Stackelberg Population Dynamics: A Predictive-Sensitivity Approach

Eduardo Mojica-Nava [1,2,*] and **Fredy Ruiz** [2]

1   Department of Electrical and Electronics Engineering, Universidad Nacional de Colombia,
    111321 Bogota, Colombia
2   Dipartimento di Elettronica, Informazione e Bioingegneria—DEIB, Politecnico di Milano, Piazza Leonardo da
    Vinci 32, 20133 Milano, Italy; fredy.ruiz@polimi.it
*   Correspondence: eamojican@unal.edu.co

Abstract: Hierarchical decision-making processes traditionally modeled as bilevel optimization problems are widespread in modern engineering and social systems. In this work, we deal with a leader with a population of followers in a hierarchical order of play. In general, this problem can be modeled as a leader–follower Stackelberg equilibrium problem using a mathematical program with equilibrium constraints. We propose two interconnected dynamical systems to dynamically solve a bilevel optimization problem between a leader and follower population in a single time scale by a predictive-sensitivity conditioning interconnection. For the leader's optimization problem, we developed a gradient descent algorithm based on the total derivative, and for the followers' optimization problem, we used the population dynamics framework to model a population of interacting strategic agents. We extended the concept of the Stackelberg population equilibrium to the differential Stackelberg population equilibrium for population dynamics. Theoretical guarantees for the stability of the proposed Stackelberg population learning dynamics are presented. Finally, a distributed energy resource coordination problem is solved via pricing dynamics based on the proposed approach. Some simulation experiments are presented to illustrate the effectiveness of the framework.

Keywords: bilevel optimization; Stackelberg games; population dynamics

1. Introduction

Hierarchical decision-making processes, where a leader is making decisions affected by the response of the follower(s), are widespread in modern engineering and social systems. Hierarchical decision-making structures or leader–follower problems appear in diverse domains of applications such as terrorism analysis [1], communications and network security [2], traffic management [3], smart grids [4–6], machine learning [7], and market systems [8]. A particular promising application domain is the intersection between machine learning and game theory to model hierarchical interactions between learning agents. Classic simultaneous play games have been extended to problems in machine learning such as robust supervised learning [9] or generative adversarial network [10] and, most recently, in multiagent system (MAS) reinforcement learning [11]. For hierarchical machine-learning approaches based on game theory, just a few works have been proposed. A couple of relevant contributions can be found in [12], where a gradient descent with time scale separation was proposed, and in [7], where Stackelberg learning dynamics was developed; both works focused on a leader with a single follower.

Traditionally, hierarchical decision-making processes have been modeled as bilevel optimization problems, which can be classified into two main theoretical domains [13]. On the one hand are models based on game theory [14], which have used bilevel programming methods to develop the concept of Stackelberg equilibria. On the other hand is mathematical programming, which proposed a first attempt to solve bilevel optimization

problems as an upper-level optimization problem with a lower-level optimization problem as a constraint [15]. Considering its nested nature, bilevel optimization problems have challenged the optimization and mathematical community ever since; they have been proven to be strongly NP-hard [16], and even the evaluation of the optimality of a solution is NP-hard [17,18]. Recent critical literature reviews recount the historical developments and current perspectives of the field [13,19,20].

In this work, we deal with a leader with a population of followers in a hierarchical order of play. In general, this problem can be modeled as a leader–follower Stackelberg equilibrium problem using a mathematical program with equilibrium constraints (MPEC) [21]. Some applications of this modeling have been recently proposed [22–25]. A preliminary MPEC model for a multiple leaders with multiple followers game was presented in [22]. A discrete-time iterative algorithm was proposed to compute a Stackelberg–Nash saddle-point in a leader with multiple followers problem. In [24], a Stackelberg differential game based on pricing for the bandwidth usage of the Internet was solved. Finally, a semi-decentralized algorithm was formulated to compute a local solution to the leader with multiple followers Stackelberg aggregate game in [25]. On the other hand, for power system applications, there is an emerging market-based control approach called transactive energy systems (TESs) [26–28]. This method is based on a privacy-preserving, market-based framework for the management of devices exchanging energy [6,29]. Hierarchical decision-making modeling has been studied as a cornerstone in this framework, where a coordinator agent is responsible for solving the market-clearing price of the distribution network. In several scenarios, the control signal of the leader agent or coordinator is not a price, but a pricing function depending on the preferences of the follower agents forming a reverse Stackelberg game [30,31]. Considering the multiple applications based on hierarchical decision-making structures, we propose a Stackelberg game learning framework to solve bilevel optimization problems as a dynamical system in a single time scale. In this paper, we focus on a single leader with a follower population of agents due to the motivation of the coordinator agent problem in TESs. For the upper-level optimization problem, we propose gradient descent dynamics in continuous-time [32,33], and for the lower-level optimization problem, we propose population games dynamics based on the replicator dynamics [34,35], in order to obtain two nested dynamical systems. This allows us to bring theoretical tools from the dynamical system theory for the design of algorithms to solve optimization problems. In contrast with traditional time scale separation methods for interconnected dynamical systems such as singular perturbation analysis [36], we used predictive-sensitivity conditioning [37,38] to integrate the two interconnected dynamical systems in a single time scale to guarantee the stability and convergence of the solution to a differential Stackelberg population equilibrium.

The main contribution of this paper is threefold. First, we propose two interconnected dynamical systems to dynamically solve a bilevel optimization problem between a leader and follower population in a single time scale by a predictive-sensitivity conditioning interconnection [37]. For the leader optimization problem, we developed a gradient descent algorithm based on the total derivative using the implicit function theorem [39]. For the follower MAS optimization problem, we used the population dynamics framework to model MAS interacting strategic agents, as shown in [34,35,40–42] and the reference therein. We extended the concept of the Stackelberg population equilibrium [43] to the differential Stackelberg population equilibrium for population dynamics. Second, theoretical guarantees for the stability of the proposed Stackelberg population learning dynamics are presented. Finally, a distributed energy resource (DER) coordination problem was solved via pricing dynamics based on the proposed approach. Some simulation experiments are presented to illustrate the effectiveness of the framework.

The rest of the paper is organized as follows. In Section 2, the formulation of the bilevel optimization problem and the main concepts of Stackelberg equilibria are introduced. Section 3 presents the population game's essential concepts, and the Stackelberg population learning dynamics is developed based on the predictive-sensitivity condition-

ing interconnection. Stability results are established to guarantee the convergence of the proposed dynamics. In Section 4, an application of the proposed framework is developed for DER coordination. Finally, the main conclusions are drawn in Section 5.

