Received 28 December 2023; revised 18 April 2024; accepted 24 May 2024. Date of publication 19 June 2024; date of current version 23 July 2024. Recommended by Guest Editor Ahmad Taha.

Digital Object Identifier 10.1109/OJCSYS.2024.3416768

Solving Decision-Dependent Games by Learning From Feedback

KILLIAN WOOD [1], AHMED S. ZAMZAM [2] (Member, IEEE), AND EMILIANO DALL'ANESE [1,3,4] (Senior Member, IEEE)

(Intersection of Machine Learning with Control)

[1]Department of Applied Mathematics, University of Colorado, Boulder, CO 80309 USA
[2]National Renewable Energy Laboratory, Golden, CO 80523 USA
[3]Department of Electrical, Computer, and Energy Engineering, University of Colorado, Boulder, CO 80309 USA
[4]Department of Electrical and Computer Engineering, Boston University, Boston, MA 02215 USA

CORRESPONDING AUTHOR: EMILIANO DALL'ANESE (e-mail: emiliano.dallanese@colorado.edu)

This work was supported in part by the National Science Foundation (NSF) under Grant 1941896 and Grant 2044946, and in part by the NSF Mathematical Sciences Graduate Internship Program.

ABSTRACT This paper tackles the problem of solving stochastic optimization problems with a decision-dependent distribution in the setting of stochastic strongly-monotone games and when the distributional dependence is unknown. A two-stage approach is proposed, which initially involves estimating the distributional dependence on decision variables, and subsequently optimizing over the estimated distributional map. The paper presents guarantees for the approximation of the cost of each agent. Furthermore, a stochastic gradient-based algorithm is developed and analyzed for finding the Nash equilibrium in a distributed fashion. Numerical simulations are provided for a novel electric vehicle charging market formulation using real-world data.

INDEX TERMS Decision-dependent distribution, learning, optimization, stochastic monotone games.

I. INTRODUCTION

The efficacy of stochastic optimization [1] and stochastic games [2], [3], [4], [5], [6] generally hinges on the premise that the underlying data distribution is stationary. This means that the distribution of the data, which parameterize the problem or the game, does not change throughout the execution of the algorithm used to solve the stochastic problem or game, and is neither influenced or dependent on time nor the optimization variables themselves. This is a common setup that has been considered when game-theoretic frameworks have been applied to problems in, for example, ride hailing [7], routing [8], charging of electric vehicles (EVs) [9], [10], power markets [11], power systems [12], and in several approaches for training of neural networks [13]. However, this assumption can be invalid in a variety of setups in which the cost to be minimized is parameterized by data that is received from populations or a collection of automated control systems, whose response is uncertain and depends on the output of the optimization problem itself. As an example, in a competitive market for electric EV charging [9], [14], the operators seek to find the charging prices (i.e., the optimization variables) to maximize the revenue from EVs; however, the expected demand (i.e., the "data" of the problem) is indeed dependent on the price itself. More broadly, power consumption in power distribution grids depends on electric prices [15]. A similar example pertains to ride hailing [7].

To accommodate this scenario, the so-called stochastic optimization with *decision-dependent distributions* (also known as *performative prediction* [16]) posits that we represent the data distribution used in optimization instead as a *distributional map* $x \mapsto D(x)$ where x are decision variables [16], [17], [18], [19], [20]. In this work, we study decision-dependent stochastic games in which players seeks to minimize their cost (based on their optimization variables) subject to other players optimization variables, and where the *data distribution of each player depends on the actions of all players* (we will use the term player and agent interchangeably).

We focus on solving the Nash equilibrium problem of a game, which is to find a decision from which no agent is incentivized by their own cost to deviate when played.

Formally, the stochastic Nash equilibrium problem with decision-dependent distributions considered in this paper is to find a point $x^* = (x_1^*, \ldots, x_n^*) \in \mathbb{R}^n$ such that

$$x_i^* \in \arg\min_{x_i \in \mathcal{X}_i} F_i(x_i, x_{-i}^*), \quad \forall\, i \in \{1, \ldots, n\} \tag{1}$$

with $F_i(x_i, x_{-i}^*)$ defined as:

$$F_i(x_i, x_{-i}^*) := \underset{z_i \sim D_i(x_i, x_{-i}^*)}{\mathbb{E}} f_i(x_i, x_{-i}^*, z_i) \tag{2}$$

where: z_i denotes a random variable supported on $\mathbb{R}^{k_i}$, $f_i : \mathbb{R}^d \times \mathbb{R}^{k_i} \to \mathbb{R}$ is a scalar valued function that is convex and continuously differentiable in x_i, $\mathcal{X}_i \subseteq \mathbb{R}^{d_i}$ is a compact convex set, and $D_i : \mathbb{R}^d \to \mathcal{P}(\mathbb{R}^{k_i})$ is a *distributional map* whose output is a probability distribution supported on $\mathbb{R}^{k_i}$.

Standard stochastic first-order methods are insufficient for solving problems of this form. As we will demonstrate later in the paper, even estimating the expected gradient from samples requires knowledge of the probability density function associated with D_i—which is not possible in a majority of practical applications.

Hereafter, we use the term "system" to refer to a population or a collection of automated controllers producing a response $z_i \in \mathbb{R}^{k_i}$ upon observing x. To illustrate our setup, consider again the example where each agent represents an EV charging provider. Here, $x_i \in \mathbb{R}^{d_i}$ represents the charging price at a station managed by provider i, expressed in \$/kWh. Correspondingly, z_i indicates demand for the service at that price, while f_i is the service cost (or the negative of the total profit) for provider i. This is an example of a competitive market in which the demand for service is a function of the price of all providers; see, for example, the game-theoretic approaches presented in [9], [21] and the Stackelberg game presented in [14]. However, compared to existing game-theoretic models for EV markets, the framework proposed in this paper allows for an uncertain response of EV owners to price variations; this randomness is difficult to model, as it related to the drivers' preferences and other externalities such as the locations of the charging stations, etc., as explained in, e.g., [21], [22], [23].

Challenges in solving problems of this form typically stem from the fact that the distributional maps D_i are often unknown [24], [25], [26], [27]. To overcome this challenge, we propose a learning-based optimization procedure – in the spirit of the methods proposed for convex optimization in [18], [28] – to tackle the multi-player decision-dependent stochastic game. The key idea behind this framework is that we first propose a parameterization for the distributional map in the system and estimate it from responses. Then, we use the estimated distributional map throughout the game without requiring further interaction with the system.

A. RELATED WORK

Our work incorporates themes from games, learning, as well as stochastic optimization with decision-dependent distributions. We highlight the relationship with this relevant literature below.

Games: Within the context of games, our work is specifically focused on solving Nash equilibrium problems using gradient-based methods and a variational inequality (VI) framework. The literature on stochastic games is extensive; for a comprehensive yet concise review of the subject, we refer the reader to the tutorials [29] and [2]; see also pertinent references therein. A common denominator of existing frameworks is that the data distribution is stationary. The work of [30] demonstrates that strictly monotone games have unique solutions and that gradient play converges to it. The modern approach of solving Nash equilibrium problems for continuous games via variational inequalities can be attributed to Facchinei and Pang [31], [32]. For solving strongly monotone variational inequalities, the projected gradient method is capable of converging linearly.

Our work introduces the additional complexity of minimizing communication between agents and hence we use a distributed gradient approach in our optimization algorithm. Distributed gradient methods have been explored extensively in the literature on convex optimization, though less so in that of variational inequalities. We refer the reader to [33] for a review in the convex optimization setting, and [34] for variational inequalities.

Decision-Dependent Data: This paper contributes to the growing body of literature that studies stochastic optimization with decision-dependent data distributions. While the concept of decision-dependent uncertainty has existed within the optimization literature for some time, the formalization via distributional maps is attributed to "Performative Prediction" and its use within the machine learning community [16]. This work posits the formulation of optimization problems in which the data distribution is explicitly dependent on the optimization variables, and proposes repeated retraining (and the limit points thereof) as a solution; these points are referred to as "performatively stable" to distinguish them from "performatively optimal" points as they solve the stationary optimization problem that they induce. Convergence of various stochastic gradient algorithms to performatively stable points are studied in [17], [35] in the batch setting, and in the time-varying setting in [36], [37]. The extension of problems to games includes two-player zero-sum games in [37], and general multiplayer games in [19]. Additional recent extensions to this line of work include distributionally robust optimization [38], [39] and time varying optimization [36], [37].

The prevailing method for finding optimal points, and by extension Nash equilibrium, involve prescribing a model of the distributional map(s) to leverage standard stochastic optimization methods. While derivative free optimization is possible, the necessary restriction to estimators using a single cost evaluation prevents any reasonable rate of convergence [18], [19], [20].

Optimization algorithms for linear parameterizations (i.e. location scale families) are discussed in [18] using a multi-phase approach, where a model is learned *before* the optimization phase, and in [19] where a model is learned during the optimization phase using an adaptive approach. A multi-phase approach suitable for a general class of parameterizations, which is referred to as "regular", is developed in [28]; here, bounds on the resulting excess risk are provided.

This work complement the technical findings of [28] by generalizing the so-called "plug-in" optimization approach of [28] to non-cooperative multiplayer games. Like [19], we focus on finding the Nash equilibrium of strongly-monotone games with decision-dependent distributions. However, i) our optimization framework works for a more general class of models, as opposed to merely linear parameterizations; additionally, ii) our algorithm learns a model before optimization rather than during. In this way, we can separate the number of required interactions with the system from the number of optimization steps.

B. CONTRIBUTIONS

In this work, we provide the following contributions to the body of work on stochastic optimization and Nash equilibrium problems with decision-dependent distributions.

i) We propose an algorithm for finding a Nash equilibrium in stochastic games with decision-dependent distributions where: (i) the distributional map for each player's cost is estimated from samples, and (ii) the estimated distributional map is used in gradient-based strategies.
ii) We provide guarantees on the approximation error of distributional maps for a class of map learning problems.
iii) We show that the parameterized cost approximates the ground-truth in high-probability.
iv) We propose a stochastic gradient-based algorithm for solving a parameterized strongly-monotone game, and we demonstrate linear convergence in expectation.
v) Finally, we provide numerical simulations of an EV charging market formulation using real-world data. The EV market formulation is new in the context of energy markets, thus providing contributions in this area.

C. ORGANIZATION

In Section II, we provide necessary notation and background for our analysis. In Section III we discuss the proposed learning algorithm in detail and present our primary result. Section IV discusses the details of the optimization stage. We provide our numerical simulations in Section V. Proofs of the results are provided in the Appendix.

II. NOTATION AND PRELIMINARIES

Throughout the paper, $\mathbb{R}^d$ denotes the d-dimensional Euclidean space with inner product $\langle \cdot, \cdot \rangle$, and Euclidean norm $\| \cdot \|$. For a matrix $X \in \mathbb{R}^{n\times m}$, $\|X\|$ denotes the spectral norm. For a given integer n, $[n]$ denotes the set $\{1, 2, \ldots, n\}$ and $\mathcal{S}^{n-1}$ denotes the Euclidean hypersphere in n dimensions, $\{x \in \mathbb{R}^n | \; \|x\|_2 = 1\}$. The symbol $\mathbb{1}_d$ is used to denote the d-dimensional vector of all ones. Given vectors $x \in \mathbb{R}^n$ and $z \in \mathbb{R}^m$, we let $(x, z) \in \mathbb{R}^{n+m}$ denote their concatenation.

For a symmetric positive definite matrix $W \in \mathbb{R}^{d\times d}$, the weighted inner product is defined by $\langle x, y\rangle_W = \langle x, Wy\rangle$ and corresponding weighted norm $\|x\|_W = \sqrt{\langle x, x\rangle_W}$ for any $x, y \in \mathbb{R}^d$. The weighted projection onto a set $\mathcal{X} \subseteq \mathbb{R}^d$ with respect to the symmetric positive definite matrix $W \in \mathbb{R}^{d\times d}$ is given by the map

$$\mathrm{proj}_{\mathcal{X},W}(x) := \arg\min_{y\in\mathcal{X}} \frac{1}{2} \|x - y\|_W^2 \tag{3}$$

for any $x \in \mathbb{R}^d$.

A. PROBABILITY MEASURES

Throughout this work, we restrict our focus to random variables drawn from continuous probability distributions supported over the Euclidean space. When random variables $X, Y \in \mathbb{R}^k$ are equal in distribution, i.e., $P(X \le x) = P(Y \le x)$ for all $x \in \mathbb{R}^k$, we write $X \overset{d}{=} Y$.

