POLITECNICO DI TORINO
Repository ISTITUZIONALE

Actively Learning Equilibria in Nash Games With Misleading Information

Original
Actively Learning Equilibria in Nash Games With Misleading Information / Franci, B., Fabiani, F., Bemporad, A.. - In: IEEE CONTROL SYSTEMS LETTERS. - ISSN 2475-1456. - 9:(2025), pp. 312-317. [10.1109/lcsys.2025.3572421]

Availability:
This version is available at: 11583/3003670 since: 2025-10-10T08:42:58Z

Publisher:
Institute of Electrical and Electronics Engineers IEEE

Published
DOI:10.1109/lcsys.2025.3572421

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
IEEE postprint/Author's Accepted Manuscript

©2025 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collecting works, for resale or lists, or reuse of any copyrighted component of this work in other works.

(Article begins on next page)

07 November 2026

Actively learning equilibria in Nash games with misleading information

Barbara Franci, Filippo Fabiani and Alberto Bemporad

arXiv:2503.13167v2 [math.OC] 7 May 2025

***Abstract*—We develop a scheme based on active learning to compute equilibria in a generalized Nash equilibrium problem (GNEP). Specifically, an external observer (or entity), with little knowledge on the multi-agent process at hand, collects sensible data by probing the agents' best-response (BR) mappings, which are then used to recursively update local parametric estimates of these mappings. Unlike [1], we consider the realistic case in which the agents share corrupted information with the external entity for, e.g., protecting their privacy. Inspired by a popular approach in stochastic optimization, we endow the external observer with an inexact proximal scheme for updating the local BR proxies. This technique will prove key to establishing the convergence of our scheme under standard assumptions, thereby enabling the external observer to predict an equilibrium strategy even when relying on masked information.**

***Index Terms*—Multi-agent systems, Active learning, Competitive decision-making, Stochastic optimization.**

I. Introduction

PREDICTING a possible outcome in problems involving self-interested and privacy-preserving agents is a key requirement for their indirect control. As a prominent example, a distribution system operator (DSO) ideally wishes to exploit the flexibility offered by the widely spread smart-home appliances and electric vehicles (EVs) for an efficient usage of the distribution grid. To this end, a DSO typically designs energy prices to induce a certain collective consumption profile of the end-users, which can be predicted in advance only if these users are willing to share sensitive information [2], [3].

Akin to [1], in this paper we take the perspective of an external observer interested in learning a so-called generalized Nash equilibrium (GNE) for a population of selfish agents taking part to a generalized Nash equilibrium problem (GNEP). Given its little knowledge on the multi-agent process at hand, such an external observer is only allowed for querying the best-response (BR) mappings held by the agents. The latter, however, may be intentionally reluctant to reveal private information exactly, could erratically change their mind when presented with the same scenarios, may optimize their individual objectives with scarce accuracy, or the communication channels might be imperfect. For these reasons, we assume that the information passed to the external observer is masked by noise, as may happen in economic models [4], competitive versions of multi-agent feedback controller synthesis [5], or signal processing [6]. We then design an active learning-based scheme [7], [8] for the external entity that, despite the misleading information collected, allows to predict a GNE via faithful approximations of the agents' BRs.

Learning an equilibrium strategy from a centralized perspective based on noisy information has been considered in that branch of literature denoted as *simulation-based game theory*. Several works [9]–[11] indeed proposed different schemes to approximate the original matrix games and associated equilibria by leveraging noisy samples of agents' costs provided by an oracle. Existing techniques addressing simulated matrix games with finite decision sets include also stochastic [12] and sample-average approximation [13], as well as methods based on Bayesian optimization [14], [15]. While the former analyze the asymptotic properties of equilibria obtained from simulation-based models, also attaching probabilistic certificates on their approximation quality, the latter leverage statistical modeling tools acting as emulators of the agents' costs. Tailored acquisition functions for equilibrium learning are then designed based on the resulting posterior distributions.

In contrast, we design an active learning procedure for an external entity that iteratively makes suitable queries to estimate the BR mappings held by the agents, aiming at an exact prediction of a GNE for the GNEP in which they take part (§II). To deal with a possibly misleading information provided by the agents, we take inspiration from a popular approach in stochastic optimization to let the external observer update the local BR proxies with noisy data by means of an inexact proximal scheme. This will prove to be a key tool for learning a GNE, as well as to accompany the overall scheme with convergence guarantees under common assumptions.

Our main contributions can be summarized as follows:

i) We propose a stochastic variant of the active learning scheme derived in [1] (§III). Our iterative algorithm is based on an inexact proximal update to learn the parameters approximating the BR mappings of the agents;
ii) Under standard assumptions [16], [17], we show how these parameters can be learned exactly. Besides improving the results of [1], where such a condition was identified as sufficient for the convergence of the overall scheme and only verified ex-post, it is instrumental for proving that the external entity can asymptotically predict a GNE of the underlying GNEP (§IV).

We finally discuss practical implementation details, which are then used to test our algorithm on a numerical case study involving the indirect control of a population of EVs that tries to optimize the collective day-ahead charging schedule (§V).

This work was partially supported by the European Research Council (ERC), Advanced Research Grant COMPACT (Grant Agreement No. 101141351). B. Franci is with the Department of Mathematical Sciences, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129, Turin, italy (barbara.franci@polito.it). F. Fabiani and A. Bemporad are with the IMT School for Advanced Studies Lucca, Piazza San Francesco 19, 55100, Lucca, Italy ({filippo.fabiani, alberto.bemporad}@imtlucca.it).

