



**HAL**  
open science

# Approximate Nash equilibria in large nonconvex aggregative games

Kang Liu, Nadia Oudjane, Cheng Wan

► **To cite this version:**

Kang Liu, Nadia Oudjane, Cheng Wan. Approximate Nash equilibria in large nonconvex aggregative games. *Mathematics of Operations Research*, 2022, 10.1287/moor.2022.1321 . hal-03023122v3

**HAL Id: hal-03023122**

**<https://hal.science/hal-03023122v3>**

Submitted on 26 Sep 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Approximate Nash equilibria in large non-convex aggregative games

Kang Liu\*, Nadia Oudjane†, Cheng Wan‡

September 26, 2022

## Abstract

This paper shows the existence of  $\mathcal{O}(\frac{1}{n^\gamma})$ -Nash equilibria in  $n$ -player noncooperative sum-aggregative games in which the players' cost functions, depending only on their own action and the average of all players' actions, are lower semicontinuous in the former while  $\gamma$ -Hölder continuous in the latter. Neither the action sets nor the cost functions need to be convex. For an important class of sum-aggregative games, which includes congestion games with  $\gamma$  equal to 1, a gradient-proximal algorithm is used to construct  $\mathcal{O}(\frac{1}{n})$ -Nash equilibria with at most  $\mathcal{O}(n^3)$  iterations. These results are applied to a numerical example concerning the demand-side management of an electricity system. The asymptotic performance of the algorithm when  $n$  tends to infinity is illustrated.

**Keywords.** Shapley-Folkman lemma, sum-aggregative games, non-convex game, large finite game,  $\epsilon$ -Nash equilibrium, gradient-proximal algorithm, congestion game

**MSC Class** Primary: 91A06; secondary: 90C26

## 1 Introduction

This paper studies approximate pure-Nash equilibria (PNE) for  $n$ -player noncooperative games involving non-convexities in players' costs or action sets. The goal is to show the existence of such approximate equilibria under certain conditions and to propose an algorithm that allows them to be calculated effectively in some specific cases. In particular, this paper focuses on a specific class of noncooperative games (which includes congestion games) referred to as sum-aggregative games (Selten [40], Corchón [9], Jensen [22]). The cost of each player depends on the weighted sum of the other players' decisions. These games have practical applications in various aspects of political science, economics, social biology, and engineering, such as voting [30, 34], market competition [29], public goods provision [2, 15], rent seeking [11], population dynamics [17], traffic analysis [10, 27], communications network control [26, 31] and electrical system management [18, 21]. However, in these real-life situations, the players' action sets and their cost functions are often *non-convex*. This paper is actually motivated by concrete applications for which it is unreasonable to neglect the

---

\*CMAP, Ecole Polytechnique; Inria Saclay; kang.liu@polytechnique.edu

†EDF R&D and FIME (Finance for Energy Market Research Centre); nadia.oudjane@edf.fr

‡EDF R&D and FIME (Finance for Energy Market Research Centre); cheng.wan.05@polytechnique.org (corresponding author)

*non-convexities* inherent to the problem. In particular, we are interested in demand-side management in electrical systems [20], where each flexible consumer is considered as a player trying to minimize her electricity bill by modulating her consumption (e.g., electric vehicle charging) which is subject to non-convex constraints.

In the convex framework, a PNE is known to exist under mild regularity conditions (see, for example, Rosen [36]). Outside the convex framework, it is generally difficult to provide *existence* results for PNEs and *approximation algorithms* with performance guarantees. Our work addresses these two issues and makes the following contributions.

(i) Theoretically, Proposition 2.4 and Theorem 2.7 show the existence of  $\mathcal{O}(\frac{1}{n^\gamma})$ -PNEs for  $n$ -player non-convex sum-aggregative games in which the players' cost functions are  $\gamma$ -Hölder continuous with respect to the aggregate. Neither the action sets nor the cost functions need to be convex.

(ii) Algorithmically, in the specific case of congestion games in which the cost functions are Lipschitz continuous with respect to the aggregate (i.e.,  $\gamma = 1$ ), we present an iterative gradient-proximal algorithm to compute an  $\mathcal{O}(\frac{1}{n})$ -PNE of the original non-convex game within at most  $\mathcal{O}(n^3)$  iterations, according to Theorem 3.3.

(iii) Practically, the usefulness of this approach is demonstrated in Section 4, where a numerical simulation with the gradient-proximal algorithm is performed for a demand-side management problem involving flexible electric vehicle charging.

