\inf_{\gamma, x \in X} \gamma \\ \text{s.t.} \quad &|D_*(\theta) - \sum_{i=1}^{M+1} x_i D_i(\theta)| \le \gamma, \forall \theta \in \Delta := [0, \frac{1}{2}\pi] \end{aligned}$$

where $x := (x_1, x_2, \cdots, x_{M+1})$ and $X := \{x : \|x\|_\infty \le 10\}$(arising from physical constraints). The optimal solution is denoted by γ^*, which represents the minimal discrepancy between a synthesized diagram and the target one. For convenience, let $f(x, \theta) := |D_*(\theta) - \sum_{i=1}^{M+1} x_i D_i(\theta)|$. Given N i.i.d. samples $\omega := (\delta_1, \delta_2, \cdots, \delta_N)$ for some $N \in \mathbb{Z}^+$, drawn from the uniform distribution $\mathbb{P}$ on

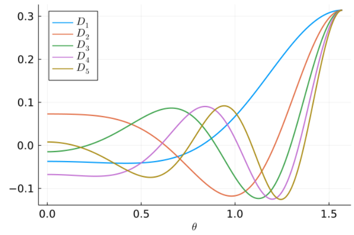
$[0, \frac{1}{2}\pi]$, the sampled problem becomes:

$$\begin{aligned} &\min_{\gamma, x \in X} \gamma \\ &\text{s.t.} \quad |D_*(\theta_j) - \sum_{i=1}^{M+1} x_i D_i(\theta_j)| \le \gamma, \forall j = 1, 2, \cdots, N. \end{aligned}$$

Let the solution be $\gamma_N(\omega)$ and the corresponding optimal set be $\phi_N(\omega)$. It is reported in Section 1.1 of [38] that the weights obtained from the sampled problem can be dramatically different from the optimal weights of the original robust optimization problem, which makes it difficult to establish the link between these two problems in terms of the weights. With the proposed method, we show below that a link can be established directly in terms of the objective function value.



(a) Array elements of equal areas.



(b) Diagrams of the elements.

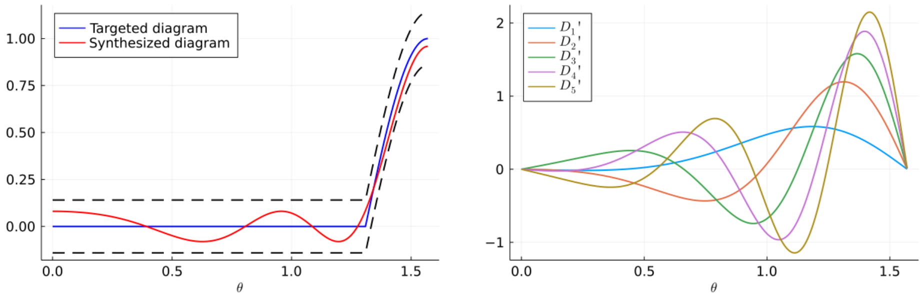
Figure 6: Antenna design with five array elements.

We consider the case where $M = 4$ and $N = 500$. The antenna elements and their diagrams are given in Figure 6. By solving the sampled problem above, we obtain the solution $\gamma(\omega) = 0.0806$ and the synthesized diagram, which is shown in Figure 7a. From Theorem 3, the exact worst-case error between the targeted and synthesized diagrams is bounded by $\gamma^* \le \gamma_N(\omega) + \min_{x \in \phi_N(\omega)} \mathcal{L}_\delta(x)\varphi^{-1}(\epsilon)$ with probability no smaller than $1 - \Psi(\epsilon; M+2, N)$, where $\varphi(\theta) = \frac{2\theta}{\pi}$ for any $\theta \in [0, \frac{1}{2}\pi]$ and $\mathcal{L}_\delta(x)$ is a Lipschitz constant satisfying $|f(x, \theta_1) - f(x, \theta_2)| \le \mathcal{L}_\delta(x)\|\theta_1 - \theta_2\|, \forall \theta_1, \theta_2 \in [0, \frac{1}{2}\pi]$. One choice of $\mathcal{L}_\delta$ can be $\mathcal{L}_\delta(x) = \max_{\theta \in [0, \frac{1}{2}\pi]} |\sum_{i=1}^{5} x_i D_i'(\theta) - \tilde{D}_*(\theta)|$, where D_i' denotes the derivative of D_i (as depicted in Figure 7b) and

$$\tilde{D}_*(\theta) = \begin{cases} 0, & \theta \in [0, \frac{5}{12}\pi) \\ 6\cos(6\theta - \frac{5}{2}\pi), & \theta \in [\frac{5}{12}\pi, \frac{1}{2}\pi] \end{cases}.$$

As the optimal solution set $\phi_N(\omega)$ is not necessarily a singleton, a reasonable choice is that we set the weights to be the $x \in \phi_N(\omega)$ of the smallest norm. Thus, our error bound is $\text{ERR}(\omega) := \gamma_N(\omega) + \mathcal{L}_\delta(\mathcal{X}(\omega))\varphi^{-1}(\epsilon)$, where $\mathcal{X}(\omega) := \arg\min_{x \in \phi_N(\omega)} \|x\|$. After these computations, we set the confidence level to be $\beta = 0.01$. By bisection, we obtain the value of ϵ such that $\Psi(\epsilon; M+2, N) = \beta$. Finally, we obtain that $\text{ERR}(\omega) = 0.141$.

