Local Stackelberg equilibrium seeking in generalized aggregative games

Fabiani, Filippo; Tajeddini, Mohammad Amin; Kebriaei, Hamed; Grammatico, Sergio

DOI
10.1109/TAC.2021.3077874

Publication date
2022

Document Version
Final published version

Published in
IEEE Transactions on Automatic Control

Citation (APA)
Fabiani, F., Tajeddini, M. A., Kebriaei, H., & Grammatico, S. (2022). Local Stackelberg equilibrium seeking in generalized aggregative games. *IEEE Transactions on Automatic Control*, *67*(2), 965-970. https://doi.org/10.1109/TAC.2021.3077874

Local Stackelberg Equilibrium Seeking in Generalized Aggregative Games

Filippo Fabiani, Mohammad Amin Tajeddini, Hamed Kebriaei, *Senior Member, IEEE*, and Sergio Grammatico, *Senior Member, IEEE*

***Abstract*—We propose a two-layer, semidecentralized algorithm to compute a local solution to the Stackelberg equilibrium problem in aggregative games with coupling constraints. Specifically, we focus on a single-leader, multiple-follower problem, and after equivalently recasting the Stackelberg game as a mathematical program with complementarity constraints (MPCC), we iteratively convexify a regularized version of the MPCC as the inner problem, whose solution generates a sequence of feasible descent directions for the original MPCC. Thus, by pursuing a descent direction at every outer iteration, we establish convergence to a local Stackelberg equilibrium. Finally, the proposed algorithm is tested on a numerical case study, a hierarchical instance of the charging coordination problem of plug-in electric vehicles.**

***Index Terms*—Game theory, hierarchical systems, optimization, Stackelberg equilibrium.**

I. Introduction

Stackelberg equilibrium problems are very popular within the system-and-control community, since they offer a multiagent, decision-making framework that enables one to model not only "horizontal" but also "vertical" interdependent relationships among heterogeneous agents, which are, hence, clustered into *leaders* and *followers*. The application domains of Stackelberg equilibrium problems are, indeed, numerous, spanning from wireless networks, telecommunications [1], and network security [2], to demand response and energy management [3]–[5], economics [6], and traffic control [7].

In its most general setting, a Stackelberg equilibrium problem between a leader and a set of followers can be formulated as a mathematical program with equilibrium constraints (MPEC) [8, Sec. 1.2] or, in some specific cases, as a mathematical program with complementarity constraints (MPCC) [9]. Both MPECs and MPCCs are usually challenging to solve. Specifically, they are inherently ill-posed, nonconvex optimization problems, since typically there are no feasible solutions strictly lying in the interior of the feasible set, which may even be disconnected, implying that any constraint qualification is violated at every feasible point [10]. It follows that, in this context, the basic convergence assumptions characterizing standard constrained optimization algorithms are not satisfied. Therefore, available solution methods are either tailored to the specific problem considered, or designed *ad hoc* for a subclass of MPECs/MPCCs.

Algorithmic solution techniques for the class of games involving dominant and nondominant strategies, i.e., leaders and followers, trace back to the 1970s. For example, open-loop and feedback-control policies for differential, hence continuous-time, unconstrained games were designed in [11] and [12], whereas in [13], a comparison between finite/infinite horizon control strategies involving discrete-time dynamics was proposed. More recently, a single-leader, multifollower differential game, modeling a pricing scheme for the Internet by basing on the bandwidth usage of the users, i.e., with congestion constraints, was solved in [14], and an iterative procedure to compute a Stackelberg–Nash saddle point for an unconstrained, single-leader, multifollower game with discrete-time dynamics was proposed in [15]. By relying on the uniqueness of the followers' equilibrium for each leader's strategy, standard fixed-point algorithms are also proposed in [16] and [17]. A first attempt to solve an MPEC by considering a more elaborated multileader, multifollower game was investigated in [18]. Specifically, the authors established the equivalence to a single-leader, multifollower game whenever the cost functions of the leaders admit a potential function and, in addition, the set of leaders has an identical conjecture or estimate on the follower equilibrium. Similar arguments are also exploited in [19] to address the same multileader, multifollower equilibrium problem. In this latter case, for each leader, the authors proposed a single-leader, multifollower game modeled as an MPEC. On the other hand, all these subgames, which are parametric in the decisions of the followers, are coupled together through a game with the leaders themselves. However, in both papers, the solution to the single-leader, multifollower game remains to be dealt with, mainly due to the presence of nonconvexities and equilibrium/complementarity constraints, which characterize MPEC/MPCC. Early algorithmic works on MPCCs to solve single-leader, multifollower Stackelberg games, such as Gauss–Seidel or Jacobi [20], [21], are computationally expensive, especially for large number of followers. Additionally, they introduce several privacy issues, since they are designed by relying on diagonalization techniques. In [22], after relaxing the complementarity conditions, a solution to an MPCC is computed through nonlinear complementarity problems, by driving the relaxation parameter to zero.

This article aims at filling the apparent lack in the aforementioned literature of scalable and privacy preserving solution algorithms for equilibrium problems with nonconvex data and complementarity conditions, i.e., MPECs/MPCCs. Specifically, we leverage on the sequential convex approximation (SCA) to design a two-layer, semidecentralized algorithm suitable to iteratively compute a local solution to the Stackelberg equilibrium problem involving a single leader and multiple followers in an aggregative form with coupling constraints. The main contributions of this article are summarized as follows.

1) We reformulate the Stackelberg game as an MPCC by embedding it into the leader nonconvex optimization problem the equivalent

Manuscript received December 5, 2019; revised June 6, 2020 and October 26, 2020; accepted May 1, 2021. Date of publication May 6, 2021; date of current version January 28, 2022. This work was supported by Innovate UK, part of UK Research and Innovation, under Project LEO 104781. Recommended by Associate Editor R. M. Jungers.
(Corresponding author: Filippo Fabiani.)

Filippo Fabiani is with the Department of Engineering Science, University of Oxford, OX1 3PJ Oxford, U.K. (e-mail: filippo.fabiani@eng.ox.ac.uk).

Mohammad Amin Tajeddini and Hamed Kebriaei are with the School of Electrical and Computer Engineering, College of Engineering, University of Tehran, Tehran 14395-515, Iran (e-mail: a.tajeddini@ut.ac.ir; kebriaei@ut.ac.ir).

Sergio Grammatico is with the Delft Center for Systems and Control, TU Delft, 2628 CD Delft, The Netherlands (e-mail: grammatico.css@gmail.com).
Digital Object Identifier 10.1109/TAC.2021.3077874

KKT conditions to compute a generalized variational Nash equilibrium (v-GNE) [23] for the followers' game (see Section II).

2) We exploit a key result provided in [24] to locally relax the complementarity constraints, obtaining the MPCC-LICQ [25, Def. 3.1], i.e., the linear independent constraint qualification (LICQ) of all the points inside a certain neighborhood of the originally formulated MPCC (see Section III).
3) Along the same lines of the work in [26] and [27], we propose to convexify the relaxed MPCC at every iteration of the outer loop, whose optimal solution, computed within the inner loop, points a descent direction for the cost function of the original MPCC. By pursuing such a descent direction, the sequence of feasible points generated by the outer loop directly leads to a local solution of the Stackelberg equilibrium problem (see Section III).
4) We analyze the performance of the proposed algorithm applied to a numerical instance of the charging coordination problem for a fleet of Plug-in Electric Vehicles (PEVs), also investigating the behavior of the leader and the followers as the regularization parameter varies (see Section IV).

To the best of our knowledge, the proposed two-layer algorithm represents the first attempt to compute a local solution to the Stackelberg equilibrium problem involving nonconvex data and equilibrium constraints by directly exploiting (and preserving) the hierarchical, multiagent structure of the original aggregative game.

Notation: $\mathbb{N}$, $\mathbb{R}$, and $\mathbb{R}_{\geq 0}$ denote the set of natural, real, and nonnegative real numbers, respectively. $\mathbf{1}$ represents a vector with all elements equal to 1. For vectors $v_1, \ldots, v_N \in \mathbb{R}^n$ and $\mathcal{I} = \{1, \ldots, N\}$, we denote $\boldsymbol{v} := (v_1^\top, \ldots, v_N^\top)^\top = \mathrm{col}((v_i)_{i\in\mathcal{I}})$ and $\boldsymbol{v}_{-i} := \mathrm{col}((v_j)_{j\in\mathcal{I}\setminus\{i\}})$. We also use $\boldsymbol{v} = (v_i, \boldsymbol{v}_{-i})$. $v \perp w$ means that v and w are orthogonal vectors. Given a matrix $A \in \mathbb{R}^{m\times n}$, $A^\top$ denotes its transpose. $A \otimes B$ represents the Kronecker product between the matrices A and B. For a function $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $f(v; \bar{v})$ denotes the approximation of f at some $\bar{v}$. For a set-valued mapping $\mathcal{F} : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, $\mathrm{gph}(\mathcal{F}) := \{(y, x) \in \mathbb{R}^n \times \mathbb{R}^m \mid x \in \mathcal{F}(y)\}$ denotes its graph.

