

# On the probabilistic feasibility of solutions in multi-agent optimization problems under uncertainty

George Pantazis<sup>1</sup>, Filiberto Fele, Kostas Margellos

Department of Engineering Science, University of Oxford, OX1 3PJ, UK

---

## Abstract

We investigate the probabilistic feasibility of randomized solutions to two distinct classes of uncertain multi-agent optimization programs. We first assume that only the constraints of the program are affected by uncertainty, while the cost function is arbitrary. Leveraging recent *a posteriori* developments of the scenario approach, we provide probabilistic guarantees for all feasible solutions of the program under study. This result is particularly useful in cases where numerical difficulties related to the convergence of the solution-seeking algorithm hinder the exact quantification of the optimal solution. **Furthermore, it allows providing guarantees for the entire solution set of optimization programs with uncertain convex constraints and a (possibly) non-convex cost function.** We then focus on optimization programs where the cost function admits an aggregate representation and depends on uncertainty while constraints are deterministic. By exploiting the structure of the program under study and leveraging the so called support rank notion, we provide agent-independent robustness certificates for the optimal solution, i.e., the constructed bound on the probability of constraint violation does not depend on the number of agents, but only on the dimension of the agents' decision. This substantially reduces the number of samples required to achieve a certain level of probabilistic robustness as the number of agents increases. All robustness certificates provided in this paper are distribution-free and can be used alongside any optimization algorithm. Our theoretical results are accompanied by a numerical case study involving a charging control problem of a fleet of electric vehicles.

**Keywords:** Scenario approach, Multi-agent problems, Optimization, Feasibility guarantees, Electric vehicles

---

## 1. Introduction

### 1.1. Background

A vast amount of today's challenges in the domains of energy systems [1], [2], traffic networks [3], economics [4] and the social sciences [5], [6] revolve around multi-agent systems, i.e., systems which comprise different entities/agents that interact with each other and make decisions, based on individual or collective criteria. Existing literature provides a plethora of methods to solve such problems. Each method is appropriately designed to fit the structure of these interactions and the agents' incentives. To address challenges, related to the computational complexity issues and privacy concerns of solving a multi-agent optimization problem in a centralised fashion, several decentralised or distributed coordination schemes have been proposed [7],[8]. In the decentralised case, agents optimize their cost function locally and then communicate their strategies to a central authority. In the distributed case, a central authority is absent and agents communicate with each other over a network, exchanging information with agents considered as neighbours given the underlying communication protocol. In either case, the presence of uncertainty in such problems constitutes a critical factor that, if not taken into account, could

lead to unpredictable behaviour, hence it is of major importance to accompany the solutions of such algorithms with robustness certificates. In this paper we assume that the probability distribution of the uncertainty is considered unknown and adopt a data-driven approach, where the uncertainty is represented by means of scenarios that could be either available as historical data, or extracted via some prediction model. To this end, we work under the framework of the so called scenario approach.

The scenario approach is a well-established mathematical technique [9], [10], [11], and still a highly active research area (see [12], [13] for some recent developments), originally introduced to provide *a priori* probabilistic guarantees for solutions of uncertain convex optimization programs. Recently, the theory has been extended to non-convex decision making problems [12], [13] where the probabilistic guarantees are obtained in an *a posteriori* fashion. The main advantage of the scenario approach is its applicability under very general conditions, since it does not require the knowledge of the uncertainty set or the probability distribution, unlike robust [14] and stochastic optimization [15]. According to the scenario approach, the original problem can be approximated by solving a computationally tractable approximate problem, the so called scenario program consisting of a finite number of constraints, each of them corresponding to a different realization of the uncertain parameter. In this realm we present some of the challenges that pertain uncertain multi-agent systems.

---

Email addresses: georgios.pantazis@lmh.ox.ac.uk (George Pantazis), filiberto.fele@eng.ox.ac.uk (Filiberto Fele), kostas.margellos@eng.ox.ac.uk (Kostas Margellos)

<sup>1</sup>Corresponding author

Type of solution	Nature of certificate	Scalability	Class of problems	Result
<b>Entire feasible set</b>	<i>A posteriori</i>	<b>Agent dependent</b>	<b>Feasibility programs with uncertain convex constraints</b>	<b>Thm. 2</b>
Subset of feasible solutions	<i>A priori</i>	Agent dependent	Feasibility programs with uncertain convex constraints	Thm. 2 of [16]
Variational inequality solution set	<i>A posteriori</i>	Agent dependent	Variational inequality problems with uncertain convex constraints	Thm. 1 of [17]
Variational inequality solution set	<i>A posteriori</i>	Agent dependent	Aggregative games with uncertain affine constraints	Thm. 1 of [18]
<b>Unique optimizer</b>	<i>A priori</i>	<b>Agent independent</b>	<b>Optimization programs with uncertain aggregative term in the cost</b>	<b>Thm. 3</b>
Unique variational inequality solution	<i>A posteriori</i> & <i>a priori</i>	Agent dependent	Variational inequalities with uncertain cost or constraints	Cor. 1 of [19], Thm. 9 of [20], Thm. 5 of [21] (only <i>a posteriori</i> )

Table 1: Classification of main results according to their main features and comparison with existing literature

## 1.2. Challenges to be addressed and main contributions

In a plethora of problems of practical interest, agents' decisions are feasible, though not necessarily optimal. This can be due to numerical difficulties related to the convergence properties of the solution-seeking algorithm that hinder the exact quantification of the optimal solution. Even under the assumption that the computed optimizer is exact, many optimization programs can have multiple (possibly infinite) optima or local optima.

Another major challenge is the dependence of the provided certificates on the number of agents. Given that we wish to obtain identical probabilistic guarantees as the population increases, a larger number of samples is required. This fact renders the provided guarantees conservative for large scale applications. As such, it is of utmost importance to show that for certain classes of problems, common in practical applications, the obtained probabilistic guarantees can be agent independent. The main contributions of this paper are as follows:

1. We address the challenges related to the lack of optimality of the agents' decisions and the presence of multiple local solutions, by considering an optimization set-up, where the choice of the cost function can be arbitrary and the uncertainty affects only the constraints. We, then, leverage recent results of [13] in order to provide a *a posteriori* robustness certificates for the entire feasibility region<sup>2</sup>. The theoretical framework of the preliminary work in [22], initially developed only for uncertain feasibility

problems with polytopic constraints, is extended in our set-up to account for general (possibly coupling) uncertain convex constraints. This result is interesting per se, as it complements the results in [13] since it allows providing guarantees collectively for a set of points thus departing from the existing stream of literature in the scenario approach that typically refers to a single solution (e.g., see [9, 10, 11, 12, 13], [23]). Moreover, it accounts for problems with multiple optimizers, or for problems with non-convex objective functions, where a different optimal, locally optimal or even suboptimal solution might be returned by a solution algorithm, hence it is of importance to provide guarantees for all of them. We also discuss in Remark 1 that, as a byproduct of our approach, computing the number of support samples for the entire feasibility region (a notion at the core of the scenario approach developments that will be formally introduced in the sequel) is computationally less sensitive compared to the approach suggested in [13] to determine the support samples of a specific solution to the underlying optimization program.

2. We then focus on a specific class of uncertain multi-agent programs, prevalent in many practical applications, where the cost is considered to be a function of the aggregate decision and affected by uncertainty. A similar problem formulation has been considered in [24] and is extended to our set-up to account for the presence of uncertainty in the cost function. Other problems whose structure shares similarities with our work can be found in [25],[26], [27], though under a purely deterministic and game-theoretic set-up. Following the recent developments in [28] we show based on the notion of the support rank [29], that the obtained probabilistic feasibility certificates are agent independent. This result directly outperforms probabilistic feasibility statements obtained by a direct application of the scenario approach theory [10], and has superior scalability properties as it does not depend on the number of

<sup>2</sup>In this paper we will focus on optimization programs affected by uncertainty. It is important to mention, though, that our results are applicable to more general uncertain feasibility programs. This allows their use for providing certificates for the feasibility region of other classes of problems, as well, such as variational inequality problems and generalised Nash equilibrium problems [22].

agents involved.

The contributions of our main results in comparison with results in the literature are summarised in Table I. Our first contribution provides probabilistic guarantees, in an *a posteriori* fashion, that hold for the entire feasibility region (or a subset of interest as detailed in Remark 2) in contrast with the *a priori* result in [16]. Therein, the feasible subset is obtained by taking the convex hull of randomized optimizers whose position is in general not known. This may in certain cases produce a "thin" subset of the feasible region. The recent contributions [18], [17] use a similar approach to [22] to provide guarantees applicable not for feasible sets, but for a set of randomized game equilibria. For our second contribution, the nature of the certificates is *a priori*. Assuming uniqueness of the solution, we exploit the aggregative structure of the cost to provide results that are agent independent. The aggregative nature of the cost has been exploited in several works (see [18], [20] and [21]), however, guarantees provided in these works are dependent on the number of agents. It should be apparent from Table I that our results are the first of their kind to provide guarantees for the entire feasibility region, as well as the first agent independent result for a particular class of optimization programs.

The rest of the paper is organized as follows: In Section 2 probabilistic guarantees for sets of feasible solutions of optimization programs with arbitrary cost and uncertain convex constraints are provided. Section 3 focuses on providing agent-independent robustness certificates for the optimal solution set of a specific class of aggregative optimization programs with uncertain cost. The aforementioned results are used in Section 4 in the context of a numerical study on the charging control problem for a fleet of electric vehicles. Section 5 concludes the paper and provides some potential future research directions.