We also investigate the impact of the number of array elements on the error bound. From a function approximation perspective, the diagrams of the array elements can be considered as the basis functions and the number of array elements is the complexity. Since the value of γ^* tends to decrease as M increases, a larger M leads to a better approximation. However, this is not the case in the framework of scenario optimization, where the error bound can deteriorate when M increases as the number of decision variables increases. We compute the error bound $\text{ERR}(\omega)$ with different values of M for several realizations of ω. Our results are shown in Figure 8. From this figure, we can see that the minimal error bound is always reached at $M = 7$. This phenomenon corresponds to the overfitting issue in machine learning which often occurs when a model is getting complex. With the proposed method, we can also provide a mathematical interpretation of this observation. The error bound consists of two parts: the first part is the numerical solution from the sampled problem; the second part is the correction term to account for the uncertainty or randomness in the sampling. When the confidence level β is fixed, as M increases, the value of ϵ satisfying $\Psi(\epsilon; M+2, N) = \beta$ increases while $\gamma_N(\omega)$ decreases. This means there is a trade-off between the numerical solution of the sampled problem and the correction term.



(a) The target and synthesized diagrams: The dashed lines denote computed bound $\mathrm{ERR}(\omega)$ on the optimality gap.

(b) Derivatives of the diagrams of the elements

Figure 7: Synthesis of antenna array with error bounds

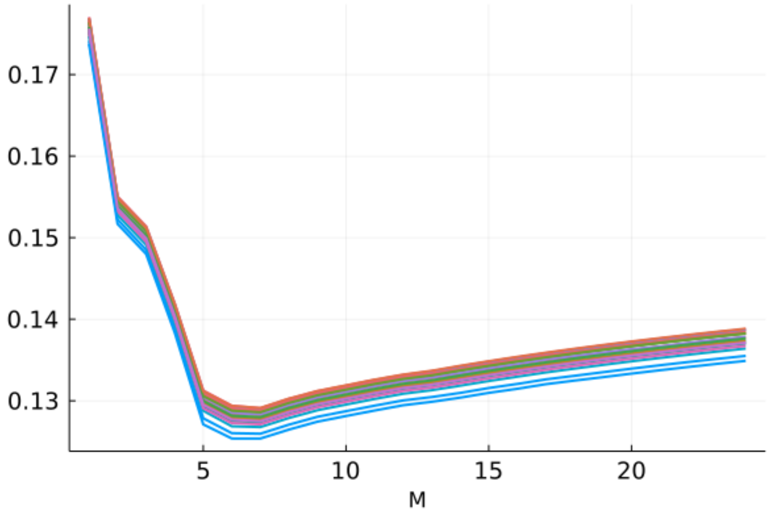


Figure 8: Error bounds with different values of M for 20 realizations: Overfitting is automatically detected when $M \geq 7$ in all the 20 experiments.

7 Conclusions

We have presented a novel method for deriving probabilistic guarantees on the objective function value for the scenario approach for robust optimization problems via sensitivity analysis. This method is based in particular on the concept of *complexity* of a robust optimization problem, which is the minimal number of constraints that define the actual solution. We show that the *complexity* of robust convex programs is bounded by the dimension of the decision variable by the use of a max-min reformulation and thereby extend classical results, which were known for scenario programs but not the original robust optimization problem, which possibly contains an infinite number of constraints. With this observation and additional regularity conditions, we are able to build probabilistic objective value bounds for the scenario approach. Our results do not rely on the chance-constrained theorems in the literature as an intermediate step, which not only enables us to release some strong technical assumptions for ensuring these chance-constrained theorems, but also to improve the value of these bounds. We illustrate these improvements using two numerical examples, among which is a practical application in planar antenna synthesis.

A Proof of Lemma 1

We prove the lemma by using standard results on parametric optimization. With Assumptions 1& 2, from Theorem 4.2.1 in [31], to show the l.s.c. property of $g_N(\omega)$, we only need to show that the point-to-set mapping $F_N : \Delta^N \to 2^X$, as defined in (8), is closed at any $\omega := (\delta_1, \delta_2, \cdots, \delta_N) \in \Delta^N$ in the sense that, for each pair of sequences $\{\omega^t = (\delta_1^t, \delta_2^t, \cdots, \delta_N^t), t \in \mathbb{Z}^+\} \subset \Delta^N$ and $\{x^t, t \in \mathbb{Z}^+\} \subset X$ with the properties $\omega^t \to \omega$, $x^t \in F_N(\omega^t)$, and $x^t \to x$, it follows that $x \in F_N(\omega)$. From the fact that $f : X \times \mathcal{M} \to \mathbb{R}$ is l.s.c. in $X \times \Delta$, $f(x, \delta_i) \le \liminf_{(x^t, \delta_i^t) \to (x, \delta_i)} f(x^t, \delta_i^t) \le 0$ for any $i = 1, 2, \cdots, N$, where the second inequality is due to the property that $x^t \in F_N(\omega^t)$. Hence, we conclude $x \in F_N(\omega)$. A similar proof can also be derived from Lemma 5.3 in [33].

B Proof of Proposition 1

(i) From the continuity of $f(\cdot, \cdot)$ and the compactness of X, we know that $g_N(\omega)$ is l.s.c. at any $\omega \in \Delta^N$ according to Lemma 1. We now show that $g_N(\omega)$ is also upper semi-continuous (u.s.c.) when $F_N(\omega) \subseteq \mathrm{cl}(F_N^o(\omega))$ is satisfied. Consider any ω satisfying $F_N(\omega) \subseteq \mathrm{cl}(F_N^o(\omega))$. From Theorem 3.1.5 in [31], we know that $F_N : \Delta^N \to 2^X$ is l.s.c. according to Berge (see Section 2.2 in [31] for the definition) at ω. Thus, from Theorem 4.2.2 in [31], g_N is also u.s.c. at ω. Hence, we prove the continuity.
(ii) Following the same argument above, we can show that g_N is l.s.c. at any $\omega \in \Delta^N$. The u.s.c property follows from Lemma 5.4 in [33]. This proves the continuity.