Notation: $\mathbb{N}$, $\mathbb{R}$ and $\mathbb{R}_{\geq 0}$ denote the set of natural, real, and nonnegative real numbers, respectively. $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. For a vector $v \in \mathbb{R}^n$, $\|v\|_2$ denotes the standard Euclidean norm. The operator $\mathrm{col}(\cdot)$ stacks its arguments in column vectors or matrices of compatible dimensions. For example, given vectors $x_1, \ldots, x_N$ with $x_i \in \mathbb{R}^{n_i}$ and $\mathcal{I} = \{1, \ldots, N\}$, we denote $\boldsymbol{x} := (x_1^\top, \ldots, x_N^\top)^\top = \mathrm{col}((x_i)_{i\in\mathcal{I}}) \in \mathbb{R}^n$, $n := \sum_{i\in\mathcal{I}} n_i$, and $\boldsymbol{x}_{-i} := \mathrm{col}((x_j)_{j\in\mathcal{I}\setminus\{i\}})$, where $(\cdot)^\top$ denotes the transpose. Abusing notation, we also use $\boldsymbol{x} = (x_i, \boldsymbol{x}_{-i})$. With $\mathbb{E}_{\mathbb{P}}[\boldsymbol{z}] = \mathrm{col}(\mathbb{E}_{\mathbb{P}_i}(z_i))_{i\in\mathcal{I}}$ we consider the stacked vector $\boldsymbol{z}$ and apply the expected value component-wise. The uniform distribution on $[a, b]$ is denoted by $\mathcal{U}(a, b)$, while the normal distribution with mean μ and variance σ^2 by $\mathcal{N}(\mu, \sigma^2)$.

II. Problem formulation

A GNEP involves N self-interested agents, indexed by the set $\mathcal{I} := \{1, \ldots, N\}$, where each of them controls a decision variable $x_i \in \mathbb{R}^{n_i}$. Their aim is to minimize a local cost function $J_i : \mathbb{R}^n \to \mathbb{R}$, $n := \sum_{i\in\mathcal{I}} n_i$, subject to both local and coupling constraints. As such, a GNEP can be written as a collection of mutually-coupled optimization problems [17]:

$$\forall i \in \mathcal{I} : \begin{cases} \min\limits_{x_i \in \mathcal{X}_i} & J_i(x_i, \boldsymbol{x}_{-i}) \\ \text{s.t.} & (x_i, \boldsymbol{x}_{-i}) \in \Omega. \end{cases} \tag{1}$$

Thus, each cost function $J_i(x_i, \boldsymbol{x}_{-i})$ depends not only on the local variable x_i, but also on the decisions of the other agents, $\boldsymbol{x}_{-i} = \mathrm{col}((x_j)_{j\in\mathcal{I}\setminus\{i\}})$. For every agent $i \in \mathcal{I}$, $\mathcal{X}_i$ represents the set of local constraints, while the coupling constraint set is $\Omega \subseteq \mathbb{R}^n$. The collective feasible set of the GNEP in (1) is then given by $\Omega \cap \mathcal{X}$, with $\mathcal{X} := \prod_{i\in\mathcal{I}} \mathcal{X}_i$, and the feasible decision set for agent $i \in \mathcal{I}$, parametric in $\boldsymbol{x}_{-i}$, is $\mathcal{X}_i(\boldsymbol{x}_{-i}) = \{x_i \in \mathcal{X}_i \mid (x_i, \boldsymbol{x}_{-i}) \in \Omega\}$. A popular solution concept for a GNEP is the so-called GNE, defined next:

Definition 1. *A collective decision vector $\boldsymbol{x}^\star \in \Omega \cap \mathcal{X}$ is a GNE of the GNEP in* (1) *if, for all $i \in \mathcal{I}$, $J_i(x_i^\star, \boldsymbol{x}_{-i}^\star) \leq \min_{x_i \in \mathcal{X}_i(\boldsymbol{x}_{-i}^\star)} J_i(x_i, \boldsymbol{x}_{-i}^\star)$.* □

Roughly speaking, at a GNE, none of the agents has incentive to deviate from the strategy currently taken. In the considered game-theoretic framework, a quantity of interest is represented by the agent's BR mapping, formally defined as:

$$f_i(\boldsymbol{x}_{-i}) := \operatorname*{argmin}_{x_i \in \mathcal{X}(\boldsymbol{x}_{-i})} J_i(x_i, \boldsymbol{x}_{-i}). \tag{2}$$

In words, each $f_i : \mathbb{R}^{n_{-i}} \rightrightarrows \mathbb{R}^{n_i}$, $n_{-i} := \sum_{j\in\mathcal{I}\setminus\{i\}} n_j$, expresses what is the best set of decisions agent i can take, given the current decision of its opponents $\boldsymbol{x}_{-i}$. It is also instrumental to characterize a GNE, since $\boldsymbol{x}^\star$ can be equivalently defined as a collective fixed point of the agents' BR mappings, i.e., $x_i^\star \in f_i(\boldsymbol{x}_{-i}^\star)$, for all $i \in \mathcal{I}$.

While not particularly restrictive, the following conditions ensure the existence of at least a GNE [17]:

Standing Assumption 1 (BR mappings and constraints)**.** *For all $i \in \mathcal{I}$, $f_i : \mathbb{R}^{n_{-i}} \to \mathbb{R}^{n_i}$ is single-valued and continuous, with each $x_i \mapsto J_i(x_i, \boldsymbol{x}_{-i})$ convex on $\mathcal{X}_i(\boldsymbol{x}_{-i})$. The collective feasible set $\Omega \cap \mathcal{X} \subseteq \mathbb{R}^n$ is nonempty, convex, and bounded.* □

In this framework, we assume that the external entity has no knowledge about the private functions J_i but can probe the agents' BR mappings in order to collect data and predict an equilibrium strategy $\boldsymbol{x}^\star$. Unlike [1], however, we assume here that instead of communicating the exact BR, each agent shares a noisy information $z_i = \tilde{f}_i(\boldsymbol{x}_{-i}, \eta_i)$ with the external entity. Specifically, $\eta_i : \Xi_i \to \mathbb{R}^d$ is a random vector defined on the probability space $(\Xi_i, \mathcal{F}_i, \mathbb{P}_i)$ with *unknown* probability distribution to all parties involved.

As commonly assumed in a stochastic framework [16], [18], we postulate next a condition on the bias associated to z_i:

Standing Assumption 2 (Unbiased noisy information)**.** *For all $i \in \mathcal{I}$ and $\boldsymbol{x}_{-i} \in \mathbb{R}^{n_{-i}}$, it holds that $\mathbb{E}_{\mathbb{P}_i}[z_i] = x_i$, i.e., $\mathbb{E}_{\mathbb{P}_i}[\tilde{f}_i(\boldsymbol{x}_{-i}, \eta_i)] = f_i(\boldsymbol{x}_{-i})$.* □

Considering noisy BRs provides a more practical setting in which the agents, intentionally or unintentionally, do not share exact best responses, for some of the reasons we described in §I. However, Standing Assumption 2 postulates, realistically, that the agents have no interest in boycotting the central entity with its prediction task, i.e., they are not intentionally malicious: it is each agent's own interest to achieve an agreement with the other agents.