### 1.3. Notation

Let  $\mathcal{N} = \{1, \dots, N\}$  be the index set of all agents, where  $N$  denotes their total number and  $x_i$  the strategy of agent  $i$  taking values in the deterministic set  $X_i \subseteq \mathbb{R}^n$ . We denote  $x = (x_i)_{i \in \mathcal{N}} \in X = \prod_{i \in \mathcal{N}} X_i \subseteq \mathbb{R}^{nN}$  the collection of all agents' strategies and  $\text{bdry}(X)$  the boundary of a set  $X$ . Similarly, the vector  $x_{-i} = (x_j)_{j \in \mathcal{N}, j \neq i} \in \prod_{j \in \mathcal{N}, j \neq i} X_j \subseteq \mathbb{R}^{n(N-1)}$  denotes the collection of the decision vectors of all other agents' strategies except for that of agent  $i$ . Let  $\theta$  be an uncertain parameter defined on the (possibly unknown) probability space  $(\Theta, \mathcal{F}, \mathbb{P})$ , where  $\Theta$  is the sample space, equipped with a  $\sigma$ -algebra  $\mathcal{F}$  and a probability measure  $\mathbb{P}$ . Furthermore, let  $\{\theta_m\}_{m \in \mathcal{M}} \in \Theta^M$ ,  $\mathcal{M} = \{1, \dots, M\}$  be a finite collection of  $M$  independent and identically distributed (i.i.d.) scenarios/realisations of the uncertain vector  $\theta$ , where  $\Theta^M$  is the cartesian product of multiple copies of the sample space  $\Theta$ . Finally,  $\mathbb{P}^M = \prod_{m \in \mathcal{M}} \mathbb{P}$  is the associated product probability measure. The symbols  $x$  and  $(x_i, x_{-i})$  are used interchangeably in this paper, depending on the context.

## 2. Optimization programs with uncertainty in the constraints

### 2.1. The convex case

Consider the following optimization program

$$P_\Theta : \min_{x \in X} J(x) \text{ subject to } x \in \bigcap_{\theta \in \Theta} X_\theta, \quad (1)$$

where the cost function  $J(x)$  can be chosen arbitrarily (even feasibility programs would be admissible),  $x$  is a decision vector taking values in  $X \subset \mathbb{R}^d$  and  $X_\theta$  is dictated by the uncertain parameter  $\theta$ . We seek to provide probabilistic guarantees for all feasible solutions of this class of programs.

For the optimization program (1), we define the following scenario program, where the multi-sample  $\{\theta_m\}_{m \in \mathcal{M}}$  is assumed to be drawn in an i.i.d. fashion from  $\Theta^M$ .

$$P_M : \min_{x \in X} J(x) \text{ subject to } x \in \bigcap_{m \in \mathcal{M}} X_{\theta_m}. \quad (2)$$

Our results depend on a convex constraint structure, thus we impose the following assumption:

- Assumption 1.** 1. *The deterministic constraint set  $X$  is a non-empty, compact and convex set.*
2. *For any random sample  $\theta \in \Theta$ , we have that  $X_\theta = \{x \in \mathbb{R}^d : u(x, \theta) \leq 0\}$ , where  $u : \mathbb{R}^d \times \Theta \rightarrow \mathbb{R}^q$  is a vector-valued convex function.*
3. *For any fixed multi-sample  $\{\theta_m\}_{m \in \mathcal{M}}$  the convex set  $C_M = \{\bigcap_{m \in \mathcal{M}} X_{\theta_m}\} \cap X = \{x \in X : u(x, \theta_m) \leq 0, \forall m \in \mathcal{M}\}$  has a non-empty interior.*

Assumption 1 guarantees that  $P_M$  admits at least one solution for any chosen multisample  $\{\theta_m\}_{m \in \mathcal{M}}$ .

The optimization program  $P_M$  can be equivalently written as  $\min J(x)$  subject to  $x \in C_M$ . Upon finding the feasibility domain  $C_M$  of the problem  $P_M$ , we are interested in investigating the robustness properties collectively for all the points of this domain to yet unseen samples, in other words in quantifying the probability that a new sample  $\theta \in \Theta$  is drawn such that the constraint  $X_\theta$  defined by this sample is not satisfied by some point  $x \in C_M$ . This concept, which is of crucial importance for our work, is known in the literature as the probability of violation and is adapted in our context to represent the probability of violation of an entire set. By Definition 1 in [9] the probability of violation of a given point  $x \in C_M$  is defined as:

$$V(x) = \mathbb{P}\{\theta \in \Theta : x \notin X_\theta\}. \quad (3)$$

By Assumption 1, the probability of violation can be equivalently written as  $V(x) = \mathbb{P}\{\theta \in \Theta : u(x, \theta) > 0\}$ . We can now define the probability of violation of the entire convex set  $C_M$ .

**Definition 1.** *Let  $\mathcal{C} \subseteq 2^X$  be the set of all non-empty, compact and convex sets contained in  $X$ . For any  $C_M \in \mathcal{C}$  we define the probability of violation of the set  $C_M$  as a mapping  $\mathbb{V} : \mathcal{C} \rightarrow [0, 1]$  given by the following relation:*

$$\mathbb{V}(C_M) = \sup_{x \in C_M} V(x).$$

In Definition 2 below two concepts of crucial importance are introduced.

- Definition 2.**
1. For any  $M$ , an algorithm is a mapping  $A_M : \Theta^M \rightarrow \mathcal{C} \subseteq 2^X$  that associates the multi-sample  $\{\theta_m\}_{m \in \mathcal{M}}$  to a unique convex set  $C_M \in \mathcal{C}$ .
  2. Given a multi-sample  $\{\theta_m\}_{m \in \mathcal{M}} \in \Theta^M$ , a set of samples  $\{\theta_m\}_{m \in I_k} \subseteq \{\theta_m\}_{m \in \mathcal{M}}$ , where  $I_k = \{m_1, m_2, \dots, m_k\}$  is called a support subsample if  $A_k(\{\theta_m\}_{m \in I_k}) = A_M(\{\theta_m\}_{m \in \mathcal{M}})$  i.e., the solution returned by the algorithm when fed with  $\{\theta_m\}_{m \in I_k}$  coincides with the one obtained when the entire multi-sample is used. The support subsample with the smallest cardinality among all the possible support subsamples, is known as the minimal support subsample.
  3. A support subsample function is a mapping of the form  $B_M : \{\theta_m\}_{m \in \mathcal{M}} \rightarrow \{m_1, m_2, \dots, m_k\}$  that takes as input all the samples and returns as output only the indices of the samples that belong to the support subsample.

Note that the notions of support subsample and support subsample function in Definitions 2.2, 2.3 are respectively referred to as compression set and compression function in [23]. The cardinality of the support subsample  $\{\theta_m\}_{m \in I_k}$  is by definition the cardinality of the output of the function  $B_M$ , namely,  $k = |B_M(\{\theta_m\}_{m \in \mathcal{M}})|$ . In our set-up the minimal support subsample consists of the indices of the samples that are of support for the entire feasibility region, according to the following definition

**Definition 3.** (Support sample for the feasibility region) A sample  $\theta_j \in \{\theta_m\}_{m \in \mathcal{M}}$  is said to be of support for the feasible set  $C_M$  of  $\mathbf{P}_M$  if its removal leads to an enlargement of the feasible region, i.e., when  $\{\bigcap_{m \in \mathcal{M}} X_{\theta_m}\} \setminus X_{\theta_j} \supset C_M$ .

The number of support samples of the feasible region or, equivalently, the cardinality of the minimal support subsample is denoted as  $F_M$ . The constraints that correspond to indices from the minimal support subsample can be alternatively viewed as an extension of the notion of the facets of a polytope (see Definition 2.1 in [30]) adapted to the more general case of compact and convex sets. Note that a single constraint may give rise to multiple "facets". Figure 1 illustrates the concept of the minimal support subsample by showing the feasible region formed by random convex constraints. Note that only the indices of the samples  $\theta_3, \theta_5, \theta_6, \theta_7$  belong to the minimal support subsample, since if we feed only these samples as input into the algorithm  $A_M$  the feasible set  $C_M$  is returned.

Another important notion used in our work is the notion of an extreme point. An extreme point can be viewed as an extension of the vertex of a polytope for arbitrary compact and convex sets and is defined as a point which is not in the interior of any line segment lying entirely in the set. This property is formally presented in the following definition.

**Definition 4.** (Extreme points) [31]. An extreme point of a convex set  $C$  is a point  $x \in C$  for which the following property holds: If  $x$  can be written as a convex combination of the form  $x = \lambda x_1 + (1 - \lambda)x_2$  with  $x_1, x_2 \in C$  and  $\lambda \in [0, 1]$ , then  $x_1 = x$  and/or  $x_2 = x$ .

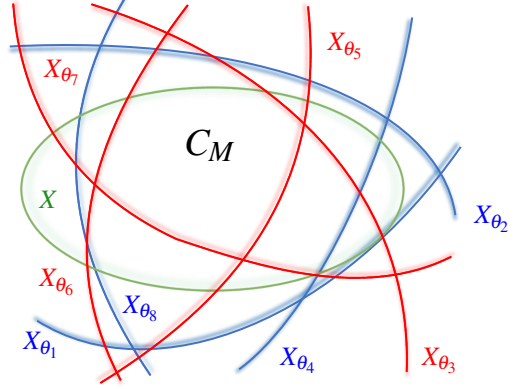


Figure 1: The feasibility region  $C_M$  and its connection with random convex constraints produced by eight i.i.d. samples  $\{\theta_1, \theta_2, \dots, \theta_8\}$ . The constraint in green corresponds to the deterministic constraint  $X$ . Note that only the indices of the samples  $\theta_3, \theta_5, \theta_6, \theta_7$  belong to the minimal support subsample since their corresponding constraints (in red) form, along with the deterministic constraint  $X$ , the boundary of  $C_M$ .