C Proof of Lemma 5

The proof consists of the following three steps:
Step I: We first show that, for any $\epsilon \in [0, 1]$, $\mathbb{P}^N\{\bar{\omega} \in \Delta^N : B_i \cap \bar{\omega} \neq \emptyset, \forall i\} \ge 1 - \Phi_a(\epsilon; k, N)$. For any i, let us define $\overline{S}_i := \{\omega \in \Delta^N : B_i \cap \bar{\omega} = \emptyset\}$, i.e., the set of ω such that none of the points in ω fails into B_i. As $\mathbb{P}\{B_i\} = \epsilon$, we know that $\mathbb{P}^N\{\overline{S}_i\} = (1-\epsilon)^N$ for any i. Hence, $\mathbb{P}^N\{\cup_i^k \overline{S}_i\} \le k(1-\epsilon)^N$. Note that $\cup_i^k \overline{S}_i$ is the set ω such that there exists at least i such that $B_i \cap \bar{\omega} = \emptyset$. Hence, $\mathbb{P}^N\{\bar{\omega} \in \Delta^N : B_i \cap \bar{\omega} \neq \emptyset, \forall i\} = 1 - \mathbb{P}^N\{\cup_i^k \overline{S}_i\} \ge 1 - k(1-\epsilon)^N = 1 - \Phi_a(\epsilon; k, N)$.
Step II: We then show that, for any $\epsilon \in [0, 1/k]$, $\mathbb{P}^N\{\bar{\omega} \in \Delta^N : B_i \cap \bar{\omega} \neq \emptyset, \forall i\} \ge 1 - \Phi(\epsilon; k, N)$ and $\Phi(\epsilon; k, N) \le \Phi_a(\epsilon; k, N)$. Let us first consider the case where $\{B_1, B_2, \cdots, B_k\}$ are disjoint. We prove by induction that $\mathbb{P}^N\{\omega \in \Delta^N : B_i \cap \bar{\omega} \neq \emptyset, \forall i\} = 1 - \Phi(\epsilon; k, N)$. For convenience, let $\Omega := \{\omega \in \Delta^N : B_i \cap \bar{\omega} \neq \emptyset, \forall i\}$. Note that the set Ω can be expressed as $\Omega = \cap_{i=1}^k S_i$, where $S_i := \{\omega \in \Delta^N : B_i \cap \bar{\omega} \neq \emptyset\}$. Hence,

$$
\begin{aligned}
\mathbb{P}^N\{\Omega\} &= \mathbb{P}^N\{\cap_{i=1}^{k-1} S_i \cap S_k\} \\
&= \mathbb{P}^N\{\cap_{i=1}^{k-1} S_i\} + \mathbb{P}^N\{S_k\} \\
&\quad - \mathbb{P}^N\{(\cap_{i=1}^{k-1} S_i) \cup S_k\} \\
&= \mathbb{P}^N\{\cap_{i=1}^{k-2} S_i\} + \mathbb{P}^N\{S_{k-1}\} + \mathbb{P}^N\{S_k\} \\
&\quad - \mathbb{P}^N\{(\cap_{i=1}^{k-2} S_i) \cup S_{k-1}\} \\
&\quad - \mathbb{P}^N\{(\cap_{i=1}^{k-1} S_i) \cup S_k\} \\
&\vdots \\
&= \sum_{i=1}^{k} \mathbb{P}^N\{S_i\} - \sum_{i=1}^{k-1} \mathbb{P}^N\{(\cap_{\ell=1}^{i} S_\ell) \cup S_{i+1}\}
\end{aligned} \tag{52}
$$

Continuing this process with $\{\mathbb{P}^N\{(\cap_{\ell=1}^i S_\ell) \cup S_{i+1}\}\}_{i=1}^{k-1}$, we will finally be able to write $\mathbb{P}^N\{\Omega\}$ in terms of $\{\mathbb{P}^N\{\cup_{\ell=i}^k S_\ell\}\}_{i=1}^k$, which allows to show that $\mathbb{P}^N\{\Omega\} = 1 - \Phi(\epsilon; k, N)$. Note that the measures of all the sets in the proof are independent of the order of $\{S_1, S_2, \cdots, S_k\}$. The details are given below. For any $i = 1, 2, \cdots, k$, the set $\cup_{\ell=i}^k S_\ell$ is in fact the set of ω such that there exists at least one point in $\cup_{\ell=i}^k B_\ell$, which means that $\mathbb{P}^N\{\cup_{\ell=i}^k S_\ell\} = 1 - (1 - (k-i+1)\epsilon)^N$ as

