Probabilistic feasibility guarantees for solution sets to uncertain variational inequalities

Filippo Fabiani, Kostas Margellos and Paul J. Goulart

Department of Engineering Science, University of Oxford, OX1 3PJ, United Kingdom

Abstract

We develop a data-driven approach to the computation of a-posteriori feasibility certificates for sets of solutions of variational inequalities affected by uncertainty. Specifically, we focus on variational inequalities with a deterministic mapping and an uncertain feasible set, and represent uncertainty by means of scenarios. Building upon recent advances in the scenario approach literature, we quantify the robustness properties of the entire set of solutions of a variational inequality, with feasibility set constructed using the scenario approach, against a new unseen realization of the uncertainty. Our results extend existing ones that typically impose that the solution set is a singleton and require certain non-degeneracy properties: hence, we thereby offer probabilistic feasibility guarantees for any feasible solution of the underlying variational inequality. We show that assessing the violation probability of an entire set of solutions requires enumeration of the support constraints that "shape" this set, and also propose a procedure to enumerate the support constraints that does not require a description of the solution set. We illustrate our results through numerical simulations on a robust game involving an electric vehicle charging coordination problem.

1 Introduction

VARIATIONAL INEQUALITIES (VIs) are a general purpose tool encompassing a broad variety of equilibrium problems such as network and traffic problems, optimal control, economics and demand-side management (see, e.g., [12,16] for an extensive discussion). VIs are formally defined by means of a feasible set $\mathcal{X} \subseteq \mathbb{R}^n$, and a mapping $F : \mathcal{X} \to \mathbb{R}^n$. We denote by $\mathrm{VI}(\mathcal{X}, F)$ the problem of finding some $x^\star \in \mathcal{X}$ such that $(y - x^\star)^\top F(x^\star) \geq 0$, for all $y \in \mathcal{X}$. In this paper we focus on stochastic approaches to uncertain VIs, in which the problem data may be affected by uncertainty. Specifically, given a random variable $\delta \in \Delta$, we adopt a worst-case formulation [25] where the feasible set is modelled as the intersection of sets $\mathcal{X}_\delta$, generated by every possible realization of δ, and a deterministic set $\mathcal{X}$. We define the worst-case VI problem, $\mathrm{VI}(\mathcal{X} \cap \mathcal{X}_\delta, F)$, as the problem of finding some $x^\star \in \mathcal{X} \cap \mathcal{X}_\delta$ that satisfies

$$(y - x^\star)^\top F(x^\star) \geq 0, \text{ for all } y \in \mathcal{X} \cap \mathcal{X}_\delta,\ \delta \in \Delta. \quad (1)$$

However, such a worst-case formulation imposes two main challenges: i) the set Δ may be unknown and the only information available may come via data/scenarios for δ; ii) even if Δ is known, it might be a set with infinite cardinality, thereby giving rise to an infinite set of constraints in (1). To address these challenges, we adopt the data-driven approach proposed in [7] to quantify a-posteriori the feasibility of the entire set of solutions to the VI against previously unseen realizations of the uncertainty. Using a set-oriented perspective, we recast our problem to the form of the abstract decision-making problems considered in [7]. This enables us to inherit the probabilistic feasibility results established in [7, Th. 1], and thereby characterize the robustness properties of the entire solution set to the uncertain VI in (1).

To the best of our knowledge, this work is the first to address the problem of evaluating the robustness of the entire set of solutions to an uncertain VI in a distribution-free fashion. The present work is indeed complementary to the one in [13, 14], where VIs arising in the computation of a Nash equilibrium problem (NEP) with an uncertain mapping and a deterministic feasible set are (indirectly) investigated. Conversely, a NEP with uncertain, yet affine, local constraints and deterministic cost functions is considered in [23]. A contribution of [23] is to provide robustness certificates for the constraint violation of any feasible point of the game considered. In contrast, we show in §3 that assessing the robustness of an equilibrium at a point inside the feasible set may lead to an over-conservative bound compared to the one derived in this paper. In [10], instead, proba-

[1] *Email addresses:* `filippo.fabiani, kostas.margellos, paul.goulart@eng.ox.ac.uk`. This work was partially supported through the Government's modern industrial strategy by Innovate UK, part of UK Research and Innovation, under Project LEO (Ref. 104781).

bilistic bounds on the feasibility of the entire set of variational generalized Nash equilibria (v-GNE) associated to a generalized Nash equilibrium problem (GNEP) in aggregative form are provided. We therefore consider a broad family of uncertain VIs in (1) rather than just VI problems arising in computing v-GNE, thus complementing the results of [10, 23]. Finally, [22] synthesized a-posteriori robustness certificates for the solution to uncertain (quasi-)VI in (1), which amounts to the problem of finding some $x^\star \in \cap_{\delta\in\Delta} \mathcal{X}_\delta(x^\star)$ such that $(y-x^\star)^\top F(x^\star) \geq 0$, for all $y \in \cap_{\delta\in\Delta}\mathcal{X}_\delta(x^\star)$. However, in [22] it is postulated that the VI admits a *unique* solution, while certain non-degeneracy assumptions, typically hard to verify [6, 14], are imposed. To conclude, we summarize our main contributions as follows:

- We provide a-posteriori robustness certificates for the entire set of solutions to an uncertain VI. Our set-oriented perspective is crucial for two reasons:
  (1) We are able to bypass the uniqueness and non-degeneracy assumptions postulated in [22];
  (2) Compared to [23], we show that our bounds are, in general, less conservative (albeit weaker than those in [22] – see Remark 1);
- In the case of affine constraints, we give a procedure to enumerate the support subsamples, a key quantity for our robustness certificates, that requires fewer iterations compared to the one in [7, 22].

We finally corroborate our findings through numerical simulations on a GNEP modelling the charging coordination of a fleet of plug-in electric vehicles (PEVs).

Notation: $\mathbb{N}$, $\mathbb{R}$, and $\mathbb{R}_{\geq 0}$ denote the set of natural, real, and nonnegative real numbers, respectively. $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. Denote vectors of appropriate dimensions with elements all equal to 1 (0) as $\mathbf{1}$ ($\mathbf{0}$). Given a matrix $A \in \mathbb{R}^{m\times n}$, $A^\top$ denotes its transpose. The operator $\otimes$ denotes the Kronecker product, $\mathrm{col}(\cdot)$ stacks its arguments in column vectors or matrices, while $\mathrm{avg}(\cdot)$ is the average operator. For vectors $v_1, \ldots, v_N \in \mathbb{R}^n$ and $\mathcal{I} = \{1, \ldots, N\}$, we denote $\boldsymbol{v} := (v_1^\top, \ldots, v_N^\top)^\top = \mathrm{col}((v_i)_{i\in\mathcal{I}})$ and $\boldsymbol{v}_{-i} := \mathrm{col}((v_j)_{j\in\mathcal{I}\setminus\{i\}})$. For a given set $\mathcal{S} \subseteq \mathbb{R}^n$, $|\mathcal{S}|$ represents its cardinality, and $\mathrm{int}(\mathcal{S})$, $\mathrm{relint}(\mathcal{S})$ and $\mathrm{bdry}(\mathcal{S})$ denote its topological interior, relative interior and boundary, respectively. The set $\mathrm{aff}(\mathcal{S})$ denotes its affine hull. We denote single-valued mappings with "$\to$" and set-valued mappings with "$\rightrightarrows$". The mapping $T : \mathcal{X} \to \mathbb{R}^n$ is pseudomonotone on $\mathcal{S}$ if for all $x, y \in \mathcal{S}$, $(x-y)^\top T(y) \geq 0 \implies (x-y)^\top T(x) \geq 0$; monotone if $(T(x) - T(y))^\top (x-y) \geq 0$ for all $x, y \in \mathcal{S}$; strongly monotone if there exists a constant $c > 0$ such that $(T(x)-T(y))^\top(x-y) \geq c\|x-y\|^2$ for all $x, y \in \mathcal{S}$.