2. Stackelberg Games and Bilevel Optimization Problems

In this section, we introduce a general formulation of a noncooperative game between an agent called a leader and a finite set of agents called followers. In this case, we have a Stackelberg game with multiple followers, where the leader plays first and then the followers make their decision. This Stackelberg game can be modeled by a bilevel optimization problem. Traditionally, the leader's optimization problem is referred to as the upper-level problem, and the followers' optimization problem is referred as the lower-level problem. Each level has its own set of variables, constraints, and objective functions, and the upper-level problem has as a constraint the lower-level problem:

$$
\begin{aligned}
\min_{\lambda \in \Lambda} \quad & F_1(\lambda, x^*) \\
\text{s.t.} \quad & x^* \in \arg\min_{x \in X}\{F_2(\lambda, x)\},
\end{aligned}
\tag{1}
$$

where $\lambda \in \Lambda \subseteq \mathbb{R}$ is the upper-level variable, $x \in X \subseteq \mathbb{R}^n$ are the lower-level variables, $F_1(\lambda, x)$ is the objective function of the leader agent, and $F_2(\lambda, x)$ is the objective function of the followers. It was assumed that each follower has an objective function $J_i(\lambda, x)$ with $i = \{1, 2, \ldots, n\}$, and then, without loss of generality, we assumed that the global objective function is:

$$
F_2(\lambda, x) = \sum_{i=1}^{n} J_i(\lambda, x).
\tag{2}
$$

Let us recall the main equilibrium concept studied in hierarchical play games: consider one leader agent and one follower agent x_f.

Definition 1 (Local Stackelberg equilibrium (SE)). *Let Λ be the set of strategies for the leader agent and X_f be the set of strategies for the follower agent. Let $BR_f(\lambda)$ be the best response function for the follower defined as:*

$$
BR_f(\lambda) = \{y \in X_f | F_2(\lambda, y) \le F_2(\lambda, x_f), \quad \forall x_f \in X_f\}.
$$

A leader strategy $\lambda^* \in \Lambda$ and a follower strategy x_f^* are in Stackelberg equilibrium if the following conditions hold:

(a) $\max_{x_f \in BR_f(\lambda^*)} F_1(\lambda^*, x_f) \le \max_{x_f \in BR_f(\lambda)} F_1(\lambda, x_f)$;

(b) $x_f^* \in BR_f(\lambda^*)$.

The local notions of the equilibrium concepts are preferred since it is standard in population games where the concavity or convexity is not assumed in the objective functions.

Some assumptions have to be considered to assure the optimality and existence of the solutions to Problem (1) such as the connectivity of the communication graph and the regularity of objective functions.

Assumption 1. *The communication graph connecting each agent of the system such as the leader and follower population agents is connected.*

Assumption 2. *Every objective function $F_1(\lambda, x)$ and $F_2(\lambda, x)$ is assumed to be differentiable everywhere and with Lipschitz continuous partial derivatives, and the Hessian matrices $\nabla^2_{\lambda\lambda} F_1(\lambda, x)$ and $\nabla^2_{xx} F_2(\lambda, x)$ are globally invertible.*

Local solutions' concepts and first- and second-order necessary and sufficient conditions for bilevel optimization problems [19] are needed to establish the Stackelberg population flow with predictive-sensitivity conditioning in the next section.

Definition 2. *A point* (λ^*, x^*) *is said to be a local solution to the bilevel optimization problem* (1) *if:*

(a) x^* *is a local minimum of* $F_2(\lambda^*, x)$;

(b) *There exists a neighborhood* $\Omega \subset \Lambda \times X$ *of* (λ^*, x^*) *such that* $F_1(\lambda^*, x^*) \leq F_1(\lambda^*, x)$ *for all local solutions* $(\lambda^*, x^*) \in \Omega$ *such that* x^* *is a local minimum of* $F_2(\lambda, x)$.

Lemma 1 (First-order optimality conditions). *If* (λ^*, x^*) *is a local solution of Problem* (1), *then it is a stationary point satisfying the first-order Karush–Kuhn–Tucker (KKT) conditions:*

$$\nabla_\lambda F_1(\lambda^*, x^*(\lambda^*)) = 0,$$
$$\nabla_x F_2(\lambda^*, x^*) = 0. \tag{3}$$

For a given λ, Assumption 2 guarantees that the lower-level problem in (1) has at most a single optimal solution x^* satisfying the first-order optimality conditions in Lemma 1 of the lower-level $\nabla_x F_2(\lambda, x^*) = 0$. The implicit function theorem [39] guarantees the local existence of the optimal solution function $x^*(\lambda)$, and its derivative is obtained as:

$$\nabla_\lambda x^*(\lambda) = -\left(\nabla_{xx}^2 F_2(\lambda, x^*(\lambda))\right)^{-1} \nabla_{x\lambda}^2 F_2(\lambda, x^*(\lambda)).$$

In addition, this optimal solution function can be used to give an expression for the directional derivative, also known as the total derivative in points $x(\lambda)$ and defined as:

$$\mathcal{D}_\lambda F_1(\lambda, x(\lambda)) := \nabla_\lambda F_1(\lambda, x(\lambda))$$
$$+ \nabla_\lambda x(\lambda)^\top \nabla_x F_2(\lambda, x(\lambda)). \tag{4}$$

Here, $(\cdot)^\top$ stands for the vector or matrix transpose.

Lemma 2 (Second-order optimality conditions). *If a stationary point* (λ^*, x^*) *satisfies:*

$$\nabla_{\lambda\lambda}^2 F_1(\lambda^*, x^*(\lambda^*)) \geq 0,$$
$$\nabla_{xx}^2 F_2(\lambda^*, x^*) \geq 0, \tag{5}$$

then (λ^*, x^*) *is a local solution of the bilevel optimization Problem* (1).

The second-order total derivative can be deduced similarly as the first-order (4) as follows,

$$\mathcal{D}_{\lambda\lambda}^2 F_1(\lambda, x(\lambda)) := \nabla_{\lambda\lambda}^2 F_1(\lambda, x(\lambda)) + \nabla_\lambda x(\lambda)^\top \nabla_{x\lambda}^2 F_1(\lambda, x(\lambda))$$
$$+ \nabla_{\lambda,x}^2 F_1(\lambda, x(\lambda)) \nabla_\lambda x(\lambda) + \nabla_\lambda x(\lambda)^\top \nabla_{xx}^2 F_1(\lambda, x(\lambda)) \nabla_\lambda x(\lambda)$$
$$+ \nabla_x F_1(\lambda, x(\lambda))^\top (\nabla_{\lambda\lambda}^2 x(\lambda) + \nabla_{xx}^2 x(\lambda) \nabla_\lambda x(\lambda)).$$

The following definition introduces the differential versions of the Nash and Stackelberg equilibria [7].

Definition 3 (Differential Nash equilibrium (DNE)). *The joint strategy* $(\lambda^*, x^*) \in \Lambda \times X$ *is a differential Nash equilibrium if* $\nabla_\lambda F_1(\lambda^*, x^*) = 0$, $\nabla_x F_2(\lambda^*, x^*) = 0$, $\nabla_{\lambda\lambda}^2 F_1(\lambda^*, x^*) > 0$, *and* $\nabla_{xx} F_2(\lambda^*, x^*) > 0$.