$\{B_1, B_2, \cdots, B_k\}$ are disjoint. For any two integers i, j with $1 \leq i < j \leq k$, we consider the set $\boldsymbol{S}_i^j := (\cap_{\ell=1}^i S_\ell) \cup (\cup_{\ell=j}^k S_\ell)$ and let $\phi(i,j) := \mathbb{P}^N\{\boldsymbol{S}_i^j\}$ with $\phi(1,j) = 1-(1-(2+k-j)\epsilon)^N$. Following the same reasoning, we obtain that, for all i, j with $1 < i < j \leq k$,

$$\begin{aligned}\mathbb{P}^N\{\boldsymbol{S}_i^j\} =& \mathbb{P}^N\{\boldsymbol{S}_{i-1}^j\} + \mathbb{P}^N\{S_i \cup (\cup_{\ell=j}^k S_\ell)\} \\ &- \mathbb{P}^N\{(\cap_{\ell=1}^{i-1} S_\ell) \cup S_i \cup (\cup_{\ell=j}^k S_\ell)\},\end{aligned}$$

which leads to the following recursive relation

$$\phi(i,j) = \phi(i-1,j) + \phi(1,j) - \phi(i-1,j-1). \tag{53}$$

After some manipulations, we can obtain that, for all $i = 1, 2, \cdots, k-1$,

$$\phi(i,k) = 1 - \sum_{\ell=1}^{i} (-1)^{\ell-1} \binom{i}{\ell} (1-(\ell+1)\epsilon)^N. \tag{54}$$

This, together with (52), yields

$$\begin{aligned}\mathbb{P}^N\{\Omega\} &= k\left(1-(1-\epsilon)^N\right) - \sum_{i=1}^{k-1} \phi(i,k) \\ &= 1 - k(1-\epsilon)^N \\ &\quad + \sum_{i=1}^{k-1} \sum_{\ell=1}^{i} (-1)^{\ell-1} \binom{i}{\ell} (1-(\ell+1)\epsilon)^N \\ &= 1 - k(1-\epsilon)^N \\ &\quad + \sum_{\ell=1}^{k-1} \left(\sum_{i=\ell}^{k-1} \binom{i}{\ell} \right) (-1)^{\ell-1} (1-(\ell+1)\epsilon)^N \\ &= 1 - k(1-\epsilon)^N \\ &\quad + \sum_{\ell=1}^{k-1} \binom{k}{\ell+1} (-1)^{\ell-1} (1-(\ell+1)\epsilon)^N \\ &= 1 - \binom{k}{1} (1-\epsilon)^N \\ &\quad - \sum_{\ell=2}^{k} \binom{k}{\ell} (-1)^{\ell-1} (1-\ell\epsilon)^N \\ &= 1 - \Phi(\epsilon; k, N),\end{aligned}$$

where the fourth equality is from the Hockey-stick identity. The case with disjoint sets $\{B_1, B_2, \cdots, B_k\}$ is in fact the worst case. In general cases where $\{B_1, B_2, \cdots, B_k\}$ are not necessarily disjoint, the probability that there exists at least one point in each set is bounded from below by $1 - \Phi(\epsilon; k, N)$. As $\mathbb{P}^N\{\omega \in \Delta^N : B_i \cap \bar{\omega} \neq \emptyset, \forall i\} = 1 - \Phi(\epsilon; k, N)$ when $\{B_1, B_2, \cdots, B_k\}$, from Step I, $\mathbb{P}^N\{\omega \in \Delta^N : B_i \cap \bar{\omega} \neq \emptyset, \forall i\} \geq 1 - k(1-\epsilon)^N$, which means that $\Phi(\epsilon; k, N) \leq \Phi_a(\epsilon; k, N) = k(1-\epsilon)^N$.

Step III: We finally show that, for any $\epsilon \in [0,1]$, $\mathbb{P}^N\{\omega \in \Delta^N : B_i \cap \bar{\omega} \neq \emptyset, \forall i\} \geq 1 - \Psi(\epsilon; k, N)$ and $\Psi(\epsilon; k, N) \leq \Phi_a(\epsilon; k, N)$, by combining Steps I & II above. When $k \leq k_\epsilon$, $\Psi(\epsilon; k, N) = \Phi(\epsilon; k, N)$ and hence this becomes **Step II**. We consider the case where $k > k_\epsilon$. Let us divide $\{B_1, B_2, \cdots, B_k\}$ into $\lfloor \frac{k}{k_\epsilon} \rfloor + 1$ groups (when $k (\mathrm{mod}\, k_\epsilon) \neq 0$): $\{B_1, \cdots, B_{k_\epsilon}\}, \{B_{1+k_\epsilon}, \cdots, B_{2k_\epsilon}\}, \cdots,$ $\{B_{1+(\lfloor \frac{k}{k_\epsilon} \rfloor - 1)k_\epsilon}, \cdots, B_{\lfloor \frac{k}{k_\epsilon} \rfloor k_\epsilon}\}, \{B_{1+\lfloor \frac{k}{k_\epsilon} \rfloor k_\epsilon}, \cdots, B_k\}$. We define the following sets: $\forall i = 1, \cdots, \lfloor \frac{k}{k_\epsilon} \rfloor$,

$$\mathcal{S}_i := \{\omega \in \Delta^N : B_j \cap \bar{\omega} \neq \emptyset, \forall j = (i-1)k_\epsilon + 1, \cdots, ik_\epsilon\}$$

with $\mathcal{S}_{\lfloor \frac{k}{k_\epsilon} \rfloor + 1} := \{\omega \in \Delta^N : B_j \cap \bar{\omega} \neq \emptyset, \forall j = ik_\epsilon + 1, \cdots, k\}$. From (ii), we know that $\mathbb{P}^N\{\mathcal{S}_i\} \geq 1 - \Phi(\epsilon; k_\epsilon, N), \forall i = 1, \cdots, \lfloor \frac{k}{k_\epsilon} \rfloor$ and $\mathbb{P}^N\{\mathcal{S}_{\lfloor \frac{k}{k_\epsilon} \rfloor + 1}\} \geq 1 - \Phi(\epsilon; k (\mathrm{mod}\, k_\epsilon), N)$. Following the same argument in (i), we conclude that $\mathbb{P}^N\{\cap_{i=1}^{\lfloor \frac{k}{k_\epsilon} \rfloor} \mathcal{S}_i \cap \mathcal{S}_{\lfloor \frac{k}{k_\epsilon} \rfloor + 1}\} \geq 1 - \lfloor \frac{k}{k_\epsilon} \rfloor \Phi(\epsilon; k_\epsilon, N) - \Phi(\epsilon; k (\mathrm{mod}\, k_\epsilon), N)$.