The external entity, equipped with some learning procedure $\mathscr{L}$, shall then predict a GNE by leveraging possibly misleading, yet non-private, information. Specifically, let us consider an estimate $\hat{f}_i : \mathbb{R}^{n_{-i}} \times \mathbb{R}^{p_i} \to \mathbb{R}^{n_i}$ of the i-th BR mapping $f_i(\cdot)$. This BR proxy is parametrized by $\theta_i \in \Theta_i \subseteq \mathbb{R}^{p_i}$, a quantity that shall be updated iteratively by integrating the data obtained from the agents through a smart query process, which will be described in the next section.

Standing Assumption 3 (Parameter set and BR proxies)**.** *For all $i \in \mathcal{I}$, it holds that:*

(i) Θ_i is a closed, compact, and convex set;
(ii) The mapping $\theta_i \mapsto \hat{f}_i(\boldsymbol{x}_{-i}, \theta_i)$ is continuous. □

While not postulated in [1], in our stochastic framework we need Standing Assumption 3.(i) to restrict the set of parameters, thereby ensuring that the learning procedure can compensate for the noise. This will be key to establishing the asymptotic convergence of the parameters, thus improving over [1], where this condition was identified as sufficient for concluding on the convergence of the overall procedure.

III. Active learning with misleading information

The proposed active GNE learning scheme is summarized in Algorithm 1. Note that in the last step the agents act as oracles, i.e., they provide samples consisting of noisy BRs that the external observer uses for learning. Specifically, at every iteration k the external entity integrates samples just collected to perform an inexact update of the BR proxies as in (3). In fact, given $\boldsymbol{\eta}^t = \mathrm{col}((\eta_i^t)_{i\in\mathcal{I}})$ at the generic t-th iteration, $\mathcal{F}^k$ is there defined according to the filtration $\mathcal{F} = \{\mathcal{F}^k\}_{k\in\mathbb{N}}$, i.e., the family of σ-algebras with $\mathcal{F}^0 = \sigma\left(X^0\right)$ and $\mathcal{F}^k = \sigma\left(X^0, \boldsymbol{\eta}^1, \boldsymbol{\eta}^2, \ldots, \boldsymbol{\eta}^k\right)$ such that $\mathcal{F}^k \subseteq \mathcal{F}^{k+1}$ for all $k \in \mathbb{N}$. In words, $\mathcal{F}^k$ contains the information up to iteration k.

Algorithm 1: Active learning-based method with misleading information

Initialization: $\boldsymbol{x}^0 \in \Omega$, $\theta_i^0 \in \mathbb{R}^{p_i}$ for all $i \in \mathcal{I}$

Iteration $(k \in \mathbb{N}_0)$**:**

- External entity computes

$$\theta_i^{k+1} \in \left\{ \xi \in \Theta_i \;\middle|\; \mathbb{E}_{\mathbb{P}_i}[\|\xi - \hat{\theta}_i(z_i^k, \hat{\boldsymbol{x}}_{-i}^k, \theta_i^k)\|^2 | \mathcal{F}^k] \leq (\alpha_i^k)^2 \text{ a.s.} \right\} \text{ for all } i \in \mathcal{I} \tag{3}$$

- External entity defines $\mathcal{M}(\theta^{k+1}) \coloneqq \operatorname{argmin}_{\boldsymbol{x} \in \Omega \cap \mathcal{X}} \sum_{i \in \mathcal{I}} \left\| x_i - \hat{f}_i(\boldsymbol{x}_{-i}, \theta_i^{k+1}) \right\|_2^2$, and computes

$$\hat{\boldsymbol{x}}^{k+1} = \operatorname*{argmin}_{\boldsymbol{x} \in \mathbb{R}^n} \left\{ \tfrac{1}{2}\|\boldsymbol{x}\|_2^2 \;\text{ s.t. }\; \boldsymbol{x} \in \mathcal{M}(\theta^{k+1}) \right\} \tag{4}$$

- External entity collects corrupted BRs, for all $i \in \mathcal{I}$:

$$z_i^{k+1} = \tilde{f}_i(\hat{\boldsymbol{x}}_{-i}^{k+1}, \eta_i^{k+1})$$

If each probability distribution $\mathbb{P}_i$ was known, the external entity would ideally implement the following proximal rule:

$$\hat{\theta}_i(z_i, \hat{\boldsymbol{x}}_{-i}, \theta_i) \in \operatorname*{argmin}_{\xi_i \in \Theta_i} \left\{ L_i(\xi_i | z_i, \hat{\boldsymbol{x}}_{-i}) + \frac{\mu}{2}\|\xi_i - \theta_i\|_2^2 \right\}, \tag{5}$$

for all $i \in \mathcal{I}$, $\mu > 0$, where in particular

$$L_i(\theta_i | z_i, \hat{\boldsymbol{x}}_{-i}) = \mathbb{E}_{\mathbb{P}_i}[\ell_i(z_i, \hat{f}_i(\hat{\boldsymbol{x}}_{-i}, \theta_i))]. \tag{6}$$

However, since each $\mathbb{E}_{\mathbb{P}_i}$ is unavailable, we propose for the external entity to focus on (3) as a viable option.

The loss function $L_i : \Theta_i \to \mathbb{R}$ measures the dissimilarity between the information received via z_i, and the estimate $\hat{f}_i$. Note that L_i does not depend explicitly on x_i and $\boldsymbol{x}_{-i}$, since those are quantities provided through samples. In addition, $\ell_i : \Theta_i \times \Xi_i \to \mathbb{R}$ depends on η_i via z_i, hence the expected value with respect to (w.r.t.) $\mathbb{P}_i$. We then impose what follows:

Standing Assumption 4 (Training loss function)**.** *For all $i \in \mathcal{I}$, the following conditions hold true:*

(i) The mapping $\theta_i \mapsto L_i(\theta_i | x_i, \boldsymbol{x}_{-i})$ is convex and twice continuously differentiable;

(ii) The mapping $\theta_i \mapsto \ell_i(z_i, \hat{f}_i(\hat{\boldsymbol{x}}_{-i}, \theta_i))$ is differentiable;