Note that the number of extreme points of a convex set depends on the geometry of the set under study and can also be infinite, e.g., in the case of a  $d$ -dimensional ball.

Our work focuses on compact and convex sets defined over a finite-dimensional space, where the following theorem can be applied.

**Theorem 1.** (Minkowski – Caratheodory Theorem) [31]. Let  $C$  be a compact convex subset of  $\mathbb{R}^d$  of dimension  $d$ . Then any point in  $C$  is a convex combination of at most  $d + 1$  extreme points.

We equip the set of extreme points of the convex set  $C_M$  with indices and denote this set of indices  $E(C_M)$ , while  $\text{bdry}(C_M)$  refers to the boundary of  $C_M$ . It is important to emphasize that the dependence of the convex set  $C_M$  on the multi-sample  $\{\theta_m\}_{m \in \mathcal{M}}$  implies that  $|B_M|$  (and  $|E(C_M)|$ ) are random variables that depend on  $\{\theta_m\}_{m \in \mathcal{M}}$ .

Next we define the set

$$\begin{aligned} \mathcal{C}_\theta &= \{C \in \mathcal{C} : u(x_j, \theta) \leq 0, \forall j \in E(C)\} \\ &= \{C \in \mathcal{C} : C \subseteq X_\theta\}, \end{aligned} \quad (4)$$

of all the non-empty, compact and convex sets where elements  $C$  satisfy the constraint associated with the sample  $\theta \in \Theta$ . Note that if all the extreme points of the set satisfy the inequality  $u(\cdot, \theta) \leq 0$ , then every point  $x \in C$  of the set satisfies it as well. To see this, note that  $x$  can always be expressed as a convex combination of the set's extreme points.

Our aim is to provide probabilistic guarantees for a non-empty, convex and compact set  $C_M$  constructed by the intersection of  $M$  random realizations of the uncertain convex constraint  $X_\theta = \{x \in \mathbb{R}^d : u(x, \theta) \leq 0\}$ , where  $u : \mathbb{R}^d \times \Theta \rightarrow \mathbb{R}^q$  is a convex function with respect to the decision variable  $x$ .

The following lemma shows how the probability of violation of a convex set is related to the probability of violation of its extreme points.



**Lemma 1.** Consider a fixed multi-sample  $\{\theta_m\}_{m \in \mathcal{M}} \in \Theta^M$  and let  $E(C_M)$  be the set of extreme points of  $C_M$ . Then,

$$\mathbb{V}(C_M) \leq \mathbb{P} \left\{ \bigcup_{j \in E(C_M)} \left\{ \theta \in \Theta : u(x_j, \theta) > 0 \right\} \right\}. \quad (5)$$

**Proof 1.** For a fixed multi-sample  $\{\theta_m\}_{m \in \mathcal{M}} \in \Theta^M$  consider an arbitrary point  $x \in C_M$ . Then, the following inequalities are satisfied

$$\begin{aligned} V(x) &= \mathbb{P} \left\{ \theta \in \Theta : x \notin X_\theta \right\} = \mathbb{P} \left\{ \theta \in \Theta : u(x, \theta) > 0 \right\} \\ &\stackrel{(i)}{=} \mathbb{P} \left\{ \theta \in \Theta : u \left( \sum_{j \in I_{d+1}} \lambda_j x_j, \theta \right) > 0 \right\} \\ &\stackrel{(ii)}{\leq} \mathbb{P} \left\{ \theta \in \Theta : \sum_{j \in I_{d+1}} \lambda_j u(x_j, \theta) > 0 \right\} \\ &\leq \mathbb{P} \left\{ \theta \in \Theta : \sum_{j \in I_{d+1}} \lambda_j \max_{j \in I_{d+1}} u(x_j, \theta) > 0 \right\} \\ &= \mathbb{P} \left\{ \theta \in \Theta : \max_{j \in I_{d+1}} u(x_j, \theta) > 0 \right\} \\ &= \mathbb{P} \left\{ \bigcup_{j \in I_{d+1}} \left\{ \theta \in \Theta : u(x_j, \theta) > 0 \right\} \right\} \\ &\stackrel{(iii)}{\leq} \mathbb{P} \left\{ \bigcup_{j \in E(C_M)} \left\{ \theta \in \Theta : u(x_j, \theta) > 0 \right\} \right\}, \end{aligned} \quad (6)$$

Equality (i) is derived from Theorem 1, where the set under study is the convex set  $C_M$ . In our case, the Minkowski-Caratheodory theorem states that any arbitrary point of the set  $x \in C_M$  can be represented as a convex combination of at most  $d+1$  extreme points of  $C_M$ , which means that there exists a subset of extreme points  $\{x_j\}_{j \in I_{d+1}} \subseteq \{x_j\}_{j \in E(C_M)}$  such that  $x = \sum_{j \in I_{d+1}} \lambda_j x_j$ , where  $\sum_{j \in I_{d+1}} \lambda_j = 1$  and  $\lambda_j \geq 0, \forall j \in I_{d+1}$ . Equality (ii) stems from the fact that  $u$  is a convex function of  $x$  for any given  $\theta \in \Theta$ . The last inequality follows from the fact that  $I_{d+1}$  is a set of indices corresponding to extreme points and as such is a subset of  $E(C_M)$ . Since (6) holds for all  $x \in C_M$ , it can equivalently be written as

$$\mathbb{V}(C_M) = \sup_{x \in C_M} V(x) \leq \mathbb{P} \left\{ \bigcup_{j \in E(C_M)} \left\{ \theta \in \Theta : u(x_j, \theta) > 0 \right\} \right\},$$

which concludes the proof.

Lemma 1 shows that the probability of constraint violation of  $C_M$  is bounded by the probability that at least one of the extreme points violates the constraint and is key towards providing probabilistic feasibility guarantees for the feasible region, a result established in the following theorem.

**Theorem 2.** Consider Assumption 1 and any  $A_M, B_M$  as in Definition 2. Fix  $\beta \in (0, 1)$  and define the violation level function  $\varepsilon : \{0, \dots, M\} \rightarrow [0, 1]$  such that

$$\varepsilon(M) = 1 \text{ and } \sum_{k=0}^{M-1} \binom{M}{k} (1 - \varepsilon(k))^{M-k} = \beta. \quad (7)$$

We have that

$$\mathbb{P}^M \left\{ \{\theta_m\}_{m \in \mathcal{M}} \in \Theta^M : \mathbb{V}(C_M) > \varepsilon(k^*) \right\} \leq \beta,$$

where  $k^* = F_M$  is the number of support samples according to Definition 3.

**Proof 2.** By Lemma 1, for any multi-sample  $\{\theta_m\}_{m \in \mathcal{M}}$  and for any cardinality (not necessarily minimal) of the support subsample  $k \in \{1, \dots, N\}$  the following inequalities are satisfied:

$$\begin{aligned} &\mathbb{P}^M \left\{ \{\theta_m\}_{m \in \mathcal{M}} \in \Theta^M : \mathbb{V}(C_M) > \varepsilon(k^*) \right\} \\ &\leq \mathbb{P}^M \left\{ \{\theta_m\}_{m \in \mathcal{M}} \in \Theta^M : \right. \\ &\quad \left. \mathbb{P} \left\{ \bigcup_{j \in E(C_M)} \left\{ \theta \in \Theta : u(x_j, \theta) > 0 \right\} > \varepsilon(k^*) \right\} \right\} \\ &= \mathbb{P}^M \left\{ \{\theta_m\}_{m \in \mathcal{M}} \in \Theta^M : \right. \\ &\quad \left. \mathbb{P} \left\{ \theta \in \Theta : \exists j \in E(C_M), u(x_j, \theta) > 0 \right\} > \varepsilon(k^*) \right\} \\ &= \mathbb{P}^M \left\{ \{\theta_m\}_{m \in \mathcal{M}} \in \Theta^M : \right. \\ &\quad \left. \mathbb{P} \left\{ \theta \in \Theta : C_M \not\subseteq X_\theta \right\} > \varepsilon(k^*) \right\}, \end{aligned} \quad (8)$$

where the last equality is due to (4). Define now an algorithm  $A_M$  as in Definition 2.1, that returns the convex set confined by the feasibility region of  $C_M$ . By construction,  $A_M$  satisfies Assumption 1 of [13], since for any multi-sample  $\{\theta_m\}_{m \in \mathcal{M}}$  it holds that  $A_M(\{\theta_m\}_{m \in \mathcal{M}}) \in \mathcal{C}_{\theta_m}$ , for all  $m \in \mathcal{M}$ . The satisfaction of Assumption 1 paves the way for the use of Theorem 1 of [13]. In particular, Theorem 1 of [13] implies that the right-hand side of (8) can be upper bounded by  $\beta$ .

As such, we have that

$$\begin{aligned} &\mathbb{P}^M \left\{ \{\theta_m\}_{m \in \mathcal{M}} \in \Theta^M : \mathbb{P} \left\{ \theta \in \Theta : C_M \not\subseteq X_\theta \right\} > \varepsilon(k^*) \right\} = \\ &\mathbb{P}^M \left\{ \{\theta_m\}_{m \in \mathcal{M}} \in \Theta^M : \mathbb{P} \left\{ \theta \in \Theta : C_M \notin \mathcal{C}_\theta \right\} > \varepsilon(k^*) \right\} \leq \beta. \end{aligned} \quad (9)$$

From (8) and (9) we obtain that:

$$\mathbb{P}^M \left\{ \{\theta_m\}_{m \in \mathcal{M}} \in \Theta^M : \mathbb{V}(C_M) > \varepsilon(k^*) \right\} \leq \beta, \quad (10)$$

thus concluding the proof.