The same reasoning also holds when $k(\text{mod } k_\epsilon) = 0$. From the results in **Step II**, we can also see that, $\Psi(\epsilon; k, N) = \lfloor \frac{k}{k_\epsilon} \rfloor \Phi(\epsilon; k_\epsilon, N) + \Phi(\epsilon; k(\text{mod } k_\epsilon), N) \leq \lfloor \frac{k}{k_\epsilon} \rfloor k_\epsilon (1-\epsilon)^N + k(\text{mod } k_\epsilon)(1-\epsilon)^N = k(1-\epsilon)^N = \Phi_a(\epsilon; k, N)$. This completes the proof.

D Proof of Lemma 6

From the condition that $h(\delta)^\top x > 0$ for any $x \in X$ and $\delta \in \Delta$, we conclude that $\lambda_N(\omega)$ is bounded for any $\omega \in \Delta^N$. In fact, an upper bound can be explicitly given:

$$\lambda_N(z) \leq \max\{0, \frac{\max_{\delta \in \Delta, x \in X} v(\delta)^\top x}{\min_{\delta \in \Delta, x \in X} h(\delta)^\top x}\}, \forall z \in \Delta^N. \tag{55}$$

Hence, λ_N^* is bounded for any $N \in \mathbb{Z}^+$. With this, we prove that $\lambda_d^* = J_{quasi}^\star$. The proof is divided into two parts. We first show that $\gamma_{d+1}^* = J_{quasi}^\star$ following the same arguments in the proof of Lemma 2. From the definitions in (41) and (42), the optimality of λ_{d+1}^* implies that the set $\{x \in X : v(\delta_i)^\top x \leq \lambda_{d+1}^* h(\delta_i)^\top x, i = 1, 2, \cdots, d+1\}$ is nonempty for any $(\delta_1, \delta_2, \cdots, \delta_{d+1}) \in \Delta^{d+1}$. Again, as X is a compact convex set, from Helly's theorem (see, e.g., Theorem 21.6 in [34]), the set $\{x \in X : v(\delta)^\top x \leq \lambda_{d+1}^* h(\delta)^\top x, \delta \in \Delta\}$ is nonempty, implying that $J_{quasi}^\star \leq \lambda_{d+1}^*$. Hence, $J_{quasi}^\star = \lambda_{d+1}^*$ as it holds trivially that $\lambda_{d+1}^* \leq J_{quasi}^\star$. We then show that $\lambda_{d+1}^* = \lambda_d^*$. The second part can be considered as a simplified proof of Theorem 2 in [26] in the sense that we do not require the uniqueness of the solution. The proof goes by contradiction. Suppose $\lambda_{d+1}^* \neq \lambda_d^*$, i.e., $\lambda_d^* < \lambda_{d+1}^*$ as the sequence $\{\lambda_k^*\}_{k \in \mathbb{Z}^+}$ is non-decreasing. This means that there exists $\omega := (\delta_1, \delta_2, \cdots, \delta_{d+1}) \in \Delta^{d+1}$ such that $\lambda_{d+1}(\omega) > \lambda_d^* \geq 0$. For any $i \in \{1, 2, \cdots, d+1\}$, let ω_i denote the samples by removing δ_i from ω . Thus, $\lambda_{d+1}(\omega) > \lambda_d(\omega_i)$ for any $i \in \{1, 2, \cdots, d+1\}$. Let (λ^*, x^*) be a corresponding optimal solution of $\lambda_{d+1}(\omega)$, i.e., $\lambda^* = \lambda_{d+1}(\omega)$. As the optimal value of Problem (41) with $\omega = (\delta_1, \delta_2, \cdots, \delta_{d+1})$ decreases when any of the $d+1$ constraints is removed, we conclude that $(v(\delta_i) - \lambda^* h(\delta_i))^\top x^* = 0$, which implies that the vectors $\{z_i := v(\delta_i) - \lambda^* h(\delta_i)\}_{i=1}^{d+1}$ all lie in a $(d-1)$-dimensional subspace orthogonal to x^*. We then define the function $Z(x) = \max_i z_i^\top x$ and the convex optimization problem $\min_{x \in X} Z(x)$. The optimality of (λ^*, x^*) implies that x^* is an optimal solution to $\min_{x \in X} Z(x)$; otherwise, there exists (λ, x) such that $x \in X$, $v(\delta_i)^\top x \leq \lambda h(\delta_i)^\top x, i = 1, 2, \cdots, d+1$ and $\lambda < \lambda^*$, which contradicts the optimality of (λ^*, x^*). With this observation, from Theorem 27.4 in [34], we conclude that $0 \in \text{conv}(\{z_i\}_{i=1}^{d+1}) + \text{Nor}_{x^*}(X)$, where $\text{Nor}_{x^*}(X)$ denotes the normal cone of X at x^*, i.e., the set of vectors u such that $u^\top(x - x^*) \leq 0$ for any $x \in X$. From Carathéodory's theorem [39], there exists a subset $\mathcal{I} \subset \{1, 2, \cdots, d+1\}$ with $|\mathcal{I}| \leq d$ such that $0 \in \text{conv}(\{z_i\}_{i \in \mathcal{I}}) + \text{Nor}_{x^*}(X)$, which implies that there exists $i \in \{1, 2, \cdots, d+1\}$ such that $\lambda_{d+1}(\omega) = \lambda_d(\bar{\omega}_i)$, leading to a contradiction. Therefore, it holds that $\lambda_{d+1}^* = \lambda_d^*$. This completes the proof.