(iii) For all $(x_i, \boldsymbol{x}_{-i}, \theta_i) \in \Omega \cap \mathcal{X} \times \Theta_i$, $0 \leq L_i(\theta_i | x_i, \boldsymbol{x}_{-i}) < \infty$, with $L_i(\theta_i | x_i, \boldsymbol{x}_{-i}) = 0 \iff x_i = \hat{f}_i(\boldsymbol{x}_{-i}, \theta_i)$. □

While Standing Assumption 4 actually turns the inclusion in (5) into equality, the following condition, postulated also in [16], [18], will be key for our convergence analysis:

Standing Assumption 5 (Proximal map)**.** *The proximal mapping $\hat{\theta}_i(\cdot, \cdot, \theta_i)$ in* (5) *is a-contractive, $a \in (0, 1)$, i.e., for all $\theta_i, \theta_i' \in \Theta_i$, $\|\hat{\theta}(\cdot, \cdot, \theta_i) - \hat{\theta}(\cdot, \cdot, \theta_i')\| \leq a\|\theta_i - \theta_i'\|$.* □

We postulate this property as an assumption, although sufficient conditions guaranteeing the contractivity of $\hat{\theta}_i(\cdot, \cdot, \theta_i)$ can be obtained similarly to [17, Prop. 12.17] and [16, §2.2]. However, since the external entity does not know the probability distribution $\mathbb{P}_i$ of the noise, the expected value in (5)–(6) can not be computed exactly. This is the reason why, inspired by [16], we propose an inexact scheme as described in (3). In fact, such an instruction is asymptotically equivalent to the exact proximal mapping in (5), but it can be computed through the iterations. The parameters α_i^k in (3), instead, form a deterministic sequence that meets the following conditions:

Standing Assumption 6 (Accuracy sequence)**.** *For all $i \in \mathcal{I}$, the sequence $\{\alpha_i^k\}_{k \in \mathbb{N}}$ is such that $\sum_{k \in \mathbb{N}_0} \alpha_i^k < \infty$ and, for all $k \in \mathbb{N}_0$, $\alpha_i^k \geq 0$ and $\lim_{k \to \infty} \alpha_i^k = 0$.* □

Remark 1. *In [16], a stochastic approximation method is used to obtain a solution to* (3) *which is α_i^k-close to the exact one of* (5)*. As a consequence of Standing Assumption 6, convergence to the exact solution holds (see Lemma 1). This consists in performing a number of stochastic proximal gradient descent steps, proportional to the outer iteration index k of Algorithm 1 [16, §3.4]. In particular, at iteration k, for all $i \in \mathcal{I}$, the external entity performs the following steps for $t > 0$:*

$$\begin{aligned} \xi_i^{t+1} = \operatorname{proj}_{\Theta_i}(\xi_i^t - \gamma^t(\tfrac{1}{S}\textstyle\sum_{j=1}^S \nabla_{\theta_i} \ell_i(z_i^{k,j}, \hat{f}_i(\hat{\boldsymbol{x}}_{-i}^k, \xi_i^t)) \\ + \mu(\xi_i^t - \theta_i^k))), \end{aligned} \tag{7}$$

with γ^t being a vanishing step-size sequence and $\{z_i^{k,j}\}_{j=1}^S$ being a collection of S samples of the noisy queries. The iterative procedure stops, say after $\bar{t}$ iterations, and sets $\theta_i^{k+1} = \xi_i^{\bar{t}}$. In this case, some further assumption on the expected-valued gradients should be considered—see, e.g., [16, Ass. 1.(c), 1.(d)]. Other algorithms can be however used and integrated with different approximation schemes. □

By leveraging the BR surrogates updated through (3), the external entity then designs the next query point $\hat{\boldsymbol{x}}^{k+1}$ to collect new information from the agents according to (4), i.e., as the minimum norm strategy profile falling into the set:

$$\mathcal{M}(\theta^{k+1}) \coloneqq \operatorname*{argmin}_{\boldsymbol{x} \in \Omega \cap \mathcal{X}} \sum_{i \in \mathcal{I}} \left\| x_i - \hat{f}_i(\boldsymbol{x}_{-i}, \theta_i^{k+1}) \right\|_2^2, \tag{8}$$

where $\mathcal{M} : \mathbb{R}^p \rightrightarrows \Omega$, $p \coloneqq \sum_{i \in \mathcal{I}} p_i$. This set contains all collective profiles that are the closest to a fixed point of each $\hat{f}_i(\cdot, \theta_i^{k+1})$, i.e., closest to a GNE as defined in Definition 1 and discussion following (2). Indeed, if each $\hat{f}_i(\bar{\boldsymbol{x}}_{-i}, \bar{\theta}_i^k)$ was exactly equal to $f_i(\bar{\boldsymbol{x}}_{-i})$, and the minimum in (8) was identically zero, then $\bar{\boldsymbol{x}} \in \mathcal{M}(\bar{\theta}^k)$ would be a GNE of the GNEP in (1). Such a smart selection of the query points amounts to the "active" part of Algorithm 1, and

allows the central entity to accumulate (noisy) information in a neighborhood of a point that is the closest to a true GNE. This will be key for the technical analysis carried out in the next section. In (8), θ^{k+1} represents the whole collection of parameters $\{\theta_i^{k+1}\}_{i\in\mathcal{I}}$ characterizing the estimate mappings, which at every iteration coincides with the argument of the corresponding parameter-to-query mapping $\mathcal{M}(\cdot)$. It then follows from the definition of $\mathcal{M}$ and from Standing Assumption 2 that $(\mathbb{E}_{\mathbb{P}}[z_i], \hat{\boldsymbol{x}}_{-i}) \in \Omega \cap \mathcal{X}$. Once obtained the minimum norm vector $\hat{\boldsymbol{x}}^{k+1}$ in (4), the external entity queries each agent with $\hat{\boldsymbol{x}}_{-i}$, which in turn reacts through a noisy BR $z_i^{k+1} = \tilde{f}_i(\hat{\boldsymbol{x}}_{-i}^{k+1}, \eta_i^{k+1})$. The observer finally collects all these data, and the process repeats.

IV. Convergence analysis

Before studying the asymptotic properties of the active learning procedure in Algorithm 1, we postulate some assumptions on the learning procedure $\mathscr{L}$. We then prove some preliminary results, functional to the asymptotic analysis.