Definition 4 (Differential Stackelberg equilibrium (DSE)). *The joint strategy* $(\lambda^*, x^*) \in \Lambda \times X$ *is a differential Stackelberg equilibrium if the total derivative* $\mathcal{D}_\lambda F_1(\lambda^*, x^*) = 0$, $\nabla_x F_2(\lambda^*, x^*) = 0$, $\mathcal{D}_{\lambda\lambda}^2 F_1(\lambda^*, x^*) > 0$, *and* $\nabla_{xx}^2 F_2(\lambda^*, x^*) > 0$.

In the next section, we present the main contribution of this work using the concept introduced above.

3. Stackelberg Population Games with Predictive-Sensitivity Conditioning

In this section, we show how the bilevel optimization problem in two time scales can be solved in a single time scale using an interconnection based on a predictive-sensitivity matrix of two dynamical systems in continuous-time, namely a gradient descent and population dynamics. First, the basic concepts of population games are introduced. Thus, the definitions and results for Stackelberg population games are derived. Finally, the interconnection using the predictive-sensitivity matrix for Stackelberg population games is shown.

3.1. Population Games' Essentials

Population games have been proposed as a multiagent model to find the Nash equilibrium for a large set of agents or a population. The main dynamical model in population games is the replicator dynamics. Replicator dynamics has been successfully implemented in various engineering applications where real-time adaptation and robustness to dynamic environmental uncertainties is of vital importance [35]. Replicator dynamics represents a mass M of players choosing strategies evolving in time. It is assumed that a finite set of pure strategies $\mathcal{S} = \{1, 2, \ldots, n\}$, and the analysis is based on a payoff function associated with the selected strategy. The population states are denoted by the vector $x = [x_1, x_2, \ldots, x_n]$, and the population states are constrained by all possible distributions of individuals among the strategies given by the simplex:

$$\Delta = \left\{ x \in \mathbb{R}^n_{\geq 0} \mid \sum_{i=1}^n x_i = M \right\} \tag{6}$$

The agents playing strategy i obtain a reward depending on the population state. A fitness function representing the reward for agents playing strategy i is defined as $f_i : \Delta \mapsto \mathbb{R}$ for $i \in \mathcal{S}$. The replicator dynamics is interpreted as an evolutionary game where the proportion of the population of agents playing the most successful strategy (higher payoff than the average) is increasing. Agents dynamically compare their fitness function with the other agents' performance through the average fitness function of the whole population $\bar{f} = \sum_i^n x_i f_i$, also understood as the expected average payoff of the population. The replicator dynamics associated with each population of agents playing the i-th strategy is given by:

$$\dot{x}_i = x_i (f_i(x) - \bar{f}(x)), \quad \text{for all } i \in \mathcal{S}. \tag{7}$$

As mentioned, the solution concept of the replicator dynamics is an equilibrium, and the Nash equilibrium is the preferred solution concept in population games. Let F be a population game: the Nash equilibrium set is defined as:

$$PNE(F) = \{x^* \in \Delta \; : \; x_i^* > 0 \Rightarrow f_i(x^*) \geq f_j(x^*), \forall \, i, j \in \mathcal{S}\}.$$

All agents obtain the same profit in a Nash equilibrium. An important type of population games is potential games. Potential games can be defined as follows.

Definition 5. *Let $F : \mathbb{R}^n_+ \mapsto \mathbb{R}^n$ be a vector of fitness functions with positive population game payoffs. Let $V : \mathbb{R}^n_+ \mapsto \mathbb{R}$ be a continuously differentiable function and the following condition hold,*

$$\nabla_x V(x) = F(x), \quad \forall x \in \mathbb{R}^n_+.$$

Then, F is a full potential game.

In potential games, a single scalar-valued function is associated with the game, called the potential function, which retains key information about the payoffs of the agents. If

there exists a continuously differentiable potential function $V : \mathbb{R}_+^n \mapsto \mathbb{R}$, then a potential game satisfies:

$$\frac{\partial V(x)}{\partial x_i} = f_i(x) \quad \text{for all } i \in \mathcal{S}, \tag{8}$$

and (8) also implies that F must satisfy the external symmetry:

$$\frac{\partial f_i}{\partial x_j} = \frac{\partial f_j}{\partial x_i} \quad \text{for all } i, j \in \mathcal{S}. \tag{9}$$

In potential games, the Nash equilibrium is related to local maximizers of potential functions [34]; this is stated in the following lemma.

Lemma 3. *In a full potential game, the Nash equilibrium is equal to the solution of the first-order KKT conditions of the maximizing $V(x)$ subject to a feasible set Δ.*

Finally, an important result on the stability of population games claims that a population game satisfying (8) is a stable game if the potential function V is concave [34].

3.2. Stackelberg Population Games

One of the main concepts in the design of population games is the selection of the fitness functions such that the replicator dynamics (7) can be used as a distributed optimization algorithm. In order to obtain fitness functions to accomplish a stable and convergent solution, we use the following Lemma 4, which characterizes an optimal solution for the lower-level optimization problem (1) satisfying Assumption 2 [44].

Lemma 4. *A solution of the bilevel optimization problem (1) (λ^*, x^*), with $\lambda^* \in \Lambda$ and x^* belonging to the feasible set Δ, is an optimal solution for the population game if and only if $\nabla_{x_i} F_2(x_i^*) = \nabla_{x_j} F_2(x_j^*)$ for all $i, j \in \mathcal{S}$.*

Definition 6. *Let Λ be the set of strategies for the leader agent and $X \in \mathbb{R}^n$ be the set of strategies for the follower population. Let the best response function for the follower population $br_f(\lambda)$ be defined as:*

$$br_f(\lambda) = \{x \in \Delta | x_i > 0 \implies f_i(\lambda, x) \geq f_j(\lambda, x), \quad \forall i, j \in S\}.$$

A leader strategy $\lambda^* \in \Lambda$ and a follower population strategy vector x^* are in a Stackelberg population equilibrium if the following conditions hold:

(a) $\max_{x \in br_f(\lambda^*)} F_1(\lambda^*, x) \leq \max_{x \in br_f(\lambda)} F_1(\lambda, x);$

(b) $x^* \in br_f(\lambda^*).$

Definition 7 (Differential Stackelberg population equilibrium (DSPE))**.** *The joint strategy $(\lambda^*, x^*) \in \Lambda \times \Delta$ is a differential Stackelberg population equilibrium if the following KKT conditions hold:*

(a)   *First-order:* $\mathcal{D}_\lambda F_1(\lambda^*, x^*) = 0, \nabla_x F_2(\lambda^*, x^*) = 0;$

(b)   *Second-order:* $\mathcal{D}_{\lambda\lambda}^2 F_1(\lambda^*, x^*) > 0, \nabla_{xx}^2 F_2(\lambda^*, x^*) > 0.$