E Computational details in Section 6.1

Given any $\alpha \in [0, 0.5)$, for any $c \leq \min\{2\alpha, 1 - 2\alpha\}$, let us define: $\forall k = 0, 1, \cdots, \lfloor \frac{2\alpha}{c} \rfloor - 1$,

$$\begin{aligned} W_k := \{\omega_N \in \Delta^N : \frac{1}{2} - \alpha \leq \frac{\overline{\eta}(\omega_N) - \underline{\eta}(\omega_N)}{2}, \\ \frac{1}{2} - 2\alpha + kc \leq \overline{\eta}(\omega_N) \leq \frac{1}{2} - 2\alpha + (k+1)c\} \end{aligned}$$

with

$$\begin{aligned} W_{\lfloor \frac{2\alpha}{c} \rfloor} := \{\omega_N \in \Delta^N : \frac{1}{2} - \alpha \leq \frac{\overline{\eta}(\omega_N) - \underline{\eta}(\omega_N)}{2}, \\ \frac{1}{2} - 2\alpha + \lfloor \frac{2\alpha}{c} \rfloor c \leq \overline{\eta}(\omega_N) \leq \frac{1}{2}\}. \end{aligned}$$

By definition, $p_N(\alpha) = \sum_{k=0}^{\lfloor \frac{2\alpha}{c} \rfloor} \mathbb{P}^N\{W_k\}$. We now construct over- and under-approximations of $\{W_k\}$ as follows: $\forall k = 0, 1, \cdots, \lfloor \frac{2\alpha}{c} \rfloor - 1$,

$$\begin{aligned}
W_k^O &\coloneqq \{\omega_N \in \Delta^N : \underline{\eta}(\omega_N) \le -\frac{1}{2} + (k+1)c, \\
&\qquad \frac{1}{2} - 2\alpha + kc \le \overline{\eta}(\omega_N) \le \frac{1}{2} - 2\alpha + (k+1)c\}, \\
W_k^I &\coloneqq \{\omega_N \in \Delta^N : \underline{\eta}(\omega_N) \le -\frac{1}{2} + kc, \\
&\qquad \frac{1}{2} - 2\alpha + kc \le \overline{\eta}(\omega_N) \le \frac{1}{2} - 2\alpha + (k+1)c\}
\end{aligned}$$

with

$$\begin{aligned}
W_{\lfloor \frac{2\alpha}{c} \rfloor}^O &\coloneqq \{\omega_N \in \Delta^N : \underline{\eta}(\omega_N) \le -\frac{1}{2} + 2\alpha, \\
&\qquad \frac{1}{2} - 2\alpha + \lfloor \frac{2\alpha}{c} \rfloor c \le \overline{\eta}(\omega_N) \le \frac{1}{2}\}, \\
W_{\lfloor \frac{2\alpha}{c} \rfloor}^I &\coloneqq \{\omega_N \in \Delta^N : \underline{\eta}(\omega_N) \le -\frac{1}{2} + \lfloor \frac{2\alpha}{c} \rfloor c, \\
&\qquad \frac{1}{2} - 2\alpha + \lfloor \frac{2\alpha}{c} \rfloor c \le \overline{\eta}(\omega_N) \le \frac{1}{2}\}.
\end{aligned}$$

It can be verified that, for any $k = 0, 1, \cdots, \lfloor \frac{2\alpha}{c} \rfloor$, $W_k^I \subseteq W_k \subseteq W_k^O$, which means that

$$\sum_{k=0}^{\lfloor \frac{2\alpha}{c} \rfloor} \mathbb{P}^N\{W_k^I\} \le p_N(\alpha) \le \sum_{k=0}^{\lfloor \frac{2\alpha}{c} \rfloor} \mathbb{P}^N\{W_k^O\}. \tag{56}$$

We then compute $\{\mathbb{P}^N\{W_k^O\}\}$ and $\{\mathbb{P}^N\{W_k^I\}\}$. Let us first consider $\{\mathbb{P}^N\{W_k^O\}\}$. For any $k = 0, 1, \cdots, \lfloor \frac{2\alpha}{c} \rfloor$, we define the following sets: $\tilde{S}_k \coloneqq \{\delta \in \Delta : D(\delta) \le \frac{1}{2} - 2\alpha + (k+1)c\}$, $\overline{S}_k \coloneqq \{\delta \in \Delta : \frac{1}{2} - 2\alpha + kc \le D(\delta) \le \frac{1}{2} - 2\alpha + (k+1)c\}$ and $\underline{S}_k \coloneqq \{\delta \in \Delta : D(\delta) \le -\frac{1}{2} + (k+1)c\}$, as illustrated in Figure 9. By simple calculation, $\mathbb{P}\{\tilde{S}_k\} = 1 - 2\alpha + (k+1)c$, $\mathbb{P}\{\overline{S}_k\} = c$ and $\mathbb{P}\{\underline{S}_k\} = (k+1)c$. Hence, $W_k^O = \tilde{\mathcal{S}}_k \cap \overline{\mathcal{S}}_k \cap \underline{\mathcal{S}}_k$, where $\tilde{\mathcal{S}}_k \coloneqq \{\omega_N \in \Delta^N : \omega_N \subset \tilde{S}_k\}, \overline{\mathcal{S}}_k \coloneqq \{\omega_N \in \Delta^N : \omega_N \cap \overline{S}_k \neq \emptyset\}$ and $\underline{\mathcal{S}}_k \coloneqq \{\omega_N \in \Delta^N : \omega_N \cap \underline{S}_k \neq \emptyset\}$. For any $k = 0, 1, \cdots, \lfloor \frac{2\alpha}{c} \rfloor - 1$, we have that