Our analysis includes study of sub-exponential random vectors. A univariate random variable $X \in \mathbb{R}$ is said to be sub-exponential with modulus $\theta > 0$ provided that the survival function satisfies $\mathbb{P}(|X| \ge t) \le 2\exp(-t/\theta)$ for all $t \ge 0$. By extension, a random vector $X \in \mathbb{R}^k$ is sub-exponential provided that $\langle u, X\rangle$ is a sub-exponential random variable for all $u \in \mathcal{S}^{k-1}$.

To compare probability distributions, we will be interested in computing the distance between their associated probability measures—for which we need a complete metric space. We let $\mathcal{P}(\mathbb{R}^k)$ denote the set of finite first moment probability measures supported on $\mathbb{R}^k$ and write the Wasserstein-1 distance as

$$W_1(\mu, \nu) = \sup_{h\in\mathcal{L}_1} \left\{ \mathbb{E}_{X\sim\mu}[h(X)] - \mathbb{E}_{Y\sim\nu}[h(Y)] \right\}$$

for any $\mu, \nu \in \mathcal{P}(\mathbb{R}^k)$, where $\mathcal{L}_1$ is the set of all 1-Lipschitz continuous functions $h : \mathbb{R}^k \to \mathbb{R}$. Under these conditions, the set $(\mathcal{P}(\mathbb{R}^k), W_1)$ forms a complete metric space [40].

B. GAMES

We consider a game that consists of n players. Each player has a cost function F_i, distributional map D_i, and decision set $\mathcal{X}_i \subseteq \mathbb{R}^{d_i}$. Hence, each player chooses a decision, or strategy $x_i \in \mathcal{X}_i \subseteq \mathbb{R}^{d_i}$. The concatenation of the decision variables is written as $x = (x_1, \ldots, x_n) \in \mathcal{X} \subseteq \mathbb{R}^d$ where $\mathcal{X} = \prod_{i=1}^n \mathcal{X}_i$ and $d = \sum_{i=1}^n d_i$. For a fixed agent i, we will decompose the decision x as $x = (x_i, x_{-i})$ where $x_{-i} \in \mathbb{R}^{d-d_i}$ is the strategy vector of all agents excluding the ith one.

The collection of costs F_i and decision sets $\mathcal{X}_i$ defines the game

$$\min_{x_i\in\mathcal{X}_i} F_i(x_i, x_{-i}), \; i \in [n]. \tag{4}$$

A Nash equilibrium of this game is a point $x^* \in \mathcal{X}$ provided that

$$x_i^* \in \arg\min_{x_i\in\mathcal{X}_i} F_i(x_i, x_{-i}^*) \tag{5}$$

for all $i \in [n]$. Intuitively, x^* is a strategy such that no agent can be incentivized by its cost to deviate from x_i^* when all other agents play x_{-i}^*. Finding Nash equilibria is the primary focus of this work.

Games of this form are commonly cast into a variational inequality framework. This is due, in part, to the observation that the Nash equilibria $x^* \in \mathcal{X}$ are the solutions to the variational inequality

$$\langle x - x^*, G(x^*) \rangle \geq 0, \quad \forall x \in \mathcal{X},$$

where the gradient map $G : \mathbb{R}^d \to \mathbb{R}^d$ is defined as

$$G(x) = (\nabla_1 F_1(x), \ldots, \nabla_n F_n(x)). \tag{6}$$

Here, the notation ∇_i is used to represent the partial gradient ∇_{x_i}. We will denote the set of Nash equilibria of a game with gradient map G and domain $\mathcal{X}$ as $\texttt{NASH}(G, \mathcal{X})$. Existence of solutions to variational inequalities of this form is guaranteed provided that the set $\mathcal{X}$ is convex and compact and the gradient map G is monotone; uniqueness is guaranteed when G is strongly-monotone [31]. We say that G is α-strongly-monotone on $\mathcal{X}$ provided that there exists $\alpha > 0$ such that

$$\langle x - y, G(x) - G(y) \rangle \geq \alpha \|x - y\|^2, \quad \forall x, y \in \mathcal{X}, \tag{7}$$

and monotone when $\alpha = 0$. In this work, we primarily focus on strongly-monotone games. While monotone games are tractable, methods for solving them with decision-dependent distributions require alternative gradient estimators—a topic we leave to future work.

C. MONOTONICITY IN DECISION-DEPENDENT GAMES

In this work, we introduce the additional complexity to the formulation in (4) that the F_i's are the expected cost over a distributional map $D_i : \mathbb{R}^d \to \mathcal{P}(\mathbb{R}^{k_i})$. In particular, we write the cost as

$$F_i(x_i, x_{-i}) := \mathop{\mathbb{E}}_{z_i \sim D_i(x_i, x_{-i})} f_i(x_i, x_{-i}, z_i). \tag{8}$$

This can be written alternatively as the integral

$$F_i(x) = \int_{\mathbb{R}^{k_i}} f_i(x, z_i) p_i(z_i, x) \mathrm{d}z_i \tag{9}$$

where p_i is the probability density function for the distribution $D_i(x)$. When the integral satisfies the Dominated Convergence Theorem, computing the gradient amounts to differentiating under the integral and using the product rule. We then obtain

$$\nabla_i F_i(x) = \mathop{\mathbb{E}}_{z_i \sim D_i(x)} \left[\nabla_{x_i} f_i(x, z_i) + f_i(x, z_i) \nabla_i \log p_i(z_i, x) \right], \tag{10}$$

where we recall that $G(x) = (\nabla_1 F_1(x), \ldots, \nabla_n F_n(x))$. In short, characterizing the gradient of this decision-dependent game requires assumptions not only on f_i, but also on the properties of the distributional map D_i. Sufficient conditions for strong monotonicity of the game in (4) are due to [19] and are stated in terms of the *decoupled* costs, given by

$$F_i(x, y) = \mathop{\mathbb{E}}_{z_i \sim D_i(y)} f_i(x, z_i) \tag{11}$$

for all $x, y \in \mathbb{R}^d$, and their associated decoupled partial gradients

$$G_i(x, y) = \mathop{\mathbb{E}}_{z_i \sim D_i(y)} \nabla_i f_i(x, z_i), \tag{12}$$

for all $x, y \in \mathbb{R}^d$ and

$$H_i(x, y) = \nabla_{y_i} \mathop{\mathbb{E}}_{z_i \sim D_i(y)} f_i(x, z_i) \tag{13}$$

for all $x, y \in \mathbb{R}^d$. A key observation used in the proof is that $G_i(x) = \nabla_i F_i(x) = G_i(x, x) + H_i(x, x)$.

Theorem 1 (Strong Monotonicity, [19]): Suppose that,

i) For all $y \in \mathcal{X}$, $x \mapsto G(x, y)$ is λ-strongly monotone,
ii) For all $x \in \mathcal{X}$, $y \mapsto H(x, y)$ is monotone, and that for all $i \in [n]$,
iii) For all $x \in \mathcal{X}, z_i \mapsto \nabla_i f_i(x, z_i)$ is L_i-Lipschitz continuous,
iv) $y \mapsto D_i(y)$ is γ_i-Lipschitz continuous on $(\mathcal{P}(\mathbb{R}^{k_i}), W_1)$.

Set $\kappa = \sqrt{\sum_{i=1}^n (\frac{\gamma_i L_i}{\lambda})^2}$. Then if $\kappa < 1/2$, $x \mapsto G(x)$ is $\alpha = (1 - 2\kappa)\lambda$-strongly monotone. □

III. LEARNING-BASED DECISION-DEPENDENT GAMES

In this work, we aim to solve the stochastic Nash equilibrium problem with decision-dependent data distributions as formulated in (1). Methods for finding Nash equilibrium for games with decision dependent data distributions either use derivative free optimization, at the expense of an extremely slow rate, or use derivative information in conjunction with a learned model of the distributional map [19].

In [28], it is shown that a "plug-in" optimization approach, whereby a model for the distributional map is learned from samples prior to optimization, yields a bounded excess risk for the convex optimization problems with decision-dependent data. In this work, we leverage the properties of the system to simplify the communication structure of our approach, which we depict in Fig. 1.

We assume that each agent i deploys a decision x_i to the system in order to receive realizations z_i. Once deployed, the decisions $\{x_i\}$ are made public and available to all other agents. Given that we assume that elements of the system are able to respond to deployed decisions $\{x_i\}$, such as a population of strategic human users of a service provided by competitors $i \in [n]$, it is reasonable to assume that other agents may respond to the decisions of agent i as well. In this way, agent i may receive x_{-i} without requiring cooperation or coordination with other agents.

To accommodate this setting, our algorithm proposes a multi-phase approach consisting of the following phases: (i) sampling; (ii) learning; (iii) optimization. It is important to note that, following the learning phase, players only need to participate in gradient play without receiving any additional feedback from the system in the form of $z_i \sim D_i(x)$. This is distinct from existing approaches in which performatively stable points can only be reached after several (even thousands of) rounds of feedback [16], [19], [20], and performatively

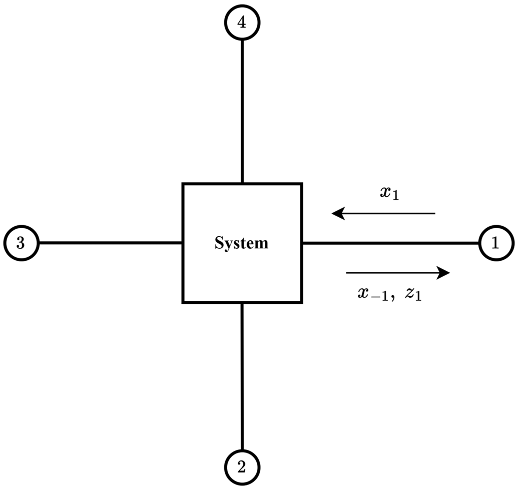


FIGURE 1. **Communication structure allows agents to interact with the system in square by sending decision x_i. After deploying, agents can receive feedback from the system in the form of other agents decisions x_{-i} and data z_i.**

Algorithm 1: Multi-phase Optimization.

Input: $m, \{D_{x_i}\}_{i=1}^{n}$
for $j \in [m]$ **do**
  for $i \in [n]$ **do**
    Draw $x_i^{(j)} \sim D_{x_i}$;
  end
  Deploy $x^{(j)}$;
  Observe $z_i^{(j)} \sim D_i(x^{(j)})$;
end
for $i \in [n]$ **do**
  Fit $\widehat{\beta}_i \in \arg\min_{\beta_i \in \mathcal{B}_i} \frac{1}{m} \sum_{j=1}^{m} R_i(x^{(j)}, z_i^{(j)}, \beta_i)$;
end
Compute $\widehat{x} \in \mathtt{Nash}(G_{\widehat{\beta}}, \mathcal{X})$;

optimal points can only be reached for models known to be location scale families a priori [18], [19].

Sampling: In the sampling phase, we require that each player to design a distribution of decisions D_{x_i} and to deploy decision samples $\{x_i^{(j)}\}_{j=1}^{m} \overset{i.i.d}{\sim} D_{x_i}$ so that they can collectively receive feedback $z_i^{(j)} \sim D_i(x^{(j)})$ from the system (in response to their deployed decisions $\{x_i^{(j)}\}_{j=1}^{m}$). The result is that each agent has access to a dataset $\{x^{(j)}, z_i^{(j)}\}_{j=1}^{m}$ which they can use to learn their distributional map D_i.

Learning: In this procedure, each player will choose a hypothesis class of parameterized functions

$$\mathcal{H}_{\mathcal{B}_i} = \left\{ D_{\beta_i} | \ \beta_i \in \mathcal{B}_i \subseteq \mathbb{R}^{\ell_i} \right\}, \tag{14}$$

as well as a suitable criterion or risk function R_i, to formulate their own expected risk minimization problem

$$\beta_i^* \in \arg\min_{\beta_i \in \mathcal{B}_i} \ \mathop{\mathbb{E}}_{x \sim D_x, z_i \sim D_i(x)} R_i(x, z_i, \beta_i) \tag{15}$$

over the random variable (x, z_i) drawn from the coupled distribution $(D_x, D_i(x))$. Then, using the set of samples from the previous sampling phase, they can formulate the corresponding empirical risk minimization (ERM) problem

$$\widehat{\beta}_i \in \arg\min_{\beta_i \in \mathcal{B}_i} \frac{1}{m} \sum_{j=1}^{m} R_i(x^{(j)}, z_i^{(j)}, \beta_i). \tag{16}$$

The result is a learned distributional map $D_{\widehat{\beta}_i}$ approximating D_i, which we can now use to solve the approximate Nash equilibrium problem.