The result of Theorem 2 implies that with confidence at least  $1 - \beta$ , the probability that there exists at least one feasible solution of  $C_M$  that violates the constraints for a new realization  $\theta \in \Theta$ , is at most equal to  $\varepsilon(k^*)$ . Note that our guarantees trivially hold for any subregion of the feasible set. However, the support subsample cannot be easily computed in the general case. Restricting our attention to programs subject to uncertain affine constraints, provides the means to quantify the support subsample.

## 2.2. The polytopic case

Assuming the presence of affine constraints only, we replace Assumption 1 with the following:

**Assumption 2.** Consider Assumption 1 and further assume that  $X$  is polytopic<sup>3</sup> and  $u(x, \theta) = a^T x - b \leq 0$ , where  $a \in \mathbb{R}^d$ ,  $b \in \mathbb{R}$  and  $\theta = (a^T b) \in \mathbb{R}^{d+1}$ .

Since all feasibility sets are now polytopic, rather than  $C$  and  $C_M$ , we use  $\Pi$  and  $\Pi_M$ , respectively. Under Assumption 2, the cardinality of the minimal support subsample coincides by definition with the number of random facets (see Definition 2.1 in [30]) of the polytope and Theorem 2 gives rise to the following corollary.

**Corollary 1.** Consider Assumption 2 and any  $A_M, B_M$  as in Definition 2. Fix  $\beta \in (0, 1)$  and define the violation level  $\varepsilon : \{0, \dots, M\} \rightarrow [0, 1]$  as a function such that

$$\varepsilon(M) = 1 \text{ and } \sum_{k=0}^{M-1} \binom{M}{k} (1 - \varepsilon(k))^{M-k} = \beta. \quad (11)$$

We have that

$$\mathbb{P}^M \left\{ \{ \theta_m \}_{m \in \mathcal{M}} \in \Theta^M : \mathbb{V}(\Pi_M) > \varepsilon(k^*) \right\} \leq \beta,$$

where  $k^* = F_M$  is the number of facets of  $\Pi_M$ .

Note that, even though during the proof of our theorem we also use the vertices of the polytope, only the number of facets is needed to provide probabilistic guarantees for the entire feasibility region. This feature is appealing from a computational point of view as, in most practical cases, the constructed polytope has a significantly smaller number of facets than extreme points. To illustrate this, consider a finite horizon multi-agent control problem with  $N$  agents, where each agent's decision is subject to upper and lower bounds at each time instance  $t \in \{1, \dots, n\}$ . Hence, for a multi-sample  $\{ \theta_m \}_{m=1}^M \in \Theta^M$ , the feasibility domain  $\prod_{t=1}^n \prod_{i=1}^N \bigcap_{m=1, \dots, M} [x_t^i(\theta_m), \bar{x}_t^i(\theta_m)]$  of the problem is a hyperrectangle whose number of facets  $F = 2Nn$  grows linearly with respect to the number of decision variables, while the number of vertices is given by  $V = 2^{Nn}$ , which grows at an exponential rate with respect to  $Nn$ . Such constraints arise in several applications including the electric vehicle scheduling problem of Section 4.

Note that, as the dimension of the decision vector increases, evaluating the minimal support subsample becomes computationally challenging. However, several efficient algorithms have been proposed for detecting redundant constraints out of the initial set of affine constraints. The currently fastest algorithm for redundancy detection is Clarkson's algorithm [32].

<sup>3</sup>A polytope  $\Pi \in \mathbb{R}^d$  can be expressed by its H-representation, i.e., the intersection of a finite number of halfspaces, and also as the convex hull of its vertex set  $v(\Pi) = \{x_1, \dots, x_Q\}$  i.e.,  $\Pi = \text{conv}(v(\Pi)) = \{ \sum_{j=1}^Q x_j \lambda_j : \sum_{j=1}^Q \lambda_j = 1, \lambda_j \geq 0, j = 1, \dots, Q \}$ , where  $v(\cdot)$  and  $\text{conv}(\cdot)$  denote the set of vertices of the polytope and the convex hull, respectively. This representation is generally known as V-representation.

Reducing the computational complexity of Clarkson's algorithm is still an active research area in computational geometry and combinatorics. One recent noteworthy attempt can be found in [33].

**Remark 1.** Corollary 1 allows retrieving the support samples of the feasibility region by computing the facets of the feasibility region. Calculating these is straightforward for a wide class of problems like those that have the structure of the example in Section 4.2 (see also [3], [26], [34], [35]). The relevance of this result can be seen by considering, for example, the case where an iterative algorithm is obtained to return an optimizer/feasible solution of the underlying problem. The quantification of the support samples using the methodology suggested in [13] can be challenging due to the numerical sensitivity of the latter and the fact that it would require calling the iterative (typically with asymptotic convergence guarantees) algorithm multiple times. This difference becomes more pronounced if the objective function of the underlying problem is non-convex, thus exhibiting possibly multiple local optima.

In case we wish to provide guarantees only for a specific subset of the feasible region, the following remark clarifies how this can be achieved leveraging our results.

**Remark 2.** Let  $A_M$  be an algorithm that returns only a subset of the feasible region. Following the same steps with the proof of Theorem 2, we can provide probabilistic feasibility guarantees for the subset of interest with  $B_M$  returning the support samples that correspond to this subset and not to the entire feasibility region. This might be useful in cases we have prior knowledge that a certain subregion of the feasible domain is of importance, e.g., a neighbourhood around the randomized optimizer  $x_M^*$  and we wish to investigate the probabilistic feasibility of this region only. In case the support samples of the region are less than the support samples of the entire feasible domain, this leads to less conservative feasibility certificates.

## 3. Optimization programs with uncertainty in the cost

### 3.1. Optimization setting

In this section we show that for a specific class of problems frequently arising in practical applications, the probabilistic feasibility guarantees for the optimizers of the problem can be substantially improved by leveraging the notion of the so called support rank [29]. Assuming an uncertain cost function of a specific form and deterministic local constraints, we consider now the following program

$$P : \min_{x \in X} J(x), \quad (12)$$

where  $J(x) = f(x) + \max_{\theta \in \Theta} g(x, \theta)$  and  $f : X \rightarrow \mathbb{R}$ ,  $g : X \times \Theta \rightarrow \mathbb{R}$  is the deterministic and the uncertain part of the cost function, respectively. In addition, the cost under study satisfies the following assumption

**Assumption 3.** 1.  $f$  is jointly convex with respect to all agents' decision vectors, and the set  $X$  is non-empty, compact and convex.

2.  $g$  takes the aggregative form

$$g(x, \theta) = \sum_{i \in \mathcal{N}} g_i(x_i, x_{-i}, \theta) \text{ and} \\ g_i(x_i, x_{-i}, \theta) = x_i^T (A(\theta) \sigma(x) + b(\theta)),$$

where  $\sigma : X \rightarrow \mathbb{R}^n$  is a mapping  $(x_i)_{i \in \mathcal{N}} \mapsto \sum_{i \in \mathcal{N}} x_i$  and  $A : \Theta \rightarrow \mathbb{R}^{n \times n}$ ,  $b : \Theta \rightarrow \mathbb{R}^n$  are uncertain mappings with  $A(\theta)$  being a symmetric positive semi-definite matrix for all  $\theta \in \Theta$ .

Under Assumption 3 the function  $J$  is convex, as the pointwise maximum of an arbitrary number of convex functions is itself a convex function [36]. From Assumption 3.2 the uncertain counterpart of the cost function under study takes the form

$$g(x, \theta) = \sigma(x)^T (A(\theta) \sigma(x) + b(\theta)).$$

The proposed structure captures a wide class of engineering problems, including the electric vehicle charging problem detailed in Section 4. Since  $g$  is convex, using an epigraphic reformulation we recast  $P$  to the equivalent semi-infinite program

$$P' : \min_{x \in X, \gamma \in \mathbb{R}} f(x) + \gamma \\ \text{subject to } h(x, \gamma, \theta) \leq 0, \forall \theta \in \Theta, \quad (13)$$

where  $h(x, \gamma, \theta) = g(x, \theta) - \gamma$ . In addition, if  $(x^*, \gamma^*)$  is the optimal solution of problem  $P'$ , then  $x^*$  is the optimal solution of the original problem  $P$ . Due to the presence of uncertainty and the possibly infinite cardinality of  $\Theta$ , problem  $P'$  is very difficult to solve, without imposing any further assumptions on the geometry of the sample set  $\Theta$  and/or the underlying probability distribution  $\mathbb{P}$ . To overcome this issue, we adopt again a scenario-based scheme [37]. The corresponding scenario program of the uncertain semi-infinite program  $P'$  is thus given by

$$P'_{SC} : \min_{x \in X, \gamma \in \mathbb{R}} f(x) + \gamma \\ \text{subject to } h(x, \gamma, \theta_m) \leq 0, \forall m \in \mathcal{M}, \quad (14)$$

where  $\{\theta_m\}_{m \in \mathcal{M}} \in \Theta^M$  is an i.i.d. multi-sample of cardinality  $M$ . For the scenario program under study, we introduce the following assumption:

**Assumption 4.** 1. For any multi-sample  $\{\theta_m\}_{m \in \mathcal{M}}$ , the scenario program  $P'_{SC}$  admits a feasible solution.

2. The optimal solution  $(x^*, \gamma^*)$  of the scenario program  $P'_{SC}$  is unique.

In case multiple optimal points exist, one can use a convex tie-break rule to select a unique solution. The following concept, at the core of the scenario approach, is important for the derivation of the results in the next subsection, where agent-independent robustness certificates are provided for the optimal solution. Note that this is similar to Definition 3, but it refers now to the optimal solution and not to the feasibility region.