For the leader upper-level optimization problem (1), we propose a continuous-time gradient descent flow using the concept of the directional derivative, also known as total derivative obtained following the traditional ideas of gradient dynamics for convex optimization problems [45] as follows:

$$\Sigma_1 : \dot{\lambda} = -\mathcal{D}_\lambda F_1(\lambda, x). \tag{10}$$

For the lower-level optimization problem, we propose the follower population game dynamics using the replicator dynamics (7). In order to guarantee the convergence properties of the population games, a full potential game introduced in Definition 5 is chosen using the objective function $F_2(\lambda, x)$ as the potential function, i.e., $V(\lambda, x) = -F_2(\lambda, x)$. Notice that the potential games are related to maximizing the potential function $V(x)$, and in this case, we seek to minimize, so we convert the problem. With this potential function, we can obtain the fitness function using (8), yielding for each agent i:

$$f_i(\lambda, x) = \frac{\partial V(\lambda, x)}{\partial x_i} = -\frac{\partial F_2(\lambda, x)}{\partial x_i},$$

and since $F_2(\lambda, x)$ is separable, then the fitness function for each agent i is:

$$f_i(\lambda, x) = -\frac{\partial}{\partial x_i}\left(\sum_{i=1}^{n} J_i(\lambda, x)\right) = -\frac{\partial J_i(\lambda, x)}{\partial x_i}. \tag{11}$$

The average fitness function $\bar{f}(\lambda, x)$ is then:

$$\bar{f}(\lambda, x) = -\sum_{i=1}^{n} x_i \frac{\partial J_i(\lambda, x)}{\partial x_i}. \tag{12}$$

With fitness functions for each agent (11) and the average fitness function (12), the replicator dynamics is the lower-level flow as follows:

$$\dot{x}_i = x_i\big(f_i(\lambda, x) - \bar{f}(\lambda, x)\big), \quad \forall i \in \mathcal{S}.$$

In compact form, we can define the vector of fitness functions $f(\lambda, x) = [f_1, f_2, \ldots, f_n]$. We obtain the replicator dynamics in compact form as:

$$\Sigma_2 : \dot{x} = x \cdot \big(f(\lambda, x) - \bar{f}(\lambda, x)\big). \tag{13}$$

We used the notation "·" to indicate an elementwise product between vectors. In this leader–follower population dynamics, we have a nested interaction model in two time scales, which could slow down the convergence of the dynamics. In order to solve these limitations in two time scales, we propose to use the prediction-sensitivity conditioning concept presented in the next section.

3.3. Predictive-Sensitivity for Stackelberg Population Games

In this section, we introduce the predictive-sensitivity conditioning approach to deal with the bilevel optimization problem (1). The regularity introduced in Assumption 2 guarantees that the bilevel problem has at most one solution, and this also ensures that the problem is well posed. Then, it is possible to define the sensitivity of x^*. The main idea is represented in Figure 1, where it is shown that the sensitivity matrix is used to integrate in one single level the dynamics associated with the upper-level by the gradient dynamics (10) and the lower-level by the replicator dynamics (13), respectively.

Considering Assumption 2, the analytic expression of (4) is well defined at any point (λ, x), so defining an extended sensitivity for any point (λ, x) is possible as:

$$S_\lambda^x(\lambda, x) := -\left(\nabla_{xx}^2 F_2(\lambda, x)\right)^{-1} \nabla_{x\lambda}^2 F_2(\lambda, x), \tag{14}$$

while:

$$S_\lambda^x(\lambda, x)|_{(\lambda, x^*(\lambda))} = \nabla_\lambda x^*(\lambda).$$

is satisfied. The total derivative using the sensitivity matrix can be extended to any point (λ, x) as:

$$\mathcal{D}_\lambda F_1(\lambda, x) := \nabla_\lambda F_1(\lambda, x) \\ + S_\lambda^x(\lambda, x)^\top \nabla_x F_2(\lambda, x). \tag{15}$$
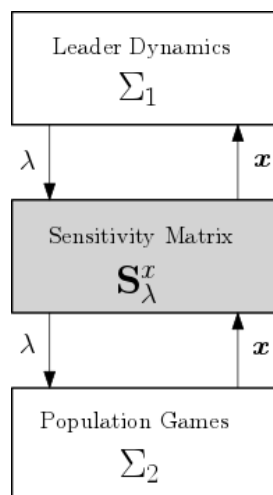

Figure 1. Predictive-sensitivity conditioning interconnection.

The two-level interconnected dynamical system represented by Σ_1 and Σ_2 can be integrated into a single time scale using the sensitivity matrix (14) and the total derivative (15). A feed-forward term is added as the predictive-sensitivity conditioning based on the optimization sensitivity to the follower population dynamics Σ_2 (13) to predict the dynamics of the leader gradient descent dynamics Σ_1 (10).

Based on a recent work [37], we used a predictive-sensitivity conditioning interconnection between the two-time-scale systems without the need for the time scale separation. This interconnection preserves the optimal solution $(\lambda^*, x^*(\lambda^*))$ and its convergence properties. A sensitivity-based conditioning matrix S_λ^x is used as the interconnection between both subsystems as shown in Figure 1, and it is defined as follows.

Definition 8. *Consider the extended sensitivity $S_\lambda^x(\lambda, x)$ defined in (14): the conditioning matrix is defined as:*

$$S = \begin{bmatrix} I & 0 \\ -S_\lambda^x(\lambda, x) & I \end{bmatrix} \tag{16}$$

where I is an identity matrix of the appropriate dimension.

Using Definition 8, we obtain the predictive-sensitivity Stackelberg learning dynamics:

$$S \begin{bmatrix} \dot\lambda \\ \dot x \end{bmatrix} = \begin{bmatrix} I & 0 \\ -S_\lambda^x(\lambda, x) & I \end{bmatrix} \begin{bmatrix} \dot\lambda \\ \dot x \end{bmatrix} = \begin{bmatrix} -\mathcal{D}_\lambda F_1(\lambda, x) \\ x \cdot (f(\lambda, x) - \bar f(\lambda, x)) \end{bmatrix} \tag{17}$$

It is observed that the extended sensitivity term modifies the dynamics of the follower population as:

$$\dot x = x \cdot (f(\lambda, x) - \bar f(\lambda, x)) - S_\lambda^x(\lambda, x)\mathcal{D}_\lambda F_1(\lambda, x),$$

where the first term drives the population state x to the optimal solution $x^*(\lambda)$, while the predictive-sensitivity feed-forward term anticipates the variation of $x^*(\lambda)$ due to the dynamics $\dot\lambda$. The complete predictive-sensitivity Stackelberg population flow is given by:

$$\dot\lambda = -\mathcal{D}_\lambda F_1(\lambda, x) \\ \dot x = x \cdot (f(\lambda, x) - \bar f(\lambda, x)) - S_\lambda^x(\lambda, x)\mathcal{D}_\lambda F_1(\lambda, x) \tag{18}$$

In the next result, the relation between the differential Stackelberg population equilibrium concepts defined in Definition 7 and the stability of the predictive-sensitivity Stackelberg population dynamics is established (18).