Optimization: Following the approximation phase, each player now has a learned model of their distributional map $D_{\widehat{\beta}_i}$, which can be used to formulate an approximation of the ground-truth cost F_i and hence an approximate Nash equilibrium problem:

$$\widehat{x}_i \in \arg\min_{x_i \in \mathcal{X}_i} F_{\widehat{\beta}_i}(x_i, \widehat{x}_{-i}) \tag{17}$$

for all $i \in [n]$, where

$$F_{\widehat{\beta}_i}(x_i, \widehat{x}_{-i}) := \mathop{\mathbb{E}}_{z_i \sim D_{\widehat{\beta}_i}(x_i, \widehat{x}_{-i})} f_i(x_i, \widehat{x}_{-i}, z_i)\,. \tag{18}$$

Hereafter, we denote the Nash equilibrium of the approximate game as $\widehat{x}$ to distinguish it from the ground truth x^*. In Algorithm 1, we write the set of Nash equilibria for the operator $G_{\widehat{\beta}}$ with domain $\mathcal{X}$ as $\mathtt{Nash}(G_{\widehat{\beta}}, \mathcal{X})$. In practice, we will assume the necessary hypotheses to guarantee uniqueness of this assignment; in which case the set inclusion is simply an equality.

By solving (17) instead of (1) we have introduced two errors: (i) the approximation error of the distributional map D_i by elements of the hypothesis class $\mathcal{H}_{\mathcal{B}_i}$, and (ii) the estimation or statistical error by solving the ERM problem instead of the expected risk minimization problem. In [28], the main result demonstrates that these two sources of error propagate through the optimization problem, and that the resulting excess risk can be bounded in terms of the sample complexity m. Our goal is to expand this result and provide additional analysis to our setting.

A. PARAMETER ESTIMATION FOR REGULAR PROBLEMS

A critical component of our analysis is the estimation or learning of the distributional map and the subsequent characterization of the estimation error. In this section, we outline a class of expected risk minimization problems, which we call *regular problems*, for which we can characterize the distance between expected risk minimization solutions and empirical risk minimization solutions. Throughout, we write $R_i(\beta_i) = \mathbb{E}_{(x,z)}[R(x, z_i, \beta_i)]$ and $\hat{R}_i(\beta_i) = (1/m) \sum_{j=1}^{m} R_i(x^{(j)}, z_i^{(j)}, \beta_i)$ for $\beta_i \in \mathbb{R}^{\ell_i}$ to denote the expected and empirical risk, respectively.

Definition 1 (Map Learning Regularity): A map learning problem, consisting of the optimization problems with costs R_i and $\hat{R}_i$ over $\mathcal{B}_i$, is regular provided that:

a) *Convexity:* The expected risk $\beta_i \mapsto R_i(\beta_i)$ is μ_i-strongly convex, and the empirical risk $\beta_i \mapsto \hat{R}_i(\beta_i)$ is convex.
b) *Smoothness:* For all realizations of $x \in \mathcal{X}$ and $z_i \in \mathbb{R}^{k_i}$, $\beta_i \mapsto \nabla_{\beta_i} R_i(x, z_i, \beta_i)$ is L_{β_i}-Lipschitz continuous.
c) *Boundedness:* The set $\mathcal{B}_i \subseteq \mathbb{R}^{\ell_i}$ is convex and compact.
d) *Sub-Exponential gradient:* For all $\beta_i \in \mathcal{B}_i$, $\nabla_{\beta_i} R_i(x, z_i, \beta_i)$ is a sub-exponential vector with parameter $\theta_i > 0$. □

Items (a) and (c), taken together, guarantee existence of $\widehat{\beta}$ and uniqueness of β^* as defined in (16) and (15), respectively. Furthermore, the inclusion of item (b) is necessary to guarantee that first-order stochastic gradient methods will converge at least sub-linearly to $\widehat{\beta}$. Lastly, the heavy-tail assumption [41] will allow us to describe the concentration of the gradient estimates. Together, they allow us to relate the solutions to the sample complexity in the following lemma.

Lemma 2 (Uniform Gradient Bound): If the smoothness and sub-exponential gradient assumptions in Definition 1 hold for player $i \in [n]$, then for any $\delta \in (0, 1/2)$ and any m such that $m/\log(m) \geq 2(\ell_i + \log(1/\delta))$, we have that:

$$\sup_{\beta \in \mathcal{B}} \left\| \nabla \hat{R}_i(\beta) - \nabla R_i(\beta) \right\| \leq C_i \sqrt{\frac{\log(m)(\ell_i + \log(1/\delta))}{m}} \tag{19}$$

with probability at least $1 - \delta$, where $C_i = 4\max\{L_{\beta_i}/15r_i, \theta_i\}$. □

The proof of this result is provided in the Appendix A1. This result offers a broad generalization of [42, Equation (19b)] to any risk with Lipschitz-continuous sub-exponential gradients over any convex and compact set. Our result is comparable to the $\mathcal{O}(\sqrt{\ell_i m})$ rate that can be found for specific problem instances such as linear least squares regression and logistic regression, but with the addition of a $\sqrt{\log m}$ factor. Indeed, the generality of the risk function requires that we enforce compactness of the domain, thus giving rise to this extra logarithmic factor. This gradient estimation result will now allow us to reach our desired bounded distance result, which we present in the following theorem.

Theorem 3 (ERM Approximation): If the map learning problem is regular for player $i \in [n]$ (i.e., it satisfies the assumptions in Definition 1), then for any $\delta \in (0, 1/2)$ and any m such that $m/\log(m) \geq 2(\ell_i + \log(1/\delta))$ we have that:

$$\left\| \widehat{\beta}_i - \beta_i^* \right\| \leq C_i' \sqrt{\frac{\log(m)(\ell_i + \log(1/\delta))}{m}} \tag{20}$$

with probability at least $1 - \delta$, where $C_i' = (4/\mu_i)\max\{L_{\beta_i}/15r_i, \theta_i\}$. □

The proof of Theorem 3 is provided in the Appendix A2. The power in this characterization lies in the fact that it holds for any statistical learning problem satisfying the assumptions listed in Definition 1, and is not specific to the setting of learning distributional maps. We note that our Definition 1, which is a property used in the Theorem 3, is different from the one in [28] and it involves conditions that are easier to check.

As an example, we provide conditions for which a linear least squares problem satisfies the regularity conditions and hence is subject to the above ERM approximation result.

Proposition 4 (Linear Least Squares Regularity): Consider the linear least squares problem with expected risk problem

$$B_i^* \in \arg\min_{B \in \mathcal{B}_i} \frac{1}{2} \mathbb{E}_{(x, z_i)} \|Bx - z_i\|^2,$$

and empirical risk minimization problem

$$\hat{B}_i \in \arg\min_{B \in \mathcal{B}_i} \frac{1}{2m} \sum_{j=1}^{m} \left\| Bx^{(j)} - z_i^{(j)} \right\|^2.$$

Let $x \sim D_x$ with zero mean and covariance matrix Σ. If

i) There exist $\gamma_i, L_i > 0$ such that $\gamma_i I \preceq \Sigma_i \preceq L_i I$,
ii) The entries of xx^T and $z_i x^T$ are sub-exponential,
iii) The constraint set $\mathcal{B}_i$ is convex and compact.

Then, the map learning problem is regular. □

The proof of Proposition 4 is provided in Appendix A3. Deriving conditions for the more general case of non-linear regression is attainable but outside the scope of this work.

B. BOUNDING THE APPROXIMATION ERROR

Finding a relationship between $\widehat{x}$ and x^* will require that we first characterize an appropriate hypothesis class of distributions for learning. Here, we formalize the notion of misspecification and sensitivity for a hypothesis class $\mathcal{H}_{\mathcal{B}_i}$.

Definition 2 (Misspecification [28]): A hypothesis class $\mathcal{H}_{\mathcal{B}_i}$ is η_i-misspecified provided that there exists a $\eta_i > 0$ such that

$$W_1(D_{\beta_i^*}(x), D_i(x)) \leq \eta_i \tag{21}$$

for all $x \in \mathcal{X}$. □

We note that, although η_i is not known to agents in practice, it is a useful conceptual quantity that can be used to represent the expressiveness of the parameterization relative to the ground truth; it also captures the ability of the chosen risk function to fit a parameterization. This is similar to the notion of approximation error used in classical statistical learning methods [43]. However, unlike this setting, we note that $\eta_i = 0$ implies that $D_{\beta_i^*}(x) = D_i(x)$ for all $x \in \mathcal{X}$; hence, $z \sim D(x)$ and $z' \sim D_{\beta_i^*}(x)$ yields $z \overset{d}{=} z'$ but not necessarily $z = z'$ almost everywhere as we might like.

Definition 3 (Sensitivity [28]): The hypothesis class $\mathcal{H}_{\mathcal{B}_i}$ is ε_i-sensitive if, for any $\beta_i, \beta_i' \in \mathcal{B}_i$,

$$W_1(D_{\beta_i}(x), D_{\beta_i'}(x)) \leq \varepsilon_i \left\| \beta_i - \beta_i' \right\| \tag{22}$$

for all $x \in \mathcal{X}$. □

Sensitivity of $\mathcal{H}_{\mathcal{B}_i}$ is merely a convenient name for the condition that $\beta \mapsto D_{\beta_i}(x)$ be ε_i-Lipschitz continuous for all realizations of $x \in \mathcal{X}$. In the result that follows, we demonstrate that an appropriately misspecified and sensitive hypothesis class induces a cost that has bounded distance to the ground truth cost in (1).

Theorem 5 (Bounded Approximation): Suppose that the following conditions hold for all $i \in [n]$:

i) The hypothesis class $\mathcal{H}_{\mathcal{B}_i}$ is η_i-misspecified, and ε_i-sensitive.
ii) The map learning problem is regular.
iii) For all $x \in \mathcal{X}_i$, $z_i \mapsto f_i(x, z_i)$ is L_{z_i}-Lipschitz continuous.

Then, the bound

$$|F_{\widehat{\beta}_i}(x) - F_i(x)| \leq \eta_i L_{z_i} + L_{z_i}\varepsilon_i C_i' \zeta_i(m, \delta), \tag{23}$$

holds with probability $1 - \delta$ for any $x \in \mathcal{X}$, where

$$\zeta_i(m, \delta) := \sqrt{\frac{\log(m)(\ell_i + \log(1/\delta))}{m}}. \tag{24}$$

and where C_i' is as in (20). □

The proof of Theorem 5 is provided in the Appendix A4. Note that since each F_{β_i} is assumed to be continuously differentiable and $\mathcal{X} \subseteq \mathbb{R}^d$ is compact, then $x \mapsto F_{\beta_i}(x)$ is L_{β_i}-Lipschitz continuous on $\mathcal{X}$ with

$$L_{\beta_i} = \max_{x \in \mathcal{X}} \left\| \nabla F_{\beta_i}(x) \right\|. \tag{25}$$

Leveraging this fact allows us to demonstrate that the excess cost can be bounded—an analog of the main result in [28].

Corollary 6: Suppose that the hypothesis of Theorem 5 holds. Then,

$$|F_i(\widehat{x}) - F_i(x^*)| \leq 2\eta_i L_{z_i} + 2L_{z_i}\varepsilon_i C_i' \zeta_i(m, \delta) + 2\max\{L_{\widehat{\beta}_i}, L_{\beta_i^*}\}\mathsf{diam}(\mathcal{X}_{-i}) \tag{26}$$

hold with probability $1 - \delta$ for any $\widehat{x} \in \mathtt{NASH}(G_{\widehat{\beta}}, \mathcal{X})$ and $x^* \in \mathtt{NASH}(G, \mathcal{X})$, where C_i' is as in (20) and $\mathcal{X}_{-i} = \prod_{j \neq i} \mathcal{X}_j$. □

The analysis in this section demonstrates that the estimation procedure in Algorithm 1 yields a cost function that approximates the original cost in (1) with an error that decreases as the number of samples increases. Furthermore, this bound exists independent of the conditioning of the Nash equilibrium problem we solve in the optimization phase. We note that (23) is similar to the result in [28], but it is based on a different definition of regular problem (see Definition 1); the bound (26) is unique to this paper.