The originality of this paper lies in the circumvention of the non-convexity through the exploitation the fact that large sum-aggregative games approach a convex framework when the number of players is large. The counterpart of this approach is to search for an  $\epsilon$ -PNE (cf. Definition 2.1) instead of an exact PNE. The main inspirations for the present work are [41] (in economics) and [46] (in optimization). Starr [41] was interested in computing general equilibria for a non-convex competitive economy in terms of price and quantity, while Wang [46] considered large-scale non-convex separable optimization problems coupled by sum-aggregative terms. In both cases, the authors proposed to convexify the problem, taking advantage of the large number of agents or subproblems to bound the error induced by the convexification thanks to the Shapley-Folkman Lemma (cf. Lemma 5.4). Roughly speaking, the Shapley-Folkman Lemma states that the Minkowski sum of a finite number of sets in a Euclidean space is close to convex when the number of sets is very large compared with their common dimension. Our first contribution consists of two novelties. First, in Proposition 2.4, we extend their approach to non-convex sum-aggregative games to show the existence of an  $\epsilon$ -PNE. Second, in Theorem 2.7 we show that one can also construct an  $\epsilon$ -PNE of the non-convex game from an  $\epsilon$ -PNE of the auxiliary convexified game provided that some *stability condition* is satisfied. This second novelty is more significant, and it is crucial for our algorithmic contribution. Our second contribution consists in proposing an algorithm returning an  $\epsilon$ -PNE of the convexified game. This  $\epsilon$ -PNE verifies the stability condition that allows us to recover an  $\epsilon$ -PNE of the original non-convex game.

**Related works.** The existence of PNEs is not guaranteed in non-convex games, except in some particular cases. For example, for games in which players have a finite number of actions, the existence of PNE is known for Rosenthal's congestion games [37] and some other specific class of congestion games [28, 42]. For games with discrete (but not necessarily finitely many) strategies, Sagratella [38] proved the existence of PNEs for a particular class

of such games and proposed an algorithm leading to one of the equilibria. However, when the players' cost functions are non-convex and/or their action sets are non-convex but not necessarily finite, there is no general result for the existence of PNEs.

Concerning algorithms for the computation of ( $\epsilon$ -)PNE, there are few results. The existing results are almost restricted to some special cases in the convex setting. A common approach is to solve the variational inequality characterizing the PNEs (cf. Facchinei and Pang [13] and the references therein). Scutari et al. [39] considered generic  $n$ -player games that need not be large nor aggregative but must have a strongly monotone inequality characterizing the PNE. They used proximal best-reply algorithms to solve this variational inequality. Paccagnan, Kamgarpour and Lygeros [33] considered a specific convex aggregative game and used a decentralized gradient projection algorithm to solve the strongly monotone variational inequality characterizing the PNE. Paccagnan et al. [32] studied  $\epsilon$ -PNEs in large convex aggregative games with coupling constraints. Their methodology is close to ours in the sense that they only look for an  $\epsilon$ -PNE (which they call the Wardrop equilibrium) instead of an exact PNE. They used, respectively, a decentralized gradient projection algorithm and a decentralized best-reply (to the aggregate term) algorithm to solve the variational inequality characterizing this Wardrop equilibrium.

Finally, the Shapley-Folkman lemma has been extensively applied in non-convex optimization for its convexification effect. Aubin and Ekeland [1] used the lemma to derive an upper bound for the duality gap in an additive, separable non-convex optimization problem. Since then, quite a few papers have extended or sharpened this result (cf. Ekeland and Temam [12], Bertsekas and coauthors [4, 7], Pappalardo [35], Kerdreux et al. [24], Bi and Tang [8]). These theoretical results have applications in engineering problems, such as the large-scale unit commitment problem [3, 25], the optimization of plug-in electric vehicle charging [45], the optimization of multicarrier communication systems [48], supply-chain management [44], and spatial graphical model estimation [14].

**Organization.** Section 2 focuses on the existence of  $\epsilon$ -PNE in large non-convex aggregative games. Section 3 presents an algorithm to compute such an  $\epsilon$ -equilibrium for congestion games. A numerical application in Section 4 illustrates the usefulness of our results. Section 5 concludes.

**Notation.** In a Euclidean space,  $\|\cdot\|$  denotes the  $l^2$ -norm. For a point  $x \in \mathbb{R}^d$  and a subset  $\mathcal{X}$  of  $\mathbb{R}^d$ ,  $d(x, \mathcal{X}) := \inf_{y \in \mathcal{X}} \{\|x - y\|\}$  is the distance from the point to the subset. For two subsets  $\mathcal{X}$  and  $\mathcal{Y}$  of  $\mathbb{R}^d$ , their Minkowski sum is the set  $\{x + y \mid x \in \mathcal{X}, y \in \mathcal{Y}\}$ . For  $x \in \mathbb{R}^d$  and  $r \in \mathbb{R}^+$ ,  $B(x, r) := \{y \in \mathbb{R}^d \mid \|y - x\| \leq r\}$ , with the  $r$ -radial ball centered on  $x$ .

For a matrix  $A \in \mathbb{R}^d \times \mathbb{R}^q$ ,  $\|\cdot\|_2$  is the 2-norm of the matrix:  $\|A\|_2 := \sqrt{\lambda_{\max}(A^T A)}$  where  $A^T$  is the transpose of  $A$  and  $\lambda_{\max}(A^T A)$  stands for the largest eigenvalue of the matrix  $A^T A$ .

The proof of Proposition 2.3 and the lemmata used for the proof are given in Appendix A. All the other proofs and intermediate results are contained in Appendix B.