2 Problem statement and main result

2.1 Uncertain VIs and scenario-based formulation

We aim at providing out-of-sample feasibility certificates for the entire set of solutions to the uncertain VI in (1) by exploiting some *scenarios* of the uncertain parameter δ. Formally, let us consider a probability space $(\Delta, \mathcal{D}, \mathbb{P})$, where $\Delta \subseteq \mathbb{R}^\ell$ represents the set of values that δ can take, $\mathcal{D}$ is a σ-algebra and $\mathbb{P}$ is a (possibly unknown) probability measure over $\mathcal{D}$. We assume to have available a finite collection of $K \in \mathbb{N}$ independent and identically distributed (i.i.d.) observed realizations of δ, i.e., $\delta_K := \{\delta^{(i)}\}_{i\in\mathcal{K}} = \{\delta^{(1)}, \ldots, \delta^{(K)}\} \in \Delta^K$, $\mathcal{K} := \{1, 2, \ldots, K\}$, henceforward called K-multisample. Note that every K-multisample is defined over the probability space $(\Delta^K, \mathcal{D}^K, \mathbb{P}^K)$, resulting from the K-fold Cartesian product of the original probability space $(\Delta, \mathcal{D}, \mathbb{P})$. Let $\mathcal{X}_{\delta^{(i)}}$ be a constraint set associated with the i-th sample, which constrains the decisions that are admissible for the situation represented by $\delta^{(i)}$. The scenario-based VI problem $\mathrm{VI}(\mathcal{X}_{\delta_K}, F)$, with $\mathcal{X}_{\delta_K} := \cap_{i\in\mathcal{K}}\mathcal{X}_{\delta^{(i)}} \cap \mathcal{X}$, is then the problem of finding an $x^\star \in \mathcal{X}_{\delta_K}$ such that

$$(y - x^\star)^\top F(x^\star) \geq 0, \text{ for all } y \in \mathcal{X}_{\delta_K}. \quad (2)$$

Let us define the set of solutions to (2) as

$$\Omega_{\delta_K} := \{x \in \mathcal{X}_{\delta_K} \mid (y-x)^\top F(x) \geq 0, \forall y \in \mathcal{X}_{\delta_K}\}. \quad (3)$$

Given the dependency on δ_K, the set Ω_{δ_K} is itself a random quantity. When $K = 0$, our problem reduces to a deterministic VI problem, $\mathrm{VI}(\mathcal{X}, F)$, with solution set Ω_{δ_0}. We introduce next a key assumption for our results.

Standing Assumption 1 *For any $K \in \mathbb{N}_0$, $\mathcal{X}_{\delta_K} \neq \emptyset$ is a compact and convex set for all $\delta_K \in \Delta^K$. The mapping $F : \mathcal{X} \to \mathbb{R}^n$ is continuous and pseudomonotone.*

Lemma 1 *For all $K \in \mathbb{N}_0$, Ω_{δ_K} is a nonempty, compact and convex set.*

PROOF. It follows by combining [12, Cor. 2.2.5, Th. 2.3.5], as $F(\cdot)$ is continuous and pseudomonotone and $\mathcal{X}_{\delta_K}$ is a finite intersection (due to Standing Assumption 1) of nonempty, compact and convex sets.

In the spirit of [7], we then introduce $\Theta_K : \Delta^K \rightrightarrows \mathcal{X}$ as the mapping that, given a set of realizations δ_K, $K \in \mathbb{N}_0$, returns the solution set to $\mathrm{VI}(\mathcal{X}_{\delta_K}, F)$, namely

$$\Theta_K(\delta^{(1)}, \ldots, \delta^{(K)}) = \Theta_K(\delta_K) := \Omega_{\delta_K}. \quad (4)$$

2.2 Robustness certificates for solution sets to VIs

Given any K-multisample δ_K, we are interested in evaluating the robustness of the entire set of solutions Ω_{δ_K} in (3) to a previously unseen realization of δ. Before stating the main result of this paper, we recall the following definition that will be crucial for the remainder:

Definition 1 (Support Subsample) [7, Def. 2] *Given any $\delta_K \in \Delta^K$, a* support subsample *$S \subseteq \delta_K$ is a p-tuple of unique elements of δ_K, i.e., $S := \{\delta^{(i_1)}, \ldots, \delta^{(i_p)}\}$, $i_1 < \ldots < i_p$, that gives the same solution as the original sample, i.e., $\Theta_p(\delta^{(i_1)}, \ldots, \delta^{(i_p)}) = \Theta_K(\delta^{(1)}, \ldots, \delta^{(K)})$.*

Here, let $\Upsilon_K : \Delta^K \rightrightarrows \mathcal{K}$ be any algorithm returning a p-tuple such that $\{\delta^{(i_1)}, \ldots, \delta^{(i_p)}\}$ is a support subsample for δ_K, and let $s_K := |\Upsilon_K(\delta_K)|$. Note that s_K is itself a random variable since it depends on δ_K. Our main result characterizes the violation probability of Ω_{δ_K}, i.e., the solution set to the scenario-based VI in (2), as follows:

Theorem 1 *Fix $\beta \in (0,1)$, and let $\varepsilon : \mathcal{K} \cup \{0\} \to [0,1]$ be a function such that $\varepsilon(K) = 1$ and $\sum_{h=0}^{K-1} \binom{K}{h} (1 - \varepsilon(h))^{K-h} = \beta$. Then, for any mappings Θ_K, Υ_K and distribution $\mathbb{P}$, it holds that*

$$\mathbb{P}^K\{\delta_K \in \Delta^K \mid V(\Omega_{\delta_K}) > \varepsilon(s_K)\} \leq \beta, \tag{5}$$

where $V(\Omega_{\delta_K}) := \mathbb{P}\{\delta \in \Delta \mid \Omega_{\delta_K} \not\subseteq \Omega_{\delta_K \cup \{\delta\}}\}$ denotes the violation probability.

Note that the bound in (5) is a distribution-free, a-posteriori statement since s_K depends on the multisample extracted. In words, Theorem 1 implies that the probability that $\Omega_{\delta_K \cup \{\delta\}}$ differs from Ω_{δ_K} (as $\Omega_{\delta_K} \subseteq \Omega_{\delta_K \cup \{\delta\}}$ necessarily implies that $\Omega_{\delta_K} = \Omega_{\delta_K \cup \{\delta\}}$ – see also Lemma 2) is at most equal to $\varepsilon(s_K)$, with confidence at least $1 - \beta$, for an arbitrarily small $\beta \in (0,1)$. We give the proof of Theorem 1 in the next section, after first stating and proving some ancillary results.

3 The scenario approach to uncertain VIs

The scenario approach theory was initially conceived to provide a-priori out-of-sample feasibility guarantees associated with the solution to an uncertain convex optimization problem [3–5]. It has recently been extended to abstract decision-making problems through an a-posteriori assessment of the feasibility risk [6,7] by relying on the following two conditions:

i) The solution to the abstract decision-making problem is assumed to be unique;
ii) The decision taken while observing K realizations of the uncertainty δ is *consistent* with respect to (w.r.t.) all the extracted scenarios [7, Ass. 1].

Thus, we aim to follow the approach of [7] by focusing on a set of decisions see also [15, §5] for further clarifications on the concept of *decision*), extending conditions i)–ii) above to the solution set for the uncertain VI in (1). Since we focus on the entire set of solutions, for any K-multisample $\delta_K \in \Delta^K$, in view of Lemma 1 there naturally exists a unique set of solutions to $\mathrm{VI}(\mathcal{X}_{\delta_K}, F)$ (not necessarily a singleton though), and therefore the uniqueness of the solution returned by Θ_K in (4) holds by definition, thus addressing i). In the spirit of [21, Def. 2], we then envision that the set-oriented counterpart of the sequence of inclusions in [7, Ass. 1] shall be translated into a consistency property of Ω_{δ_K}, as defined next.