In particular, our analysis will be based on the possibility of matching pointwise the BR mapping f_i of each agent. To this aim, for all $i \in \mathcal{I}$ and for all $(x_i, \boldsymbol{x}_{-i}) \in \Omega \cap \mathcal{X}$, let

$$\mathcal{A}_i(x_i, \boldsymbol{x}_{-i}) = \{\tilde{\theta}_i \in \Theta_i | L_i(\tilde{\theta}_i | x_i, \boldsymbol{x}_{-i}) = 0\}.$$

This set is instrumental to prove the following crucial result:

Lemma 1. *For all $i \in \mathcal{I}$, let $\{\theta_i^k\}_{k\in\mathbb{N}}$ be the sequence generated by* (3) *in Algorithm 1. If $\lim_{k\to\infty} \mathbb{E}_{\mathbb{P}_i}[z_i^k] = \bar{x}_i$, $\lim_{k\to\infty} \hat{\boldsymbol{x}}_{-i}^k = \bar{\boldsymbol{x}}_{-i}$ so that $(\bar{x}_i, \bar{\boldsymbol{x}}_{-i}) \in \Omega \cap \mathcal{X}$, then for all $i \in \mathcal{I}$, $\lim_{k\to\infty} \theta_i^k = \bar{\theta}_i$, and $\lim_{k\to\infty} \mathbb{E}_{\mathbb{P}_i}\left[\left\|\theta_i^k - \bar{\theta}_i\right\|\right] = 0$ a.s.. Moreover, $\bar{\theta}_i \in \mathcal{A}_i(\bar{x}_i, \bar{\boldsymbol{x}}_{-i})$, i.e.,* (3) *converges to a solution of the exact proximal scheme in* (5). □

Proof. The fact that $\lim_{k\to\infty} \theta_i^k = \bar{\theta}_i$ for all $i \in \mathcal{I}$ follows analogously to [16, Prop. 1] by using contractivity (Standing Assumption 5), the unbiased noise (Standing Assumption 2) and convexity of Θ_i (Standing Assumption 3) on $\|\theta_i^{k+1} - \bar{\theta}_i\|$, together with the vanishing property of $\{\alpha_i^k\}$ (Standing Assumption 6). With the same assumptions, it follows similarly that $\lim_{k\to\infty} \mathbb{E}_{\mathbb{P}_i}\left[\left\|\theta_i^k - \bar{\theta}^k\right\|\right] = 0$ [16, Prop. 2.(b)]. The last statement, instead, follows analogously to [1, Lemma 4.2] from the properties of ℓ_i (Standing Assumption 4). ■

Lemma 1 says that when all the ingredients involved in Algorithm 1 converge, then the pointwise approximation of f_i shall be exact at $\boldsymbol{x}_{-i}$, i.e., $\hat{f}_i(\boldsymbol{x}_{-i}, \bar{\theta}_i) = \mathbb{E}_{\mathbb{P}_i}[\tilde{f}_i(\boldsymbol{x}_{-i}, \eta_i)] = f_i(\boldsymbol{x}_{-i})$, with $\bar{\theta}_i \in \mathcal{A}_i(x_i, \boldsymbol{x}_{-i})$. We will then prove in Proposition 1 that the conditions required in Lemma 1 are verified. To simplify the notation, let $r(\boldsymbol{x}, \theta) = \sum_{i\in\mathcal{I}} \|x_i - \hat{f}_i(\boldsymbol{x}_{-i}, \theta_i)\|_2^2$. Next, we impose some conditions on $r(\boldsymbol{x}, \theta)$:

Standing Assumption 7. *The following conditions hold true:*

(i) For all $x \in \Omega, \theta \mapsto r(\boldsymbol{x}, \theta)$ is convex and differentiable;
(ii) For all $\theta \in \mathbb{R}^p, \boldsymbol{x} \mapsto r(\boldsymbol{x}, \theta)$ is convex and continuous;
(iii) For all $\theta \in \mathbb{R}^p$, the vector $\partial r(\boldsymbol{x}, \theta)/\partial\theta \in \mathbb{R}^p$ of partial derivatives is bounded w.r.t. $\boldsymbol{x}$. □

The following technical result characterizes the properties of the sequence of query points $\{\hat{\boldsymbol{x}}^k\}_{k\in\mathbb{N}}$ produced by the central entity in the second step (4) of Algorithm 1.

Proposition 1. *Let $\mathcal{M}(\tilde{\theta}) = \{\tilde{\boldsymbol{x}}\}$. Then the sequence $\{\hat{\boldsymbol{x}}^k\}_{k\in\mathbb{N}}$ generated by* (4) *is feasible, i.e., $\hat{\boldsymbol{x}}^k \in \Omega \cap \mathcal{X}$ for all $k \in \mathbb{N}$, and satisfies $\lim_{k\to\infty} \hat{\boldsymbol{x}}^k = \tilde{\boldsymbol{x}}$.* □

Proof. In view of Lemma 1, $\lim_{k\to\infty} \theta_i^k = \tilde{\theta}_i$ a.s. for all $i \in \mathcal{I}$. The proof then follows from Standing Assumptions 1 and 7 which imply that $\mathcal{M}(\theta^k)$ is a convex set [1, Lemma 3.3]. From the same assumptions then follows that the sequence $\{\hat{\boldsymbol{x}}^k\}_{k\in\mathbb{N}}$ is bounded and its cluster points $\bar{\boldsymbol{x}}$ belong to $\mathcal{M}(\tilde{\theta})$ [1, Lemma 3.6, 3.7]. By contradiction we can instead show that it can not happen that $\bar{\boldsymbol{x}} \neq \tilde{\boldsymbol{x}}$ [1, Lemma 3.7]. ■

We are now ready to state the asymptotic properties of the active learning-based technique summarized in Algorithm 1:

Theorem 1. *Let $\mathcal{M}(\tilde{\theta}) = \{\tilde{\boldsymbol{x}}\}$. Then, $\lim_{k\to\infty} \|\mathbb{E}_{\mathbb{P}}[\boldsymbol{z}^k - \hat{\boldsymbol{x}}^k]\|_2 = 0$, and the sequences $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}}$ and $\{\hat{\boldsymbol{x}}^k\}_{k\in\mathbb{N}}$ generated by Algorithm 1 converge to the same GNE of the GNEP in* (1). □

Proof. It follows from Standing Assumptions 1, 4 and 7 by noting that, in view of the consistency property in Lemma 1, and by Proposition 1 the pointwise approximation shall be exact, namely each $\tilde{\theta}_i$ is so that, for all $i \in \mathcal{I}$, $\|\hat{f}_i(\tilde{\boldsymbol{x}}_{-i}, \tilde{\theta}_i) - \mathbb{E}_{\mathbb{P}_i}[\tilde{f}_i(\tilde{\boldsymbol{x}}_{-i}, \eta_i)]\|_2 = 0$ [1, Th. 4.5]. ■

Theorem 1 establishes that the external entity achieves convergence to the true values, i.e., it predicts both an exact GNE of the game and the BR mappings, despite the possibly misleading information passed by the agents.