In the section that follows, we examine a family of hypothesis classes that allows the approximated game to be monotone, and provide suitable algorithms for solving them with convergence guarantees.

IV. SOLVING STRONGLY-MONOTONE DECISION-DEPENDENT GAMES

Since the agents lack full knowledge of the system and hence the ground truth distributional map D_i in (1), we cannot hope to enforce that D_i satisfy any assumptions to encourage tractability of our optimization problem. We can however impose conditions on the hypothesis class $\mathcal{H}_{\mathcal{B}_i}$, which is chosen by the agents. To successfully find a Nash equilibrium of the approximate problem in (17), it will be crucial that agents choose a class that balances expressiveness of the system (thereby making η_i small) with tractability of the optimization.

Perhaps the simplest model capable of achieving this goal is the location-scale family [18], [19], [20]. In our setting, a location scale family parameterization for agent i is a distributional map D_{B_i} having matrix parameter $B_i \in \mathbb{R}^{k_i \times d}$ where $z_i \sim D_{B_i}$ if and only if

$$z_i \overset{d}{=} \xi_i + B_i x \tag{27}$$

for stationary random variable $\xi_i \sim D_{\xi_i}$. We note that this parameterization can be written alternatively as $z_i \overset{d}{=} \xi_i + B_i^i x_i + B_{-i}^i x_{-i}$, where $B_i^i \in \mathbb{R}^{k_i \times d_i}$ and $B_{-i}^i \in \mathbb{R}^{k_i \times (d-d_i)}$ are block matrices such that $B_i x = B_i^i x_i + B_{-i}^i x_{-i}$ due to linearity. The resulting partial gradient has the form

$$\nabla_i F_{B_i}(x) = \mathbb{E}_{z_i \sim D(x)}\left[\nabla_i f_i(x, z_i) + (B_i^i)^T \nabla_{z_i} f_i(x, z_i)\right],$$

which is typically much simpler to analyze than alternative models. Intuitively, this model allows us to express z_i as the sum of a stationary random variable from a base distribution with a linear factor depending on x, where the matrix parameter B_i weights the responsiveness of the population to the agents decisions.

This model is particularly appealing since guarantees for learning B_i are known and established in Proposition 4. Moreover, the matter of expressiveness is due to the fact that location scale families are a particular instance of strategic regression [16], [28], in which member of the population interact with agents by modifying their stationary data (such as features in a learning task) ξ_i in an optimal way upon observing x:

$$z_i \overset{d}{=} \arg\min_y \left[-u_{\beta_i}(x, y) + \frac{1}{2}\|y - \xi_i\|^2\right],$$

where u_{β_i} is a utility function parameterized by $\beta_i \in \mathcal{B}_i$ corresponding to the utility that members of the population derive from changing their data in response to the decisions in x; and the quadratic term $1/2\|y - \xi_i\|^2$ is the cost of changing their data from ξ_i to y. Indeed when $u_{\beta_i}(x,) = \langle y, B_i x\rangle$ for $\beta_i = B_i \in \mathbb{R}^{k_i \times d}$, we recover the form in (27).

Furthermore, location scale families immediately satisfies several of the assumption required for further analysis. In particular, it is known that Sensitivity (Definition 3) holds with $\varepsilon_i = \max_{x \in \mathcal{X}} \|x\|^2$, Lipschitz continuity of $x \mapsto D_{B_i}$ holds with $\gamma_i = \|B_i\|^2$, and Lipschitz continuity of G_{B_i} holds due to the following result.

Lemma 7 (Lipschitz Gradient, [19]): Suppose that D_{β_i} is such that $z_i \overset{d}{=} B_i x + \xi_i$ with $\beta_i = B_i$, and that for each $i \in [n]$ there exists $\zeta_i \geq 0$ such that $(x, z_i) \mapsto \nabla_{i, z_i} f_i(x, z_i)$ is ζ_i-Lipschitz continuous. Then G_{β_i} is L-Lipschitz continuous with

$$L := \sqrt{\sum_{i=1}^n \zeta_i^2 \max\{1, \left\|B_i^i\right\|^2\}(1 + \|B_i\|^2)}. \tag{28}$$

□

Strong monotonicity will follow from Theorem 1 provided that G_{β_i} satisfy the remaining hypothesis on the f_{β_i}—which tends to be on a case-by-case basis. We will not require that G_{β_i} use this parameterization in our analysis, however we can proceed with the knowledge that a model class satisfying our hypotheses does exist.

A. DISTRIBUTED GRADIENT-BASED METHOD

In our optimization phase, we seek to use a gradient-based algorithm that respects the agent's communication structure with the system. For the sake of readability, we will suppress the β_i subscript and instead refer to quantities G_i keeping in mind that they will correspond to the approximate Nash equilibrium problem in (17) with solution $\widehat{x}$.

We will assume that each agent has access to an estimator of the gradient $\nabla_i F_i$ and is capable of projecting onto their decision set $\mathcal{X}_i$. In the constant step-size setup, each agent chooses a rate $\omega_i > 0$ and performs the update

$$x_i^{t+1} = \mathsf{proj}_{\mathcal{X}_i}\left(x_i^t - \omega_i^{-1} g_i^t\right),$$

where g_i^t is a stochastic gradient estimator for $\nabla_i F_i$ used at iteration t, which is then reported to the system and made available to all agents. For the sake of analysis, we will assume without loss of generality that the step-sizes satisfy the ordering

$$\omega_1 \geq \omega_2 \geq \ldots \geq \omega_n$$

and hence $\omega_1 = \max_{i\in[n]} \omega_i$ and $\omega_n = \min_{i\in[n]} \omega_i$. The collective update can be written compactly as

$$x^{t+1} = \mathsf{proj}_{\mathcal{X},W}\left(x^t - W^{-1} g^t\right), \tag{29}$$

where $W = \mathsf{diag}(\omega_1 \mathbb{1}_{d_1}, \ldots, \omega_n \mathbb{1}_{d_n})$ and g^t is an estimator for $G(x^t)$ at iteration t. Convergence of this procedure hinges on the following assumptions.

Assumption 1: The gradient function $G : \mathcal{X} \subseteq \mathbb{R}^d \to \mathbb{R}^d$ is α-strongly monotone and L-Lipschitz continuous.

Assumption 2 (Stochastic Framework): Let $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ with elements

$$\mathcal{F}_t = \sigma(g^\tau, \tau \leq t) \tag{30}$$

be the natural filtration of the Borel σ-algebra over $\mathbb{R}^d$ with respect to g^t, and use the short-hand notation $\mathbb{E}_t \cdot := \mathbb{E}_{z\sim D(x^t)}[\cdot|\mathcal{F}_t]$ as the conditional expectation over the product distribution $D(x^t) = \prod_{i=1}^n D_i(x^t)$. There exist bounded sequences $\{\rho^t\}_{t\geq 0}, \{\sigma^t\}_{t\geq 0} \subseteq \mathbb{R}_+$ such that

$$(\textbf{\textit{Bias}}) \qquad \left\|\mathbb{E}_t g^t - G(x^t)\right\| \leq \rho^t$$

$$(\textbf{\textit{Variance}}) \qquad \mathbb{E}_t \left\|g^t - \mathbb{E}_t g^t\right\|^2 \leq (\sigma^t)^2$$

where $\rho^t \leq \rho$ and $\sigma^2 \leq \sigma$ for all $t \geq 0$.

Assumption 1 is standard for guaranteeing convergence of gradient play [31], and the uniformly bounded variance component of Assumption 2 is standard for convergence for stochastic algorithms. As we will show shortly, convergence with bias is possible and the result reduces to the unbiased case when $\rho^t = 0$ for all t. The next result will quantify the one-step improvement of (33).

Lemma 8 (One-step Improvement): Let Assumptions 1 and 2 hold. Then, the sequence generated by iteration (29) satisfies:

$$\mathbb{E}_t \left\|x^{t+1} - \widehat{x}\right\|_W^2 \leq \frac{\omega_1}{\omega_n + \alpha}\left\|x^t - \widehat{x}\right\|_W^2 + \frac{2\omega_1\left(\omega_1\rho^2 + \alpha\sigma^2\right)}{\alpha\omega_n\left(\omega_n + \alpha\right)}$$

for all $t \geq 0$, provided that $\omega_1/\omega_n^2 \leq \alpha/(4L^2)$. □

The proof of Lemma 8 is provided in the Appendix A6. We note that setting $\omega_i = \omega$ for some $\omega > 0$ recovers the result in [19, Theorem 15]. Following this one-step analysis, we can show convergence to a neighborhood of the Nash equilibrium.

Theorem 9 (Neighborhood Convergence): Let Assumptions 1 and 2 hold, and suppose that $(\omega_1 - \omega_n) < \alpha$. Then,

$$\limsup_{t\to\infty} \mathbb{E}\left\|x^t - \widehat{x}\right\|^2 \leq \frac{2\omega_1\left(\omega_1\rho^2 + \alpha\sigma^2\right)}{\alpha\omega_n\left(\omega_1 - \omega_n + \alpha\right)}. \tag{31}$$

The proof can be found in Appendix A7. The result shows that the algorithm converges linearly to a neighborhood of the Nash equilibrium $\widehat{x}$, where the radius of the neighborhood is dictated by the step-size, variance, and bias bounds. When $\rho = \sigma = 0$, we retrieve linear convergence. In order to converge to $\widehat{x}$ directly, we will require a decaying step-size policy. For example, we consider the following policy:

$$\omega^t = \frac{\alpha(r + t - 2)}{2} \tag{32}$$

for fixed constant $r > 2$, which we assumed to be shared by all agents. Hence, the decaying step-size update is given by

$$x^{t+1} = \mathsf{proj}_{\mathcal{X}}\left(x^t - (\omega^t)^{-1} g^t\right). \tag{33}$$

In the theorem that follows, we show that this sequence converges to $\widehat{x}$ provided that the bias shares an asymptotic rate with $(\omega^t)^{-1}$.

Theorem 10 (Convergence): Suppose that Assumptions 1 and 2 hold and that there exists $\bar{\rho}, s > 0$ such that

$$\left\|\mathbb{E}_t g^t - G(x^t)\right\| \leq \frac{\bar{\rho}}{s + t} \tag{34}$$

for all $t \geq 0$. Then,

$$\mathbb{E}\left\|x^t - \widehat{x}\right\|^2 \leq \frac{A}{\alpha^2(r + t)} \tag{35}$$

where

$$A = \max\left\{\alpha^2 r\left\|x^0 - \widehat{x}\right\|^2, 4\bar{\rho}^2 \max\left\{\frac{r}{s}, 1\right\} + \frac{8r\sigma^2}{r - 2}\right\}.$$

□

The proof of this result follows by a standard induction argument, and can be found in Appendix A8.

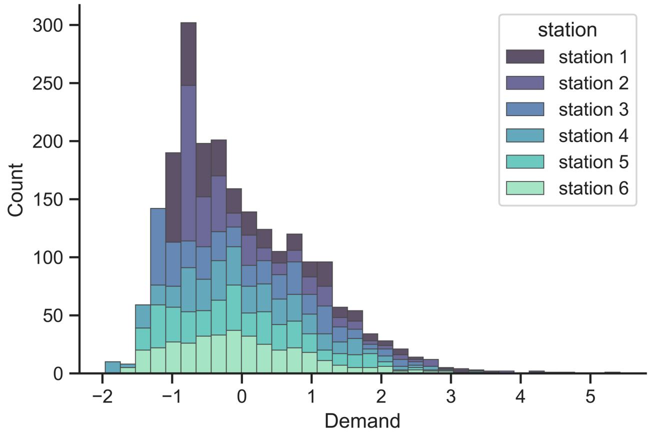


FIGURE 2. Standardized demand data for six medium demand EVCS's consisting of either 2 or 6 ports and port power values of 50, 150, and 350 kWh. Standardization maps raw demand instances to instances of demand that are deviations from the average at each station.