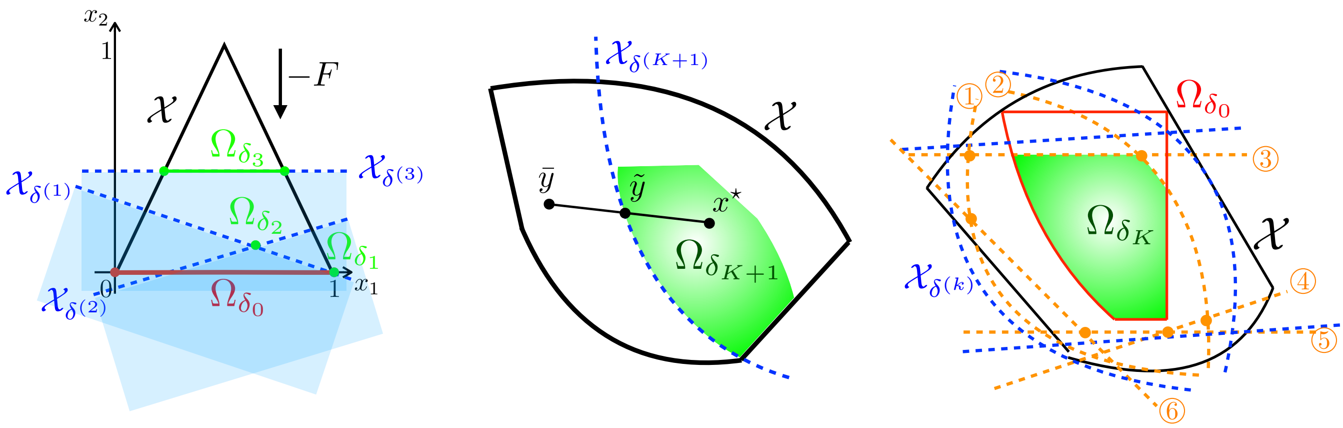
Definition 2 (Consistency of Solution Sets) *Given some $K \in \mathbb{N}$ and $\delta_K \in \Delta^K$, the solution set to $\mathrm{VI}(\mathcal{X}_{\delta_K}, F)$ is* consistent *with the collected scenarios if $\Theta_K(\delta_K) = \Omega_{\delta_K} \subseteq \mathcal{X}_{\delta^{(i)}}$, for all $i \in \mathcal{K}$.*

Definition 2 establishes that the set of solutions to $\mathrm{VI}(\mathcal{X}_{\delta_K}, F)$, Ω_{δ_K}, which is based on K scenarios, should be feasible for each of the sets $\mathcal{X}_{\delta^{(i)}}$, $i \in \mathcal{K}$, corresponding to each of the K realizations of the uncertain parameter. Thus, aiming to apply the bound in [7, Th. 1], we note that the mapping $\Theta_K(\cdot)$ in (4) is consistent with the realizations observed in the scenario-based VI in (2). For any $K \in \mathbb{N}$ and associated K-multisample $\delta_K \in \Delta^K$, indeed, we have that $\Theta_K(\delta_K) := \Omega_{\delta_K} \subseteq \cap_{i \in \mathcal{K}} \mathcal{X}_{\delta^{(i)}} \cap \mathcal{X}$, which implies that $\Theta_K(\delta_K) \subseteq \mathcal{X}_{\delta^{(i)}}$, for all $i \in \mathcal{K}$, thus directly falling within Definition 2. We will make use of these considerations in the proof of Theorem 1, along with the following assumption on the solution set Ω_{δ_K}.

Standing Assumption 2 *For all $K \in \mathbb{N}$ and $\delta_K \in \Delta^K$, $\mathrm{aff}(\Omega_{\delta_K}) = \mathrm{aff}(\Omega_{\delta_0})$.*

If the uncertain VI in (1) is defined in $\mathbb{R}^n$ and Ω_{δ_K} is a convex, m-dimensional set, then Standing Assumption 2 allows for $m < n$. In this sense, assuming $\mathrm{aff}(\Omega_{\delta_K}) = \mathrm{aff}(\Omega_{\delta_0})$ for any $\delta_K \in \Delta^K$, $K \in \mathbb{N}$, is weaker than, e.g., assuming $\mathrm{int}(\Omega_{\delta_K}) \neq \emptyset$ for every possible realization of δ_K. To clarify the role of Standing Assumption 2, we then introduce and discuss the following example.

Example 1 *Let us consider the two-dimensional case shown in Fig. 1a, where $F = \mathrm{col}(0,1)$, is monotone and $\mathcal{X}$ has a triangular shape. Here, $\Omega_{\delta_0} = \{x \in \mathbb{R}^2 \mid x_1 \in [0,1],\, x_2 = 0\}$, and its affine hull corresponds to the entire x_1-axis. After observing the first realization of δ, i.e., $\delta^{(1)}$, which introduces the set $\mathcal{X}_{\delta^{(1)}} = \{x \in \mathbb{R}^2 \mid -[1/3\ \ 1]^\top x \leq -1/3\}$, the solution set reduces to a singleton $\Omega_{\delta_1} = \{x \in \mathbb{R}^2 \mid x = \mathrm{col}(1,0)\}$. Here, Ω_1 has a smaller dimension compared to Ω_{δ_0}, despite its affine hull, i.e., the singleton itself, being a subset of the x_1-axis. Then, drawing a new sample $\delta^{(2)}$, which introduces the set $\mathcal{X}_{\delta^{(2)}} = \{x \in \mathbb{R}^2 \mid [1/3\ \ -1]^\top x \leq 1/15\}$, we have $\Omega_{\delta_2} = \{x \in \mathbb{R}^2 \mid x = \mathrm{col}(3/5, 2/15)\}$, which has the same dimension as Ω_{δ_1} but its affine hull is not a subset of*



(a) (b) (c)

Fig. 1. (a) Compared to Ω_{δ_0} (red line), every realization of δ (dashed blue lines, while the shaded cyan area denotes a region excluded by any $\mathcal{X}_{\delta^{(i)}}$, $i = 1, 2, 3$) results in a solution set Ω_{δ_i}, $i = 1, 2, 3$, that belongs to a different affine hull and/or on a space of lower dimension (green dots or line); (b) Schematic two-dimensional construction of the proof of Lemma 2, part (ii). Due to the convexity, there always exists some $\tilde{y} \in \mathcal{X}_{\delta_{K+1}}$, but $\tilde{y} \notin \mathrm{int}(\Omega_{\delta_{K+1}})$, that allows to construct a contradiction. In this case, $\tilde{y} \in \mathrm{bdry}(\Omega_{\delta_{K+1}})$; (c) The solution set to $\mathrm{VI}(\mathcal{X}_{\delta_K}, F)$, Ω_{δ_K} (green region), can be "shaped" by the set of constraints, $\mathcal{X}_{\delta^{(i)}}$, $i \in \mathcal{K}$ (dashed blue lines). According to Definition 1, the number of support subsamples for δ_K w.r.t. to Ω_{δ_K} is, in general, smaller compared to $|\mathcal{X}_{\delta_K}|$ (dashed orange lines, whose intersections are defined by orange dots).

$\mathrm{aff}(\Omega_{\delta_0})$. *Finally, the third sample,* $\delta^{(3)}$*, introduces the set* $\mathcal{X}_{\delta^{(3)}} = \{x \in \mathbb{R}^2 \mid [0 \ -1]^\top x \leq -1/2\}$*, and hence we have* $\Omega_{\delta_3} = \{x \in \mathbb{R}^2 \mid x_1 \in [1/4, 3/4], x_2 = 1/2\}$*. Here,* Ω_{δ_3} *has the same dimension of* Ω_{δ_0} *but a different affine hull, i.e., the* x_1*-axis translated to* $x_2 = 1/2$*. Standing Assumption 2 is meant to rule out all these possible scenarios, allowing only for samples that "shape"* $\mathrm{aff}(\Omega_{\delta_0})$ *without altering its dimension.*

Example 1 provides insight on translating Standing Assumption 2 to a condition on the probability space Δ, as it represents situations that can generally happen with non-zero probability. Specifically, let Δ be a subset of $\mathbb{R}^2$ with $\delta = \mathrm{col}(a, b)$, and let $a \in \mathbb{R}$ parametrize the slope and $b \in \mathbb{R}$ the offset of the halfspaces introduced by every scenario, i.e., $\mathcal{X}_\delta = \{x \in \mathbb{R}^2 \mid [a \ 1]\, x \leq b\}$. Then, for any distribution that admits a density, we can find non-zero intervals for a and b such that the i.i.d. scenarios δ can be extracted from some $\Delta' \subseteq \Delta$, determined by the values of a and b themselves, in order to meet Standing Assumption 2, thus ruling out the pathological cases shown in Example 1. In the case the samples are extracted from Δ', note that the guarantees would hold for the probability measure that is induced by this restriction. Alternatively, if Standing Assumption 2 is not satisfied for all multisamples, then we can still claim that with confidence at most β, if Standing Assumption 2 is satisfied, then the probability of violation is greater than $\varepsilon(s_K)$. To achieve this, in the statement of Theorem 1, we can restrict the space of multisamples to the ones for which Standing Assumption 2 is satisfied (see how to treat infeasible problem instances in [3,4]). Moreover, by adopting restrictions on Δ, Standing Assumption 2 allows us to address the strongly monotone case, where $\mathrm{VI}(\mathcal{X} \cap \mathcal{X}_\delta, F)$ has a unique solution, for all $\delta \in \Delta$.