II. Mathematical Setup

A. Stackelberg Game

We consider a hierarchical noncooperative game with one leader, controlling its decision variable $y_0 \in \mathbb{R}^{n_0}$, and N followers, indexed by the set $\mathcal{I} := \{1, \ldots, N\}$, where each follower $i \in \mathcal{I}$ controls its own variable $x_i \in \mathcal{X}_i := \{x_i \in \mathbb{R}^{n_i} \mid F_i x_i \leq g_i\}$, $F_i \in \mathbb{R}^{p_i\times n_i}$, $g_i \in \mathbb{R}^{p_i}$, and aims at solving the following optimization problem:

$$\forall i \in \mathcal{I} : \begin{cases} \min\limits_{x_i\in\mathcal{X}_i} & J_i(y_0, x_i, \boldsymbol{x}_{-i}) \\ \text{s.t.} & A_i x_i + \sum_{j\in\mathcal{I}\setminus\{i\}} A_j x_j \leq b \end{cases} \tag{1}$$

for some cost function $J_i : \mathbb{R}^{n_0} \times \mathbb{R}^n \to \mathbb{R}$. Let $\boldsymbol{x} := \mathrm{col}((x_i)_{i\in\mathcal{I}}) \in \mathbb{R}^n$, $n = \sum_{i\in\mathcal{I}} n_i$, be the collective vector of strategies of the followers, whereas $\boldsymbol{x}_{-i} \in \mathbb{R}^{n-n_i}$ stacks all the local decision variables except the ith one. We postulate the following standard assumptions on the followers' data in (1).

Standing Assumption 1: For each $i \in \mathcal{I}$, the function $J_i(y_0, \cdot)$ is convex and continuously differentiable, for fixed y_0. □

Standing Assumption 2: For each $i \in \mathcal{I}$, $\mathrm{rank}(F_i) = p_i$. □

In (1), each matrix $A_i \in \mathbb{R}^{m\times n_i}$ stacks m linear coupling constraints, whereas $b \in \mathbb{R}^m$ is the vector of shared resources among the followers. Let $A := [A_1 \ldots A_N] \in \mathbb{R}^{m\times n}$. Then, we preliminary define the sets $\mathcal{X} := \prod_{i\in\mathcal{I}} \mathcal{X}_i$ and $\Theta := \{\boldsymbol{x} \in \mathcal{X} \mid A\boldsymbol{x} \leq b\}$.

For a fixed strategy of the leader y_0, the followers aim to solve a generalized Nash equilibrium problem (GNEP). Specifically, by focusing on v-GNEs, such problem is equivalent to solve $\mathrm{VI}(\Theta, H(y_0, \cdot))$ [23], where, in view of Standing Assumption 1, $H : \mathbb{R}^{n_0} \times \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a continuously differentiable set-valued mapping defined as $H(y_0, \boldsymbol{x}) := \mathrm{col}((\nabla_{x_i} J_i(y_0, \boldsymbol{x}))_{i\in\mathcal{I}})$. This fact, along with the properties of Θ, guarantees the nonemptiness of the set of v-GNE that, for any $y_0 \in \mathcal{Y}_0$, corresponds to the set

$$\mathcal{S}(y_0) := \{\boldsymbol{x} \in \Theta \mid (\boldsymbol{z} - \boldsymbol{x})^\top H(y_0, \boldsymbol{x}) \geq 0 \quad \forall \boldsymbol{z} \in \Theta\}. \tag{2}$$

On the other hand, the optimization problem of the leader reads as

$$\begin{cases} \min\limits_{y_0,\boldsymbol{x}} & J_0(y_0, \boldsymbol{x}) \\ \text{s.t.} & (y_0, \boldsymbol{x}) \in \mathrm{gph}(\mathcal{S}) \cap (\mathcal{Y}_0 \times \mathbb{R}^n) \end{cases} \tag{3}$$

for some cost function $J_0 : \mathbb{R}^{n_0} \times \mathbb{R}^n \to \mathbb{R}$ and local constraint set $\mathcal{Y}_0$ characterized by the following standard conditions.

Standing Assumption 3: The set $\mathcal{Y}_0$ is nonempty, closed, and convex. □

Standing Assumption 4: The function J_0 is coercive, its gradient ∇J_0 is Lipschitz continuous on $\Phi := \mathcal{Y}_0 \times \mathcal{X}$ with constant κ_0. □

We note that (3) defines an MPEC where $\boldsymbol{x}$ is not strictly within the leader's control, but it corresponds to an optimistic conjecture [18]. In view of [8, Th. 1.4.1], the MPEC in (3) admits an optimal solution, since the coerciveness of J_0 implies compactness of its level sets, and the feasible set, $\mathrm{gph}(\mathcal{S}) \cap (\mathcal{Y}_0 \times \mathbb{R}^n)$, is closed under the postulated assumptions. Therefore, this ensures existence of a solution to the hierarchical game, according to the following notion of local generalized Stackelberg equilibrium, inspired by Kulkarni and Shanbhag [18] and Hu and Ralph [28].

Definition 1: A pair $(y_0^*, \boldsymbol{x}^*) \in \mathrm{gph}(\mathcal{S}) \cap (\mathcal{Y}_0 \times \mathbb{R}^n)$, with $\mathcal{S}$ as in (2), is a local Stackelberg equilibrium (ℓ-SE) of the hierarchical game in (1)–(3) if there exist open neighborhoods $\mathcal{O}_{y_0^*}$ and $\mathcal{O}_{\boldsymbol{x}^*}$ of y_0^* and $\boldsymbol{x}^*$, respectively, such that

$$J_0(y_0^*, \boldsymbol{x}^*) \leq \inf_{(y_0,\boldsymbol{x})\in\mathrm{gph}(\mathcal{S})\cap\mathcal{O}} J_0(y_0, \boldsymbol{x})$$

where $\mathcal{O} := (\mathcal{Y}_0 \cap \mathcal{O}_{y_0^*}) \times \mathcal{O}_{\boldsymbol{x}^*}$. □

Informally speaking, at an ℓ-SE, the leader and the followers locally fulfill the set of mutually coupling constraints and none of them can gain by unilaterally deviating from their current strategy. Note that we refer to an SE if Definition 1 holds true with $\mathcal{O} = \mathcal{Y}_0 \times \mathbb{R}^n$, i.e., $\mathcal{O}_{y_0^*} = \mathbb{R}^{n_0}$ and $\mathcal{O}_{\boldsymbol{x}^*} = \mathbb{R}^n$, thus coinciding with [18, Def. 1.1].

B. Aggregative Game Formulation

For computational purposes, we consider the cost function of the followers and leader to be in aggregative form, i.e.,

$$J_i := \tfrac{1}{2} x_i^\top Q_i x_i + \left(\tfrac{1}{N} \sum_{j\in\mathcal{I}} C_{i,j} x_j + C_{i,0} y_0\right)^\top x_i \quad \forall i \in \mathcal{I}$$

$$J_0 := f_0(y_0) + \left(\sum_{i\in\mathcal{I}} f_{0,i}(x_i)\right)^\top y_0 \tag{4}$$

where $Q_i \succcurlyeq 0$, $C_{i,j} \in \mathbb{R}^{n_i\times n_j}$, and $C_{i,0} \in \mathbb{R}^{n_i\times n_0}$. In view of Standing Assumption 1, given any feasible $y_0 \in \mathcal{Y}_0$, it follows from [29, Th. 3.1] that a set of strategies is a v-GNE of the followers game in (1) if and only if the following coupled KKT conditions hold true:

$$\begin{cases} \nabla_{x_i} J_i(y_0, x_i, \boldsymbol{x}_{-i}) + A_i^\top \lambda + F_i^\top \lambda_i = 0 \quad \forall i \in \mathcal{I} \\ 0 \leq \lambda \perp -(A\boldsymbol{x} - b) \geq 0 \\ 0 \leq \lambda_i \perp -(F_i x_i - g_i) \geq 0 \quad \forall i \in \mathcal{I} \end{cases}$$