$$\begin{aligned}
\mathbb{P}^N\{W_k^O\} &= \mathbb{P}^N\{\tilde{\mathcal{S}}_k \cap \overline{\mathcal{S}}_k\} + \mathbb{P}^N\{\tilde{\mathcal{S}}_k \cap \underline{\mathcal{S}}_k\} \\
&\quad - \mathbb{P}^N\{\tilde{\mathcal{S}}_k \cap (\overline{\mathcal{S}}_k \cup \underline{\mathcal{S}}_k)\} \\
&= \mathbb{P}\{\tilde{S}_k\}^N [1 - (1 - \frac{\mathbb{P}\{\overline{S}_k\}}{\mathbb{P}\{\tilde{S}_k\}})^N \\
&\quad + 1 - (1 - \frac{\mathbb{P}\{\underline{S}_k\}}{\mathbb{P}\{\underline{S}_k\}})^N \\
&\quad - \left(1 - \left(1 - \frac{\mathbb{P}\{\overline{S}_k\} + \mathbb{P}\{\underline{S}_k\}}{\mathbb{P}\{\underline{S}_k\}}\right)^N\right)] \\
&= (1 - 2\alpha + (k+1)c)^N - (1 - 2\alpha + kc)^N \\
&\quad - (1 - 2\alpha)^N + (1 - 2\alpha - c)^N.
\end{aligned}$$

When $k = \lfloor \frac{2\alpha}{c} \rfloor$, $\mathbb{P}^N\{W_k^O\} = 1-(1-(2\alpha-\lfloor\frac{2\alpha}{c}\rfloor c))^N-(1-2\alpha)^N+\big(1-(4\alpha-\lfloor\frac{2\alpha}{c}\rfloor c))\big)^N$. Combing the equalities above yields

$$\begin{aligned}\sum_{k=0}^{\lfloor\frac{2\alpha}{c}\rfloor} \mathbb{P}^N\{W_k^O\} =& (1-2\alpha+\lfloor\frac{2\alpha}{c}\rfloor c)^N - (1-2\alpha)^N \\ &+ \lfloor\frac{2\alpha}{c}\rfloor \left((1-2\alpha-c)^N - (1-2\alpha)^N\right) \\ &+ 1-(1-(2\alpha-\lfloor\frac{2\alpha}{c}\rfloor c))^N - (1-2\alpha)^N \\ &+ \left(1-(4\alpha-\lfloor\frac{2\alpha}{c}\rfloor c))\right)^N\end{aligned}$$

From the fact that $\lim_{c\to 0}\lfloor\frac{2\alpha}{c}\rfloor c = 2\alpha$, we obtain that

$$\lim_{c\to 0}\sum_{k=0}^{\lfloor\frac{2\alpha}{c}\rfloor} \mathbb{P}^N\{W_k^O\} = 1-(1-2\alpha)^N - 2\alpha N(1-2\alpha)^{N-1}.$$

Following the same reasoning, it also holds that

$$\lim_{c\to 0}\sum_{k=0}^{\lfloor\frac{2\alpha}{c}\rfloor} \mathbb{P}^N\{W_k^I\} = 1-(1-2\alpha)^N - 2\alpha N(1-2\alpha)^{N-1}.$$

Since (56) holds for any $c \leq \min\{2\alpha, 1-2\alpha\}$, we conclude that

$$p_N(\alpha) = 1-(1-2\alpha)^N - 2\alpha N(1-2\alpha)^{N-1}.$$

It is easy to see that $p_N(\alpha) = 1$ when $\alpha = 0.5$. Hence, $\mathcal{U}(\cdot)$ as defined above is the tail probability.

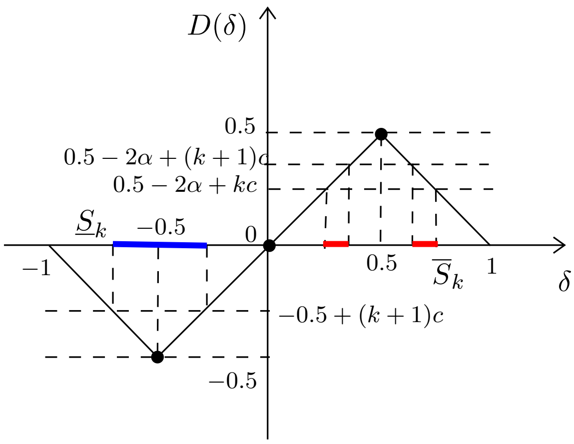


Figure 9: Illustration of $\overline{S}_k$ and$\underline{S}_k$.

References

[1] R. Tempo, G. Calafiore, and F. Dabbene. *Randomized algorithms for analysis and control of uncertain systems: with applications.* Springer, 2013.

[2] A. Prékopa. *Stochastic programming*, volume 324. Springer Science & Business Media, 1995.

[3] A. Ruszczyński and A. Shapiro. Stochastic programming models. *Handbooks in operations research and management science*, 10:1–64, 2003.

[4] J. R. Birge and F. Louveaux. *Introduction to stochastic programming.* Springer Science & Business Media, 2011.

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

[6] Aharon Ben-Tal and Arkadi Nemirovski. Selected topics in robust convex optimization. *Mathematical Programming*, 112(1):125–158, 2008.