Given some $K \in \mathbb{N}$, let $\Omega_{\delta_{K+1}} := \Omega_{\delta_K \cup \{\delta^{(K+1)}\}}$ be the solution set to the scenario-based VI in (2) after observing the $(K+1)$-th realization of δ, i.e., the feasible set of the VI shrinks to $\mathcal{X}_{\delta_{K+1}} := \mathcal{X}_{\delta_K} \cap \mathcal{X}_{\delta^{(K+1)}}$, for some $\delta^{(K+1)} \in \Delta$. We have the following preliminary result.

Lemma 2 *For all* $K \in \mathbb{N}_0$ *and for all* $\delta_K \in \Delta^K$, $\Omega_{\delta_{K+1}} = \Omega_{\delta_K} \cap \mathcal{X}_{\delta^{(K+1)}}$.

PROOF. We split the proof into two different inclusions. Specifically, we first prove (i) that $\Omega_{\delta_K} \cap \mathcal{X}_{\delta^{(K+1)}} \subseteq \Omega_{\delta_{K+1}}$, and then (ii) that $\Omega_{\delta_K} \cap \mathcal{X}_{\delta^{(K+1)}} \supseteq \Omega_{\delta_{K+1}}$.
(i) We show that if $x^\star \in \Omega_{\delta_K}$ and $x^\star \in \mathcal{X}_{\delta^{(K+1)}}$, then $x^\star \in \Omega_{\delta_{K+1}}$. In view of Standing Assumptions 1, given an arbitrary $K \in \mathbb{N}_0$ and related $\delta_K \in \Delta^K$, $\mathcal{X}_{\delta_K}$ is a compact and convex set, as it is finite intersection of convex sets. Then, a vector $x^\star \in \mathcal{X}_{\delta_K}$ is a solution to $\mathrm{VI}(\mathcal{X}_{\delta_K}, F)$ if and only if $x^\star \in \mathrm{argmin}_{y \in \mathcal{X}_{\delta_K}} y^\top F(x^\star)$ [12, §1.2]. Since the uncertain parameter enters in the constraints only, every sample $\delta^{(K+1)} \in \Delta$ introduces an additional set of convex constraints, i.e., $\mathcal{X}_{\delta_{K+1}} = \mathcal{X}_{\delta_K} \cap \mathcal{X}_{\delta^{(K+1)}} \subseteq \mathcal{X}_{\delta_K}$, which is compact and convex as well. Thus, it follows immediately that, if $x^\star \in \mathcal{X}_{\delta^{(K+1)}}$, then $x^\star \in \mathcal{X}_{\delta_{K+1}}$. Therefore, $x^\star \in \mathrm{argmin}_{y \in \mathcal{X}_{\delta_K} \cap \mathcal{X}_{\delta^{(K+1)}}} y^\top F(x^*)$, which by definition implies that $x^\star \in \Omega_{\delta_{K+1}}$.
(ii) We first prove that, if $x^\star \in \mathrm{relint}(\Omega_{\delta_{K+1}})$, then $x^\star \in \Omega_{\delta_K}$. The case where $x^\star \in \mathrm{bdry}(\Omega_{\delta_{K+1}})$ will be treated in the sequel. Let us recall that, in view of [24, Cor. 1.6.1], for any given m-dimensional convex set $\mathcal{S}$ in $\mathbb{R}^n$, $m \leq n$, there always exists an affine transforma-

tion which carries $\mathrm{aff}(\mathcal{S})$ onto the subspace $V := \{x = (z_1, \ldots, z_m, z_{m+1}, \ldots, z_n)^\top \in \mathbb{R}^n \mid z_{m+1} = \ldots = z_n = 0\}$. Therefore, as closures and relative interiors are preserved under one-to-one affine transformations of $\mathbb{R}^n$ onto itself, we can limit our attention to the case where $\Omega_{\delta_{K+1}}$, and hence Ω_{δ_K} (since $\mathrm{aff}(\Omega_{\delta_{K+1}}) = \mathrm{aff}(\Omega_{\delta_K}) = \mathrm{aff}(\Omega_{\delta_0})$ from Standing Assumption 2), is n-dimensional so that $\mathrm{relint}(\Omega_{\delta_{K+1}}) = \mathrm{int}(\Omega_{\delta_{K+1}})$. Now, for the sake of contradiction, let $x^\star \in \mathcal{X}_{\delta_K} \cap \mathcal{X}_{\delta^{(K+1)}}$ be any point such that $x^\star \in \mathrm{int}(\Omega_{\delta_{K+1}})$, but $x^\star \notin \Omega_{\delta_K}$. Since $x^\star \in \mathcal{X}_{\delta_K}$, $x^\star \notin \Omega_{\delta_K}$ implies that there exists some $\bar{y} \in \mathcal{X}_{\delta_K}$, with $\bar{y} \neq x^\star$, such that the VI is not satisfied, i.e., $(\bar{y} - x^\star)^\top F(x^\star) < 0$. Given the convexity of the sets involved, there must exist some $\lambda \in (0, 1)$ that allows one to construct some $\tilde{y} = \lambda x^\star + (1-\lambda)\bar{y}$ such that $\tilde{y} \in \mathcal{X}_{\delta_K} \cap \mathcal{X}_{\delta^{(K+1)}}$, but $\tilde{y} \notin \mathrm{int}(\Omega_{\delta_{K+1}})$ (see Fig. 1b for a graphical representation). Therefore, since $x^\star \in \mathrm{int}(\Omega_{\delta_{K+1}})$, it shall satisfy $(\tilde{y} - x^\star)^\top F(x^\star) \geq 0$, which leads to $(1-\lambda)(\bar{y} - x^\star)^\top F(x^\star) \geq 0$ that clearly generates a contradiction, since $(1-\lambda) > 0$. It remains to show the claim when $x^\star \in \mathrm{bdry}(\Omega_{\delta_{K+1}})$. Since $\mathrm{relint}(\Omega_{\delta_{K+1}}) \neq \emptyset$ as $\Omega_{\delta_{K+1}}$ is nonempty, and since the involved sets are closed and convex, for any $x^\star \in \mathrm{bdry}(\Omega_{\delta_{K+1}})$ we can construct a convergent sequence $\{x_t\}_{t \in \mathbb{N}}$ such that, for all $t \in \mathbb{N}$, $x_t \in \mathrm{relint}(\Omega_{\delta_{K+1}}) \subseteq \Omega_{\delta_K}$, and $\{x_t\}_{t\in\mathbb{N}} \to x^\star$, implying that $x^\star \in \Omega_{\delta_K}$. Specifically, given any $\bar{x} \in \mathrm{relint}(\Omega_{\delta_{K+1}})$, in view of [24, Th. 6.1], for all $t \geq 1$, any term of the sequence $x_t := \frac{1}{t}\bar{x} + (1 - \frac{1}{t})x^\star \in \Omega_{\delta_K} \cap \mathcal{X}_{\delta^{(K+1)}}$ belongs to $\mathrm{relint}(\Omega_{\delta_{K+1}})$. Therefore, the inclusion $\Omega_{\delta_{K+1}} \subseteq \Omega_{\delta_K}$ directly follows.

A consequence of Lemma 2 is that $\Theta_0 =: \Omega_{\delta_0} \supseteq \Omega_{\delta_1} \supseteq \ldots \supseteq \Omega_{\delta_K} =: \Theta_K(\delta_K)$. Moreover, the intrinsic consistency of the set Ω_{δ_K} implies that by introducing additional constraints, the effect of the uncertain parameter is to shrink the feasible set $\mathcal{X}_{\delta_K}$ of the scenario-based VI in (2), and therefore the set of solutions can only shrink, accordingly (see Fig. 1c for a graphical illustration).