which, in our aggregative setup, can be compactly rewritten as

$$\begin{cases} Q\boldsymbol{x} + C y_0 + A^\top \lambda + F^\top \boldsymbol{\lambda} = 0 \\ 0 \leq \lambda \perp -(A\boldsymbol{x} - b) \geq 0 \\ 0 \leq \lambda_i \perp -(F_i x_i - g_i) \geq 0 \quad \forall i \in \mathcal{I} \end{cases} \tag{5}$$

where $F := \mathrm{diag}((F_i)_{i\in\mathcal{I}})$, $\lambda \in \mathbb{R}^m_{\geq 0}$ is the dual variable associated with $A\boldsymbol{x} \leq b$, $\lambda_i \in \mathbb{R}^{p_i}_{\geq 0}$ is the (local) dual variable associated with the local constraints defining $\mathcal{X}_i$, $\boldsymbol{\lambda} := \mathrm{col}((\lambda_i)_{i\in\mathcal{I}})$, and

$$Q := \begin{bmatrix} Q_1 + \frac{1}{N} C_{1,1} & \cdots & \frac{1}{N} C_{1,N} \\ \vdots & \ddots & \vdots \\ \frac{1}{N} C_{N,1} & \cdots & Q_N + \frac{1}{N} C_{N,N} \end{bmatrix}, \quad C := \begin{bmatrix} C_{10} \\ \vdots \\ C_{N0} \end{bmatrix}.$$

Finally, by substituting back the KKT conditions in (5) into the optimization problem of the leader in (3), the problem of finding an SE of the hierarchical game in (1)–(3) can be equivalently written as

$$\begin{cases} \min\limits_{y_0,\boldsymbol{x},\lambda,\boldsymbol{\lambda}} & J_0(y_0,\boldsymbol{x}) \\ \text{s.t.} & Q\boldsymbol{x}+Cy_0+A^\top\lambda+F^\top\boldsymbol{\lambda}=0 \\ & 0\le\lambda_i\perp-(F_ix_i-g_i)\ge 0 \quad \forall i\in\mathcal{I} \\ & 0\le\lambda\perp-(A\boldsymbol{x}-b)\ge 0,\ y_0\in\mathcal{Y}_0. \end{cases} \tag{6}$$

C. Complementarity Constraints Relaxation

We note that the leader nonconvex optimization problem in (6) is an MPCC and, in general, it does not satisfy any standard constraint qualification. Therefore, we propose to study a regularized version by introducing slack variables $\mu\in\mathbb{R}^m_{\ge 0}$ and $\mu_i\in\mathbb{R}^{p_i}_{\ge 0}$, $i\in\mathcal{I}$, together with parameters θ, $\theta_i>0$, $i\in\mathcal{I}$, which enable us to replace the complementarity constraints in (6) with the nonlinear constraints $\lambda^\top\mu\le\theta$ and $\lambda_i^\top\mu_i\le\theta_i$, for all $i\in\mathcal{I}$ [24]. Thus, after defining $\nu:=\mathrm{col}(\lambda,\mu)\in\mathbb{R}^{2m}$, $\nu_i:=\mathrm{col}(\lambda_i,\mu_i)\in\mathbb{R}^{2p_i}$, $\boldsymbol{y}:=\mathrm{col}(\boldsymbol{x},(\nu_i)_{i\in\mathcal{I}})$, the regularized version of (6) reads as

$$R(\theta):\begin{cases} \min\limits_{y_0,\boldsymbol{y},\nu} & J_0(y_0,\boldsymbol{x}) \\ \text{s.t.} & A_{\mathrm{f}}\,\boldsymbol{y}+A_\ell\,y_0+A_{\mathrm{c}}\,\nu=d \\ & \lambda_i^\top\mu_i\le\theta_i,\lambda_i,\mu_i\ge 0 \quad \forall i\in\mathcal{I} \\ & \lambda^\top\mu\le\theta,\lambda,\mu\ge 0,\ y_0\in\mathcal{Y}_0 \end{cases} \tag{7}$$

where $d:=\mathrm{col}(0,b,g)$, $g:=\mathrm{col}((g_i)_{i\in\mathcal{I}})$, $A_\ell:=\mathrm{col}(C,0,0)$, and

$$A_{\mathrm{f}}:=\begin{bmatrix} Q & ([F_i^\top\ \ 0])_{i\in\mathcal{I}} \\ A & 0 \\ F & [0\ \ I]\otimes\mathbf{1} \end{bmatrix} \qquad A_{\mathrm{c}}:=\begin{bmatrix} A^\top & 0 \\ 0 & I \\ 0 & 0 \end{bmatrix}.$$

For any given θ, $\theta_i>0$, $i\in\mathcal{I}$, let us now introduce the sets

$$\mathcal{C}(\theta):=\{\nu\in\mathbb{R}^{2m}_{\ge 0}\mid\tfrac{1}{2}\nu^\top P\nu\le\theta\}$$

$$\mathcal{C}_i(\theta_i):=\{\nu_i\in\mathbb{R}^{2p_i}_{\ge 0}\mid\tfrac{1}{2}\nu_i^\top P_i\nu_i\le\theta_i\} \quad \forall i\in\mathcal{I}. \tag{8}$$

Here, each P and P_i, $i\in\mathcal{I}$, is a symmetric matrix with identities of suitable dimension on the antidiagonal. Furthermore, we define $\Omega(\theta):=\mathcal{Y}_0\times\mathcal{Y}\times\mathcal{C}(\theta)$, where, for brevity, we omit the dependence from θ_i, explicated in $\mathcal{Y}:=\mathcal{X}\times\prod_{i\in\mathcal{I}}\mathcal{C}_i(\theta_i)$. Finally, by introducing $\boldsymbol{\omega}:=\mathrm{col}(y_0,\boldsymbol{y},\nu)$ and $A_\omega:=[A_\ell\ A_{\mathrm{f}}\ A_{\mathrm{c}}]$, the closed, nonconvex feasible set of $R(\theta)$ in (7) reads as

$$\mathcal{R}(\theta):=\{\boldsymbol{\omega}\in\Omega(\theta)\mid A_\omega\,\boldsymbol{\omega}-d=0\}. \tag{9}$$

We recall now the notion of MPCC-LICQ for the MPCC in (6), which is characterized by the result stated immediately below.

Definition 2: The MPCC in (6) satisfies the MPCC-LICQ at $\tilde{\boldsymbol{\omega}}\in\mathcal{R}(0)$ if $R(0)$ in (7) satisfies the LICQ at $\tilde{\boldsymbol{\omega}}$. □

Lemma 1: (see [24, Lemma 2.1]) Let $\tilde{\boldsymbol{\omega}}\in\mathcal{R}(0)$. If $\tilde{\boldsymbol{\omega}}$ satisfies the MPCC-LICQ for the MPCC in (6), then there exists an open neighborhood $\mathcal{O}$ of $\tilde{\boldsymbol{\omega}}$ and scalars $\tilde{\theta}$, $\tilde{\theta}_i>0$, for all $i\in\mathcal{I}$, such that, for every $\theta\in(0,\tilde{\theta})$ and $\theta_i\in(0,\tilde{\theta}_i)$, for all $i\in\mathcal{I}$, the LICQ holds true at every point $\boldsymbol{\omega}\in\mathcal{O}$ of $R(\theta)$. □

Then, let us introduce the following fundamental assumption.

Standing Assumption 5: There exists some $\tilde{\boldsymbol{\omega}}\in\mathcal{R}(0)$ that satisfies the MPCC-LICQ for the MPCC in (6). The regularization parameters are chosen so that $\theta\in(0,\tilde{\theta})$ and $\theta_i\in(0,\tilde{\theta}_i)$, for all $i\in\mathcal{I}$. □

In view of Standing Assumption 5, there exists a neighborhood such that $R(\theta)$ locally satisfies the LICQ. As shown in Section IV-B, the coefficients θ, θ_i, $i\in\mathcal{I}$, play a tradeoff role between the distance from a v-GNE for the followers and a lower cost for the leader. To conclude the section, we stress that an optimal solution to (7), whose existence follows by its local LICQ and the coerciveness of J_0, generates a pair $(y_0^*,\boldsymbol{x}^*)$ that corresponds to an ℓ-SE of the original hierarchical game in (1)–(3).