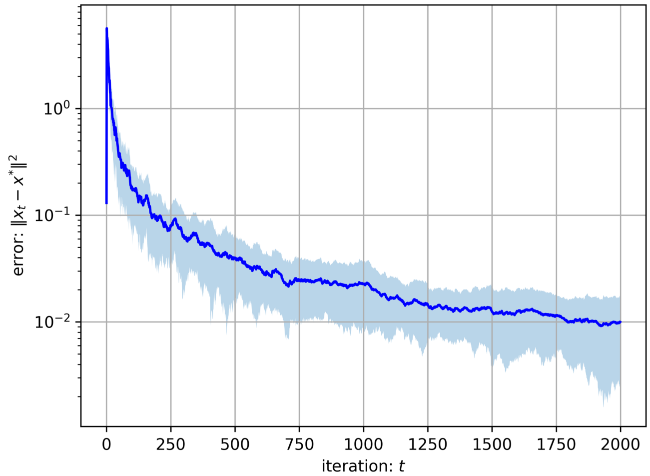


FIGURE 3. Expected error curve and confidence interval for regularized stochastic gradient descent with decaying step size for a location-scale model.

V. NUMERICAL EXPERIMENTS ON ELECTRIC VEHICLE CHARGING

In this section, we consider a competitive game between n distinct electric vehicle charging station operators, where stations are equipped with renewable power sources. The goal of each player is to set prices to maximize their own profit in a system where demand for their station will change in response to the prices set by other competing stations as well. The cost function (negative profit) takes the form

$$f_i(x, z_i) = \underbrace{-z_i x_i + \frac{\lambda_i}{2} x_i^2}_{\text{service profit}} - \underbrace{p_w \phi(w_i - z_i)}_{\text{renewable profit}} + \underbrace{p_r \phi(z_i - w_i)}_{\text{operational cost}}$$

where $\phi(y) = \log(1 + \exp(y))$ for all $y \in \mathbb{R}$. The renewable profit and operational cost terms allow us to describe the trade-off between profit from renewable power generation sold to the grid at rate p_w, and surplus power required from the grid to meet demand at rate p_r. To set prices, we can formulate a Nash equilibrium problem over the expected costs $F_i(x) = \mathbb{E}_{z_i \sim D_i(x)}[f_i(x, z_i)]$ for $i \in [n]$ and $x \in \mathcal{X} = \Pi_{i=1}^n \mathcal{X}_i$, where $\mathcal{X}_i = [p_w, p_r]$ is the interval of price values between the wholesale and retail price.

Since the set of reasonable prices will be quite small, we hypothesize that the the price and demand have a linear relationship of the form $z_i \overset{d}{=} \xi_i + \langle b_i, x_i \rangle$ where $b_i \in \mathbb{R}^n$ with $\xi_i \sim D_{\xi_i}$ corresponding to the base demand. Since we have a simple model, the first and second derivatives can be computed in closed form, and the relevant constants can be computed directly. Indeed, we find that the hypothesis of Theorem 1 are satisfied with $\lambda = \min_i \lambda_i$ which we set to 1, $L_i = 1$,and $\gamma_i = \|b_i\|^2$. We conclude that $G : \mathbb{R}^n \to \mathbb{R}^n$ is $\alpha = (1 - 2\|B\|_F)$-strongly monotone with where B is the parameter matrix whose columns are b_i.

Our data in Fig. 2 depicts the demand of electricity across an hour-long period for 6 ports of varying power profiles for each day in year. We standardize the data to be zero mean and unit variance across each station. Solutions are calculated by performing expected gradient play with constant step size; the expected mean is estimated via the empirical mean over the data set.

We set $b_{ii} = -1/18 + \nu$ and $b_{ij} = 1/18 + \nu$, where we use $\nu \sim \mathcal{N}(0, 10^{-5})$ to simulate learning B from samples. Hence demand for agent i decreases as their own price increases, and increases as the price of other agents decreases. We run the stochastic gradient play algorithm initialized at $x^0 = p_r \mathbb{1}_n$ with a single sample at each round and a decaying step size policy $\omega_t = \alpha(r + t - 2)/2$ for $r = 3$. In Fig. 3 we plot the mean error trajectory an confidence interval over 50 trials of 2000 iterations.

VI. CONCLUSION

In this work, we studied a class of stochastic Nash equilibrium problems, characterized by data distributions that are dependant on the decisions of all involved players. We showed that a learning-based approach enables the formulation of an approximate Nash equilibrium problem that is solvable using a stochastic gradient play algorithm. The results of this procedure is a cost that can be related to cost of the original Nash equilibrium problem via an error that depends on both our approximation and estimation error. To demonstrate the flexibility of these findings, we simulated these techniques in an electric vehicle charging market problem in which service providers set prices, and users modify their demand based on prices set by providers. Future research will look at a scenario where the estimate of the distributional map is improved during the operation of the algorithm, based on the feedback received. Future applications will demonstrate the efficacy of more complex models (beyond linear)

APPENDIX

A1 PROOF OF LEMMA 2

For the sake of notation convenience, and visual clarity, we will suppress the i index throughout the proof. We denote the gradient error by $J(\beta) = \nabla\hat{R}(\beta) - \nabla R(\beta)$ for all $\beta \in \mathbb{R}^\ell$.

To begin, we will generate coverings for the unit sphere in $\mathbb{R}^\ell$ and $\mathcal{B} \subseteq \mathbb{R}^\ell$ and use a discretization argument to create bounds over these finite sets. Fix $\beta \in \mathcal{B}$ and $u \in \mathcal{S}^{\ell-1}$. Let $\{u_j\}_{j=1}^N$ be an arbitrary 1/2-covering of the sphere $\mathbb{S}^{d_{\ell_i}}$ with respect to the Euclidean norm. From [44, Lemma 5.7], we know that $N \leq 5^\ell$. From our covering, we have that there exists u_j in the covering such that $\|u - u_j\| \leq 1/2$. Hence,

$$\begin{aligned}\langle u, J(\beta)\rangle &= \langle u_j + (u - u_j), J(\beta)\rangle \\ &= \langle u_j, J(\beta)\rangle + \langle u - u_j, J(\beta)\rangle \\ &\leq \langle u_j, J(\beta)\rangle + \|u - u_j\| \, \|J(\beta)\| \\ &\leq \langle u_j, J(\beta)\rangle + \frac{1}{2}\|J(\beta)\| \\ &\leq \max_{j\in[N]}\langle u_j, J(\beta)\rangle + \frac{1}{2}\|J(\beta)\| \, .\end{aligned}$$

Since this is true for any $u \in \mathcal{S}^{d-1}$, then it holds for $u = J(\beta)/\|J(\beta)\|$. Thus the above becomes

$$\|J(\beta)\| \leq 2\langle u_j, J(\beta)\rangle \leq 2\max_{j\in[N]}\langle u_j, J(\beta)\rangle. \tag{36}$$

Now we fix $\varepsilon \in (0, 1]$, and choose and ε-covering for the set $\mathcal{B}$, which we will write as $\{\beta_k\}_{k=1}^M$. Recall that $\mathcal{B}$ is bounded, so there exists a constant $r > 0$ such that for all $\beta \in \mathcal{B}$, $\|\beta\| \leq r$. Hence $\mathcal{B} \subseteq B(r)$. From [41, Proposition 4.2.12], we have that

$$M \leq \frac{\mathrm{vol}\left(\mathrm{B(r)} + \frac{\varepsilon}{2}\mathrm{B(1)}\right)}{\mathrm{vol}\left(\frac{\varepsilon}{2}\mathrm{B(1)}\right)} = \frac{\mathrm{vol}\left(\frac{3}{2}\mathrm{B(r)}\right)}{\mathrm{vol}\left(\frac{\varepsilon}{2}\mathrm{B(1)}\right)} = \left(\frac{3r}{\varepsilon}\right)^\ell. \tag{37}$$

Thus, we conclude that $M \leq (3r/\varepsilon)^\ell$.

Now by our discretization argument, there exists $k \in [M]$ such that $\|\beta - \beta_k\| \leq \varepsilon$ and hence

$$\begin{aligned}\max_{j\in[N]}\langle u_j, J(\beta)\rangle &= \max_{j\in[N]}\langle u_j, J(\beta_k) + (J(\beta) - J(\beta_k)\rangle \\ &= \max_{j\in[N]}\langle u_j, J(\beta_k)\rangle + \langle u_j, J(\beta) - J(\beta_k)\rangle \\ &\leq \max_{j\in[N]}\langle u_j, J(\beta_k)\rangle + \max_{j\in[N]}\langle u_j, J(\beta) - J(\beta_k)\rangle \\ &\leq \max_{k\in[M]}\max_{j\in[N]}\langle u_j, J(\beta_k)\rangle \\ &\quad + \sup_{\|\alpha-\alpha'\|\leq\varepsilon}\max_{j\in[N]}\langle u_j, J(\alpha) - J(\alpha')\rangle.\end{aligned}$$

We observe that if $\alpha, \alpha' \in \mathcal{B}$ are such that $\|\alpha - \alpha'\| \leq \varepsilon$, then applying our smoothness assumption yields

$$\begin{aligned}&\langle u_j, J(\alpha) - J(\alpha')\rangle \\ &= \langle u_j, (\nabla\hat{R}(\alpha) - \nabla R(\alpha)) - (\nabla\hat{R}(\alpha') - \nabla R(\alpha'))\rangle \\ &= \langle u_j, \nabla\hat{R}(\alpha) - \nabla\hat{R}(\alpha')\rangle + \langle u_j, \nabla R(\alpha') - \nabla R(\alpha)\rangle \\ &\leq \|u_j\| \left\|\nabla\hat{R}(\alpha) - \nabla\hat{R}(\alpha')\right\| + \|u_j\| \left\|\nabla R(\alpha) - \nabla R(\alpha')\right\| \\ &\leq L_{\beta_i}\left\|\alpha - \alpha'\right\| + L_{\beta_i}\left\|\alpha - \alpha'\right\| \\ &\leq 2L_\beta\varepsilon,\end{aligned}$$

where the second-to-last inequality uses $\|u_j\| = 1$.

To bound the remaining term, we use the concentration of sub-exponential random variables, due to Bernstein's Inequality combined with the Union Bound. We have that

$$\mathbb{P}\left(\langle u_j, J(\beta_k)\rangle \geq t\right) \leq 2\exp\left(-\frac{mt^2}{2\theta^2}\right)$$

for all $t \leq \theta$, and hence

$$\begin{aligned}&\mathbb{P}\left(\max_{k\in[M]}\max_{j\in[N]}\langle u_j, J(\beta_k\rangle \geq t\right) \\ &= \mathbb{P}\left(\bigcup_{k\in[M]}\bigcup_{j\in[N]}\{\langle u_j, J(\beta_k\rangle \geq t\}\right) \\ &\leq \sum_{k\in[M]}\sum_{j\in[N]}\mathbb{P}\left(\{\langle u_j, J(\beta_k\rangle \geq t\}\right) \\ &\leq \sum_{k\in[M]}\sum_{j\in[N]} 2\exp\left(-\frac{mt^2}{2\theta^2}\right) \\ &= M\cdot N\cdot 2\exp\left(-\frac{mt^2}{2\theta^2}\right) \\ &\leq 2\left(\frac{15r}{\varepsilon}\right)^\ell \exp\left(-\frac{mt^2}{2\theta^2}\right)\end{aligned}$$

for all $t \leq \theta$, where we used the fact that $M \leq (3r/\varepsilon)^\ell$ and $N \leq 5^\ell$. Setting the right hand side equal to 2δ yields

$$t = \sqrt{2}\theta\sqrt{\frac{\ell\log(15r/\varepsilon) + \log(1/\delta)}{m}}. \tag{38}$$

Next we choose $\varepsilon = \frac{1}{15r}\sqrt{\frac{\ell+\log(1/\delta)}{m}}$ so that

$$\begin{aligned}t &= \sqrt{2}\theta\sqrt{\frac{\ell\log(15r/\varepsilon) + \log(1/\delta)}{m}} \\ &= \sqrt{2}\theta\sqrt{\frac{\frac{\ell}{2}\log(m) - \frac{\ell}{2}\log(\ell + \log(1/\delta)) + \log(1/\delta)}{m}} \\ &\leq \sqrt{2}\theta\sqrt{\frac{\ell\log(m) + \log(1/\delta)}{m}} \\ &\leq \sqrt{2}\theta\sqrt{\frac{\log(m)(\ell + \log(1/\delta))}{m}}.\end{aligned}$$

By requiring that m satisfy $m/\log(m) \geq 2(\ell + \log(1/\delta))$, we enforce that $t \leq \theta$. In combining, we observe that

$$\begin{aligned}&t + 2\varepsilon L \\ &\leq \sqrt{2}\theta\sqrt{\frac{\log(m)(\ell + \log(1/\delta))}{m}} + \frac{2L}{15r}\sqrt{\frac{\ell + \log(1/\delta)}{m}} \\ &\leq 2\left(\theta + \frac{L}{15r}\right)\sqrt{\frac{\log(m)(\ell + \log(1/\delta))}{m}}\end{aligned}$$

$$\leq 4\max\left\{\frac{L}{15r},\theta\right\}\sqrt{\frac{\log(m)(\ell+\log(1/\delta))}{m}},$$

and the result follows.