3.1 Proof of Theorem 1 and discussion

PROOF. For any $K \in \mathbb{N}$, $\delta_K \in \Delta^K$, we know that Ω_{δ_K} is consistent w.r.t. the collected scenarios, δ_K. In view of the definition in (4), indeed, we have that $\Omega_{\delta_K} \subseteq \cap_{i\in\mathcal{K}} \mathcal{X}_{\delta^{(i)}} \cap \mathcal{X}$, which implies that $\Omega_{\delta_K} \subseteq \mathcal{X}_{\delta^{(i)}}$, for all $i \in \mathcal{K}$, thus directly satisfying Definition 2. Then, by applying [7, Th. 1], we have that $\mathbb{P}^K\{\delta_K \in \Delta^K \mid \mathbb{P}\{\delta \in \Delta \mid \Omega_{\delta_K} \not\subseteq \mathcal{X}_\delta\} > \varepsilon(s_K)\} \leq \beta$. However, by Lemma 2, $\Omega_\delta = \Omega_{\delta_K} \cap \mathcal{X}_\delta$. Therefore, $\Omega_{\delta_K} \not\subseteq \mathcal{X}_\delta$ is equivalent to $\Omega_\delta \neq \Omega_{\delta_K}$, and since the set of solutions can only shrink once a new scenario is added, this is in turn equivalent to $\Omega_\delta \not\subseteq \Omega_{\delta_K}$. Thus, in view of the definition of $V(\cdot)$, we finally have that $\mathbb{P}^K\{\delta_K \in \Delta^K \mid V(\Omega_{\delta_K}) > \varepsilon(s_K)\} \leq \beta$.

A more direct expression of the critical parameter $\varepsilon(\cdot)$ can be obtained by splitting the confidence parameter β evenly among the K terms within the summation, i.e.,

$$\varepsilon(h) = \begin{cases} 1 & \text{if } h = K, \\ 1 - \left(\beta / \left(K \binom{K}{h}\right)\right)^{1/K-h} & \text{otherwise.} \end{cases} \tag{6}$$

Remark 1 *In the case of a non-degenerate* VI*, the bound $\varepsilon(\cdot)$ could be improved by means of the* wait-and-judge *analysis in [6]. Specifically, in view of [6, Th. 2], we can replace the expression for $\varepsilon(\cdot)$ in (6) with $\varepsilon(h) = 1 - t(h)$, where $t(h)$ is shown to be the unique solution in $(0, 1)$ to $\frac{\beta}{K+1}\sum_{m=h}^{K}\binom{m}{h}t^{m-h} - \binom{K}{h}t^{K-h} = 0$. However, note that the non-degeneracy condition is in general difficult to verify even in convex optimization settings [6, 14], a challenge that becomes more involved for* VIs.

Similarly to $\Upsilon(\cdot)$, let us suppose to have available an algorithm that allows us to compute a support subsamples for δ_K associated with the feasible set $\mathcal{X}_{\delta_K}$.

Proposition 1 *Given any $K \in \mathbb{N}_0$ and $\delta_K \in \Delta^K$, let s_K and v_K be the cardinality of the support subsample for δ_K w.r.t. Ω_{δ_K} and $\mathcal{X}_{\delta_K}$, respectively. Then, $s_K \leq v_K$.*

PROOF. For every $K \in \mathbb{N}_0$ and $\delta_K \in \Delta^K$, by Definition 1 a sample $\delta^{(k)}$ is of support for δ_K w.r.t. $\mathcal{X}_{\delta_K}$ if $\mathcal{X}_{\delta^{(k)}}$ is active on $\mathrm{bdry}(\mathcal{X}_{\delta_K})$, i.e., $\mathrm{bdry}(\mathcal{X}_{\delta^{(k)}}) \cap \mathrm{bdry}(\mathcal{X}_{\delta_K}) \neq \emptyset$. On the other hand, $\delta^{(k)}$ is of support w.r.t. Ω_{δ_K} if $\mathrm{bdry}(\mathcal{X}_{\delta^{(k)}}) \cap \Omega_{\delta_K} \neq \emptyset$ (see, e.g., Fig. 1c). Since $\Omega_{\delta_K} \subseteq \mathcal{X}_{\delta_K} := \cap_{k\in\mathcal{K}}\mathcal{X}_{\delta^{(k)}} \cap \mathcal{X}$, those samples that are of support for δ_K w.r.t. Ω_{δ_K}, are necessarily of support w.r.t. $\mathcal{X}_{\delta_K}$, but not vice-versa, and hence $s_K \leq v_K$.

Under Proposition 1, Theorem 1 improves over [23], where $V(\Omega_{\delta_K}) > \varepsilon(v_K)$ was claimed with confidence at most β. The latter is since $\varepsilon(s_K) \leq \varepsilon(v_K)$ as $\varepsilon(\cdot)$ is nondecreasing. Moreover, within the set-oriented framework proposed in §2, as evident from (5), to bound the feasibility risk $V(\cdot)$ of the entire set of solutions Ω_{δ_K}, one does not need an explicit characterization of Ω_{δ_K}, namely some mapping $\Theta_K(\cdot)$, but rather the number of support subsamples s_K, computed through an algorithm $\Upsilon(\cdot)$.

3.2 Computation of the support subsample: the case of affine constraints

The general setting considered so far, i.e., pseudomonotone mapping F and convex constraint set $\mathcal{X}_{\delta_K}$, for any $\delta_K \in \Delta^K$, poses several challenges in designing an efficient procedure to compute the number of support subsamples w.r.t. Ω_{δ_K}. We therefore introduce the following additional assumption that restricts attention to the class of linearly constrained, pseudomonotone VIs.

Algorithm 1 Computation of the cardinality of the support subsample w.r.t. Ω_{δ_K}

Initialization: Set $s_K := 0$, identify $\mathcal{A}_K := \{k \in \mathcal{K} \mid \mathrm{bdry}(\mathcal{X}_{\delta^{(k)}}) \cap \mathrm{bdry}(\mathcal{X}_{\delta_K}) \neq \emptyset\}$

Iteration ($i \in \mathcal{A}_K$)**:**

(S1) Run $\Phi(\delta_{K,i})$ to solve $\mathrm{VI}(\mathcal{X}_{\delta_K} \cap \mathrm{bdry}(\mathcal{X}_{\delta^{(i)}}), F)$

(S2) If $\Phi(\delta_{K,i}) \neq \emptyset$, set $s_K := s_K + 1$

Assumption 3 *Let $\mathcal{X} := \{\boldsymbol{x} \in \mathbb{R}^n \mid C\boldsymbol{x} \leq d\}$, $C \in \mathbb{R}^{m\times n}$ and $d \in \mathbb{R}^m$, with $\mathrm{rank}(C) = n$, and, for all $\delta \in \Delta$, $\mathcal{X}_\delta := \{\boldsymbol{x} \in \mathbb{R}^n \mid A(\delta)\boldsymbol{x} \leq b(\delta)\}$, $A : \Delta \to \mathbb{R}^{r\times n}$ and $b : \Delta \to \mathbb{R}^r$.*

Then, given any K-multisample δ_K, let $\Phi : \Delta^K \rightrightarrows \Omega_{\delta_K}$ be any mapping that returns a (set of) solution(s) to $\mathrm{VI}(\mathcal{X}_{\delta_K}, F)$. The procedure $\Phi(\cdot)$ is run in (S1) to verify whether (at least) one solution to the VI with constraints involving $\mathcal{X}_{\delta_K} \cap \mathrm{bdry}(\mathcal{X}_{\delta^{(i)}})$ exists, thus increasing s_K in case of affirmative answer in (S2). The next result states that, even without knowing Ω_{δ_K}, Algorithm 1 returns the cardinality of a support subsample for δ_K w.r.t. Ω_{δ_K}.

Proposition 2 *Let Assumption 3 hold true. Given any $K \in \mathbb{N}$ and $\delta_K \in \Delta^K$, Algorithm 1 returns $s_K^\star$, the cardinality of a support subsample δ_K w.r.t. Ω_{δ_K}.*

PROOF. First note that, in view of Assumption 3, $\mathcal{A}_K$ forms a support subsample for δ_K w.r.t. $\mathcal{X}_{\delta_K}$. Then, by following the considerations adopted within the proof of Proposition 1, every $\delta^{(k)}$, $k \in \mathcal{A}_K$, is of support also w.r.t. to Ω_{δ_K} if and only if $\mathrm{bdry}(\mathcal{X}_{\delta^{(k)}}) \cap \Omega_{\delta_K} \neq \emptyset$. To determine this, it is sufficient to compute a solution (if any) on the active region of $\mathcal{X}_{\delta_K}$ associated with $\mathcal{X}_{\delta^{(k)}}$. Then, s_K increases only if $\Phi(\delta_{K,k}) \neq \emptyset$, excluding all those samples such that $\mathcal{X}_{\delta^{(k)}}$ does not intersect Ω_{δ_K}.