III. Local Stackelberg Equilibrium Seeking via Sequential Convex Approximation

A. Two-Layer Algorithm

In the spirit of the work in [26] and [27], we then investigate how to solve (7) in a decentralized fashion by means of a two-layer algorithm while preserving the hierarchical structure of the game (1)–(3). First, we linearize the nonlinear terms appearing in the cost function around some $\bar{\boldsymbol{\omega}}\in\mathcal{R}(\theta)$. Specifically, with $\boldsymbol{\varphi}:=(y_0,\boldsymbol{x})$, J_0 is linearized by following a first-order Taylor expansion as $J_0(\boldsymbol{\varphi})\simeq J_0(\bar{\boldsymbol{\varphi}})+\nabla J_0(\bar{\boldsymbol{\varphi}})^\top(\boldsymbol{\varphi}-\bar{\boldsymbol{\varphi}})$ where, for our aggregative game, we have

$$\nabla J_0(\boldsymbol{\varphi})=\mathrm{col}(\nabla_{y_0}f_0(y_0)+\sum_{j\in\mathcal{I}}f_{0,j}(x_j),(\nabla_{x_j}f_{0,j}(x_j)^\top y_0)_{j\in\mathcal{I}})$$

$$:=\mathrm{col}(c_\ell(y_0,\boldsymbol{x}),c_{\mathrm{f}}(y_0,\boldsymbol{x})).$$

According to [27, Sec. III.A], for the nonlinear constraints defining the sets in (8), we compute an upper approximation by observing that, e.g., $\frac{1}{2}\nu^\top P\nu=\lambda^\top\mu=\frac{1}{2}(\lambda+\mu)^\top(\lambda+\mu)-\frac{1}{2}(\lambda^\top\lambda+\mu^\top\mu)$. Thus, after linearizing the concave term around some $\bar{\nu}\in\mathcal{C}(\theta)$, we define

$$\tilde{\mathcal{C}}(\theta;\bar{\boldsymbol{\omega}}):=\{\nu\in\mathbb{R}^{2m}_{\ge 0}\mid\tfrac{1}{2}([I\ I]\,\nu)^\top([I\ I]\,\nu)-\bar{\nu}^\top\nu+\tfrac{1}{2}\bar{\nu}^\top\bar{\nu}\le\theta\}.$$

The same procedure can be applied to each $\mathcal{C}_i(\theta_i)$ to obtain $\tilde{\mathcal{C}}_i(\theta_i;\bar{\boldsymbol{\omega}})$. Accordingly, $\Omega(\theta)$ is approximated by $\tilde{\Omega}(\theta;\bar{\boldsymbol{\omega}}):=\mathcal{Y}_0\times\tilde{\mathcal{Y}}(\bar{\boldsymbol{\omega}})\times\tilde{\mathcal{C}}(\theta;\bar{\boldsymbol{\omega}})$, with $\tilde{\mathcal{Y}}(\bar{\boldsymbol{\omega}}):=\mathcal{X}\times\prod_{i\in\mathcal{I}}\tilde{\mathcal{C}}_i(\theta_i;\bar{\boldsymbol{\omega}})$ while $\mathcal{R}(\theta)$ by

$$\tilde{\mathcal{R}}(\theta;\bar{\boldsymbol{\omega}}):=\{\boldsymbol{\omega}\in\tilde{\Omega}(\theta;\bar{\boldsymbol{\omega}})\mid A_\omega\,\boldsymbol{\omega}-d=0\}. \tag{10}$$

Finally, by discarding constant terms and introducing $c_\omega(\bar{\boldsymbol{\omega}}):=\mathrm{col}(\nabla J_0(\bar{\boldsymbol{\varphi}}),0)$, the convexified version of $R(\theta)$ in (7) reads as

$$\tilde{R}\,(\theta;\bar{\boldsymbol{\omega}}):\begin{cases} \min\limits_{\boldsymbol{\omega}\in\tilde{\Omega}(\theta;\bar{\boldsymbol{\omega}})} & c_\omega(\bar{\boldsymbol{\omega}})^\top\boldsymbol{\omega}+\dfrac{\sigma}{2}\|\boldsymbol{\omega}-\bar{\boldsymbol{\omega}}\|^2 \\ \text{s.t.} & A_\omega\,\boldsymbol{\omega}=d \end{cases} \tag{11}$$

where we add a "proximal-like" term to the linearized cost function in (7) with $\sigma>0$. Hence, the cost function in (11), namely $\tilde{J}_0(\boldsymbol{\omega};\bar{\boldsymbol{\omega}}):=c_\omega(\bar{\boldsymbol{\omega}})^\top\boldsymbol{\omega}+\frac{\sigma}{2}\|\boldsymbol{\omega}-\bar{\boldsymbol{\omega}}\|^2$, is characterized as follows.

Lemma 2: The following statements hold true.

i) Given any $\bar{\boldsymbol{\omega}}\in\mathcal{R}(\theta)$, $\tilde{J}_0(\cdot;\bar{\boldsymbol{\omega}})$ is uniformly strongly convex on $\Phi\times\mathbb{R}^{2(m+p)}_{\ge 0}$, $p:=\sum_{i\in\mathcal{I}}p_i$, with coefficient σ.

ii) Given any $\boldsymbol{\omega}\in\mathcal{R}(\theta)$, $\nabla\tilde{J}_0(\boldsymbol{\omega};\cdot)$ is uniformly Lipschitz continuous on $\mathcal{R}(\theta)$ with coefficient $\tilde{\kappa}_0:=\kappa_0+\sigma$. □

Proof: (i) The statement directly follows by applying the definition of uniform strong convexity on the set $\Phi\times\mathbb{R}^{2(m+p)}_{\ge 0}$.

(ii) Let $\boldsymbol{\omega}_1,\boldsymbol{\omega}_2\in\mathcal{R}(\theta)$. For any given $\boldsymbol{\omega}\in\mathcal{R}(\theta)$, we have

$$\begin{aligned}\|\nabla\tilde{J}_0(\boldsymbol{\omega};\boldsymbol{\omega}_1)-\nabla\tilde{J}_0(\boldsymbol{\omega};\boldsymbol{\omega}_2)\|&=\|c_\omega(\boldsymbol{\omega}_1)-c_\omega(\boldsymbol{\omega}_2)+\sigma(\boldsymbol{\omega}_2-\boldsymbol{\omega}_1)\|\\&\le\|\mathrm{col}(\nabla J_0(\boldsymbol{\varphi}_1),0)-\mathrm{col}(\nabla J_0(\boldsymbol{\varphi}_2),0)\|+\sigma\|\boldsymbol{\omega}_1-\boldsymbol{\omega}_2\|\\&\le(\kappa_0+\sigma)\,\|\boldsymbol{\omega}_1-\boldsymbol{\omega}_2\|.\end{aligned}$$

■

Remark 1: According to the structure of the vector $\boldsymbol{\omega}$, the coefficient σ may be replaced with locally defined σ_0, σ_{c}, $\sigma_i>0$, $i\in\mathcal{I}$, without affecting the results given in the remainder, see [26, Sec. III.A]. For simplicity, we adopt a unique, globally known parameter σ. □

Thus, given any $\bar{\boldsymbol{\omega}}\in\mathcal{R}(\theta)$, $\tilde{R}(\theta;\bar{\boldsymbol{\omega}})$ in (11) admits a unique optimal solution associated with the mapping $\hat{\boldsymbol{\omega}}:\mathbb{R}^s\to\mathbb{R}^s$, with $s:=n_0+n+2(p+m)$, defined as follows:

$$\hat{\boldsymbol{\omega}}(\bar{\boldsymbol{\omega}}):=\underset{\boldsymbol{\omega}\in\tilde{\mathcal{R}}(\theta;\bar{\boldsymbol{\omega}})}{\mathrm{argmin}}\ \tilde{J}_0(\boldsymbol{\omega};\bar{\boldsymbol{\omega}}). \tag{12}$$

For computing an ℓ-SE, we propose the iterative procedure summarized in Algorithm 1, which is composed of two main loops and

Algorithm 1: Two-Layer SCA Computation of ℓ-SE.