V. Implementation details and simulation results

We now discuss several implementation details related to Algorithm 1 that will be employed to perform numerical experiments on a charging coordination problem for EVs.

A. Practical considerations

A distinct feature of the proposed active learning-based scheme is represented by the inexact proximal step in (3), which can be accomplished as discussed in Remark 1. To this end, performing for instance the stochastic proximal gradient descent in (7) requires one the availability of a batch of S samples $\{z_i^{k,j}\}_{j=1}^S$ at every outer iteration k. The latter can be obtained by the central entity either probing the i-th BR mapping S times with the same $\hat{\boldsymbol{x}}_{-i}^{k-1}$, or producing synthetic samples. While the former may not represent a viable approach in a realistic case involving, e.g., human agents, the latter can be always pursued on the basis of the data collected up to iteration k, i.e., $\{z_i^j\}_{j=1}^k$. Among the simplest approaches, a maximum likelihood estimation (MLE) method [19] allows one to estimate the (possibly time-varying) measurement noise covariance matrix R_i^k using measurement residuals (innovations), i.e., $e_i^k = z_i^k - \hat{f}_i(\boldsymbol{z}_{-i}^k, \theta_i^k)$. The likelihood function associated to the measurements $\{z_i^j\}_{j=1}^k$ given R^k is then: $\mathcal{L}_i(R_i^k) = \prod_{j=1}^k \exp\left(-\frac{1}{2}(e_i^j)^\top (P_i^j)^{-1} e_i^j\right) / \sqrt{|2\pi P_i^j|}$, where P_i^j denotes the measurement covariance at the j-th outer iteration. Taking the logarithm of the likelihood function, we obtain $\log \mathcal{L}_i(R_i^k) = -\frac{k}{2}\log|P_i^k| - \frac{1}{2}\sum_{j=1}^k (e_i^j)^\top (P_i^j)^{-1} e_i^j$, which, in case the measurements are affected by Gaussian

Table I
INDIRECT CONTROL OF EVS – SIMULATION PARAMETERS

Parameters	Description	Value
T	Time interval	14
N	Number of EVs	10
$Q_i = q_i I_T$	Degradation cost – quadratic term	$q_i \sim \mathcal{U}(0.006, 0.01)$
c_i	Degradation cost – affine term	$\sim \mathcal{U}(0.055, 0.095)^T$
d	Normalized inflexibility demand	from [20, Fig. 1]
ρ_i	Local charging requirement	$\sim \mathcal{U}(1.2, 1.8)$
$\bar{c}_i$	Upper bound - power injection	0.25
$\bar{c}$	Grid capacity	0.2
a	Inverse price elasticity of demand	0.8
b	Baseline price	0.02
η_i	Additive noise on each BR	$\sim \mathcal{N}(0, 0.1)$
Θ_i	Parameters' set	$[-10, 10]^{p_i}$
K	Number of iterations (Alg. 1)	200
γ^t	Step-size in (7)	10^{-3t}
μ	Proximal parameter	10
$\bar{t}$	Iterations performed in (7)	$10k$

noise, allows one to estimate R_i^k by maximizing $\log \mathcal{L}_i(R_i^k)$ w.r.t. R_i^k itself. Thus, setting the derivative of the above to zero and solving for R_i^k yields: $\hat{R}_i^k = \frac{1}{k}\sum_{j=1}^{k} e_i^j (e_i^j)^\top$, which is the sample covariance of the residuals, and can be employed to produce synthetic samples $\{z_i^{k,j}\}_{j=1}^S$ for (7) through, e.g., multivariate normal sampling. Specifically, one generates data $z_i^{k,j} = z_i^k + V_i^k \nu_i^j$, where $\nu_i^j \sim \mathcal{N}(0, I_{n_i})$ and V_i^k is a matrix obtained from the Cholesky or singular value decomposition of $\hat{R}_i^k$. We will later exploit this empirical approach in §V-B.

Note that the convergence property of our active learning-based scheme requires only a pointwise exact approximation of the BR mappings held by the agents for the external observer to successfully accomplish the prediction task, despite noisy data. For this reason, as observed in [1] it is convenient for the external entity to adopt affine BR proxies $\hat{f}_i(\cdot, \theta_i)$, i.e.,

$$\hat{f}_i(\boldsymbol{x}_{-i}, \theta_i) = \Lambda_i \begin{bmatrix} \boldsymbol{x}_{-i} \\ 1 \end{bmatrix}, \tag{9}$$

for $\Lambda_i \in \mathbb{R}^{n_i \times (n_{-i}+1)}$—note that θ_i is the vectorization of Λ_i, with $p_i = n_i(n_{-i}+1)$. In case one adopts a standard mean squared error (MSE) for the training, i.e., $\ell_i(z_i, \hat{f}_i(\boldsymbol{x}_{-i}, \theta_i)) = \frac{1}{2}\|z_i - \Lambda_i \begin{bmatrix} \boldsymbol{x}_{-i}^\top & 1 \end{bmatrix}^\top\|^2$ such a design choice allows to automatically satisfy Standing Assumption 4, as well as the requirements in Standing Assumption 7. Besides all these technical motivations, affine BR surrogates also yield important practical consequences. Specifically, each gradient in (7) reads as $(\Lambda_i^t[(\hat{\boldsymbol{x}}_{-i}^k)^\top \quad 1]^\top - z_i^{k,j})[(\hat{\boldsymbol{x}}_{-i}^k)^\top \quad 1]$, while solving (8) turns out to be a constrained least-squares (LS) problem, which is convex in view of Standing Assumption 1.