**Definition 5.** (Support constraint [10]) Fix any i.i.d. multi-sample  $\{\theta_m\}_{m \in \mathcal{M}} \in \Theta^M$  and let  $x_0^* = x_0^*(\{\theta_m\}_{m \in \mathcal{M}})$  be the unique optimal solution of the corresponding scenario program of  $P$ , when all the  $M$  samples are taken into account. Let  $x_{-s}^* = x_{-s}^*(\{\theta_m\}_{m \in \mathcal{M}} \setminus \theta_s)$  be the optimal solution obtained after removal of sample  $\theta_s$ . If  $x_0^* \neq x_{-s}^*$  we say that the constraint that corresponds to sample  $\theta_s$  is a support constraint.

3.2. Agent independent probabilistic feasibility guarantees for a unique solution

In many practical applications there are cases where a random constraint may leave a linear subspace unconstrained for any possible sample  $\theta \in \Theta$ . This observation motivated the concept of the support rank as introduced in [29], which allows us to provide tighter probabilistic guarantees for the problem under study. Let  $y \in \mathbb{Y} \subseteq \mathbb{R}^d$  and consider the following semi-infinite optimization program

$$\min_{y \in \mathbb{Y}} c^T y \\ \text{subject to } l(y, \theta) \leq 0, \forall \theta \in \Theta. \quad (15)$$

Notice that the objective function is linear without loss of generality and in the opposite case an epigraphic reformulation could be introduced. Denoting the collection of all the linear subspaces of  $\mathbb{R}^d$  as  $\mathcal{L}$ , we consider all the linear subspaces  $L \in \mathcal{L}$  that, under the presence of the random constraint (15), remain unconstrained for any uncertainty realization  $\theta \in \Theta$  and any point  $y \in \mathbb{Y}$ , i.e., the set

$$\mathcal{U} = \bigcap_{\theta \in \Theta} \bigcap_{y \in \mathbb{Y}} \{L \in \mathcal{L} : L \subset F(y, \theta)\}, \\ \text{where } F(y, \theta) = \{\xi \in \mathbb{R}^d : l(y + \xi, \theta) = l(y, \theta)\}.$$

**Definition 6.** (Support rank [29])

The support rank  $\rho \in \{0, \dots, d\}$  of a random constraint equals to the dimension of the problem  $d$  minus the dimension of the maximal unconstrained linear subspace  $L_{\max}$ , i.e.  $\rho = d - \dim(L_{\max})$ . By maximal unconstrained subspace we mean the unique maximal element  $L_{\max} \in \mathcal{U}$  for which  $L \subseteq L_{\max}$ , for all  $L \in \mathcal{U}$ .

From the support rank definition (see Lemma 3.8 in [29]) we have that Helly's dimension can be upper bounded by the support rank instead of the dimension  $d$  of the problem, which is a more conservative upper bound [10],  $\zeta \leq \rho \in \{0, \dots, d\}$ , where  $\zeta$  denotes the support dimension, a notion similar to that of Helly's dimension extended for the case of multiple random constraints (see [29]). Keeping this relation in mind, our main goal is to obtain a bound for the support rank of the random constraint of problem  $P'$ , thus improving the robustness certificates of its optimal solution. The following proposition aims at finding an upper bound for the support rank  $\rho$  of the random constraint (13), that is independent from the number of agents involved in the optimization program.

**Proposition 1.** Under Assumptions 3 and 4, the support rank  $\rho$  of the random constraint (13) in  $P'$ , has an agent independent upper bound, and in particular,  $\rho \leq n + 1$ .

**Proof 3.** The dimension of the problem under study is  $d = nN + 1$ , due to the presence of the epigraphic variable. Let  $\mathcal{L}$  be the collection of all linear subspaces in  $\mathbb{R}^{nN+1}$ . We aim at finding the dimension of the subspace that remains unconstrained for a scenario program subject to the random constraint  $h(x, \gamma, \theta) \leq 0$  for any uncertain realisation  $\theta \in \Theta$  and any decision vector  $(x, \gamma) \in X \times \mathbb{R}$ . We first define the collection of linear subspaces that are contained in all the sets  $F(x, \gamma, \theta)$ :

$$\mathcal{U} = \bigcap_{\theta \in \Theta} \bigcap_{(x, \gamma) \in \mathbb{R}^{nN+1}} \{L \in \mathcal{L} : L \subset F(x, \gamma, \theta)\}, \text{ where}$$

$$F(x, \gamma, \theta) = \{(\xi, \xi') \in \mathbb{R}^{nN+1} : h(x + \xi, \gamma + \xi', \theta) = h(x, \gamma, \theta)\}$$

In our case,  $h(x + \xi, \gamma + \xi', \theta) = h(x, \gamma, \theta)$  yields:

$$\begin{aligned} & \sigma(x + \xi)^T (A(\theta)\sigma(x + \xi) + b(\theta)) - (\gamma + \xi') \\ &= \sigma(x)^T (A(\theta)\sigma(x) + b(\theta)) - \gamma, \\ \iff & \sigma^T(x)A(\theta)\sigma(\xi) + \sigma^T(\xi)A(\theta)\sigma(x) + \\ & \sigma^T(\xi)A(\theta)\sigma(\xi) + \sigma^T(\xi)b(\theta) - \xi' = 0, \\ \iff & \sigma^T(\xi)A^T(\theta)\sigma(x) + \sigma^T(\xi)(A(\theta)\sigma(x) + b(\theta)) + \\ & \sigma^T(\xi)A(\theta)\sigma(\xi) - \xi' = 0, \end{aligned}$$

where the first equivalence stems from the fact that  $\sigma(x + \xi)$  is linear with respect to its arguments, and the second one after some algebraic rearrangement. Note that each of the terms above is scalar, which means that it is equal to its transpose for all  $x \in X$  and  $\theta \in \Theta$ , while by Assumption 3,  $A^T(\theta) = A(\theta)$  for any  $\theta \in \Theta$ . As such,

$$\sigma^T(\xi)(2A(\theta)\sigma(x) + b(\theta)) + \sigma^T(\xi)A(\theta)\sigma(\xi) - \xi' = 0. \quad (16)$$

Using the equalities  $\sigma^T(\xi)(2A(\theta)\sigma(x) + b(\theta)) = (\mathbb{1}_{1 \times N} \otimes (2A(\theta)\sigma(x) + b(\theta))^T)\xi$  and  $\sigma^T(\xi)A(\theta)\sigma(\xi) = \xi^T(\mathbb{1}_{N \times N} \otimes A(\theta))\xi$ , where  $\mathbb{1}_{1 \times N}$  denotes a row vector with all elements being equal to one and  $\otimes$  denotes the Kronecker product, (16) can be written in the following form:

$$\begin{aligned} & (\mathbb{1}_{1 \times N} \otimes (2\sigma^T(x)A(\theta) + b^T(\theta)))\xi \\ & + \xi^T(\mathbb{1}_{N \times N} \otimes A(\theta))\xi - \xi' = 0, \end{aligned} \quad (17)$$

Let  $\tilde{C} : X \times \Theta \rightarrow \mathbb{R}^{nN}$ ,  $\tilde{A} : \Theta \rightarrow \mathbb{R}^{nN \times nN}$ , where  $\tilde{C}(x, \theta) = \mathbb{1}_{1 \times N} \otimes (2\sigma^T(x)A(\theta) + b^T(\theta))$  and  $\tilde{A}(\theta) = \mathbb{1}_{N \times N} \otimes A(\theta)$ , respectively. Then, equation (17) can be written as:

$$\begin{aligned} & \tilde{C}(x, \theta)\xi + \xi^T \tilde{A}(\theta)\xi - \xi' = 0, \\ \iff & (\tilde{C}(x, \theta) \quad -1) \begin{pmatrix} \xi \\ \xi' \end{pmatrix} \\ & + \begin{pmatrix} \xi & \xi' \end{pmatrix} \begin{pmatrix} \tilde{A}(\theta) & 0_{nN \times 1} \\ 0_{1 \times nN} & 0 \end{pmatrix} \begin{pmatrix} \xi \\ \xi' \end{pmatrix} = 0, \\ \iff & V(x, \theta)w + w^T P(\theta)w = 0, \end{aligned}$$

where,  $V(x, \theta) = (\tilde{C}(x, \theta) \quad -1)$ ,  $P(\theta) = \begin{pmatrix} \tilde{A}(\theta) & 0_{nN \times 1} \\ 0_{1 \times nN} & 0 \end{pmatrix}$  and  $w = \begin{pmatrix} \xi \\ \xi' \end{pmatrix}$ . We need to find an unconstrained linear subspace that is a subset of  $F(x, \gamma, \theta)$ . We define

$$L(x, \gamma, \theta) = \{w \in \mathbb{R}^{nN+1} : \begin{pmatrix} P(\theta) \\ V(x, \theta) \end{pmatrix} w = 0\}.$$

We can easily see that  $L(x, \gamma, \theta) \subset F(x, \gamma, \theta)$ . As such, the random constraint  $h(x, \gamma, \theta) \leq 0$  cannot constrain any of the dimensions of  $L(x, \gamma, \theta)$  (also denoted as  $L$  for simplicity). Let  $Q(x, \gamma, \theta) = \begin{pmatrix} P(\theta) \\ V(x, \theta) \end{pmatrix}$ . Then  $L(x, \gamma, \theta) = \text{nullspace}(Q(x, \gamma, \theta))$  and from nullity-rank theorem [38] we have that  $\dim(L(x, \gamma, \theta)) = nN + 1 - \text{rank}(Q(x, \gamma, \theta))$ . Since  $\text{rank}(P(\theta)) = n$  and  $\text{rank}(V(x, \theta)) = 1$ , this means that  $\text{rank}(Q(x, \gamma, \theta)) = n + 1$ , which implies that  $\dim(L(x, \gamma, \theta)) = nN + 1 - (n + 1)$ . Notice that the unconstrained subspace that we chose may not be the maximal one. This means that the support dimension is  $\rho = nN + 1 - \dim(L_{\max}) \leq nN + 1 - \dim(L) = n + 1$ , thus concluding the proof.