Initialization: $\boldsymbol{\omega}^0 \in \mathcal{R}(\theta)$, $\alpha > 0$
Iteration $(k \in \mathbb{N})$**:**
(S1) Convexify $R(\theta)$ to obtain $\tilde{R}(\theta; \boldsymbol{\omega}^k)$ as in (11)
(S2) Compute $\hat{\boldsymbol{\omega}}^k$, solution to $\tilde{R}(\theta; \boldsymbol{\omega}^k)$
(S3) Update $\boldsymbol{\omega}^{k+1} = (1-\alpha)\boldsymbol{\omega}^k + \alpha \hat{\boldsymbol{\omega}}^k$

resorts on the so-called SCA method. Specifically, once fixed the coefficients θ, $\theta_i > 0$, for all $i \in \mathcal{I}$, at each iteration $k \in \mathbb{N}$, the outer loop is in charge of providing a feasible set of strategies $\boldsymbol{\omega}^k$, which are used to convexify $R(\theta)$ (S1). Then, after solving the inner loop by computing the optimal solution $\hat{\boldsymbol{\omega}}^k := \hat{\boldsymbol{\omega}}(\boldsymbol{\omega}^k)$ to $\tilde{R}(\theta; \boldsymbol{\omega}^k)$ (S2), the outer loop updates the strategies $\boldsymbol{\omega}^{k+1}$ (S3) to find a new approximation $\tilde{R}(\theta; \boldsymbol{\omega}^{k+1})$, and the procedure repeats until a certain stopping criterion is met.

B. Convergence Analysis

First, we characterize the sequence $(\boldsymbol{\omega}^k)_{k\in\mathbb{N}}$ generated by Algorithm 1 in terms of iterate feasibility. Then, we establish a key property of the mapping $\hat{\boldsymbol{\omega}}(\cdot)$, and finally, we prove that $(\boldsymbol{\omega}^k)_{k\in\mathbb{N}}$ converges to an optimal solution to (7), generating an ℓ-SE of the hierarchical aggregative game (1)–(3), according to Definition 1.

Lemma 3: The following inclusions hold true.

i) $\tilde{\mathcal{R}}(\theta; \bar{\boldsymbol{\omega}}) \subseteq \mathcal{R}(\theta)$, for all $\bar{\boldsymbol{\omega}} \in \mathcal{R}(\theta)$.

ii) $\boldsymbol{\omega}^k \in \mathcal{R}(\theta)$. □

Proof:

i) The upper approximation of the nonlinear constraints, which holds true for all $\bar{\boldsymbol{\omega}} \in \mathcal{R}(\theta)$, implies $\tilde{\mathcal{C}}(\theta; \bar{\boldsymbol{\omega}}) \subseteq \mathcal{C}(\theta)$ and $\tilde{\mathcal{C}}_i(\theta_i; \bar{\boldsymbol{\omega}}) \subseteq \mathcal{C}_i(\theta_i)$, $i \in \mathcal{I}$. Therefore, $\tilde{\Omega}(\theta; \bar{\boldsymbol{\omega}}) \subseteq \Omega(\theta)$, and in view of the definitions in (9) and (10), inclusion (i) can be deduced.

ii) First, in view of the approximation of the constraints, note that $\boldsymbol{\omega}^k \in \tilde{\mathcal{R}}(\theta; \boldsymbol{\omega}^k)$, for all $k \in \mathbb{N}$, with $\tilde{\mathcal{R}}(\theta; \boldsymbol{\omega}^k)$ convex subset of $\mathcal{R}(\theta)$. Then, the proof follows by induction by considering that $\boldsymbol{\omega}^{k+1}$ is a convex combination of $\hat{\boldsymbol{\omega}}^k \in \tilde{\mathcal{R}}(\theta; \boldsymbol{\omega}^k)$ and $\boldsymbol{\omega}^k$. ■

Lemma 4: For every $\bar{\boldsymbol{\omega}} \in \mathcal{R}(\theta)$, the vector $(\hat{\boldsymbol{\varphi}}(\bar{\boldsymbol{\omega}}) - \bar{\boldsymbol{\varphi}})$ is a descent direction for $J_0(\boldsymbol{\varphi})$ in $R(\theta)$, evaluated at $\bar{\boldsymbol{\varphi}}$, i.e., $(\bar{\boldsymbol{\varphi}} - \hat{\boldsymbol{\varphi}}(\bar{\boldsymbol{\omega}}))^\top \nabla J_0(\bar{\boldsymbol{\varphi}}) \geq \sigma \|\bar{\boldsymbol{\omega}} - \hat{\boldsymbol{\omega}}(\bar{\boldsymbol{\omega}})\|^2 > 0$. □

Proof: Given any $\bar{\boldsymbol{\omega}} \in \mathcal{R}(\theta)$, by definition, $\hat{\boldsymbol{\omega}}(\bar{\boldsymbol{\omega}})$ satisfies the minimum principle for (11), i.e., $(\boldsymbol{\zeta} - \hat{\boldsymbol{\omega}}(\bar{\boldsymbol{\omega}}))^\top \nabla \tilde{J}_0(\hat{\boldsymbol{\omega}}(\bar{\boldsymbol{\omega}}); \bar{\boldsymbol{\omega}}) \geq 0$ for all $\boldsymbol{\zeta} \in \tilde{\mathcal{R}}(\theta; \bar{\boldsymbol{\omega}})$. From Lemma 3(ii), we choose $\boldsymbol{\zeta} = \bar{\boldsymbol{\omega}}$, and by adding and subtracting the term $(\bar{\boldsymbol{\omega}} - \hat{\boldsymbol{\omega}}(\bar{\boldsymbol{\omega}}))^\top \nabla \tilde{J}_0(\bar{\boldsymbol{\omega}}; \bar{\boldsymbol{\omega}})$, we obtain

$$(\bar{\boldsymbol{\omega}} - \hat{\boldsymbol{\omega}}(\bar{\boldsymbol{\omega}}))^\top \nabla \tilde{J}_0(\bar{\boldsymbol{\omega}}; \bar{\boldsymbol{\omega}}) \geq (\bar{\boldsymbol{\omega}} - \hat{\boldsymbol{\omega}}(\bar{\boldsymbol{\omega}}))^\top (\nabla \tilde{J}_0(\bar{\boldsymbol{\omega}}; \bar{\boldsymbol{\omega}}) - \nabla \tilde{J}_0(\hat{\boldsymbol{\omega}}(\bar{\boldsymbol{\omega}}); \bar{\boldsymbol{\omega}})).$$

By directly replacing $\nabla \tilde{J}_0(\bar{\boldsymbol{\omega}}; \bar{\boldsymbol{\omega}})$ with $c_\omega(\bar{\boldsymbol{\omega}}) = \mathrm{col}(\nabla J_0(\bar{\boldsymbol{\varphi}}), 0)$, the term on the left-hand side is equal to $(\bar{\boldsymbol{\varphi}} - \hat{\boldsymbol{\varphi}}(\bar{\boldsymbol{\omega}}))^\top \nabla J_0(\bar{\boldsymbol{\varphi}})$, whereas the one on the right-hand side, in view of Lemma 2(i), is bounded from below by $\sigma \|\bar{\boldsymbol{\omega}} - \hat{\boldsymbol{\omega}}(\bar{\boldsymbol{\omega}})\|^2$, leading to

$$(\bar{\boldsymbol{\varphi}} - \hat{\boldsymbol{\varphi}}(\bar{\boldsymbol{\omega}}))^\top \nabla J_0(\bar{\boldsymbol{\varphi}}) \geq \sigma \|\bar{\boldsymbol{\omega}} - \hat{\boldsymbol{\omega}}(\bar{\boldsymbol{\omega}})\|^2.$$

■

Before establishing the convergence to an ℓ-SE for the sequence generated by Algorithm 1, we recall a key result provided in [27].

Lemma 5: (see [27, Th. 14]) Let $(\boldsymbol{\omega}^k)_{k\in\mathbb{N}}$ be the sequence generated by Algorithm 1 and assume that $\lim_{k\to\infty} \|\hat{\boldsymbol{\omega}}(\boldsymbol{\omega}^k) - \boldsymbol{\omega}^k\| = 0$. Then, every limit point of $(\boldsymbol{\omega}^k)_{k\in\mathbb{N}}$ generated by Algorithm 1 is a stationary solution to $R(\theta)$. □

Theorem 1: Let α in Algorithm 1 be chosen so that $\alpha \in (0, 2\sigma/\kappa_0)$. Then, the sequence $(\boldsymbol{\omega}^k)_{k\in\mathbb{N}}$ generated by Algorithm 1 converges to an optimal solution $\boldsymbol{\omega}^*$ to $R(\theta)$ in (7), where subvector $(y_0^*, \boldsymbol{x}^*)$ is an ℓ-SE of the hierarchical game in (1)–(3). □

Algorithm 2: ADAL for (S2) of Algorithm 1.