B. PROOF OF THEOREM 3

We suppress the subscript i for notational simplicity. We recall that that the μ-strong convexity of the map $\beta \mapsto R(x,z;\beta)$ implies μ-strong monotonicity of $\nabla R(\beta)$, and $\nabla \hat{R}(\beta)$. It follows that

$$\begin{aligned}\mu\left\|\widehat{\beta}-\beta^*\right\|^2 &\leq \langle\widehat{\beta}-\beta^*, \nabla R(\widehat{\beta})-\nabla R(\beta^*)\rangle\\ &= \langle\widehat{\beta}-\beta^*, \nabla R(\widehat{\beta})\rangle - \langle\widehat{\beta}-\beta^*, \nabla R(\beta^*)\rangle\\ &\leq \langle\widehat{\beta}-\beta^*, \nabla R(\widehat{\beta})\rangle\\ &\leq \langle\widehat{\beta}-\beta^*, \nabla R(\widehat{\beta})\rangle + \langle\beta^*-\widehat{\beta}, \nabla \hat{R}(\widehat{\beta})\rangle\\ &= \langle\widehat{\beta}-\beta^*, \nabla R(\widehat{\beta})-\nabla\hat{R}(\widehat{\beta})\rangle\\ &\leq \left\|\widehat{\beta}-\beta^*\right\|\sup_{\beta\in\mathcal{B}}\left\|\nabla R(\beta)-\nabla\hat{R}(\beta)\right\|\end{aligned}$$

and hence

$$\left\|\widehat{\beta}-\beta^*\right\| \leq \frac{1}{\mu}\sup_{\beta\in\mathcal{B}}\left\|\nabla R(\beta)-\nabla\hat{R}(\beta)\right\|. \quad (39)$$

The result now follows by applying Lemma 2. □

C. PROOF OF PROPOSITION 4

We suppress the i index throughout. The associated risk function is $R(x,z,B)=\frac{1}{2}\|Bx-z\|^2$, so that $\nabla R(x,z,B)=(Bx-z)x^T=Bxx^T-zx^T$ and $\nabla^2 R(x,z,B)=xx^T$ are the corresponding gradient and hessian. We observe that enforcing $\gamma I \leq \mathbb{E}[xx^T]\leq LI$ for some $\gamma, L>0$ ensures γ-strong convexity and L-smoothness of the expected risk. Similarly, the empirical risk has gradient $\nabla R_m(B)=1/m(BXX^T-ZX^T)$, and hessian $\nabla^2 R_m(B)=(1/m)XX^T$. Thus R_m is convex the hessian is symmetric, then it is positive semi-definite and thus R_m is convex. Furthermore, smoothness of R_m follows with constant $\max\{L,\|XX^T\|_2\}$. Lastly, since zx^T and xx^T have sub-exponential entries, the gradient is sub-exponential and the result follows. □

D. PROOF OF THEOREM 5

We observe that for any fixed $x\in\mathcal{X}$, we have that $|F_{\widehat{\beta}_i}(x)-F_i(x)|\leq|F_{\widehat{\beta}_i}(x)-F_{\beta_i^*}(x)|+|F_{\beta_i^*}(x)-F_i(x)|$. The first term describes our statistical error at x. We denote $\Pi(D_{\widehat{\beta}_i},D_{\beta_i^*})$ as a coupling on $\mathcal{P}(\mathbb{R}^{m_i})$ so that

$$\begin{aligned}&|F_{\widehat{\beta}_i}(x)-F_{\beta_i^*}(x)|\\ &=\left|\inf_{\Pi(D_{\widehat{\beta}_i}(x),D_{\beta_i^*}(x))}\mathbb{E}_{(z,z')\sim\Pi(D_{\widehat{\beta}_i}(x),D_{\beta_i^*}(x))}\left(f(x,z)-f(x,z')\right)\right|\\ &\leq\inf_{\Pi(D_{\widehat{\beta}_i}(x),D_{\beta_i^*}(x))}\mathbb{E}_{(z,z')\sim\Pi(D_{\widehat{\beta}_i}(x),D_{\beta_i^*}(x))}\left|f(x,z)-f(x,z')\right|\\ &\leq L_{z_i}\left(\inf_{\Pi(D_{\widehat{\beta}_i}(x),D_{\beta_i^*}(x))}\mathbb{E}_{(z,z')\sim\Pi(D_{\widehat{\beta}_i}(x),D_{\beta_i^*}(x))}\left\|z_i-z_i'\right\|\right)\\ &=L_{z_i}W_1(D_{\widehat{\beta}_i}(x),D_{\beta_i^*}(x))\\ &\leq L_{z_i}\varepsilon_i\left\|\widehat{\beta}_i-\beta_i^*\right\|.\end{aligned}$$

By similar argument, we find that $|F_{\beta_i^*}(x)-F_i(x)|\leq L_{z_i}W_1(D_{\beta_i^*}(x),D_i(x))\leq L_{z_i}\gamma_i$. In combining, we get $|F_{\widehat{\beta}_i}(x)-F_i(x)|\leq L_{z_i}\varepsilon_i\|\widehat{\beta}_i-\beta_i^*\|+L_{z_i}\eta_i$. Lastly, $\|\widehat{\beta}_i-\beta_i^*\|$ can be bounded as in Theorem 3. □

E. PROOF OF COROLLARY 6

Observe that

$$\begin{aligned}&F_i(\widehat{x})-F_i(x^*)\\ &=\left[F_i(\widehat{x})-F_{\beta_i^*}(\widehat{x})\right]+\left[F_{\beta_i^*}(\widehat{x})-F_{\hat{\beta}_i}(\widehat{x})\right]\\ &\quad+\left[F_{\hat{\beta}_i}(\widehat{x})-F_{\hat{\beta}_i}(x^{**})\right]+\left[F_{\hat{\beta}_i}(x^{**})-F_{\beta_i^*}(x^{**})\right]\\ &\quad+\left[F_{\beta_i^*}(x^{**})-F_{\beta_i^*}(x^*)\right]+\left[F_{\beta_i^*}(x^*)-F_i(x^*)\right]\\ &\leq 2\left\|F_i-F_{\beta_i^*}\right\|_\infty+2\left\|F_{\beta_i^*}-F_{\hat{\beta}_i}\right\|_\infty\\ &\quad+\left[F_{\hat{\beta}_i}(\widehat{x})-F_{\hat{\beta}_i}(x^{**})\right]+\left[F_{\beta_i^*}(x^{**})-F_{\beta_i^*}(x^*)\right]\end{aligned}$$

where $x^{**}\in\mathcal{X}$ is the Nash equilibrium satisfying

$$x_i^{**}\in\arg\min_{x_i\in\mathcal{X}_i}F_{\beta_i^*}(x_i,x_{-i}^{**}),\quad i\in[n]. \quad (40)$$

It follows from (25) that

$$\begin{aligned}F_{\hat{\beta}_i}(\widehat{x})-F_{\hat{\beta}_i}(x^{**})&\leq\left[F_{\hat{\beta}_i}(\widehat{x}_i,\widehat{x}_{-i})-F_{\hat{\beta}_i}(x_i^{**},\widehat{x}_{-i})\right]\\ &\quad+\left[F_{\hat{\beta}_i}(x_i^{**},\widehat{x}_{-i})-F_{\hat{\beta}_i}(x_i^{**},x_{-i}^{**})\right]\\ &\leq F_{\hat{\beta}_i}(x_i^{**},\widehat{x}_{-i})-F_{\hat{\beta}_i}(x_i^{**},x_{-i}^{**})\\ &\leq L_{\widehat{\beta}_i}\left\|\widehat{x}_{-i}-x_{-i}^{**}\right\|.\end{aligned}$$

Similarly,

$$\begin{aligned}F_{\beta_i^*}(x^{**})-F_{\beta_i^*}(x^*)&=\left[F_{\beta_i^*}(x_i^{**},x_{-i}^{**})-F_{\beta_i^*}(x_i^{**},x_{-i}^*)\right]\\ &\quad+\left[F_{\beta_i^*}(x_i^{**},x_{-i}^*)-F_{\beta_i^*}(x_i^*,x_{-i}^*)\right]\\ &\leq F_{\beta_i^*}(x_i^{**},x_{-i}^*)-F_{\beta_i^*}(x_i^*,x_{-i}^*)\\ &\leq L_{\beta^*}\left\|x_{-i}^{**}-x_{-i}^*\right\|.\end{aligned}$$

Combining the bounds yields

$$\begin{aligned}&F_i(\widehat{x})-F_i(x^*)\\ &\leq 2\left\|F_i-F_{\beta_i^*}\right\|_\infty+2\left\|F_{\beta_i^*}-F_{\hat{\beta}_i}\right\|_\infty\end{aligned}$$

$$
\begin{aligned}
&+ L_{\widehat{\beta}} \left\| \widehat{x}_{-i} - x^{**}_{-i} \right\| + L_{\beta^*} \left\| x^*_{-i} - x^{**}_{-i} \right\| \\
&\leq 2\gamma_i L_{z_i} + 2\varepsilon L_{z_i} \left\| \widehat{\beta}_i - \beta_i^* \right\| \\
&+ L_{\widehat{\beta}} \mathsf{diam}(\mathcal{X}_{-i}) + L_{\beta^*} \mathsf{diam}(\mathcal{X}_{-i}) \\
&\leq 2\gamma_i L_{z_i} + 2\varepsilon L_{z_i} \left\| \widehat{\beta}_i - \beta_i^* \right\| + 2 \max\{L_{\widehat{\beta}_i}, L_{\beta_i^*}\} \mathsf{diam}(\mathcal{X}_{-i})
\end{aligned}
$$

Then, (26) follows using the bound on $\|\widehat{\beta}_i - \beta_i^*\|$ from Theorem 3. □

F. PROOF OF LEMMA 8

Consider the function $\varphi : \mathbb{R}^d \to \mathbb{R}$ defined by $\varphi(y) = \frac{1}{2}\|x^t - W_i^{-1} g^t - y\|_W^2$ for all $y \in \mathcal{X}$. Then, φ is ω_n-strongly convex over $\mathcal{X}$ and has a unique minimizer $x^{t+1} \in \mathcal{X}$. This implies that:

$$
\begin{aligned}
\varphi(x^*) \geq \varphi(x^{t+1}) &+ \langle x^* - x^{t+1}, \\
\nabla\varphi(x^{t+1})\rangle &+ \frac{\omega_n}{2} \left\| x^{t+1} - x^* \right\|^2 .
\end{aligned}
$$

Since $\langle x - x^{t+1}, \nabla\varphi(x^{t+1})\rangle \geq 0$ for all $x \in \mathcal{X}$, we obtain

$$
\begin{aligned}
&\omega_n \left\| x_i^{t+1} - x_i^* \right\| \\
&\leq \left\| x_i^t - \eta_i g_i^t - x_i^* \right\|_W^2 - \left\| x_i^t - \eta_i g_i^t - x_i^{k+1} \right\|_W^2 .
\end{aligned}
$$

It follows that

$$
\begin{aligned}
\frac{\omega_n}{\omega_1} \left\| x^{t+1} - \widehat{x} \right\|_W^2 \leq &\left\| x^t - \widehat{x} \right\|_W^2 - \left\| x^t - x_i^{t+1} \right\|_W^2 \\
&- 2\langle x^t - \widehat{x}, g^t \rangle + 2\eta_i \langle x^t - x^{t+1}, g^t \rangle.
\end{aligned}
$$