[7] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust optimization.* Princeton university press, 2009.

[8] Dimitris Bertsimas, David B Brown, and Constantine Caramanis. Theory and applications of robust optimization. *SIAM review*, 53(3):464–501, 2011.

[9] G. Calafiore and M. C. Campi. Uncertain convex programs: randomized solutions and confidence levels. *Mathematical Programming*, 102(1):25–46, 2005.

[10] G. Calafiore and M. C. Campi. The scenario approach to robust control design. *IEEE Transactions on Automatic Control*, 51(5):742–753, 2006.

[11] M. C. Campi and S. Garatti. *Introduction to the scenario approach.* SIAM, 2018.

[12] M.C. Campi, A. Carè, and S. Garatti. The scenario approach: A tool at the service of data-driven decision making. *Annual Reviews in Control*, 52:1–17, 2021.

[13] M. C. Campi and S. Garatti. The exact feasibility of randomized solutions of uncertain convex programs. *SIAM Journal on Optimization*, 19(3):1211–1230, 2008.

[14] G. C. Calafiore. Random convex programs. *SIAM Journal on Optimization*, 20(6):3427–3464, 2010.

[15] M. C. Campi and S. Garatti. Wait-and-judge scenario optimization. *Mathematical Programming*, 167(1):155–189, 2018.

[16] S. Garatti and M.C. Campi. Risk and complexity in scenario optimization. *Mathematical Programming*, pages 1–37, 2019.

[17] M. C. Campi, S. Garatti, and F. A. Ramponi. A general scenario theory for nonconvex optimization and decision making. *IEEE Transactions on Automatic Control*, 63(12):4067–4078, 2018.

[18] S. Garatti and M. C. Campi. The risk of making decisions from data through the lens of the scenario approach. *IFAC-PapersOnLine*, 54(7):607–612, 2021.

[19] P. M. Esfahani, T. Sutter, and J. Lygeros. Performance bounds for the scenario approach and an extension to a class of non-convex programs. *IEEE Transactions on Automatic Control*, 60(1):46–58, 2014.

[20] J. Kapinski, J. V. Deshmukh, S. Sankaranarayanan, and N. Arechiga. Simulation-guided lyapunov analysis for hybrid dynamical systems. In *Proceedings of the 17th international conference on Hybrid systems: computation and control*, pages 133–142. ACM, 2014.

[21] A. Kozarev, J. Quindlen, J. How, and U. Topcu. Case studies in data-driven verification of dynamical systems. In *Proceedings of the 19th International Conference on Hybrid Systems: Computation and Control*, pages 81–86. ACM, 2016.

[22] N. M. Boffi, S. Tu, N. Matni, J.-J. E. Slotine, and V. Sindhwani. Learning stability certificates from data. *arXiv preprint arXiv:2008.05952*, 2020.

[23] G. Mamakoukas, O. Xherija, and T. Murphey. Memory-efficient learning of stable linear dynamical systems for prediction and control. *Advances in Neural Information Processing Systems*, 33:13527–13538, 2020.

[24] J. Kenanian, A. Balkan, R. M. Jungers, and P. Tabuada. Data driven stability analysis of black-box switched linear systems. *Automatica*, 109:108533, 2019.

[25] Z. Wang and R. M. Jungers. Scenario-based set invariance verification for black-box nonlinear systems. *IEEE Control Systems Letters*, 5(1):193–198, 2021.

[26] G. O. Berger, R. M. Jungers, and Z. Wang. Chance-constrained quasi-convex optimization with application to data-driven switched systems control. In *Learning for Dynamics and Control*, pages 571–583. PMLR, 2021.

[27] A. Rubbens, Z. Wang, and R. M. Jungers. Data-driven stability analysis of switched linear systems with sum of squares guarantees. In *The 7th IFAC Conference on Analysis and Design of Hybrid Systems*. IFAC, 2021.

[28] Z. Wang and R. M. Jungers. A data-driven method for computing polyhedral invariant sets of black-box switched linear systems. *IEEE Control Systems Letters*, 5(5):1843 – 1848, 2021.

[29] Z. Wang and R. M. Jungers. Probabilistic guarantees on the objective value for the scenario approach via sensitivity analysis. In *Proceedings of the 61st IEEE Conference on Decision and Control*, Cancún, Mexico, 2022.

[30] H. Djelassi, A. Mitsos, and O. Stein. Recent advances in nonconvex semi-infinite programming: Applications and algorithms. *EURO Journal on Computational Optimization*, 9:100006, 2021.

[31] B. Bank, J. Guddat, D. Klatte, B. Kummer, and K. Tammer. *Non-linear parametric optimization*. Birkhäuser Verlag, Basel, 1983.

[32] R. T. Rockafellar and R. J-B Wets. *Variational analysis*, volume 317. Springer Science & Business Media, 1998.

[33] G. Still. Lectures on parametric optimization: An introduction. *Optimization Online*, 2018.

[34] R. T. Rockafellar. *Convex analysis*. Princeton university press, 1970.

[35] T. Kanamori and A. Takeda. Worst-case violation of sampled convex programs for optimization with uncertainty. *Journal of Optimization Theory and Applications*, 152(1):171–197, 2012.

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

[37] Antonio Cuevas and Alberto Rodríguez-Casal. On boundary estimation. *Advances in Applied Probability*, 36(2):340–354, 2004.

[38] A. Nemirovski. Lectures on robust convex optimization. *Georgia Institute of Technology*, 2012.

[39] J. Eckhoff. Helly, radon, and carathéodory type theorems. In *Handbook of convex geometry*, pages 389–448. Elsevier, 1993.