Initialization: $\eta(0) \in \mathbb{R}^s$, τ, $\rho > 0$
Iteration $(t \in \mathbb{N})$**:**
Leader:
$$\begin{cases} y_0^\star(t) = \underset{y_0 \in \mathcal{Y}_0}{\operatorname{argmin}}\ \hat{\mathcal{L}}_\ell^k(y_0, \eta(t), z_{\mathrm{f}}(t), z_{\mathrm{c}}(t)) \\ z_\ell(t+1) = z_\ell(t) + \tau(A_\ell\, y_0^\star(t) - z_\ell(t)) \end{cases}$$
Followers:
$$\begin{cases} \boldsymbol{y}^\star(t) = \underset{\boldsymbol{y} \in \tilde{\mathcal{Y}}^k}{\operatorname{argmin}}\ \hat{\mathcal{L}}_{\mathrm{f}}^k(\boldsymbol{y}, \eta(t), z_\ell(t), z_{\mathrm{c}}(t)) \\ z_{\mathrm{f}}(t+1) = z_{\mathrm{f}}(t) + \tau(A_{\mathrm{f}}\, \boldsymbol{y}^\star(t) - z_{\mathrm{f}}(t)) \end{cases}$$
Coordinator:
$$\begin{cases} \nu^\star(t) = \underset{\nu \in \tilde{\mathcal{C}}^k(\theta)}{\operatorname{argmin}}\ \hat{\mathcal{L}}_{\mathrm{c}}^k(\nu, \eta(t), z_{\mathrm{f}}(t), z_\ell(t)) \\ z_{\mathrm{c}}(t+1) = z_{\mathrm{c}}(t) + \tau(A_{\mathrm{c}}\, \nu^\star(t) - z_{\mathrm{c}}(t)) \end{cases}$$
$$\eta(t+1) = \eta(t) + \rho\tau\,(z_{\mathrm{f}}(t+1) + z_\ell(t+1) + z_{\mathrm{c}}(t+1) - d)$$

Proof: By combining the descent lemma [30, Prop. A.24] and Lemma 4, the step (S3) in Algorithm 1 leads to

$$\begin{aligned} J_0(\boldsymbol{\varphi}^{k+1}) &\leq J_0(\boldsymbol{\varphi}^k) + \alpha \nabla^\top J_0(\boldsymbol{\varphi}^k)(\hat{\boldsymbol{\varphi}}(\boldsymbol{\omega}^k) - \boldsymbol{\varphi}^k) \\ &\quad + \alpha^2 \tfrac{\kappa_0}{2} \|\hat{\boldsymbol{\varphi}}(\boldsymbol{\omega}^k) - \boldsymbol{\varphi}^k\|^2 \\ &\leq J_0(\boldsymbol{\varphi}^k) - \alpha\left(\sigma - \alpha \tfrac{\kappa_0}{2}\right) \|\hat{\boldsymbol{\omega}}(\boldsymbol{\omega}^k) - \boldsymbol{\omega}^k\|^2 \end{aligned}$$

where the second inequality follows from $\|\hat{\boldsymbol{\omega}}(\boldsymbol{\omega}^k) - \boldsymbol{\omega}^k\| \geq \|\hat{\boldsymbol{\varphi}}(\boldsymbol{\omega}^k) - \boldsymbol{\varphi}^k\|$. If $\alpha < 2\sigma/\kappa_0$, then $(J_0(\boldsymbol{\varphi}^k))_{k\in\mathbb{N}}$ shall converge to a finite value, since $J_0(\boldsymbol{\varphi}^k) \to -\infty$ cannot happen in view of Standing Assumption 4. Thus, the convergence of $(J_0(\boldsymbol{\varphi}^k))_{k\in\mathbb{N}}$ implies $\lim_{k\to\infty} \|\hat{\boldsymbol{\omega}}(\boldsymbol{\omega}^k) - \boldsymbol{\omega}^k\| = 0$, and, therefore, the bounded sequence $(\boldsymbol{\omega}^k)_{k\in\mathbb{N}} \in \mathcal{R}(\theta)$ in view of Lemma 3, and has a limit point in $\mathcal{R}(\theta)$. From Lemma 5, such a limit point is a stationary solution to $R(\theta)$, and since $(J_0(\boldsymbol{\varphi}^k))_{k\in\mathbb{N}}$ is a strictly decreasing sequence, no limit point can be a local maximum of J_0. Thus, $(\boldsymbol{\omega}^k)_{k\in\mathbb{N}}$ converges to an optimal solution $\boldsymbol{\omega}^*$ to (7), where subvector $(y_0^*, \boldsymbol{x}^*)$ is an ℓ-SE of the original hierarchical game in (1)–(3). ■

Remark 2: If the parameters σ and κ_0 are not globally known, Theorem 1 can be equivalently restated according to a vanishing step-size rule, i.e., $\alpha = \alpha^k$ that shall be chosen so that $\alpha^k \in (0, 1]$, for all $k \in \mathbb{N}$, $\alpha^k \to 0$ and $\sum_{k\in\mathbb{N}} \alpha^k = +\infty$. □

C. Augmented Lagrangian Approach to Solve the Inner Loop

A scalable and privacy-preserving algorithm, suitable to solve (S2) in Algorithm 1 by exploiting the hierarchical structure of the original game, is the accelerated distributed augmented Lagrangian (ADAL) method proposed in [31]. Since we are interested in finding the optimal solution to $\tilde{R}(\theta; \boldsymbol{\omega}^k)$, from now on, we omit the dependence on $\boldsymbol{\omega}^k$ (unless differently specified) to alleviate the notation.

Thus, at every iteration $k \in \mathbb{N}$ of the outer loop, the Lagrangian function associated to (11) is defined as

$$\mathcal{L}^k(\boldsymbol{\omega}, \nu) = (c_\omega^k)^\top \boldsymbol{\omega} + \frac{\sigma}{2}\|\boldsymbol{\omega} - \boldsymbol{\omega}^k\|^2 + \eta^\top (A_\omega\, \boldsymbol{\omega} - d) \qquad (13)$$

where $c_\omega^k := c_\omega(\boldsymbol{\omega}^k)$, and $\nu \in \mathbb{R}^l$, $l := n + m + p$, is the dual variable associated with the linear equality constraints. We note that the Lagrangian in (13) can be rewritten as the sum of terms associated to different entities, which happens to correspond to leader, the set of followers, and a central coordinator, respectively. In details, we define $\mathcal{L}_\ell^k := (c_\ell^k)^\top y_0 + \frac{\sigma}{2}\|y_0 - y_0^k\|^2 + \eta^\top A_\ell\, y_0$, $\mathcal{L}_{\mathrm{f}}^k := (c_{\mathrm{f}}^k)^\top \boldsymbol{y} + \frac{\sigma}{2}\|\boldsymbol{y} - \boldsymbol{y}^k\|^2 + \eta^\top A_{\mathrm{f}}\, \boldsymbol{y}$, and $\mathcal{L}_{\mathrm{c}}^k := \frac{\sigma}{2}\|\nu - \nu^k\|^2 + \eta^\top A_{\mathrm{c}}\, \nu$. In the light of the work in [31], we augment each one of these terms as, e.g., $\hat{\mathcal{L}}_{\mathrm{f}}^k := \mathcal{L}_{\mathrm{f}}^k + \frac{\rho}{2}\|A_{\mathrm{f}}\,\boldsymbol{y} + A_\ell\, y_0 + A_{\mathrm{c}}\,\nu - d\|^2$ ($\hat{\mathcal{L}}_\ell^k$ and $\hat{\mathcal{L}}_{\mathrm{c}}^k$ are identical), where $\rho > 0$ is a penalty term to be designed freely.

The main steps of the proposed semidecentralized procedure are summarized in Algorithm 2, where we emphasize that each augmented Lagrangian term depends on the linearization at the current outer iteration $k \in \mathbb{N}$. Specifically, at every iteration $t \in \mathbb{N}$ of the inner loop, the ADAL requires that the followers, the leader, and the central coordinator compute in parallel a minimization step of the local augmented Lagrangian. Here, $z_\ell := A_\ell\, y_0$, $z_{\mathrm{f}} := A_{\mathrm{f}}\,\boldsymbol{y}$ and $z_{\mathrm{c}} := A_{\mathrm{c}}\,\nu$ are auxiliary variables introduced for privacy purposes and, given some $\tau > 0$, are locally updated. Finally, the central coordinator, which in some practical applications may eventually coincide with the leader, gathers $z_\ell(t+1)$ and $z_{\mathrm{f}}(t+1)$ from the leader and followers, and updates the dual variable.