B. Case study: Indirect control of smart grids

We test our technique on an indirect control problem faced by DSOs, which design price signals enabling the energy flexibility offered by price-sensitive end-users [21].

In particular, we consider a set of N EVs populating a distribution grid [20], [22], where every selfish agent aims at determining an optimal EV charging schedule over a certain discrete time interval $\{1, \dots, T\}$ by controlling the energy injection $x_i \in \mathbb{R}^T_{\geq 0}$. The underlying problem is typically modeled as a GNEP, consisting of the following collection of mutually coupled optimization problems:

$$\forall i \in \mathcal{I} : \begin{cases} \min\limits_{x_i} & \|x_i\|_{Q_i}^2 + c_i^\top x_i + (a(\sigma(\boldsymbol{x}) + d) + b\mathbf{1}_T)^\top x_i \\ \text{s.t.} & \mathbf{1}_T^\top x_i \geq \rho_i,\ x_i \in [0, \bar{x}_i]^T,\ \sigma(\boldsymbol{x}) \leq \bar{c}. \end{cases} \tag{10}$$

Each private cost function is composed of two terms: $\|x_i\|_{Q_i}^2 + c_i^\top x_i$, which models the battery degradation cost, and $(a(\sigma(\boldsymbol{x}) + d) + b\mathbf{1}_T)^\top x_i$, which is associated to the electricity pricing. Here, $\sigma(\boldsymbol{x})$ denotes the aggregate demand of the whole population of EVs, defined as $\sigma(\boldsymbol{x}) = \frac{1}{N}\sum_{i=1}^N x_i \in \mathbb{R}^T_{\geq 0}$, where $a > 0$ represents the inverse of the price elasticity of demand, $b > 0$ the baseline price, and $d \in \mathbb{R}^T_{\geq 0}$ the normalized average inflexible demand. In addition, each user has to satisfy both local and shared constraints due for instance to a minimum charging amount over the interval, $\mathbf{1}_T^\top x_i \geq \rho_i \geq 0$, a cap on the power injection $x_i \in [0, \bar{x}_i]^T$, or accounting for intrinsic grid limitations, i.e., $\sigma(\boldsymbol{x}) + d \in [0, \bar{c}]^T$.

In this framework, an equilibrium strategy $\boldsymbol{x}^\star$, which produces the aggregate consumption $\frac{1}{N}\sum_{i=1}^N x_i^\star$, heavily depends on the values of a and b. It is then clear how a suitable design of a and b, based on an accurate prediction of the resulting $\boldsymbol{x}^\star(a, b)$, allows for an efficient usage of the distribution grid. Thus, a DSO is interested in making accurate forecasts on the aggregate electricity consumption of end-users in response to price-signals, aimed at enabling flexibility offered by the users themselves. On the other hand, the smart query process proposed in [1] does not account for the possible malice of end-users, who may not be willing to provide correct information, are uncertain or even contradictory about it.

We conduct numerical experiments by using the values reported in Tab. I. Specifically, we assume the DSO endowed with affine BR proxies as in (9), and additive noise affecting the agents' BRs, i.e., $z_i = f_i(\boldsymbol{x}_{-i}) + \eta_i$ for all $i \in \mathcal{I}$. While Algorithm 1 is initialized as described in [1, §VI.A], we exploit the procedure in Remark 1 for solving (5) with increasing accuracy at every outer iteration k. With this regard, we preliminary analyze the impact that the size of S has on the computation of a GNE with misleading information. For each agent, to generate synthetic samples $\{z_i^{k,j}\}_{j=1}^S$ at each iteration, we have adopted the MLE-based approach discussed in §V-A. Then, for each $S \in \{1, 5, 10, 20, 50\}$, we have generated 20 numerical instances of (10), run [1, Alg. 1] with noise-free BR samples for computing a reference GNE, and then Algorithm 1. In Fig. 1, which illustrates the box plot associated to the relative distance from a GNE for each case, we observe that, as expected, a larger batch of samples allows for a better accuracy in the GNE computation and reduces the related variance. On the other hand, we have also experienced a significant increase in the computational time, since each iteration of Algorithm 1 with $S = 1$ takes 2.79[s] as worst-case average (i.e., with $k = 200$), up to 52.8[s] for $S = 50$. Motivated by these considerations, we have then set $S = 10$ and compared the query point sequences generated by Algorithm 1 with a naïve implementation of [1, Alg. 1]. Also in this case, we have considered 20 different numerical instances, with reference GNE computed through [1, Alg. 1] by relying on noiseless data. From Fig. 2, it is clear that,

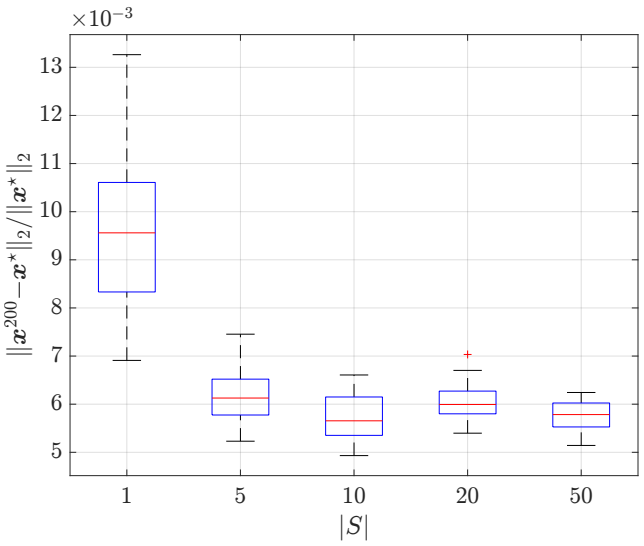

Figure 1. Relative distance between the GNE computed by relying on misleading information through Algorithm 1 (i.e., $\boldsymbol{x}^{200}$), and the one obtained with noiseless data (i.e., $\boldsymbol{x}^\star$), averaged over 20 numerical instances of (10).