Theorem 1. *The joint strategy* (λ^*, x^*) *is a DSPE of the Stackelberg population game* (1) *satisfying the condition in Definition 7 if and only if it is a locally exponentially stable point of the predictive-sensitivity Stackelberg learning dynamics* (18).

Proof. Consider that (18) satisfies Assumptions 1 and 2. Since we are proving local stability, it is sufficient to prove that the eigenvalues of the Jacobian of the system at the joint strategy (λ^*, x^*) have a strictly negative real part [36]. For this, observe that by Definition 7, if the point (λ^*, x^*) is a DSPE, then the second-order conditions are satisfied at this point, and in addition, it has been proven that if the population dynamics are a full potential game, this implies that the Stackelberg follower population (see Definition 6) is asymptotically stable [35], then it is guaranteed that the Jacobian would have a strictly negative part at point (λ^*, x^*). Conversely, if a point (λ^*, x^*) of the system (18) is locally exponentially stable, then the corresponding Jacobian must have eigenvalues with a negative real part; thus, it is a strictly local solution satisfying the conditions in Definition 7. □

In the next section, an interesting application to the coordination of distributed energy resources via pricing dynamics is presented to illustrate the theoretical results of the proposed framework.

4. DER Coordination via Pricing Dynamics

A distributed energy resource (DER) coordination problem via pricing dynamics [6] is presented to illustrate the applicability of the proposed predictive-sensitivity Stackelberg learning dynamics. A dynamic transactive control involving a feedback loop is implemented to solve a bilevel optimization problem via a pricing dynamics. The DERs consist of a set of distributed generators (follower population) interacting with a distribution system operator (DSO) (leader) to reach a Stackelberg equilibrium.

Consider a set of distributed generators $G = \{1, 2, \ldots, M\}$: each distributed generator has a payoff function:

$$U_i(\lambda, p_i) = \lambda p_i - C_i(p_j) \tag{19}$$

where p_i is the amount of power produced by each generator i (this power is stacked in a vector as $p = [p_1, p_2, \ldots, p_M]^\top$), λ is the energy price defined by the DSO, and $C_i(p_i)$ is the cost function for producing energy for each generator. The cost function is traditionally chosen as a convex quadratic function such as $C_i(p_i) = 1/2\alpha_i p_i^2 + \beta_i p_i$, where α_i, β_i are the cost coefficients. The power resource allocation is constrained by the power balance equation $\sum_{i=1}^{M} p_i = P_L$, where P_L is the total power demand of the distribution network.

On the other hand, the main goal of the pricing problem is to determine the market clearing price (MCP) for the energy producer. The DSO solves an optimization problem that maximize the welfare of the population over a market cycle. It is observed that the coordination scheme is cast as a Stackelberg game, which is a challenging bilevel optimization problem. In the objective function, it is assumed that the DERs will have chosen the optimal generation profile for a particular price λ. The optimal market clearing price is then obtained by solving the following bilevel optimization problem:

$$\max_{\lambda} \quad U(\lambda, p) = \sum_{i=1}^{M} U_i(\lambda, p_i)$$

$$\text{s.t.} \quad p = \arg\max_{p}\{U(\lambda, p) : \sum_{i=1}^{M} p_i = P_L\}. \tag{20}$$

Since we posed the proposed dynamics for minimization problems, we define the objective functions as $F_1(\lambda, p) = F_2(\lambda, p) = -U(\lambda, p) = \sum_{i=1}^{M}(C_i(p_j) - \lambda p_i)$. Following the population game's description, the fitness function for each generator is given by:

$$f_i(\lambda, \boldsymbol{p}) = \frac{dU}{dp_i} = \frac{dU_i}{dp_i} = \lambda - (\alpha_i p_i + \beta_i).$$

In order to obtain the pricing dynamics, we need the sensitivity matrix (14):

$$S_\lambda^x(\lambda, x) = -\left(\nabla_{xx}^2 F_2(\lambda, x)\right)^{-1} \nabla_{x\lambda}^2 F_2(\lambda, x)$$

$$= -\begin{bmatrix} -\alpha_1 & 0 & \cdots & 0 \\ 0 & -\alpha_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -\alpha_M \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1/\alpha_1 \\ 1/\alpha_2 \\ \vdots \\ 1/\alpha_M \end{bmatrix}$$

With this sensitivity matrix, we can obtain the total derivative (15) for the gradient descent for the upper-level pricing variable λ. Now, we are ready to introduce the predictive-sensitivity Stackelberg learning dynamics (18) for the DER coordination problem:

$$\dot{\lambda} = \sum_{i=1}^{M} p_i + \sum_{i=1}^{M}\left(p_i + \frac{\beta_i}{\alpha_i} - \frac{\lambda}{\alpha_i}\right) = \sum_{i=1}^{M}\left(2p_i + \frac{(\beta_i - \lambda)}{\alpha_i}\right),$$

$$\dot{p} = (1/P_L)p \cdot \left(f(\lambda, p) - \bar{f}(\lambda, p)\right) - \frac{1}{\alpha}\left(\sum_{i=1}^{M} p_i + \sum_{i=1}^{M}\left(p_i + \frac{\beta_i}{\alpha_i} - \frac{\lambda}{\alpha_i}\right)\right),$$

(21)

where $\frac{1}{\alpha}$ is the vector with elements $1/\alpha_i$.

The distributed generators' (DGs) capacities and coefficients for the simulation experiments are presented in Table 1. The numerical experiments were simulated in a IEEE 9-bus test feeder adapted from [41], where each DG was modeled using a voltage–source inverter with local controllers. In Figure 2 is presented the distribution network with three distributed generators and three loads. The test case scenario is as follows. Initially, the system has to supply a total load of $P_L = 4950$ W, corresponding to three loads as $L_1 = 1500$ W, $L_2 = 1250$ W, and $L_3 = 2200$ W in the distribution network. The distributed generators in the distribution network dynamically interact with each other and with the DSO to reach the optimal solution via a pricing dynamics that determines the optimal price and the optimal energy allocation for each DG based on the proposed predictive-sensitivity Stackelberg learning dynamics (18). Then, at $t = 0.8$ s of simulation, an additional load is added to the grid, summing up the total to $P_L = 5850$ W.

In Figure 3 is presented the frequency response. It is shown that the system maintains the frequency stable around the reference of 60 Hz. The variations of the load are observed as small overshoots, but the system returns to the equilibrium point with a rapid response.
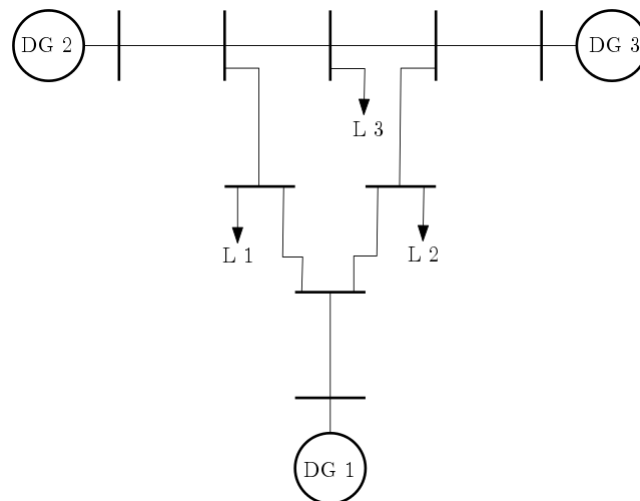
Figure 2. Distribution network with DERs adapted from an IEEE 9-bus test feeder.