Remark 2 *Algorithm 1 requires one to run the adopted solution algorithm $\Phi(\delta_K)$ a total of $|\mathcal{A}_K|$-times, with $|\mathcal{A}_K| \leq K$. This improves w.r.t. the greedy algorithms proposed in [7, §II] and [22, §III], which would require one to run $\Phi(\delta_K)$ K-times. On the other hand, we remark that the greedy algorithm applies more generally, i.e., not necessarily only in the case of affine constraints. In addition, if reducing computation time is a consideration one may skip (S1) and (S2) in Algorithm 1 altogether, since by construction of Algorithm 1 $s_K^\star \leq |\mathcal{A}_K|$, thus exploiting Theorem 1 with $|\mathcal{A}_K|$ in place of $s_K^\star$. The latter, however, may result in a more conservative bound.*

From a computational point of view, we note that Assumption 3 is needed for two main reasons: i) Evaluating a solution to the VI on the boundary of a convex set, i.e., (S1), may require solution of a VI defined over a nonconvex domain; ii) The initialization step requires one to identify the minimal number of active constraints. While item i) prevents us from trivially extending Algorithm 1 to the case of general convex constraints (the literature on solution algorithms with convergence guarantees for the nonconvex case is not extensive), item ii) can be equivalently seen as a problem of removing redundant halfspaces, an offline task that can be efficiently accomplished in polynomial time (see, e.g., [1, 2, 28]).

4 Case study: Charging coordination of PEVs

The problem of coordinating the day-ahead charging of a fleet of PEVs, originally introduced in [20], can be modelled as a noncooperative GNEP [8, 9]. Specifically, for each PEV $j \in \mathcal{J}$, we consider a discrete-time linear dynamical system $s_j(t+1) = s_j(t) + b_j x_j(t)$, $t \in \mathbb{N}$, where $s_j \in [0,1]$ is the State of Charge (SoC), i.e., $s_j = 1$ represents a fully charged battery, while $s_j = 0$ a completely discharged one; $x_j(t) \in [0,1]$ is the charging control input at the specific time instant t, and $b_j > 0$ denotes the charging efficiency. The goal of each PEV is to acquire a charge amount above γ_j within a finite charging horizon $\mathcal{T} := \{0, \ldots, T-1\}$, with $T = 24$, thus satisfying $\sum_{t\in\mathcal{T}} x_j(t) = \mathbf{1}_T^\top x_j \geq \gamma_j$, $x_j := \mathrm{col}((x_j(t))_{t\in\mathcal{T}}) \in \mathbb{R}^T$, while minimizing its charging cost, $p(\boldsymbol{x})^\top x_j$. Here, $p : \mathbb{R}^T_{\geq 0} \to \mathbb{R}^T_{\geq 0}$, denotes the electricity price function over $\mathcal{T}$, which for simplicity we assume to be affine in the aggregate demand of energy associated with the set of PEVs, i.e., $p(\boldsymbol{x}) := \alpha\sigma(\boldsymbol{x}) + \eta$, with $\sigma(\boldsymbol{x}) := \sum_{j\in\mathcal{J}} x_j \in \mathbb{R}^T$, for some $\alpha > 0$ and $\eta \in \mathbb{R}^T_{\geq 0}$. Moreover, due to the intrinsic limitations of the grid capacity $d_{\max} > 0$, we assume that the amount of energy required in each single time period by both the PEVs and uncertain non-PEV loads should not be greater than $d_{\max}$. This translates into a constraint on the total demand of the PEVs, i.e., $\delta(t) + \sum_{j\in\mathcal{J}} x_j(t) \in [0, d_{\max}]$, for all $t \in \mathcal{T}$, where δ is a random variable characterized by an unknown support set $\Delta \subseteq \mathbb{R}^T$ and probability distribution $\mathbb{P}$, although we have direct access to K-multisamples from $2010-2019$ daily energy profiles [17]. The (uncertain) GNEP thus coincides with the following set of optimization problems

$$\forall j \in \mathcal{J} : \begin{cases} \min\limits_{x_j \in [0,1]^T} & (\alpha\sigma(\boldsymbol{x}) + \eta)^\top x_j \\ \quad \text{s.t.} & \delta + \sigma(\boldsymbol{x}) \leq \mathbf{1}_T d_{\max},\ \forall \delta \in \Delta, \\ & A_j x_j \leq c_j,\ d_{\mathrm{nom}} + \sigma(\boldsymbol{x}) \leq \mathbf{1}_T d_{\max}, \end{cases} \tag{7}$$

where $A_j := \mathrm{col}(-B_j, B_j, -\mathbf{1}_T^\top) \in \mathbb{R}^{(2T+1)\times T}$, $B_j \in \mathbb{R}^{T\times T}$ is matrix with all entries in the lower triangular part equal to b_j, $c_j := \mathrm{col}(\mathbf{1}_T s_j(0), \mathbf{1}_T(1 - s_j(0)), -\gamma_j) \in \mathbb{R}^{2T+1}$, $s_j(0) \in [0,1]$ is a given initial SoC, and $d_{\mathrm{nom}} \in \mathbb{R}^T_{\geq 0}$ represents the nominal non-PEV daily energy demand that is inferred from available data [17]. We note that the game mapping $F(\boldsymbol{x}) := \mathrm{col}(\nabla_{x_j}((\alpha\sigma(\boldsymbol{x}) + \eta)^\top x_j)_{j\in\mathcal{J}})$, which allows us to define the VI whose solution set determines the

Table 1
Simulation parameters

Name	Description	Value
b_j	Charging efficiency	$[0.075, 0.25]$
$s_j(0)$	Initial SoC of battery	$[0.1, 0.4]$
$s_j(T)$	Desired SoC of battery	$[0.7, 1]$
γ_j	Required charge amount	$[1.62, 7.49]$
α	Inverse of price elasticity	0.01
d_{nom}	Non-PEV demand	Average over daily profiles in 2019 [17]
$d_{\max}$	Grid power capacity	$2 \cdot \max_{t\in\mathcal{T}} d_{\text{nom}}(t)$

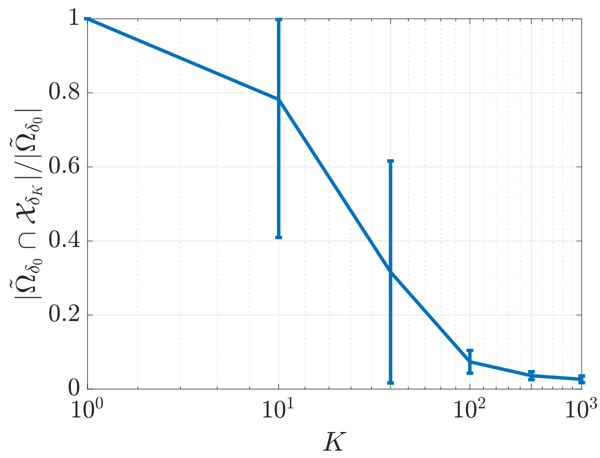


Fig. 2. The solid blue line represents the average of $|\tilde{\Omega}_{\delta_0} \cap \mathcal{X}_{\delta_K}|/|\tilde{\Omega}_{\delta_0}|$ over 100 numerical experiments, while the vertical blue lines the standard deviation.