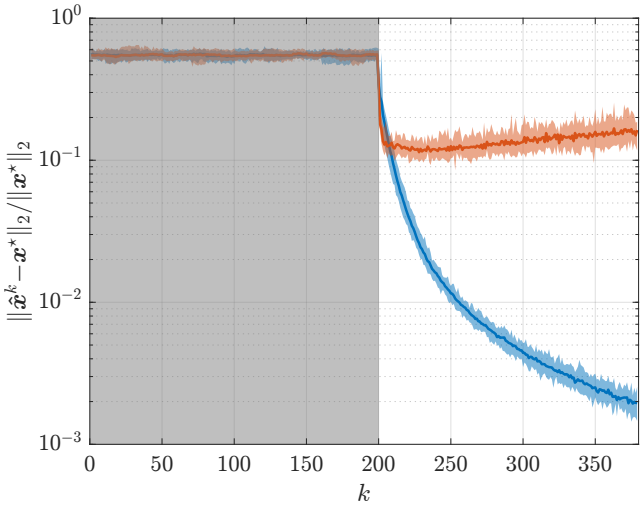

Figure 2. Relative distance sequence produced by Algorithm 1 (solid blue line) and by [1, Alg. 1] exploiting noisy data (solid red line), averaged over 20 numerical instances. The shaded colored areas represent the standard deviation over the different numerical trials, while the shaded black region corresponds to the random initialization of both procedures.

whether the procedure in Algorithm 1 can cope with noisy BR samples provided by the agents, running [1, Alg. 1] blindly with inexact data produces a non-convergent behavior.

VI. Conclusion

We have proposed a novel procedure based on active learning to let an external observer learn faithful local proxies of BR mappings privately held by a population of agents taking part to a GNEP. With the goal of predicting a GNE of the underlying game, we have adopted an inexact proximal update of those surrogates that allows to integrate possible misleading information provided by the agents. We have shown that this technique guarantees the convergence of the BR estimates and, at the same time, of the overall active learning scheme, ensuring that the external entity succeeds in its prediction task.

References

[1] F. Fabiani and A. Bemporad, "An active learning method for solving competitive multi-agent decision-making and control problems," *IEEE Transactions on Automatic Control*, pp. 1–16, 2024.

[2] O. Corradi, H. Ochsenfeld, H. Madsen, and P. Pinson, "Controlling electricity consumption by forecasting its response to varying prices," *IEEE Transactions on Power Systems*, vol. 28, no. 1, pp. 421–429, 2012.

[3] O. Bilenne, B. Franci, P. Jacquot, N. Oudjane, M. Staudigl, and C. Wan, "A privacy-preserving decentralized algorithm for distribution locational marginal prices," in *2022 IEEE 61st Conference on Decision and Control (CDC)*. IEEE, 2022, pp. 4143–4148.

[4] J. Renault and T. Tomala, "Communication equilibrium payoffs in repeated games with imperfect monitoring," *Games and Economic Behavior*, vol. 49, no. 2, pp. 313–344, 2004.

[5] C. A. R. Crusius and A. Trofino, "Sufficient LMI conditions for output feedback control problems," *IEEE Transactions on Automatic Control*, vol. 44, no. 5, pp. 1053–1057, 1999.

[6] R. Mochaourab and E. A. Jorswieck, "Robust beamforming in interference channels with imperfect transmitter channel information," *Signal Processing*, vol. 92, no. 10, pp. 2509–2518, 2012.

[7] B. Settles, "Active learning," in *Synthesis Lectures on Artificial Intelligence and Machine Learning*. Morgan & Claypool Publishers, 2012.

[8] A. Bemporad, "Active learning for regression by inverse distance weighting," *Information Sciences*, vol. 626, pp. 275–292, 2023.

[9] E. A. Viqueira and C. Cousins, "Learning simulation-based games from data," in *Proceedings of the 18th International Conference on Autonomous Agents and Multi-Agent Systems*, vol. 2019, 2019.

[10] E. A. Viqueira, C. Cousins, and A. Greenwald, "Improved algorithms for learning equilibria in simulation-based games," in *Proceedings of the 19th International Conference on Autonomous Agents and Multi-Agent Systems*, 2020, pp. 79–87.

[11] A. Marchesi, F. Trovò, and N. Gatti, "Learning probably approximately correct maximin strategies in simulation-based games with infinite strategy spaces," in *Proceedings of the 19th International Conference on Autonomous Agents and Multi-Agent Systems*, 2020, pp. 834–842.

[12] Y. Vorobeychik and M. P. Wellman, "Stochastic search methods for Nash equilibrium approximation in simulation-based games," in *Proceedings of the 7th International Conference on Autonomous Agents and Multi-Agent Systems*, 2008, pp. 1055–1062.

[13] Y. Vorobeychik, "Probabilistic analysis of simulation-based games," *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, vol. 20, no. 3, pp. 1–25, 2010.

[14] A. Al-Dujaili, E. Hemberg, and U. M. O'Reilly, "Approximating Nash equilibria for black-box games: A bayesian optimization approach," *arXiv preprint arXiv:1804.10586*, 2018.

[15] V. Picheny, M. Binois, and A. Habbal, "A Bayesian optimization approach to find Nash equilibria," *Journal of Global Optimization*, vol. 73, no. 1, pp. 171–192, 2019.

[16] J. Lei, U. V. Shanbhag, J.-S. Pang, and S. Sen, "On synchronous, asynchronous, and randomized best-response schemes for stochastic Nash games," *Mathematics of Operations Research*, vol. 45, no. 1, pp. 157–190, 2020.

[17] F. Facchinei and J.-S. Pang, "Nash equilibria: The variational approach," *Convex Optimization in Signal Processing and Communications*, p. 443, 2010.

[18] J. Lei and U. V. Shanbhag, "Distributed variable sample-size gradient-response and best-response schemes for stochastic nash equilibrium problems," *SIAM Journal on Optimization*, vol. 32, no. 2, pp. 573–603, 2022.

[19] C. M. Bishop, *Pattern Recognition and Machine Learning*. Berlin, Heidelberg: Springer-Verlag, 2006.

[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*, vol. 21, no. 1, pp. 67–78, 2011.

[21] F. D'Ettorre, M. Banaei, R. Ebrahimy, S. A. Pourmousavi, E. Blomgren, J. Kowalski, Z. Bohdanowicz, B. Łopaciuk-Gonczaryk, C. Biele, and H. Madsen, "Exploiting demand-side flexibility: State-of-the-art, open issues and social perspective," *Renewable and Sustainable Energy Reviews*, vol. 165, p. 112605, 2022.

[22] 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)*, 2019, pp. 4904–4909.