Table 1. System parameters' simulation.

Distributed Generators				
i	$\overline{p}_i$	$\underline{p}_i$	α_i	β_i
1	6000 (W)	200 (W)	2.5	3.5
2	5500 (W)	200 (W)	2.5	3.6
3	4000 (W)	200 (W)	1.9	2.5
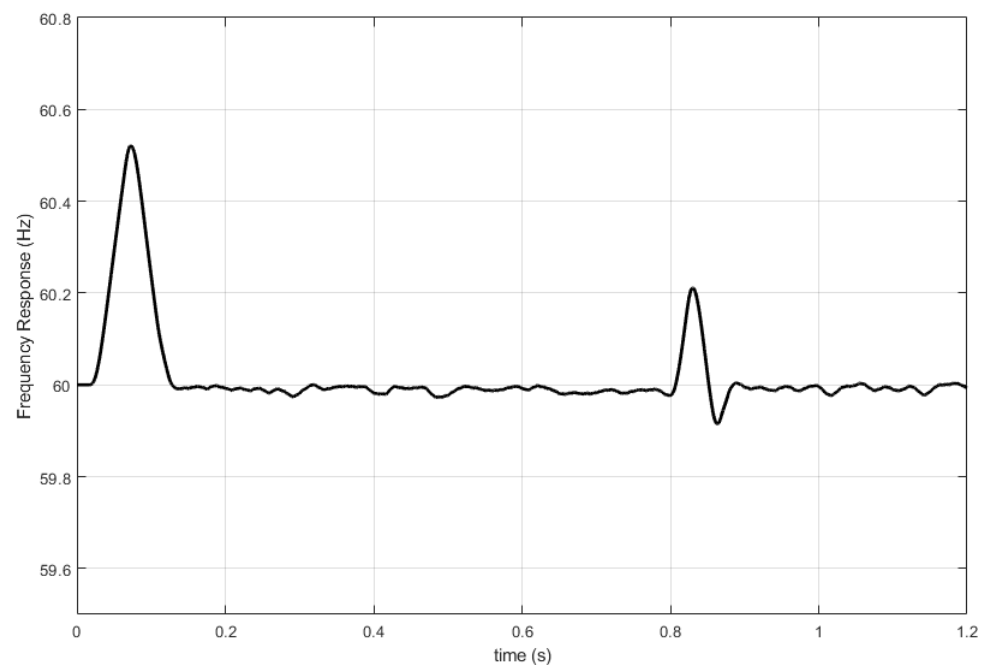
Figure 3. Frequency response of the distribution network.

In Figure 4 is presented the power behavior of each DG. It is shown that the less expensive DG3 (with small α) dispatches more power to the system, while the more expensive DG2 (with the bigger α) dispatches less power. The response to variations of the load (see Figure 5) at $t = 0.8$ s is observed, and the generators dynamically adapt to the new demand with a small overshoot.
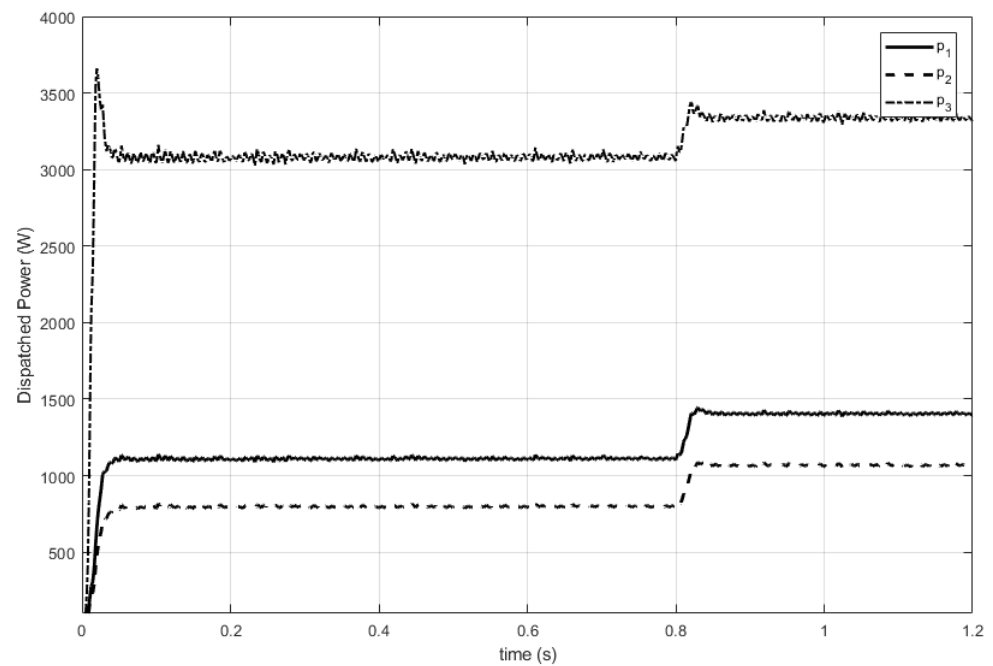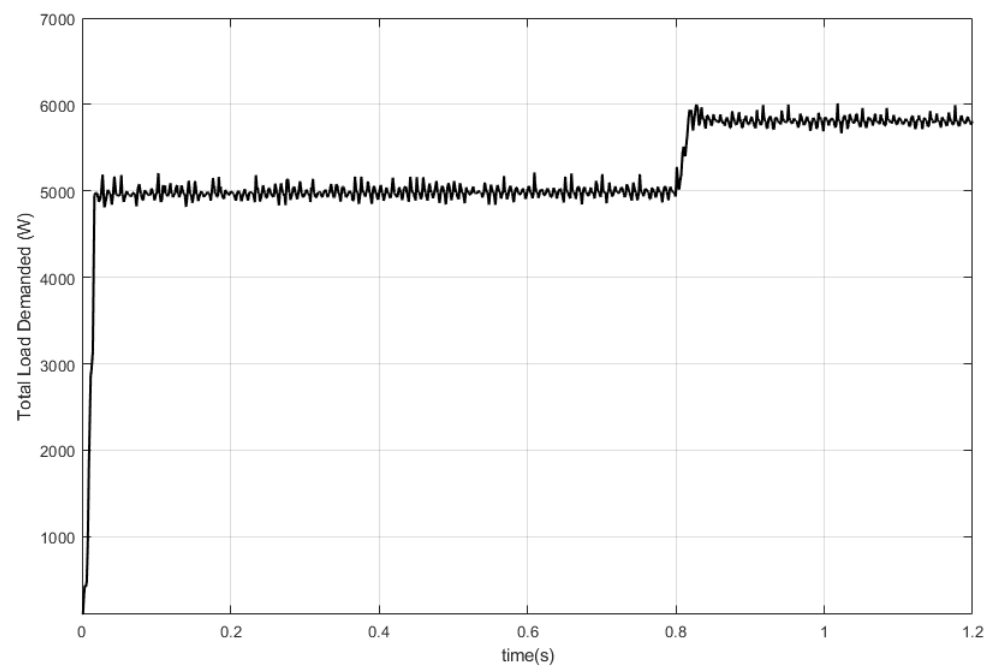
Figure 4. Power response of each DG.