v-GNE of the game [11, Def. 3], turns out to be affine in $\boldsymbol{x}$. Specifically, $F(\boldsymbol{x}) = M\boldsymbol{x} + q$, where $M \in \mathbb{R}^{NT\times NT}$ has entries all equal to α, while $q := \mathbf{1}_N \otimes \eta \in \mathbb{R}^{NT}$. Note that, for any $\alpha > 0$, $F(\cdot)$ is a monotone mapping. Thus, based on K observations of historical data, the GNEP in (7) admits a scenario-based counterpart, i.e.,

$$\forall j \in \mathcal{J}: \begin{cases} \min\limits_{x_j\in[0,1]^T} & (\alpha\sigma(\boldsymbol{x}) + \eta)^\top x_j \\ \text{s.t.} & \delta^{(i)} + \sigma(\boldsymbol{x}) \leq \mathbf{1}_T d_{\max},\ \forall i \in \mathcal{K}, \\ & A_j x_j \leq c_j,\ d_{\text{nom}} + \sigma(\boldsymbol{x}) \leq \mathbf{1}_T d_{\max}, \end{cases} \tag{8}$$

for which we aim at quantifying the robustness of Ω_{δ_K}, solution set to $\mathrm{VI}(\mathcal{X}_{\delta_K}, F)$. Here, $\mathcal{X}_{\delta_K} := \mathcal{X} \cap_{i\in\mathcal{K}} \mathcal{X}_{\delta^{(i)}}$, $\mathcal{X} := \prod_{j\in\mathcal{J}} \mathcal{X}_j \cap \{\boldsymbol{x} \in \mathbb{R}^{NT} \mid d_{\text{nom}} + \sigma(\boldsymbol{x}) \leq \mathbf{1}_T d_{\max}\}$, $\mathcal{X}_j := \{x_j \in [0,1]^T \mid A_j x_j \leq c_j\}$, and $\mathcal{X}_{\delta^{(i)}} := \{\boldsymbol{x} \in \mathbb{R}^{NT} \mid \delta^{(i)} + \sigma(\boldsymbol{x}) \leq \mathbf{1}_T d_{\max}\}$, $i \in \mathcal{K}$.

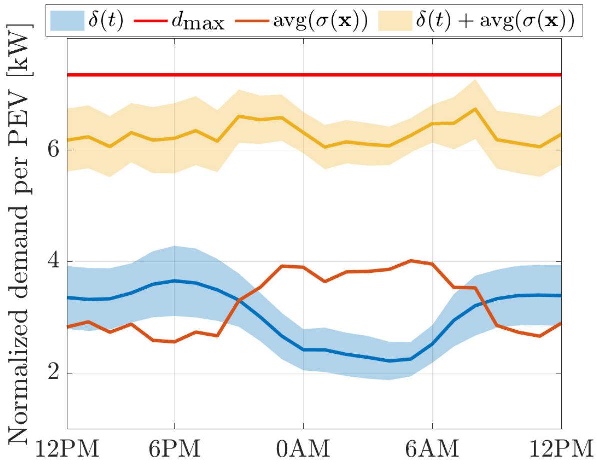


Fig. 3. Average behaviour of the fleet of PEVs, computed across the estimated set of solutions $\tilde{\Omega}_{\delta_{10^3}}$. The overall demand, affected by the uncertainty, meets the grid limitations.

Table 2
Robustness certificate (5) and empirical violation probability

K	$\lvert\mathcal{A}_K\rvert$	$s^\star_K$	$\varepsilon(s^\star_K)$	$V_{\max}(\tilde{\Omega}_{\delta_K})$	$\mathrm{avg}(V_{\max}(\tilde{\Omega}_{\delta_K}))$
10^1	24	18	0.305	0.016	0.013
10^2	24	20	0.055	$1.9 \cdot 10^{-3}$	$1.3 \cdot 10^{-3}$
10^3	24	24	$7 \cdot 10^{-3}$	$0.8 \cdot 10^{-3}$	$0.5 \cdot 10^{-3}$

4.1 Numerical simulations

The numerical simulations are run in Matlab by using Gurobi [18] as a solver, on a laptop with a Quad-Core Intel Core i5 2.4 GHz CPU and 8 Gb RAM.

We first support the consistency of Ω_{δ_K} numerically. Specifically, we estimate Ω_{δ_0} by computing 10^3 different solutions to $\mathrm{VI}(\mathcal{X}_{\delta_0}, F)$, thus obtaining $\tilde{\Omega}_{\delta_0}$, with the numerical parameters reported in Table 1, $N = 20$ and $\eta = \mathbf{0}_T$. Every solution is computed by means of the projection-type method in [26], initialized with different conditions, which takes around 6.19[s] on average to compute a solution with a precision in norm of 10^{-6}. Given the linearity of the constraints, this value is representative for solving (`S1`) in Algorithm 1. Thus, as illustrated in Fig. 2, the average number of solutions contained in $\tilde{\Omega}_{\delta_K}$ over 100 numerical experiments, normalized w.r.t. $\tilde{\Omega}_{\delta_0}$, shrinks as K grows. An example of aggregate behaviour for the PEVs is reported in Fig. 3.

For any $K \in \mathbb{N}_0$, the feasible set of the scenario-based counterpart of (7) satisfies Assumption 3. Thus, in Table 2 we compare the output of the procedure summarized in Algorithm 1 to compute the cardinality $s^\star_K$ of the support subsample w.r.t. Ω_{δ_K}, for different values of K, where $\mathcal{A}_K$ gathers the indices of the active hyperplanes of $\cap_{i\in\mathcal{K}} \mathcal{X}_{\delta^{(i)}}$. The bound on the violation probability is then computed by the function $\varepsilon(\cdot)$ in (6) with

$\beta = 10^{-6}$. Note that Algorithm 1 requires us to run $\Phi(\cdot)$ only $|\mathcal{A}_K|$-times, which represents a noticeable improvement compared to the greedy algorithm proposed in [7, 22], which would require running $\Phi(\cdot)$ TK-times. On the other hand, the offline initialization step with $K = 10^3$, which translates into 25964 linear inequalities, takes around 5893[s] to identify the set of constraints defining $\mathcal{X}_{\delta_{10^3}}$, for a total of 6041[s] to return $s^\star_K$. The last two columns of Table 2 involve estimating empirically the violation probability. To achieve this, we first estimate Ω_{δ_0}, by computing 10^3 different solutions of $\mathrm{VI}(\mathcal{X}_{\delta_0}, F)$ to get $\tilde{\Omega}_{\delta_0}$, which takes around 6010[s]. We then intersect this discrete set with the constraints sets obtained with the K samples of the offline initialization step, thus creating the estimate $\tilde{\Omega}_{\delta_K} = \tilde{\Omega}_{\delta_0} \cap \mathcal{X}_{\delta_K}$. We then generate new sets of constraints associated to validation samples and intersect them with $\tilde{\Omega}_{\delta_K}$ to estimate the empirical violation probability associated to such a set, which is, as expected, lower than the theoretical bound in Theorem 1. In the last two columns of Table 2, we report the maximum and the average value of this empirical quantity as computed across 100 repetitions of this procedure.

4.2 *A comparison with Monte Carlo type validations*

We now compare the proposed theoretical guarantees with Monte Carlo-like considerations. Specifically, given any $K \in \mathbb{N}$ and associated K-multisample, $\delta_K \in \Delta^K$, the bound in Theorem 1

$$\begin{aligned} \mathbb{P} &:= \mathbb{P}\{\delta \in \Delta \mid \Omega_{\delta_K} \subseteq \Omega_{\delta_K \cup \{\delta\}}\} \\ &= 1 - V(\Omega_{\delta_K}) \geq 1 - \varepsilon(s_K) \end{aligned} \tag{9}$$

holds true with confidence at least $1 - \beta$. The left-hand side of (9) admits a sample-based counterpart $\hat{\mathbb{P}}_{K_v} = 1 - \hat{V}_{K_v}(\Omega_{\delta_K}) := (1/K_v)\sum_{i=1}^{K_v} \iota_{\Omega_{\delta_K}}(\mathcal{X}_{\delta^{(i)}})$ (see [27, §3.1]), which is computed on the basis of K_v *new* validation samples (i.e., in addition to the K needed to compute Ω_{δ_K}), with $\iota_{\Omega_{\delta_K}}(\mathcal{X}_{\delta^{(i)}}) = 1$ if $\Omega_{\delta_K} \subseteq \mathcal{X}_{\delta^{(i)}}$, 0 otherwise. Note that $\hat{V}_{K_v}(\cdot)$ coincides with the empirical violation probability, estimated using the same procedure as the one employed for the last two columns of Table 2, using K_v validation samples. Then, by adopting, e.g., the Chernoff bound for some $\hat{\varepsilon}$, $\hat{\beta} \in (0, 1)$, selecting $K_v = \lceil \frac{\hat{\varepsilon}^{-2}}{2} \ln \frac{2}{\hat{\beta}} \rceil$ leads to the following bound characterizing the empirical distribution $\hat{\mathbb{P}}_{K_v}$:

$$\begin{aligned} &\mathbb{P}^{K_v}\{\delta_{K_v} \in \Delta^{K_v} \mid |\hat{\mathbb{P}}_{K_v} - \mathbb{P}| \leq \hat{\varepsilon}\} \geq 1 - \hat{\beta} \\ &\Leftrightarrow \mathbb{P}^{K_v}\{\delta_{K_v} \in \Delta^{K_v} \mid |V(\Omega_{\delta_K}) - \hat{V}_{K_v}(\Omega_{\delta_K})| \leq \hat{\varepsilon}\} \geq 1 - \hat{\beta} \\ &\Rightarrow \mathbb{P}^{K_v}\{\delta_{K_v} \in \Delta^{K_v} \mid V(\Omega_{\delta_K}) \leq \hat{V}_{K_v}(\Omega_{\delta_K}) + \hat{\varepsilon}\} \geq 1 - \hat{\beta} \end{aligned} \tag{10}$$

namely the empirical violation probability $\hat{V}_{K_v}(\Omega_{\delta_K})$ differs at most $\hat{\varepsilon}$ from the actual violation probability $V(\Omega_{\delta_K})$, with confidence at least $1 - \hat{\beta}$. We are now able to contrast the bound provided in Theorem 1 and the one in (10). Specifically, Theorem 1 ensures that $\mathbb{P}^K\{\delta_K \in \Delta^K \mid V(\Omega_{\delta_K}) \leq \varepsilon(s_K)\} \geq 1 - \beta$, whereas (10) results in $\mathbb{P}^{K_v}\{\delta_{K_v} \in \Delta^{K_v} \mid V(\Omega_{\delta_K}) \leq \hat{V}_{K_v}(\Omega_{\delta_K}) + \hat{\varepsilon}\} \geq 1 - \hat{\beta}$ For instance, let us choose $K = 10^3$, $\hat{\beta} = \beta = 10^{-6}$ and $\hat{\varepsilon} = 6 \cdot 10^{-3}$ such that $\hat{V}_{K_v}(\Omega_{\delta_{10^3}}) + \hat{\varepsilon} = \varepsilon(s^\star_{10^3}) = 7 \cdot 10^{-3}$ from Table 2. It turns out that the Monte Carlo approach provides the same probabilistic statement but requires $K_v = 201510$ additional validation samples to provide a bound on $V(\Omega_{\delta_K})$ comparable to the theoretical one in (9). We conclude by summarizing the main differences:

i) To compute $\hat{\mathbb{P}}_{K_v}$, one needs a formal characterization of Ω_{δ_K}, which is rarely available. Even if it is, since Ω_{δ_K} is a continuous set one would need to compute the probability of violation for an uncountable number of points. This is intractable, hence we can approximate $\hat{\mathbb{P}}_{K_v}$ numerically by computing some estimate of Ω_{δ_K} by gridding the space, resulting in the discrete set $\tilde{\Omega}_{\delta_K}$, and then computing the probability of violation for each grid point, as performed to fill the last two columns of Table 2. This is not required in the probabilistic certificate of Theorem 1 however, which requires computing $x^\star \in \Omega_{\delta_K}$, and then applying the bound in (9) [19] (which holds for any point in Ω_{δ_K});

ii) Different set of samples are needed to construct $\hat{\mathbb{P}}_{K_v}$ and hence a bound on $\mathbb{P}$, while in (9) the same set of samples are adopted for both decision-making and validation. Note that the ability to use real-world data is a distinct feature of adopting [7]. In some cases, a probabilistic model to generate samples of the uncertainty might be available, though we would only guarantee the bound in Theorem 1 not w.r.t. the true probability with which data was generated, but w.r.t. the probability induced by the choice of such a model;

iii) For $\hat{\beta} = \beta$, and to achieve $\hat{V}_{K_v}(\Omega_{\delta_K}) + \hat{\varepsilon} = \varepsilon(s^\star_K)$, i.e., to offer the same probabilistic statement, the Monte Carlo approach tends to be more conservative, requiring a higher number of samples. For the day-ahead charging coordination of PEVs, $K_v = 201510$ amounts to 552 years of non-PEV daily energy profiles for the validation process.

5 Conclusion

The scenario approach paradigm applied to uncertain VIs provides a numerically tractable framework to compute solutions with quantifiable robustness properties in a distribution-free fashion. In the specific family of uncertain VIs considered, we are able to evaluate the robustness properties of the entire set of solutions, thereby relaxing the requirement of a unique solution as often imposed in the literature. We have shown that this requires us to enumerate the active constraints that "shape" that set. Future research directions involve synthesizing algorithms to enumerate the number of support subsamples in a convex setting, as well as investigating extensions of the proposed approach to quasi-variational inequalities.

References

[1] D. Avis and K. Fukuda. A pivoting algorithm for convex hulls and vertex enumeration of arrangements and polyhedra. *Discrete & Computational Geometry*, 8(3):295–313, 1992.

[2] D. Bremner, K. Fukuda, and A. Marzetta. Primal-dual methods for vertex and facet enumeration. *Discrete & Computational Geometry*, 20(3):333–357, 1998.

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

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

[5] M. C. Campi and S. Garatti. *Introduction to the scenario approach*, volume 26. SIAM, 2018.

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

[7] 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.

[8] C. Cenedese, F. Fabiani, M. Cucuzzella, J. M. A. Scherpen, M. Cao, and S. Grammatico. Charging plug-in electric vehicles as a mixed-integer aggregative game. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pages 4904–4909, 2019.

[9] L. Deori, K. Margellos, and M. Prandini. Price of anarchy in electric vehicle charging control games: when Nash equilibria achieve social welfare. *Automatica*, 96:150–158, 2018.

[10] F. Fabiani, K. Margellos, and P. J. Goulart. On the robustness of equilibria in generalized aggregative games. In *2020 59th IEEE Conference on Decision and Control (CDC)*, pages 3725–3730. IEEE, 2020.

[11] F. Facchinei and C. Kanzow. Generalized Nash equilibrium problems. *4OR*, 5(3):173–210, 2007.

[12] F. Facchinei and J. S. Pang. *Finite-dimensional variational inequalities and complementarity problems*. Springer Science & Business Media, 2007.

[13] F. Fele and 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)*, pages 5026–5031, 2019.

[14] F. Fele and K. Margellos. Probably approximately correct Nash equilibrium learning. *IEEE Transactions on Automatic Control*, 66(9):4238–4245, 2021.

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

[16] F. Giannessi and A. Maugeri. *Variational inequalities and network equilibrium problems*. Springer, 1995.

[17] National Grid. Historical demand data, 2019. [Online] `https://www.nationalgrideso.com/data-explorer`.

[18] Gurobi Optimization, LLC. Gurobi Optimizer Reference Manual, 2021.

[19] P. Khargonekar and A. Tikku. Randomized algorithms for robust control analysis and synthesis have polynomial complexity. In *Proceedings of 35th IEEE Conference on Decision and Control*, volume 3, pages 3470–3475. IEEE, 1996.

[20] Z. Ma, D. S. Callaway, and I. A. Hiskens. Decentralized charging control of large populations of plug-in electric vehicles. *IEEE Transactions on Control Systems Technology*, 21(1):67–78, 2011.

[21] K. Margellos, M. Prandini, and J. Lygeros. On the connection between compression learning and scenario based single-stage and cascading optimization problems. *IEEE Transactions on Automatic Control*, 60(10):2716–2721, 2015.

[22] D. Paccagnan and M. C. Campi. The scenario approach meets uncertain game theory and variational inequalities. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pages 6124–6129, 2019.

[23] G. Pantazis, F. Fele, and K. Margellos. A posteriori probabilistic feasibility guarantees for nash equilibria in uncertain multi-agent games. *IFAC-PapersOnLine*, 53(2):3403–3408, 2020.

[24] R. T. Rockafellar. *Convex analysis*. Number 28. Princeton university press, 1970.

[25] U. V. Shanbhag. Stochastic variational inequality problems: applications, analysis, and algorithms. In *Theory Driven by Influential Applications*, pages 71–107. INFORMS, 2013.

[26] M. V. Solodov and P. Tseng. Modified projection-type methods for monotone variational inequalities. *SIAM Journal on Control and Optimization*, 34(5):1814–1830, 1996.

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

[28] G. M. Ziegler. *Lectures on polytopes*, volume 152. Springer Science & Business Media, 2012.