An immediate consequence of Proposition 1 when combined with Theorem 4.1 of [29] is the following theorem.

**Theorem 3.** Let  $(x^*, \gamma^*)$  denote the optimal solution of the scenario program  $P'_{SC}$ . Under Assumptions 3 and 4 we have that

$$\mathbb{P}^M \{\{\theta_m\}_{m \in \mathcal{M}} \in \Theta^M : \mathbb{P}(\theta \in \Theta : h(x^*, \gamma^*, \theta) > 0) > \varepsilon\} \leq \beta, \quad (18)$$

$$\text{where } \beta = \sum_{j=0}^n \binom{M}{j} \varepsilon^j (1 - \varepsilon)^{M-j}. \quad (19)$$

The bound obtained from Theorem 3 constitutes a major improvement for this class of problems, since, irrespective of the number of agents  $N$ , the same number of samples  $M$  is required to provide identical robustness certificates given a local decision vector of size  $n$ . The proof of this theorem, is a direct application of the scenario approach theory (see Theorem 2.4 in [10], and Theorem 4.1 in [29], where the number of support constraints is replaced by the obtained bound for the support rank, namely,  $n + 1$ ). Note that in the absence of Proposition 1, a direct application of the scenario approach theory [10] to the problem under consideration would still result in (18), however, (19) would be replaced by

$$\beta = \sum_{j=0}^{nN} \binom{M}{j} \varepsilon^j (1 - \varepsilon)^{M-j}, \quad (20)$$

where the dependence of the guarantees on the number of agents  $N$  is apparent.

Corollary 2 establishes a link between Theorem 3 and the initial program under study  $P$  by providing probabilistic performance guarantees for the optimal solution. Specifically, it quantifies, in an *a priori* fashion, the probability that the cost



that corresponds to the optimal value  $x^*$  of  $P'_{SC}$  will deteriorate, when a new sample  $\theta \in \Theta$  is encountered. To formalise this, with a slight abuse of notation let  $J(x) = J(x, \{\theta_m\}_{m \in \mathcal{M}})$  be the cost function of the corresponding scenario program of program  $P$  and  $J^+(x) = J(x, \{\theta_m\}_{m \in \mathcal{M}} \cup \{\theta\})$  the cost defined over  $M + 1$  scenarios by taking into account the new sample  $\theta$ .

**Corollary 2.** *Under Assumptions 3 and 4 we have that*

$$\mathbb{P}^M\{\{\theta_m\}_{m \in \mathcal{M}} \in \Theta^M : \mathbb{P}(\theta \in \Theta : J^+(x^*) > J(x^*)) > \varepsilon\} \leq \beta, \quad (21)$$

$$\text{where } \beta = \sum_{j=0}^n \binom{M}{j} \varepsilon^j (1 - \varepsilon)^{M-j} \quad (22)$$

**Proof 4.** Let  $(x^*, \gamma^*)$  be the optimal solution of program  $P'$ , which implies that  $\gamma^* = \max_{m \in \mathcal{M}} g(x^*, \theta_m)$ . As such,

$$\begin{aligned} \mathbb{P}(\theta \in \Theta : h(x^*, \gamma^*, \theta) > 0) &= \mathbb{P}(\theta \in \Theta : g(x^*, \theta) > \gamma^*) = \\ \mathbb{P}(\theta \in \Theta : g(x^*, \theta) > \max_{m \in \mathcal{M}} g(x^*, \theta_m)) &= \\ \mathbb{P}(\theta \in \Theta : \max\{g(x^*, \theta), \max_{m \in \mathcal{M}} g(x^*, \theta_m)\} > \max_{m \in \mathcal{M}} g(x^*, \theta_m)) &= \\ \mathbb{P}(\theta \in \Theta : J^+(x^*) > J(x^*)), \end{aligned} \quad (23)$$

where the second equality follows from the fact that  $\gamma^* = \max_{m \in \mathcal{M}} g(x^*, \theta_m)$ , and the last one from the definitions of  $J$  and  $J^+$ . Direct substitution of (23) in (18) of Theorem 3 concludes then the proof.

## 4. Numerical Study

### 4.1. Probabilistic guarantees for all feasible electric vehicle charging schedules

In the following set-up, a cooperative scheme is considered, where agents-vehicles minimize a common electricity cost, while their charging schedules are subject to constraints. However, most of the work up to this point assumed that these constraints are purely deterministic [34], [3], [35]. We extend this framework by imposing uncertainty on the constraints by considering the following program

$$\begin{aligned} \min_{x \in \mathbb{R}^{nN}} J(x) \text{ subject to} \\ x \in \bigcap_{\theta \in \Theta} \bigcap_{i \in \mathcal{N}} \{x_i \in [\underline{x}_i(\theta), \bar{x}_i(\theta)] : \sum_{t=1}^n x_i^{(t)} \geq E_i(\theta)\}. \end{aligned} \quad (24)$$

The two main requirements for the operation of the system under study, namely, the lower and upper bounds imposed on the power rate of each vehicle and the total energy level to be achieved at the end of charging, can be modelled as constraints of affine form. The variables  $x_i = (x_i^{(t)})_{t=1}^n$  denote the charging schedule for all time instances  $t \in \{1, \dots, n\}$ .

The corresponding scenario program of (24) is given by

$$\begin{aligned} \min_{x \in \mathbb{R}^{nN}} J(x) \text{ subject to} \\ x \in \bigcap_{m \in \mathcal{M}} \bigcap_{i \in \mathcal{N}} \{x_i \in [\underline{x}_i(\theta_m), \bar{x}_i(\theta_m)] : \sum_{t=1}^n x_i^{(t)} \geq E_i(\theta_m)\}. \end{aligned} \quad (25)$$

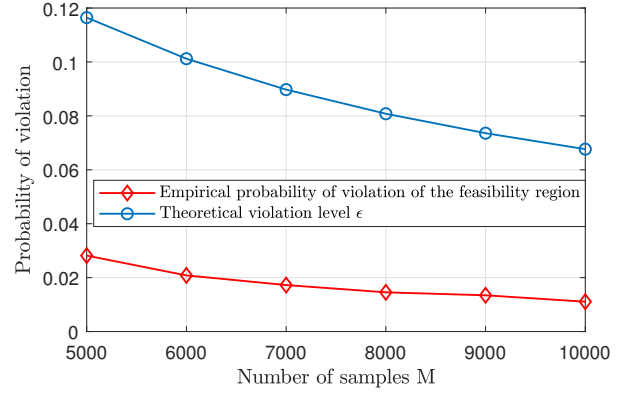


Figure 2: Empirical probability of violation of the feasibility region (red line) versus the theoretical violation level for a different number of samples  $M = \{5000, 6000, \dots, 10000\}$ . To calculate the probability of violation a total of  $M_{test} = 40000$  samples is used.

The cost function  $J$  is allowed to have any arbitrary form and can even be affected by uncertainty. In our set-up we assume that the upper and lower bounds of the charging rate,  $(\bar{x}_i(\theta_u))_{i \in \mathcal{N}}, (\underline{x}_i(\theta_l))_{i \in \mathcal{N}} \in \mathbb{R}^{nN}$  are affected by the uncertain parameters  $\theta_u, \theta_l \in \mathbb{R}^{nN}$ , respectively, with uncertainty representing volatile grid power restrictions. Each of the parameters' elements is extracted according to the same probability distribution  $\mathcal{N}(0, 0.5)$ , where  $\mathcal{N}(\mu, \sigma)$  is a Gaussian distribution with mean  $\mu$  and standard deviation  $\sigma$ . The distribution is truncated by a prespecified quantity to avoid infeasibility issues. We further assume that the uncertainty is additive, i.e.,  $\bar{x}(\theta_u) = \bar{x}^{nom} + \theta_u$  and  $\underline{x}(\theta_l) = \underline{x}^{nom} + \theta_l$ , where each element of  $\bar{x}^{nom}$  is drawn from a uniform probability distribution with support  $[10, 20]$  kW and  $\underline{x}^{nom}$  is set to 2 kW. Finally, the energy capacity of the battery can be affected by a variety of factors such as battery aging and lithium plating for Li-ion batteries. These phenomena can have an important effect on the amount of energy required by each vehicle to fully charge, thus imposing uncertainty on the final energy level to be achieved by each of them by the end of the charging cycle. In our set-up, the uncertainty in the total energy  $E = (E_i)_{i \in \mathcal{N}}$  (in kWh) of each vehicle at the end of charging is yet again assumed to be additive, i.e.,  $E = E^{nom} + \theta_e$ , where  $\theta_e \in \mathbb{R}^N$  and its elements are extracted according to the probability  $\mathcal{N}(0, 1)$  and  $E_i^{nom} \in \mathbb{R}$  is the nominal final energy demand of each agent  $i \in \mathcal{N}$ . The uncertainty vector is given by  $\theta = [\theta_u, \theta_l, \theta_e] \in \mathbb{R}^{N(2n+1)}$ .