Figure 5. Total load demanded response.

Finally, Figure 6 presents the price dynamics $\lambda(t)$. Behavior towards the stable point is observed and, then, the dynamic response to the variation of the load at $t = 0.8$ s, as expected.
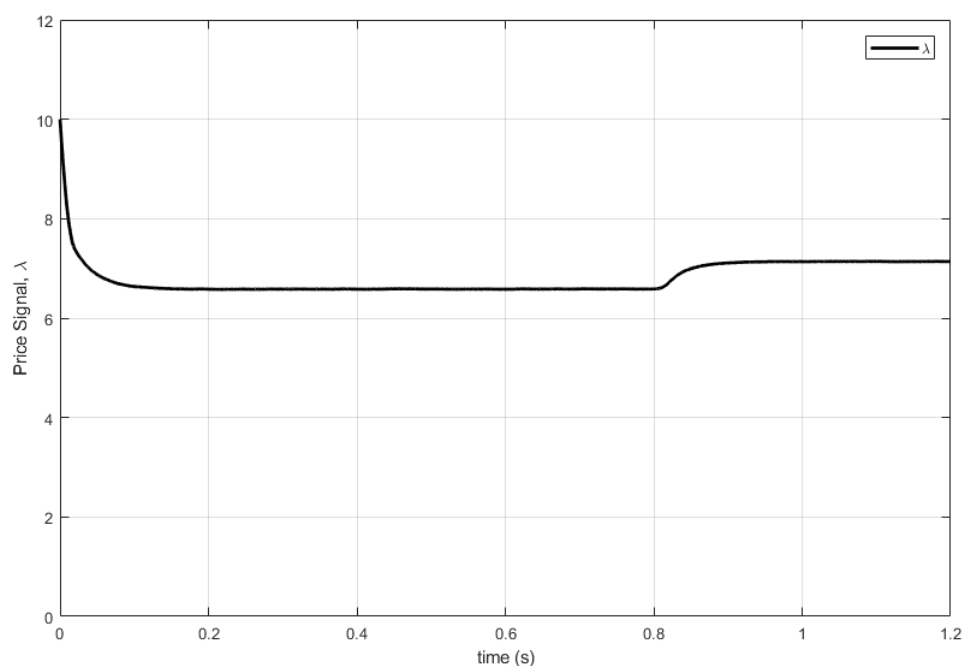
Figure 6. Price dynamics response.

5. Conclusions

To solve bilevel optimization problems between a leader and a follower population, we proposed two dynamical systems interconnected by a predictive-sensitivity conditioning matrix in a single time scale. For the leader optimization problem, we developed a gradient descent algorithm based on the total derivative, and for the followers' optimization problem, we used the population dynamics framework to model a population of interacting strategic agents. We extended the concept of the Stackelberg population equilibrium to the differential Stackelberg population equilibrium for population dynamics. Theoretical guarantees for the stability of the proposed Stackelberg population learning dynamics were presented. A distributed energy resource coordination problem was solved via a pricing dynamics based on the proposed approach. Some simulation experiments were presented to illustrate the effectiveness of the framework. As future work, several research avenues are available. For instance, extensions of this work to include uncertainty would be very useful for real-life problems.