We now consider the above in the conditional expectation $\mathbb{E}_t \cdot := \mathbb{E}_{z_i \sim D(x_t)}[\,\cdot\,|\mathcal{F}_t]$ with $\mathcal{F}_t = \sigma(g^t, \tau \geq t)$. We find that

$$
\begin{aligned}
&\frac{\omega_n}{\omega_1} \mathbb{E}_t \left\| x^{t+1} - \widehat{x} \right\|_W^2 \\
&\leq \mathbb{E}_t \left\| x^t - \widehat{x} \right\|_W^2 - \mathbb{E}_t \left\| x^t - x^{t+1} \right\|_W^2 \\
&- 2\mathbb{E}_t \langle x^t - \widehat{x}, g^t \rangle - 2\mathbb{E}_t \langle x^{t+1} - x^t, g^t \rangle \\
&= \left\| x^t - \widehat{x} \right\|_W^2 - \mathbb{E}_t \left\| x^t - x^{t+1} \right\|_W^2 \\
&- 2\langle x^t - \widehat{x}, \mu^t \rangle - 2\mathbb{E}_t \langle x^{t+1} - x^t, g^t \rangle \\
&= \left\| x^t - \widehat{x} \right\|_W^2 - \mathbb{E}_t \left\| x^t - x^{t+1} \right\|_W^2 \\
&+ 2\mathbb{E}_t \langle x^t - x^{t+1}, g^t - \mu^t \rangle + 2\mathbb{E}_t \langle \widehat{x} - x^{t+1}, \mu^t \rangle \\
&= \left\| x^t - \widehat{x} \right\|_W^2 - \mathbb{E}_t \left\| x^t - x^{t+1} \right\|_W^2 - 2\langle x^{t+1} - \widehat{x}, G(x^{t+1}) \rangle \\
&+ 2\mathbb{E}_t \langle \widehat{x} - x^{t+1}, \mu^t - G(x^{t+1}) \rangle \\
&+ 2\mathbb{E}_t \langle x^t - x^{t+1}, g^t - \mu^t \rangle.
\end{aligned}
$$

To proceed, we bound the inner product terms. Using strong monotonicity, we have that

$$
\begin{aligned}
\mathbb{E}_t \langle \widehat{x} - x^{t+1}, G(x^{t+1}) \rangle &\geq \alpha \mathbb{E}_t \left\| x^{t+1} - \widehat{x} \right\|^2 \\
&\geq \frac{\alpha}{\omega_1} \mathbb{E}_t \left\| x^{t+1} - \widehat{x} \right\|_W^2 .
\end{aligned}
$$

Furthermore, we observe that

$$
\begin{aligned}
&\mathbb{E}_t \langle \widehat{x} - x_i^{t+1}, \mu^t - G(x^{t+1}) \rangle \\
&\quad = \mathbb{E}_t \langle \widehat{x} - x^{t+1}, \mu^t - G(x^t) \rangle \\
&\quad + \mathbb{E}_t \langle \widehat{x} - x^{t+1}, G(x^t) - G(x^{t+1}) \rangle.
\end{aligned}
$$

To bound the remaining terms, we use arguments based on a weighted Young's inequality. Let $\Delta_1, \Delta_2, \Delta_3 > 0$ be fixed constants. It follows that

$$
\begin{aligned}
&2\mathbb{E}_t \langle x^t - x^{t+1}, g^t - \mu^t \rangle \\
&\leq \Delta_1 \mathbb{E}_t \left\| x^{t+1} - x^t \right\|^2 + \frac{1}{\Delta_1} \mathbb{E}_t \left\| g^t - \mu^t \right\|^2 \\
&\leq \frac{\Delta_1}{\omega_n} \mathbb{E}_t \left\| x^{t+1} - x^t \right\|_W^2 + \frac{1}{\Delta_1} \sum_{i=1}^n \mathbb{E}_t \left\| g^t - \mu^t \right\|^2 \\
&\leq \frac{\Delta_1}{\omega_n} \mathbb{E}_t \left\| x^{t+1} - x^t \right\|_W^2 + \frac{1}{\Delta_1} \sum_{i=1}^n \sigma_i^2 \\
&\leq \frac{\Delta_1}{\omega_n} \mathbb{E}_t \left\| x^{t+1} - x^t \right\|_W^2 + \frac{\sigma^2}{\Delta_1},
\end{aligned}
$$

and

$$
\begin{aligned}
&2\mathbb{E}_t \langle \widehat{x} - x^{t+1}, \mu^t - G(x^t) \rangle \\
&\leq \Delta_2 \mathbb{E}_t \left\| x^{t+1} - \widehat{x} \right\|^2 + \frac{1}{\Delta_2} \mathbb{E}_t \left\| \mu^t - G(x^t) \right\|^2 \\
&\leq \frac{\Delta_2}{\omega_n} \mathbb{E}_t \left\| x^{t+1} - \widehat{x} \right\|_W^2 + \frac{1}{\Delta_2} \sum_{i=1}^n \mathbb{E}_t \left\| \mu^t - G(x^t) \right\|^2 \\
&\leq \frac{\Delta_2}{\omega_n} \mathbb{E}_t \left\| x^{t+1} - \widehat{x} \right\|_W^2 + \frac{1}{\Delta_2} \sum_{i=1}^n \rho_i^2 \\
&\leq \frac{\Delta_2}{\omega_n} \mathbb{E}_t \left\| x^{t+1} - \widehat{x} \right\|_W^2 + \frac{\rho^2}{\Delta_2}.
\end{aligned}
$$

Additionally, we have that

$$
\begin{aligned}
&2\mathbb{E}_t \langle \widehat{x} - x^{t+1}, G(x^t) - G(x^{t+1}) \rangle \\
&\leq \Delta_3 \mathbb{E}_t \left\| x^{t+1} - \widehat{x} \right\|^2 + \frac{1}{\Delta_3} \mathbb{E}_t \left\| G(x^t) - G(x^{t+1}) \right\|^2 \\
&\leq \frac{\Delta_3}{\omega_n} \mathbb{E}_t \left\| x^{t+1} - \widehat{x} \right\|_W^2 + \frac{L^2}{\Delta_3} \mathbb{E}_t \left\| x^{t+1} - x^t \right\|^2 \\
&\leq \frac{\Delta_3}{\omega_n} \mathbb{E}_t \left\| x^{t+1} - \widehat{x} \right\|_W^2 + \frac{L^2}{\omega_n \Delta_3} \mathbb{E}_t \left\| x^{t+1} - x^t \right\|_W^2 .
\end{aligned}
$$

Combining these estimates yields

$$
\begin{aligned}
&\frac{\omega_n}{\omega_1} \mathbb{E}_t \left\| x^{t+1} - \widehat{x} \right\|_W^2 \\
&\leq \left\| x^t - \widehat{x} \right\|_W^2 - \mathbb{E}_t \left\| x^{t+1} - x^t \right\|_W^2 - \frac{2\alpha}{\omega_1} \mathbb{E}_t \left\| x^{t+1} - \widehat{x} \right\|_W^2 \\
&+ \left(\frac{\Delta_1}{\omega_n} \mathbb{E}_t \left\| x^{t+1} - x^t \right\|_W^2 + \frac{\sigma^2}{\Delta_1} \right)
\end{aligned}
$$

$$+\left(\frac{\Delta_2}{\omega_n}\mathbb{E}_t\left\|x^{t+1}-\widehat{x}\right\|_W^2+\frac{\rho^2}{\Delta_2}\right)$$
$$+\left(\frac{\Delta_3}{\omega_n}\mathbb{E}_t\left\|x^{t+1}-\widehat{x}\right\|_W^2+\frac{L^2}{\omega_n\Delta_3}\mathbb{E}_t\left\|x^{t+1}-x^t\right\|_W^2\right)$$
$$=\left\|x^t-\widehat{x}\right\|_W^2+\left(\frac{\Delta_1}{\omega_n}+\frac{L^2}{\omega_n\Delta_3}-1\right)\mathbb{E}_t\left\|x^{t+1}-x^t\right\|_W^2$$
$$+\left(\frac{\Delta_2}{\omega_n}+\frac{\Delta_3}{\omega_n}-\frac{2\alpha}{\omega_1}\right)\mathbb{E}_t\left\|x^{t+1}-\widehat{x}\right\|_W^2$$
$$+\left(\frac{\sigma^2}{\Delta_1}+\frac{\rho^2}{\Delta_2}\right)$$

and simplifying gives

$$\left(\frac{\omega_n}{\omega_1}+\frac{2\alpha}{\omega_1}-\frac{\Delta_2}{\omega_n}-\frac{\Delta_3}{\omega_n}\right)\mathbb{E}_t\left\|x^{t+1}-\widehat{x}\right\|_W^2$$
$$\leq\left\|x^t-\widehat{x}\right\|_W^2+\left(\frac{\sigma^2}{\Delta_1}+\frac{\rho^2}{\Delta_2}\right)$$
$$+\left(\frac{\Delta_1}{\omega_n}+\frac{L^2}{\omega_n\Delta_3}-1\right)\mathbb{E}_t\left\|x^{t+1}-x^t\right\|_W^2.$$

To proceed, we choose $\Delta_2=\Delta_3=\frac{\alpha\omega_n}{2\omega_1}$ and $\Delta_1=\omega_n-2\omega_1L^2/(\alpha\omega_n)$ to ensure that the coefficient on the $\mathbb{E}_t\|x^{t+1}-x^t\|_W^2$ term is zero. Furthermore, enforcing that $\frac{\omega_1}{\omega_n^2}\leq\frac{\alpha}{4L^2}$ guarantees that $\Delta_1^{-1}\leq 2\omega_n^{-1}$. Hence the variance term is finite. Substituting these values and simplifying yields the result. □

G. PROOF OF THEOREM 9

For notational convenience, we will use the short-hand notation $e^t:=\|x^t-\widehat{x}\|_W^2$, $c=\omega_1/(\alpha+\omega_n)$, and

$$A=2\frac{\alpha\sigma^2+\omega_1\rho^2}{\alpha\omega_n}.$$

Hence, the result in Lemma 8 can be written compactly as

$$\mathbb{E}_{t-1}e^t\leq ce^{t-1}+cA.$$

By recursively applying this result and applying the law of total expectation, we find that

$$\mathbb{E}e^t\leq c^te^0+cA\sum_{j=1}^{t-1}c^j\leq c^te^0+cA\frac{1-c^t}{1-c}.$$

Furthermore, if $(\omega_1-\omega_n)<\alpha$, then $c<1$ and the geometric series converges and is equal to its limit $1/(1-c)$. Hence $\mathbb{E}e^t\leq c^te^0+A\frac{c}{1-c}$. □

H. PROOF OF THEOREM 10

Fix $t\geq 0$. For notational convenience, we will denote $e^t=\|x^t-\widehat{x}\|^2$. Replacing the step-size matrix in Lemma 8 with $W=\omega^tI_{d\times d}$ yields

$$\mathbb{E}_te^{t+1}\leq\frac{\omega^t}{\omega^t+\alpha}e^t+\frac{2\sigma^2}{\omega^t(\omega^t+\alpha)}+\frac{2(\rho^t)^2}{\alpha(\omega^t+\alpha)}. \tag{41}$$

To proceed, we will use the observation that

$$\frac{1}{(s+t)(r+t)}=\frac{r+t}{(s+t)(r+t)^2}\leq\frac{\max\{\frac{r}{s},1\}}{(r+t)^2} \tag{42}$$

and

$$\frac{1}{(r+t)(r+t-2)}\leq\frac{\frac{r}{r-2}}{(r+t)^2}. \tag{43}$$

By substituting our expression for ω^t, ρ^t, and e^t into (41) we obtain

$$\mathbb{E}_te^{t+1}\leq\frac{r+t-2}{\alpha^2(r+t)^2}A+\frac{8\sigma^2}{\alpha^2(r+t-2)(r+t)}$$
$$+\frac{4\bar{\rho}}{\alpha^2(s+t)(r+t)}$$
$$\leq\frac{r+t-2}{\alpha^2(r+t)^2}A+\frac{8\sigma^2\left(\frac{r}{r-2}\right)}{\alpha^2(r+t)^2}+\frac{4\bar{\rho}\max\left\{\frac{r}{s},1\right\}}{\alpha^2(r+t)^2}$$
$$=\frac{r+t-1}{\alpha^2(r+t)^2}A$$
$$+\frac{-A+8\sigma^2\left(\frac{r}{r-2}\right)+4\bar{\rho}\max\left\{\frac{r}{s},1\right\}}{\alpha^2(r+t)^2}$$
$$\leq\frac{r+t-1}{\alpha^2(r+t)^2}A$$
$$\leq\frac{A}{\alpha^2(r+t+1)}.$$

Here, the last steps follow from construction of A, and the fact that $(r+t+1)(r+t-1)\leq(r+t)^2$. □

ACKNOWLEDGMENT

The MSGI program is administered by the Oak Ridge Institute for Science and Education (ORISE) through an inter-agency agreement between the U.S. Department of Energy (DOE) and NSF. ORISE is managed for DOE by ORAU. All opinions expressed in this paper are the author's and do not necessarily reflect the policies and views of NSF, ORAU/ORISE, or DOE.