Proposition 1: Let $\rho > 0$ be sufficiently large and $\tau \in (0, r_{\max}^{-1})$, where $r_{\max}$ corresponds to the maximum degree among the constraints in (10). Then, the sequence $(\boldsymbol{\omega}(t))_{t\in\mathbb{N}}$ generated by Algorithm 2 converges to the minimizer of $\tilde{R}(\theta; \boldsymbol{\omega}^k)$, for all $k \in \mathbb{N}$. □

Proof: The proof follows by noticing that $\tilde{R}(\theta; \boldsymbol{\omega}^k)$ satisfies the assumptions in [31, Th. 2], for all $k \in \mathbb{N}$. Specifically, $\tilde{\mathcal{R}}(\theta; \boldsymbol{\omega}^k)$ is a closed and convex set, $\tilde{J}_0(\boldsymbol{\omega}; \boldsymbol{\omega}^k)$ is inf-compact and each one of its terms is twice continuously differentiable. Finally, Lemma 1 provides the local LICQ for $R(\theta)$, directly inherited by $\tilde{R}(\theta; \boldsymbol{\omega}^k)$. ■

Remark 3: For simplicity, we adopt a common τ to update the auxiliary variables z_ℓ, z_{f}, and z_{c}. In principle, each entity involved within the ADAL in Algorithm 2 can locally set its own step size according to the degree of each constraint in (11), see [31, Sec. II.A]. □

IV. Numerical Case Study: Charging Coordination of Plug-In Electric Vehicles

A. Numerical Simulation Setup

We consider a set of PEVs (followers), $\mathcal{I} := \{1, 2, \ldots, N\}$, which must be charged over a certain horizon $\mathcal{T} := \{1, \ldots, T\}$. All PEVs are connected to an aggregator (leader, e.g., a retailer), which manages the energy requirements of the fleet by purchasing the electricity from the wholesale energy market. Let us define $x_i := \mathrm{col}((x_i^j)_{j\in\mathcal{T}})$, and $p := \mathrm{col}((p^j)_{j\in\mathcal{T}})$ as the amount of requested energy by the fleet and the price of energy over time, i.e., the strategy of the ith follower and of the leader, respectively. For every PEV $i \in \mathcal{I}$, we consider the cost function $J_i(p, \boldsymbol{x}) = q_i x_i^\top x_i + c_i^\top x_i - (-s_i x_i^\top x_i + \kappa_i^\top x_i + p^\top x_i) + \delta\|x_i - \sigma(\boldsymbol{x})\|^2$, where $\boldsymbol{x} := \mathrm{col}((x_i)_{i\in\mathcal{I}})$, $q_i, c_i > 0$ depend on the nominal voltage and on the capacity loss of each battery, whereas $\kappa_i, s_i > 0$ model the battery size and the satisfaction of the ith PEV for charging the amount x_i. Moreover, the term $(q_i x_i^\top x_i + c_i^\top x_i)$ denotes the battery-degradation cost, $(-s_i x_i^\top x_i + \kappa_i^\top x_i + p^\top x_i)$ the benefit for charging [32], and $\delta\|x_i - \sigma(\boldsymbol{x})\|^2$ a penalty for deviating from the average charging profile, $\sigma(\boldsymbol{x}) := \frac{1}{N}\mathbf{1}^\top\boldsymbol{x}$, with $\delta > 0$. On the other hand, the leader aims at maximizing the following cost function:

$$J_0(p, \boldsymbol{x}) = -p^\top (D + \sigma(\boldsymbol{x})) \tag{14}$$

which represents the economic benefit for charging the PEVs, where $D \in \mathbb{R}^T$ is the total non-PEV demand over time. We assume that the net energy available for the PEVs is fixed, and therefore, the overall PEV demand shall meet the capacity constraint $\frac{1}{N}\mathbf{1}^\top\boldsymbol{x} \le C$, for some $C > 0$. Furthermore, we assume that, at every time step, $x_i \in [\underline{x}_i, \overline{x}_i]$, for all $i \in \mathcal{I}$. Thus, given the amount of energy requested by the PEVs, the retailer chooses a price p per unit of energy, with $p \in [0, \overline{p}]$, aiming at maximizing its revenue in (14).

For the numerical simulations, we consider $N = 10^4$ PEVs, a charging horizon discretized into $T = 24$ time intervals, $q_i = 1.2 \times 10^{-3}$, $c_i = 0.11$, whereas κ_i and s_i are randomly drawn from $\mathcal{N}(12, 2)$ and $\mathcal{U}(0.02, 0.1)$ respectively, and the capacity upper bound C is equal to 1.5 from 11 P.M. to 8 A.M., and to 0.5 for the rest of the day. The convergence behavior of Algorithm 1 over 10 experiments is shown in Fig. 1. During the numerical simulations, the inner loop takes between 50 and 75 iterations (on average) to meet a predefined stopping condition, and above 10^2 experiments, we did not experienced any influence on the outer-loop convergence behavior. For this latter, in view of the fact that $\lim_{k\to\infty}\|\hat{\boldsymbol{\omega}}(\boldsymbol{\omega}^k) - \boldsymbol{\omega}^k\| = 0$, we have chosen $\|\hat{\boldsymbol{\omega}}^k - \hat{\boldsymbol{\omega}}^{k-1}\| \le 10^{-4}$ as a stopping criterion.

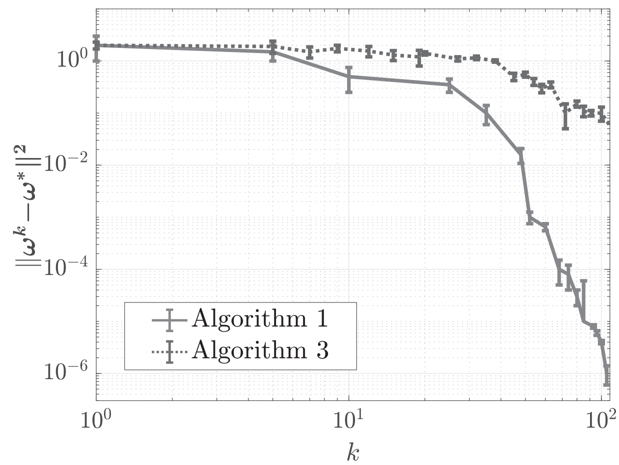

Fig. 1. Comparison of the convergence behavior between Algorithm 1 (solid blue line) and 3 (dotted red line).

Algorithm 3: Two-layer Naïve Heuristic for ℓ-SE Computation.

Initialization: $y_0(0) \in \mathcal{Y}_0$
Iteration $(k \in \mathbb{N})$**:**
(S1) Compute an v-GNE, $\boldsymbol{x}(k)$, for the game in (1)
(S2) Compute $y_0^*(k)$, solution to (3)
(S3) Update $y_0(k) := (1 - \beta(k))y_0(k-1) + \beta(k)y_0^*(k)$

The procedure proposed in Algorithm 1 is, then, compared with the simplest naïve method for possibly computing an ℓ-SE, whose main steps are summarized in Algorithm 3. Specifically, given the strategy of the leader at the previous step, the followers compute an v-GNE of the game in (1), and send their strategy back to the leader (S1). In turn, the leader first solves its optimization problem in (3) with solution $y_0^*(k)$ (S2) and, then, updates its strategy taking a convex combination between $y_0^*(k)$ and the strategy at the previous step, where the parameter $\beta(k) \in [0, 1]$ introduces a possible inertia (S3). Note that, albeit rather intuitive, this naïve algorithm has no convergence guarantees. However, in our numerical experience, by considering the cost function in (14) for the leader and setting $\beta(k) = 1/k$, Algorithm 3 apparently shows a slower convergent behavior compared with the proposed Algorithm 1, as depicted in Fig. 1 over 10 numerical experiments.