Considering  $N = 5$  vehicles and  $n = 12$  time slots we construct the feasibility region of the corresponding scenario program (25). Using  $M_{test} = 40000$  test samples we empirically compute the probability that a new yet unseen constraint will be violated by at least one element of the feasible region and compare it with the theoretical violation level  $\varepsilon(k)$  from Theorem 2. The results are shown in Figure 2, where the red line corresponds to the empirical probability and the blue line corresponds to the theoretical bound. Note that an upper bound for the theoretical violation level  $\varepsilon(k)$  can be obtained by counting the number of the facets  $k = F_M$  of the feasible set or by leveraging the geometry of our numerical example to provide an upper

bound for  $F_N$ , that is  $F_N \leq 2nN + N$ . This bound can be easily derived noticing that the feasible region is in fact a cartesian product of  $N$  rectangles intersected by a halfspace that corresponds to the energy constraint. As such the worst case number of facets for the entire polytope in our example is  $N(2n + 1)$ .

In general, for samples that give rise to more than one affine constraints, the number of facets constitutes only an upper bound for the cardinality of the minimal support subsample. This bound is tight only in the case when there is a one to one correspondence between a sample realization and a scalar-valued constraint. This implies that the guarantees for the feasible subset of the problem under study can be significantly tighter. The reason behind the use of the looser bound  $k = 2nN + N$  in our example lies in the fact that its quantification is straightforward and the use of a support subsample function is thus not required. In other cases, however, where an upper bound for  $k$  is absent, the methodologies for the detection of redundant affine constraints provided by [33] and references therein can be used as minimal support subsample functions.

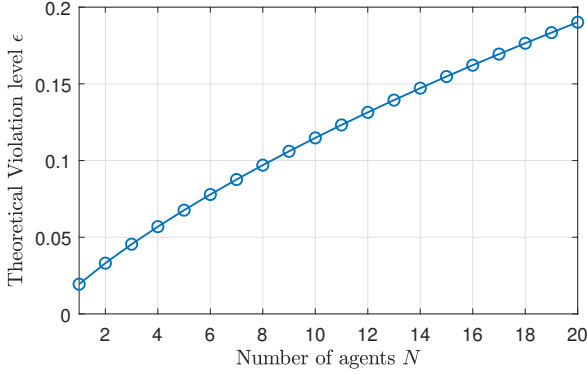


Figure 3: The violation level  $\epsilon$  with respect to the number of agents  $N$  for  $\beta = 10^{-6}$  and a multi-sample of size  $M = 10000$ . Note how the violation level for a fixed number of samples is agent dependent, i.e., the guarantees become more conservative as the number of agents increases.

Figure 3 illustrates the violation level  $\epsilon(k)$  with respect to the number of agents  $N$  for a fixed confidence level  $1 - \beta$ , a fixed multi-sample of size  $M = 10000$  and  $k^* = 2Nn + N$ . It is clear that for the same number of samples the probabilistic guarantees provided by Theorem 2 for all feasible solutions become more conservative as the number of vehicles in the fleet increases. This means that more samples are required to provide the same probabilistic guarantees as the dimension grows.

#### 4.2. Agent independent robustness certificates for the optimal electric vehicle charging profile

In this set-up the cost to be minimized is influenced by the electricity price, which in turn is considered to be a random variable affected by uncertainty. Uncertainty here refers to price volatility. All electric vehicles cooperate with each other choosing their charging schedules so as to minimize the total uncertain electricity cost, while satisfying their own deterministic constraints. To this end, we consider the following uncertain

electric vehicle charging problem

$$P_{EV} : \min_{x \in \mathbb{R}^{nN}} f(x) + \max_{\theta \in \Theta} g(x, \theta),$$

$$\text{subject to } x_i \in [\underline{x}_i, \bar{x}_i], \sum_{t=1}^n x_i^{(t)} \geq E_i \text{ for all } i \in \mathcal{N}, \quad (26)$$

where  $f(x) = \sum_{i \in \mathcal{N}} f_i(x_i, x_{-i}) = \sigma(x)^T p_0(\sigma(x))$  is the deterministic part of the electricity cost that depends on a nominal electricity price  $p_0(\sigma(x)) = A_0 \sigma(x) + b_0$  that is, in turn, a function of the aggregate consumption of the vehicles.  $g(x, \theta) = \sum_{i \in \mathcal{N}} g_i(x_i, x_{-i}, \theta) = \sigma(x)^T p(\sigma(x), \theta)$  constitutes the uncertain part of the electricity cost, where the price  $p(\sigma(x), \theta) = A(\theta) \sigma(x) + b(\theta)$  is additionally affected by the uncertain parameter  $\theta$  extracted from the support set  $\Theta$  according to a probability distribution  $\mathbb{P}$ , where  $\Theta$  and  $\mathbb{P}$  are considered unknown. The elements of  $A_0 \in \mathbb{R}^{n \times n}$  and  $b_0 \in \mathbb{R}^n$  are deterministic with  $A_0$  being a symmetric positive semi-definite matrix, while the uncertain mappings  $A : \Theta \rightarrow \mathbb{R}^{n \times n}$  and  $b : \Theta \rightarrow \mathbb{R}^n$  are defined as in Section 3. The vectors  $\underline{x}_i, \bar{x}_i \in \mathbb{R}^n$  constitute the lower and upper bound of the charging rate of vehicle  $i \in \mathcal{N}$ , respectively, while  $E_i \in \mathbb{R}$  is the final energy to be achieved by each vehicle  $i \in \mathcal{N}$  by the end of the charging cycle.

Following the same lines as in Section 3, we apply an epigraphic reformulation and use samples for  $\Theta$  to obtain the following scenario program

$$P_{EV}^{sc} : \min_{(x, \gamma) \in \mathbb{R}^{nN+1}} f(x) + \gamma, \quad (27)$$

$$\text{subject to } x_i \in [\underline{x}_i, \bar{x}_i], \sum_{t=1}^n x_i^{(t)} \geq E_i, \text{ for all } i \in \mathcal{N},$$

$$g(x, \theta_m) \leq \gamma, \text{ for all } m \in \mathcal{M}.$$

In our set-up  $A(\theta) \in \mathbb{R}^{n \times n}$  is assumed to be a diagonal matrix with non-negative diagonal elements for any uncertain realization  $\theta \in \Theta$ . The diagonal elements of  $A(\theta)$  and the elements of  $b(\theta) \in \mathbb{R}^n$  are extracted according to uniform distributions. For each agent  $i \in \mathcal{N}$  the upper bound  $\bar{x}_i$  takes a random value in the set  $[6, 15]$  kW, the lower bound  $\underline{x}_i$  is set to 2 kW and the final energy to be achieved by the end of the charging cycle is appropriately chosen to be feasible, considering the number of timesteps  $n$  and the upper bound of the power rate of each agent.  $A_0 \in \mathbb{R}^{n \times n}$  is assumed to be a diagonal matrix, whose diagonal entries are all set to 0.01 and  $b_0$  is derived by rescaling a winter weekday demand profile in the UK.

Note that our results can be used alongside any optimization algorithm irrespective of its nature, i.e., centralised, decentralised or distributed; here we solved the problem in a centralised fashion. The number of samples we use for each problem is  $M = 500$ . By fixing  $\beta = 10^{-6}$  and using the bound

$$\epsilon = \frac{2}{M} (\ln \frac{1}{\beta} + n \ln 2), \quad (28)$$

which is a sufficient condition (see [37, (p.42)]) for satisfaction of (19), we obtain the theoretical violation level  $\epsilon = 0.0885$ . Note that the dimension we use to provide probabilistic guarantees for the optimal solution is set, in accordance to Theorem

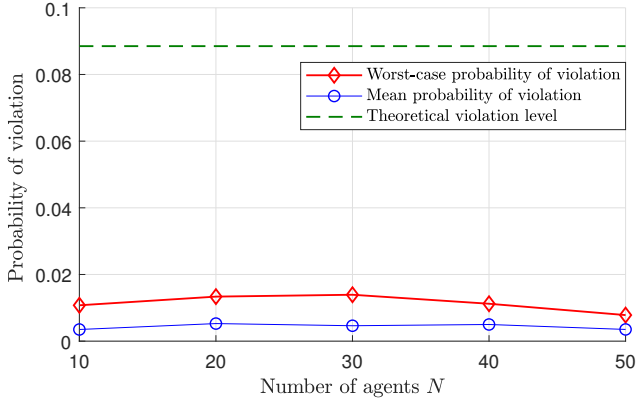


Figure 4: Mean and worst-case empirical probability of violation of the optimal solution with respect to the number of agents versus the theoretical violation level  $\varepsilon = 0.0885$ . The number of samples used is  $M = 500$  and  $\beta = 10^{-6}$ . By drawing a different multi-sample for each choice of the number of agents  $N = \{10, 20, 30, 40, 50\}$ , we solve the corresponding scenario program for a fixed number of time slots  $n = 12$ . We then repeat this process 20 times (note that the multi-sample used for each repetition is also different) and compute the empirical probability of violation of the obtained optimal solutions, using  $M_{test} = 100000$  test samples.

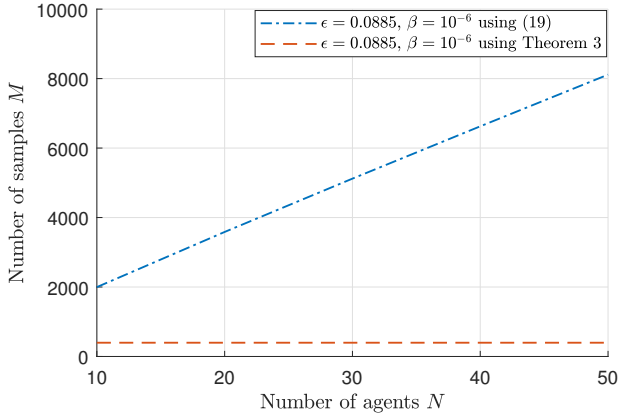


Figure 5: The number of samples required with respect to the number of agents  $N = \{10, \dots, 50\}$  using the results of Theorem 3 versus the one that would have been obtained if (19) is used instead. We consider a charging cycle of duration  $n = 12$ . The red line corresponds to Theorem 3, while the blue line corresponds to (20).