REFERENCES

[1] A. Nemirovski, A. Juditsky, G. Lan, and A. Shapiro, "Robust stochastic approximation approach to stochastic programming," *SIAM J. Optim.*, vol. 19, no. 4, pp. 1574–1609, 2009.
[2] J. Lei and U. V. Shanbhag, "Stochastic Nash equilibrium problems: Models, analysis, and algorithms," *IEEE Control Syst. Mag.*, vol. 42, no. 4, pp. 103–124, Aug. 2022.
[3] J. Koshal, A. Nedić, and U. V. Shanbhag, "Single timescale regularized stochastic approximation schemes for monotone Nash games under uncertainty," in *Proc. IEEE Conf. Decis. Control*, 2010, pp. 231–236.
[4] S.-J. Liu and M. Krstić, "Stochastic Nash equilibrium seeking for games with general nonlinear payoffs," *SIAM J. Control Optim.*, vol. 49, no. 4, pp. 1659–1679, 2011.
[5] J. Lei, U. V. Shanbhag, J.-S. Pang, and S. Sen, "On synchronous, asynchronous, and randomized best-response schemes for stochastic Nash games," *Math. Operations Res.*, vol. 45, no. 1, pp. 157–190, 2020.
[6] B. Franci and S. Grammatico, "Stochastic generalized Nash equilibrium-seeking in merely monotone games," *IEEE Trans. Autom. Control*, vol. 67, no. 8, pp. 3905–3919, Aug. 2022.

[7] F. Fabiani and B. Franci, “A stochastic generalized Nash equilibrium model for platforms competition in the ride-hail market,” in *Proc. IEEE Conf. Decis. Control*, 2022, pp. 4455–4460.

[8] B. G. Bakhshayesh and H. Kebriaei, “Decentralized equilibrium seeking of joint routing and destination planning of electric vehicles: A constrained aggregative game approach,” *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 8, pp. 13265–13274, Aug. 2022.

[9] F. Fele and K. Margellos, “Scenario-based robust scheduling for electric vehicle charging games,” in *Proc. IEEE Int. Conf. Environ. Elect. Eng. IEEE Ind. Commercial Power Syst. Europe*, 2019, pp. 1–6.

[10] D. Paccagnan, B. Gentile, F. Parise, M. Kamgarpour, and J. Lygeros, “Nash and wardrop equilibria in aggregative games with coupling constraints,” *IEEE Trans. Autom. Control*, vol. 64, no. 4, pp. 1373–1388, Apr. 2019.

[11] A. Kannan, U. V. Shanbhag, and H. M. Kim, “Strategic behavior in power markets under uncertainty,” *Energy Syst.*, vol. 2, no. 2, pp. 115–141, 2011.

[12] X. Zhou, E. Dall’Anese, and L. Chen, “Online stochastic optimization of networked distributed energy resources,” *IEEE Trans. Autom. Control*, vol. 65, no. 6, pp. 2387–2401, Jun. 2020.

[13] B. Franci and S. Grammatico, “Training generative adversarial networks via stochastic Nash games,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 3, pp. 1319–1328, Mar. 2023.

[14] W. Tushar, W. Saad, H. V. Poor, and D. B. Smith, “Economics of electric vehicle charging: A game theoretic approach,” *IEEE Trans. Smart Grid*, vol. 3, no. 4, pp. 1767–1778, Dec. 2012.

[15] J. L. Mathieu, D. S. Callaway, and S. Kiliccote, “Examining uncertainty in demand response baseline models and variability in automated responses to dynamic pricing,” in *Proc. 50th IEEE Conf. Decis. Control Eur. Control Conf.*, 2011, pp. 4332–4339.

[16] J. Perdomo, T. Zrnic, C. Mendler-Dünner, and M. Hardt, “Performative prediction,” in *Proc. Int. Conf. Mach. Learn.*, 2020, pp. 7599–7609.

[17] D. Drusvyatskiy and L. Xiao, “Stochastic optimization with decision-dependent distributions,” *Math. Operations Res.*, vol. 48, pp. 954–998, 2022.

[18] J. P. Miller, J. C. Perdomo, and T. Zrnic, “Outside the echo chamber: Optimizing the performative risk,” in *Proc. Int. Conf. Mach. Learn.*, 2021, pp. 7710–7720.

[19] A. Narang, E. Faulkner, D. Drusvyatskiy, M. Fazel, and L. Ratliff, “Learning in stochastic monotone games with decision-dependent data,” in *Proc. Int. Conf. Artif. Intell. Statist.*, 2022, pp. 5891–5912.

[20] K. Wood and E. Dall’Anese, “Stochastic saddle point problems with decision-dependent distributions,” *SIAM J. Optim.*, vol. 33, no. 3, pp. 1943–1967, 2023.

[21] C. P. Mediwaththe and D. B. Smith, “Game-theoretic electric vehicle charging management resilient to non-ideal user behavior,” *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 11, pp. 3486–3495, Nov. 2018.

[22] C. Latinopoulos, A. Sivakumar, and J. Polak, “Response of electric vehicle drivers to dynamic pricing of parking and charging services: Risky choice in early reservations,” *Transp. Res. Part C: Emerg. Technol.*, vol. 80, pp. 175–189, 2017.

[23] N. Daina, A. Sivakumar, and J. W. Polak, “Electric vehicle charging choices: Modelling and implications for smart charging services,” *Transp. Res. Part C: Emerg. Technol.*, vol. 81, pp. 36–56, 2017.

[24] Z. Sun, A. C. Hupman, H. I. Ritchey, and A. E. Abbas, “Bayesian updating of the price elasticity of uncertain demand,” *IEEE Syst. J.*, vol. 10, no. 1, pp. 136–146, Mar. 2016.

[25] A. V. D. Boer, “Dynamic pricing and learning: Historical origins, current research, and new directions,” *Surv. Operations Res. Manage. Sci.*, vol. 20, no. 1, pp. 1–18, 2015.

[26] W. C. Cheung, D. Simchi-Levi, and H. Wang, “Dynamic pricing and demand learning with limited price experimentation,” *Operations Res.*, vol. 65, no. 6, pp. 1722–1731, 2017.

[27] B. Chen, X. Chao, and C. Shi, “Nonparametric learning algorithms for joint pricing and inventory control with lost sales and censored demand,” *Math. Operations Res.*, vol. 46, no. 2, pp. 726–756, 2021.

[28] L. Lin and T. Zrnic, “Plug-in performative optimization,” 2023, *arXiv:2305.18728v1*.

[29] G. Scutari, D. P. Palomar, F. Facchinei, and J.-S. Pang, “Convex optimization, game theory, and variational inequality theory,” *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 35–49, May 2010.

[30] J. B. Rosen, “Existence and uniqueness of equilibrium points for concave n-person games,” *Econometrica: J. Econometric Soc.*, vol. 33, pp. 520–534, 1965.

[31] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*. Berlin, Germany: Springer, 2003.

[32] P. T. Harker and J.-S. Pang, “Finite-dimensional variational inequality and nonlinear complementarity problems: A survey of theory, algorithms and applications,” *Math. Program.*, vol. 48, no. 1–3, pp. 161–220, 1990.

[33] A. Nedic, “Distributed gradient methods for convex machine learning problems in networks: Distributed optimization,” *IEEE Signal Process. Mag.*, vol. 37, no. 3, pp. 92–101, May 2020.

[34] D. Kovalev, A. Beznosikov, A. Sadiev, M. Persiianov, P. Richtárik, and A. Gasnikov, “Optimal algorithms for decentralized stochastic variational inequalities,” in *Proc. Annu. Conf. Neural Inf. Process. Syst.*, 2022, pp. 31073–31088.

[35] C. Mendler-Dünner, J. Perdomo, T. Zrnic, and M. Hardt, “Stochastic optimization for performative prediction,” in *Proc. Annu. Conf. Neural Inf. Process. Syst.*, 2020, pp. 4929–4939.

[36] J. Cutler, D. Drusvyatskiy, and Z. Harchaoui, “Stochastic optimization under time drift: Iterate averaging, step decay, and high probability guarantees,” 2021, *arXiv:2108.07356*.

[37] K. Wood, G. Bianchin, and E. Dall’Anese, “Online projected gradient descent for stochastic optimization with decision-dependent distributions,” *IEEE Control Syst. Lett.*, vol. 6, pp. 1646–1651, 2022.

[38] B. Basciftci, S. Ahmed, and S. Shen, “Distributionally robust facility location problem under decision-dependent stochastic demand,” *Eur. J. Oper. Res.*, vol. 292, no. 2, pp. 548–561, 2021.

[39] F. Luo and S. Mehrotra, “Distributionally robust optimization with decision dependent ambiguity sets,” *Optim. Lett.*, vol. 14, no. 8, pp. 2565–2594, 2020.

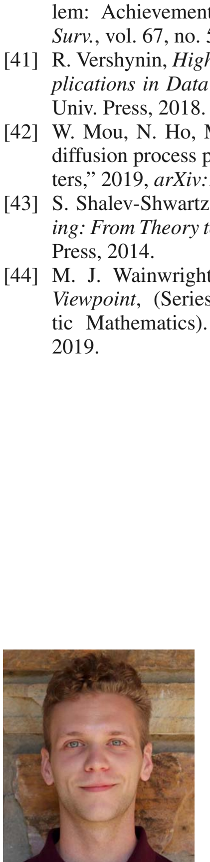
[40] V. I. Bogachev and A. V. Kolesnikov, “The Monge-Kantorovich problem: Achievements, connections, and perspectives,” *Russian Math. Surv.*, vol. 67, no. 5, 2012, Art. no. 785.

[41] R. Vershynin, *High-Dimensional Probability: An Introduction With Applications in Data Science*, vol. 47. New York, NY, USA: Cambridge Univ. Press, 2018.

[42] W. Mou, N. Ho, M. J. Wainwright, P. Bartlett, and M. I. Jordan, “A diffusion process perspective on posterior contraction rates for parameters,” 2019, *arXiv:1909.00966*.

[43] S. Shalev-Shwartz and S. Ben-David, *Understanding Machine Learning: From Theory to Algorithms*. New York, NY, USA: Cambridge Univ. Press, 2014.

[44] M. J. Wainwright, *High-Dimensional Statistics: A Non-Asymptotic Viewpoint*, (Series Cambridge Series in Statistical and Probabilistic Mathematics). New York, NY, USA: Cambridge Univ. Press, 2019.



KILLIAN WOOD received the B.A. degree in mathematics from California State University Fullerton, Fullerton, CA, USA, in 2019, and the M.S. degree in applied mathematics in 2022 from the University of Colorado Boulder, Boulder, CO, USA, where he is currently working toward the Ph.D. degree in applied mathematics, advised by Prof. Emiliano Dall’Anese. His research interests include the development and analysis of stochastic optimization for decision-dependent data distributions, as well as randomized derivative-free methods in scientific computing.

AHMED S. ZAMZAM (Member, IEEE) received the B.Sc. degree from Cairo University, Al Giza, Egypt, in 2013, and the Ph.D. degree in electrical engineering from the University of Minnesota, Minneapolis, MN, USA, in 2019. He is currently a Senior Research Scientist with the National Renewable Energy Laboratory, where he is part of the Power Systems Engineering Center. His research interests include machine learning and optimization for smart grid applications, large-scale complex energy systems optimization, and grid data analytics. He received the Louis John Schnell Fellowship in 2015 and the Doctoral Dissertation Fellowship in 2018 from the University of Minnesota.

EMILIANO DALL'ANESE (Senior Member, IEEE) received the Ph.D. degree in information engineering from the Department of Information Engineering, University of Padova, Padua, Italy, in 2011. He was with the University of Minnesota, Minneapolis, MN, USA, as a postdoc from 2011 to 2014, and the National Renewable Energy Laboratory as a Senior Researcher from 2014 to 2018. He is currently an Associate Professor with the Department of Electrical, Computer, and Energy Engineering, the University of Colorado Boulder, Boulder, CO, USA, where he is also an Affiliate Faculty with the Department of Applied Mathematics. His research interests include optimization, control, and learning; current applications include power systems and autonomous systems. He was the recipient of the National Science Foundation CAREER Award in 2020, the IEEE PES Prize Paper Award in 2021, and the IEEE TRANSACTIONS ON CONTROL OF NETWORK SYSTEMS Best Paper Award in 2023.