B. Tradeoff Between the Leader and the Followers

Finally, we highlight the tradeoff role played by the relaxation parameter θ in (7). In fact, for θ sufficiently large, the leader has a larger feasible set, whereas, on the other hand, the followers are farther away from an v-GNE, since the complementarity condition is not exactly satisfied. Therefore, the larger θ, the lower the optimal cost of the leader, and possibly the higher the optimal cost of each follower. Vice versa, the smaller θ, the higher the optimal cost of the leader, because his feasible set shrinks, and possibly the lower the optimal cost of each follower, since the equilibrium condition is closer to being satisfied. This behavior is essentially confirmed in Fig. 2 where, for ease of

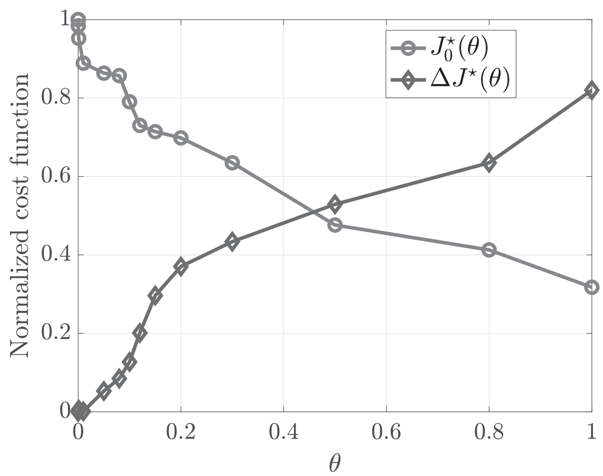

Fig. 2. Tradeoff role played by the regularization parameter θ.

visualization, we show the normalized benefit of the leader ($J_0^\star(\theta)$) and the normalized maximum disadvantage among the followers ($\Delta J^\star(\theta)$) as θ increases. Specifically, for each $\theta \in [\underline{\theta}, 1]$, we compute an ℓ-SE, and we denote with $J_0^\star(\theta)$ the corresponding optimal cost for the leader. For the followers, we introduce and show the maximum relative disadvantage with respect to a near-equilibrium condition, i.e., $\Delta J^\star(\theta) := \max_{i\in\mathcal{I}} \, (J_i^\star(\theta) - J_i^\star(\underline{\theta}))$, where, for a given θ, $J_i^\star(\theta)$ is the optimal cost for the ith follower, whereas in this case, we set $\underline{\theta}$ equal to 10^{-6}.

V. Conclusion

We have considered a multiagent, hierarchical equilibrium problem with one leader and multiple followers, with possibly nonconvex data for the leader, convex-quadratic objective functions and linear constraints for the followers, and overall an aggregative structure. In this setup, a local Stackelberg equilibrium can be approximated arbitrarily close via the relaxation of the complementarity condition that represents the equilibrium among the followers. In turn, the relaxed problem can be solved via a two-layer algorithm, which—thanks to the aggregative structure—requires semidecentralized computations and information exchange.

References

[1] H. Liu, "SINR-based multi-channel power schedule under DoS attacks: A Stackelberg game approach with incomplete information," *Automatica*, vol. 100, pp. 274–280, 2019.

[2] Y. Li, D. Shi and T. Chen, "False data injection attacks on networked control systems: A Stackelberg game analysis," *IEEE Trans. Autom. Control*, vol. 63, no. 10, pp. 3503–3509, Oct. 2018.

[3] M. Motalleb, P. Siano, and R. Ghorbani, "Networked Stackelberg competition in a demand response market," *Appl. Energy*, vol. 239, pp. 680–691, 2019.

[4] J. Chen and Q. Zhu, "A Stackelberg game approach for two-level distributed energy management in smart grids," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 6554–6565, Nov. 2018.

[5] F. G. Mendoza, D. Bauso, and G. Konstantopoulos, "Online pricing via Stackelberg and incentive games in a micro-grid," in *Proc. 18th Eur. Control Conf.*, 2019, pp. 3520–3525.

[6] K. Hirose and T. Matsumura, "Comparing welfare and profit in quantity and price competition within Stackelberg mixed duopolies," *J. Econ.*, vol. 126, no. 1, pp. 75–93, 2019.

[7] N. Groot, G. Zaccour, and B. De Schutter, "Hierarchical game theory for system-optimal control: Applications of reverse Stackelberg games in regulating marketing channels and traffic routing," *IEEE Control Syst. Mag.*, vol. 37, no. 2, pp. 129–152, Apr. 2017.

[8] Z.-Q. Luo, J.-S. Pang, and D. Ralph, *Mathematical Programs With Equilibrium Constraints*. Cambridge, U.K.: Cambridge Univ. Press, 1996.

[9] H. Scheel and S. Scholtes, "Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity," *Math. Oper. Res.*, vol. 25, no. 1, pp. 1–22, 2000.

[10] H. T. Jongen and G.-W. Weber, "Nonlinear optimization: Characterization of structural stability," *J. Glob. Optim.*, vol. 1, no. 1, pp. 47–64, 1991.

[11] M. Simaan and J. Cruz, "A Stackelberg solution for games with many players," *IEEE Trans. Autom. Control*, vol. AC-18, no. 3, pp. 322–324, Jun. 1973.

[12] F. Kydland, "Equilibrium solutions in dynamic dominant-player models," *J. Econ. Theory*, vol. 15, no. 2, pp. 307–324, 1977.

[13] F. Kydland, "Noncooperative and dominant player solutions in discrete dynamic games," *Int. Econ. Rev.*, vol. 16, pp. 321–335, 1975.

[14] T. Başar and R. Srikant, "A Stackelberg network game with a large number of followers," *J. Optim. Theory Appl.*, vol. 115, no. 3, pp. 479–490, 2002.

[15] H. Kebriaei and L. Iannelli, "Discrete-time robust hierarchical linear-quadratic dynamic games," *IEEE Trans. Autom. Control*, vol. 63, no. 3, pp. 902–909, Mar. 2018.

[16] 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.

[17] S. Zou, I. Hiskens, and Z. Ma, "Decentralized coordination of controlled loads and transformers in a hierarchical structure," *IFAC World Congr.*, vol. 50, no. 1, pp. 5560–5566, 2017.

[18] A. A. Kulkarni and U. V. Shanbhag, "An existence result for hierarchical Stackelberg v/s Stackelberg games," *IEEE Trans. Autom. Control*, vol. 60, no. 12, pp. 3379–3384, Dec. 2015.

[19] S. Leyffer and T. Munson, "Solving multi-leader-common-follower games," *Optim. Methods Softw.*, vol. 25, no. 4, pp. 601–623, 2010.

[20] B. Hobbs, "Linear complementarity models of Nash-Cournot competition in bilateral and POOLCO power markets," *IEEE Trans. Power Syst.*, vol. 16, no. 2, pp. 194–202, May 2001.

[21] A. Ehrenmann and K. Neuhoff, "A comparison of electricity market designs in networks," *Oper. Res.*, vol. 57, no. 2, pp. 274–286, 2009.

[22] C.-L. Su, "A sequential NCP algorithm for solving equilibrium problems with equilibrium constraints," Dept. Manage. Sci. Eng., Stanford Univ., Stanford, CA, USA, Tech. Rep., 2004.

[23] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*. Berlin, Germany: Springer, 2007.

[24] S. Scholtes, "Convergence properties of a regularization scheme for mathematical programs with complementarity constraints," *SIAM J. Optim.*, vol. 11, no. 4, pp. 918–936, 2001.

[25] R. Fletcher, S. Leyffer, D. Ralph, and S. Scholtes, "Local convergence of SQP methods for mathematical programs with equilibrium constraints," *SIAM J. Optim.*, vol. 17, no. 1, pp. 259–286, 2006.

[26] G. Scutari, F. Facchinei, P. Song, D. P. Palomar, and J.-S. Pang, "Decomposition by partial linearization: Parallel optimization of multi-agent systems," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 641–656, Feb. 2014.

[27] G. Scutari, F. Facchinei, and L. Lampariello, "Parallel and distributed methods for constrained nonconvex optimization—Part I: Theory," *IEEE Trans. Signal Process.*, vol. 65, no. 8, pp. 1929–1944, Apr. 2017.

[28] X. Hu and D. Ralph, "Using EPECs to model bilevel games in restructured electricity markets with locational prices," *Oper. Res.*, vol. 55, no. 5, pp. 809–827, 2007.

[29] F. Facchinei, A. Fischer, and V. Piccialli, "On generalized Nash games and variational inequalities," *Oper. Res. Lett.*, vol. 35, no. 2, pp. 159–164, 2007.

[30] D. P. Bertsekas, "Nonlinear programming," *J. Oper. Res. Soc.*, vol. 48, no. 3, pp. 334–334, 1997.

[31] N. Chatzipanagiotis and M. M. Zavlanos, "On the convergence of a distributed augmented Lagrangian method for nonconvex optimization," *IEEE Trans. Autom. Control*, vol. 62, no. 9, pp. 4405–4420, Sep. 2017.

[32] M. A. Tajeddini and H. Kebriaei, "A mean-field game method for decentralized charging coordination of a large population of plug-in electric vehicles," *IEEE Syst. J.*, vol. 13, no. 1, pp. 854–863, Mar. 2019.