References

1. Novak, A.; Feichtinger, G.; Leitmann, G. A differential game related to terrorism: Nash and Stackelberg strategies. *J. Optim. Theory Appl.* **2010**, *144*, 533–555. [CrossRef]
2. Li, Y.; Shi, D.; Chen, T. False data injection attacks on networked control systems: A Stackelberg game analysis. *IEEE Trans. Autom. Control* **2018**, *63*, 3503–3509. [CrossRef]
3. Groot, N.; Zaccour, G.; De Schutter, B. Hierarchical game theory for system-optimal control: Applications of reverse Stackelberg games in regulating marketing channels and traffic routing. *IEEE Control Syst. Mag.* **2017**, *37*, 129–152.
4. Motalleb, M.; Siano, P.; Ghorbani, R. Networked Stackelberg competition in a demand response market. *Appl. Energy* **2019**, *239*, 680–691. [CrossRef]
5. Chen, J.; Zhu, Q. A Stackelberg game approach for two-level distributed energy management in smart grids. *IEEE Trans. Smart Grid* **2017**, *9*, 6554–6565. [CrossRef]
6. Baron-Prada, E.; Mojica-Nava, E. A population games transactive control for distributed energy resources. *Int. J. Electr. Power Energy Syst.* **2021**, *130*, 106874. [CrossRef]
7. Fiez, T.; Chasnov, B.; Ratliff, L. Implicit learning dynamics in Stackelberg games: Equilibria characterization, convergence analysis, and empirical study. In Proceedings of the International Conference on Machine Learning, PMLR 2020, Online, 13–18 July 2020; pp. 3133–3144.
8. Hirose, K.; Matsumura, T. Comparing welfare and profit in quantity and price competition within Stackelberg mixed duopolies. *J. Econ.* **2019**, *126*, 75–93. [CrossRef]
9. Madry, A.; Makelov, A.; Schmidt, L.; Tsipras, D.; Vladu, A. Towards deep learning models resistant to adversarial attacks. In Proceedings of the International Conference on Learning Representations, Vancouver, BC, Canada, 30 April–3 May 2018.
10. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. *Adv. Neural Inf. Process. Syst.* **2014**, *27*, pp. 2672–2680.
11. Zhang, K.; Yang, Z.; Başar, T. Multi-agent reinforcement learning: A selective overview of theories and algorithms. In *Handbook of Reinforcement Learning and Control*; Studies in Systems, Decision and Control; Springer, Cham, Switzerland, 2021; Volume 325, pp. 321–384..
12. Jin, C.; Netrapalli, P.; Jordan, M. What is local optimality in nonconvex-nonconcave minimax optimization? In Proceedings of the International Conference on Machine Learning. PMLR 2020, Online, 13–18 July 2020; pp. 4880–4889.
13. Sinha, A.; Malo, P.; Deb, K. A Review on Bilevel Optimization: From Classical to Evolutionary Approaches and Applications. *IEEE Trans. Evol. Comput.* **2018**, *22*, 276–295. [CrossRef]
14. Stackelberg, H.V. *Theory of the Market Economy*; Oxford University Press: Oxford, UK, 1952.
15. Bracken, J.; McGill, J.T. Mathematical programs with optimization problems in the constraints. *Oper. Res.* **1973**, *21*, 37–44. [CrossRef]
16. Hansen, P.; Jaumard, B.; Savard, G. New branch-and-bound rules for linear bilevel programming. *SIAM J. Sci. Stat. Comput.* **1992**, *13*, 1194–1217. [CrossRef]
17. Vicente, L.; Savard, G.; Júdice, J. Descent approaches for quadratic bilevel programming. *J. Optim. Theory Appl.* **1994**, *81*, 379–399. [CrossRef]
18. Deng, X. Complexity issues in bilevel linear Programming. In *Multilevel Optimization: Algorithms and Applications*; Springer: Berlin/Heidelberg, Germany, 1998; pp. 149–164.
19. Bard, J.F. *Practical Bilevel Optimization: Algorithms and Applications*; Springer Science & Business Media; Boston, MA, USA, 1998; Volume 30.
20. Li, T.; Sethi, S.P. A review of dynamic Stackelberg game models. *Discret. Contin. Dyn. Syst.-B* **2017**, *22*, 125. [CrossRef]
21. Luo, Z.Q.; Pang, J.S.; Ralph, D. *Mathematical Programs with Equilibrium Constraints*; Cambridge University Press: Cambridge, UK, 1996.
22. Kulkarni, A.A.; Shanbhag, U.V. An existence result for hierarchical Stackelberg v/s Stackelberg games. *IEEE Trans. Autom. Control* **2015**, *60*, 3379–3384. [CrossRef]
23. Kebriaei, H.; Iannelli, L. Discrete-time robust hierarchical linear-quadratic dynamic games. *IEEE Trans. Autom. Control* **2017**, *63*, 902–909. [CrossRef]
24. Başar, T.; Srikant, R. A Stackelberg network game with a large number of followers. *J. Optim. Theory Appl.* **2002**, *115*, 479–490. [CrossRef]
25. Fabiani, F.; Tajeddini, M.A.; Kebriaei, H.; Grammatico, S. Local Stackelberg equilibrium seeking in generalized aggregative games. *IEEE Trans. Autom. Control* **2021**, 1–6. [CrossRef]
26. Kok, K.; Widergren, S. A Society of Devices: Integrating Intelligent Distributed Resources with Transactive Energy. *IEEE Power Energy Mag.* **2016**, *14*, 34–45. [CrossRef]
27. Bejestani, A.K.; Annaswamy, A.; Samad, T. A hierarchical transactive control architecture for renewables integration in smart grids: Analytical modeling and stability. *IEEE Trans. Smart Grid* **2014**, *5*, 2054–2065. [CrossRef]
28. Hu, J.; Yang, G.; Kok, K.; Xue, Y.; Binder, H.W. Transactive control: a framework for operating power systems characterized by high penetration of distributed energy resources. *J. Mod. Power Syst. Clean Energy* **2017**, *5*, 451–464. [CrossRef]
29. Li, S.; Lian, J.; Conejo, A.J.; Zhang, W. Transactive Energy Systems: The Market-Based Coordination of Distributed Energy Resources. *IEEE Control. Syst. Mag.* **2020**, *40*, 26–52. [CrossRef]

30. Barreto, C.; Mojica-Nava, E.; Quijano, N. Incentive mechanisms to prevent efficiency loss of non-profit utilities. *Int. J. Electr. Power Energy Syst.* **2019**, *110*, 523–535. [CrossRef]
31. Groot, N.; De Schutter, B.; Hellendoorn, H. On Systematic Computation of Optimal Nonlinear Solutions for the Reverse Stackelberg Game. *IEEE Trans. Syst. Man Cybern. Syst.* **2014**, *44*, 1315–1327. [CrossRef]
32. Arrow, K.J.; Azawa, H.; Hurwicz, L.; Uzawa, H. *Studies in Linear and Non-Linear Programming*; Stanford University Press: Palo Alto, CA, USA, 1958; Volume 2.
33. Helmke, U.; Moore, J.B. *Optimization and Dynamical Systems*; Springer Science & Business Media: Boston, MA, USA, 2012.
34. Sandholm, W.H. *Population Games and Evolutionary Dynamics*; MIT Press: Boston, MA, USA, 2010.
35. Quijano, N.; Ocampo-Martinez, C.; Barreiro-Gomez, J.; Obando, G.; Pantoja, A.; Mojica-Nava, E. The Role of Population Games and Evolutionary Dynamics in Distributed Control Systems. *IEEE Control Syst.* **2017**, *37*, 70–97.
36. Khalil, H.K. *Nonlinear Systems*, 3rd ed.; Patience Hall: Upper Saddle River, NJ, USA, 2002.
37. Picallo, M.; Bolognani, S.; Dörfler, F. Predictive-sensitivity: Beyond Singular Perturbation for Control Design on Multiple Time Scales. *arXiv* **2021**, arXiv:2101.04367.
38. Dempe, S.; Mordukhovich, B.S.; Zemkoho, A.B. Sensitivity analysis for two-level value functions with applications to bilevel programming. *SIAM J. Optim.* **2012**, *22*, 1309–1343. [CrossRef]
39. Krantz, S.G.; Parks, H.R. *The Implicit Function Theorem: History, Theory, and Applications*; Springer Science & Business Media: Boston, MA, USA, 2012.
40. Mojica-Nava, E.; Macana, C.A.; Quijano, N. Dynamic Population Games for Optimal Dispatch on Hierarchical Microgrid Control. *IEEE Trans. Syst. Man Cybern. Syst.* **2014**, *44*, 306–317. [CrossRef]
41. Mojica-Nava, E.; Barreto, C.; Quijano, N. Population Games Methods for Distributed Control of Microgrids. *IEEE Trans. Smart Grid* **2015**, *6*, 2586–2595. [CrossRef]
42. Mojica-Nava, E.; Rivera, S.; Quijano, N. Game-theoretic dispatch control in microgrids considering network losses and renewable distributed energy resources integration. *IET Gener. Transm. Distrib.* **2017**, *11*, 1583–1590. [CrossRef]
43. Zhan, Y.T.; Li, X.S.; Huang, N.J. A Stackelberg population competition model via variational inequalities and fixed points. *Carpathian J. Math.* **2020**, *36*, 331–339. [CrossRef]
44. Lakshmanan, H.; De Farias, D.P. Decentralized resource allocation in dynamic networks of agents. *SIAM J. Optim.* **2008**, *19*, 911–940. [CrossRef]
45. Bertsekas, D.; Nedic, A.; Ozdaglar, A. *Convex Analysis and Optimization*; Athena Scientific Belmont: Nashua, NH, USA, 2003.