3 to  $n + 1$  instead of  $nN + 1$ , which circumvents the computational issues related to the rapid surge in dimension due to the multiplication of the number of agents with the number of time slots.

By drawing a different multi-sample for each choice of the number of agents  $N \in \{10, 20, 30, 40, 50\}$  we solve the corresponding scenario program for a fixed number of time slots  $n = 12$ . We then repeat this process 20 times (note that the multi-sample used for each repetition is also different) and compute the empirical probability of violation of the obtained optimal solutions, using  $M_{test} = 100000$  test samples each time. The mean and worst-case empirical probability of violation is depicted in Figure 4 in comparison with the theoretical violation level  $\varepsilon$ . The empirical values are always below the theoretical level of violation, which is constant with the number of agents due to the agent independent nature of our Theorem 3. In addition, the trend in Figure 4 shows, as expected by Theorem 3, that the number of agents does not affect the empirical probability of violation.

This result highlights the fact that, for fixed number of time periods  $n$ , the number of samples  $M$  required to provide identical probabilistic guarantees, as the size of the fleet of electric vehicles increases, remains constant. This is illustrated in Figure 5, where we show the number of samples required (for  $\varepsilon = 0.0885$ ,  $\beta = 10^{-6}$  and  $n = 12$ ) using the results of Theorem 3 versus the number of samples needed to provide the same robustness certificates using the classic results in scenario approach for a different number of agents  $N = 10, \dots, 50$ . The red line corresponds to Theorem 3 and shows the agent independent nature of our guarantees, while the blue line corresponds to the conservative agent dependent result of (20).

## 5. Concluding remarks

We first considered a general class of optimization programs with an arbitrary cost function and uncertain convex constraints and provided *a posteriori* bounds for the probability of violation of all feasible solutions. We then focused on a different class of multi-agent programs that involved an uncertain aggregative term and deterministic constraints. For such problems we provided agent independent probabilistic guarantees for the optimal solution in an *a priori* fashion.

Effort is being made towards extending our results to provide agent independent probabilistic guarantees in a non-cooperative set-up, giving rise to aggregative games. In addition, we aim at generalising the methodology of our first contribution to provide in an *a posteriori* fashion tighter robustness certificates that depend only on a subset of the feasible region, which circumscribes the game solutions.

## References

- [1] D. Ioli, L. Deori, A. Falsone, M. Prandini, A two-layer decentralized approach to the optimal energy management of a building district with a shared thermal storage, IFAC-PapersOnLine 50 (2017) 8844–8849. doi:10.1016/j.ifacol.2017.08.1540.

- [2] B. Gharesifard, T. Başar, A. Domínguez-García, Price-based coordinated aggregation of networked distributed energy resources, *IEEE Transactions on Automatic Control* 61 (10) (2016) 2936–2946. doi:10.1109/TAC.2015.2504964.
- [3] D. Paccagnan, B. Gentile, F. Parise, M. Kamgarpour, J. Lygeros, Nash and wardrop equilibria in aggregative games with coupling constraints, *IEEE Transactions on Automatic Control* 64 (4) (2019) 1373–1388. doi:10.1109/TAC.2018.2849946.
- [4] D. Acemoglu, M. K. Jensen, Aggregate comparative statics, *Games and Economic Behavior* 81 (2013) 27–49.
- [5] J. Ghaderi, R. Srikant, Opinion dynamics in social networks with stubborn agents: Equilibrium and convergence rate, *Automatica* 50 (12) (2014) 3209–3215.
- [6] D. Acemoglu, A. Ozdaglar, Opinion dynamics and learning in social networks, *Dynamic Games and Applications* 1 (1) (2011) 3–49.
- [7] D. Bertsekas, J. Tsitsiklis, Parallel and distributed computation : numerical methods / dimitri p. bertsekas, john n. tsitsiklis, Athena Scientific (01 1989).
- [8] N. Parikh, S. Boyd, Proximal algorithms, *Foundations and Trends in Optimization* 1 (3) (2014) 127–239.
- [9] M. C. Campi, G. C. Calafiore, The scenario approach to robust control design, *IEEE Transactions on automatic control* 51 (5) (2006) 742–753.
- [10] M. C. Campi, S. Garatti, The exact feasibility of randomized solutions of uncertain convex programs, *SIAM Journal on Optimization* 19 (3) (2008) 1211–1230.
- [11] M. C. Campi, S. Garatti, M. Prandini, The scenario approach for systems and control design, *Annual Reviews in Control* 33 (2008) 149–157.
- [12] M. C. Campi, S. Garatti, Wait-and-judge scenario optimization, *Mathematical Programming* 167 (1) (2018) 155–189. doi:10.1007/s10107-016-1056-9. URL <https://doi.org/10.1007/s10107-016-1056-9>
- [13] M. C. Campi, S. Garatti, F. A. Ramponi, A general scenario theory for nonconvex optimization and decision making, *IEEE Transactions on Automatic Control* 63 (12) (2018) 4067–4078.
- [14] D. Bai, T. Carpenter, J. Mulvey, Making a case for robust optimization models, *Management Science* 43 (1997) 895–907. doi:10.1287/mnsc.43.7.895.
- [15] J. R. Birge, F. Louveaux, Introduction to Stochastic Programming, Springer-Verlag, New York, NY, USA, 1997.
- [16] S. Grammatico, X. Zhang, K. Margellos, P. Goulart, J. Lygeros, A scenario approach for non-convex control design, *IEEE Transactions on Automatic Control* 61 (2) (2016) 334–345.
- [17] F. Fabiani, K. Margellos, P. Goulart, Probabilistic feasibility guarantees for solution sets to uncertain variational inequalities, *Automatica* (May 2020).
- [18] F. Fabiani, K. Margellos, P. Goulart, On the robustness of equilibria in generalized aggregative games, *Proceedings of the 59th IEEE Conference on Decision and Control* (May 2020).
- [19] D. Paccagnan, M. Campi, The scenario approach meets uncertain variational inequalities and game theory, *IEEE Transactions on Automatic Control* (2019).
- [20] F. Fele, K. Margellos, Probably approximately correct nash equilibrium learning, *IEEE Transactions on Automatic Control*, (conditionally accepted) (2020).
- [21] F. Fele, K. Margellos, Probabilistic sensitivity of nash equilibria in multi-agent games: a wait-and-judge approach, in: 2019 IEEE 58th Conference on Decision and Control (CDC), 2019, pp. 5026–5031.
- [22] G. Pantazis, F. Fele, K. Margellos, A posteriori probabilistic feasibility guarantees for nash equilibria in uncertain multi-agent games, *IFAC World Congress* (2020).
- [23] K. Margellos, M. Prandini, J. Lygeros, On the connection between compression learning and scenario based single-stage and cascading optimization problems, *IEEE Transactions on Automatic Control* 60 (10) (2015) 2716–2721.
- [24] L. Deori, K. Margellos, M. Prandini, Regularized jacobi iteration for decentralized convex quadratic optimization with separable constraints, *IEEE Transactions on Control Systems Technology* 27 (2016) 1636–1644.
- [25] S. Grammatico, Aggregative control of competitive agents with coupled quadratic costs and shared constraints, in: 2016 IEEE 55th Conference on Decision and Control (CDC), 2016, pp. 3597–3602.
- [26] S. Grammatico, Dynamic control of agents playing aggregative games with coupling constraints, *IEEE Transactions on Automatic Control* 62 (9) (2017) 4537–4548.
- [27] S. Grammatico, F. Parise, M. Colombino, J. Lygeros, Decentralized convergence to nash equilibria in constrained deterministic mean field control, *IEEE Transactions on Automatic Control* 61 (11) (2016) 3315–3329.
- [28] G. Pantazis, F. Fele, K. Margellos, Agent independent probabilistic robustness certificates for robust optimization programs with uncertain quadratic cost, *Proceedings of the 59th IEEE Conference on Decision and Control* (2020).
- [29] G. Schildbach, L. Fagiano, M. Morari, Randomized solutions to convex programs with multiple chance constraints, *SIAM Journal on Optimization* 23 (05 2012). doi:10.1137/120878719.
- [30] M. Ziegler, Lectures on polytopes, Springer-Verlag New York, Inc (1995).
- [31] B. Simon, Convexity: An analytic viewpoint, Cambridge University Press (2011).
- [32] K. L. Clarkson, More output-sensitive geometric algorithms, in: *Proceedings of the 35th Annual Symposium on Foundations of Computer Science, SFCS '94*, IEEE Computer Society, USA, 1994, p. 695–702.
- [33] K. Fukuda, B. Gärtner, M. Szedlak, Combinatorial redundancy detection, *Annals of Operations Research* 265 (06 2018).
- [34] Z. Ma, D. S. Callaway, I. Hiskens, Decentralized charging control of large populations of plug-in electric vehicles, *IEEE Transactions on Control Systems Technology* 21 (2013) 67–78.
- [35] L. Deori, K. Margellos, M. Prandini, Price of anarchy in electric vehicle charging control games: When nash equilibria achieve social welfare, *Automatica* 96 (2018) 150–158.
- [36] S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge University Press, USA, 2004.
- [37] M. Campi, S. Garatti, Introduction to the scenario approach, *Society of industrial and applied mathematics (SIAM)* (Nov. 2018).
- [38] S. Axler, Linear algebra done right, Springer (1997).