## 2 Existence of $\epsilon$ -PNE in large non-convex sum-aggregative games

### 2.1 A non-convex sum-aggregative game and its convexification

Consider an  $n$ -player noncooperative game  $\Gamma$ . The players are indexed over  $N = \{1, 2, \dots, n\}$ . Each player  $i \in N$  has an action set  $\mathcal{X}_i \subset \mathbb{R}^d$ , which is closed and bounded but not necessarily convex. Let  $\tilde{\mathcal{X}}_i := \text{conv}(\mathcal{X}_i)$  be the convex hull of  $\mathcal{X}_i$  (which is also closed and bounded) and denote  $\mathcal{X} := \prod_{i \in N} \mathcal{X}_i$ ,  $\tilde{\mathcal{X}} := \prod_{i \in N} \tilde{\mathcal{X}}_i$ ,  $\tilde{\mathcal{X}}_{-i} := \prod_{j \in N_{-i}} \tilde{\mathcal{X}}_j$ , where  $N_{-i} := N \setminus \{i\}$ . Let the constant  $\Delta > 0$  be such that, for all  $i \in N$ , the compact set  $\mathcal{X}_i$  has a diameter  $|\mathcal{X}_i| := \max_{x_i, y_i \in \mathcal{X}_i} \|x_i - y_i\|$  that is not greater than  $\Delta$ .

As usual, let  $x_{-i}$  denote the profile of actions of all the players except player  $i$ . Each player  $i$  has a real-valued cost function  $f_i$  defined on  $\mathcal{X}_i \times \tilde{\mathcal{X}}_{-i}$ , which has the following specific form:

$$f_i(x_i, x_{-i}) := \theta_i \left( x_i, \frac{1}{n} \sum_{j \in N} A_j x_j \right), \quad \text{for any } x_i \in \mathcal{X}_i, x_{-i} \in \tilde{\mathcal{X}}_{-i}, \quad (2.1)$$

where each  $A_j$  is a  $q \times d$  matrix for all  $j \in N$ , and  $\theta_i$  is a real-valued function defined on  $\mathcal{X}_i \times \Omega$ , with  $\Omega \subset \mathbb{R}^q$  being a neighborhood of  $\{\frac{1}{n} \sum_{j \in N} A_j y_j \mid y_j \in \tilde{\mathcal{X}}_j, \forall j \in N\}$ .

Let the constant  $M > 0$  be such that  $\|A_i\|_2 \leq M$  for each  $i \in N$ .

The game  $\Gamma$  is a sum-aggregative game because each player's cost function depends on her own action and an aggregate of all the players' actions.

**Definition 2.1** ( $\epsilon$ -pure Nash equilibrium). For a constant  $\epsilon \geq 0$ , an  $\epsilon$ -pure Nash equilibrium ( $\epsilon$ -PNE)  $x^\epsilon \in \mathcal{X}$  in game  $\Gamma$  is a profile of actions of the  $n$  players such that, for each player  $i \in N$ ,

$$f_i(x_i^\epsilon, x_{-i}^\epsilon) \leq f_i(x_i, x_{-i}^\epsilon) + \epsilon, \quad \text{for any } x_i \in \mathcal{X}_i.$$

If  $\epsilon = 0$ , then  $x^\epsilon$  is a *pure Nash equilibrium (PNE)*.

This definition of  $\epsilon$ -PNE corresponds to the notion of *additively  $\epsilon$ -PNE* in the literature.

For non-convex games (in which either action sets or cost functions are not convex), the existence of a PNE is not established for the general cases. This paper uses an auxiliary convexified version of the non-convex game, which is helpful both in the proof of the existence of an  $\epsilon$ -PNE of the non-convex game and in the construction of such an approximate PNE.

**Definition 2.2** (Convexified game and generators). The *convexified game*  $\tilde{\Gamma}$  associated with  $\Gamma$  is a noncooperative game played by  $n$  players. Each player  $i \in N$  has an action set  $\tilde{\mathcal{X}}_i$  and a real-valued cost function  $\tilde{f}_i$  defined on  $\tilde{\mathcal{X}}$  as follows: for all  $x \in \tilde{\mathcal{X}}$ ,

$$\tilde{f}_i(x_i, x_{-i}) = \inf_{(\alpha^k)_{k=1}^{d+1} \in \mathcal{S}_d; (z^k)_{k=1}^{d+1} \in \mathcal{X}_i^{d+1}} \left\{ \sum_{k=1}^{d+1} \alpha^k f_i(z^k, x_{-i}) \mid x_i = \sum_{k=1}^{d+1} \alpha^k z^k \right\}, \quad (2.2)$$

where  $\mathcal{S}_d := \{\alpha = (\alpha^k)_{k=1}^{d+1} \in \mathbb{R}^{d+1} \mid \forall k, \alpha^k \geq 0, \sum_{k=1}^{d+1} \alpha^k = 1\}$  denotes the probability simplex of dimension  $d$ .

For all  $i \in N$  and  $x \in \tilde{\mathcal{X}}$ , let a minimizer in (2.2) be generically denoted by  $(\alpha(i, x), z(i, x)) \in \mathcal{S}_d \times \mathcal{X}_i^{d+1}$ . For such a vector  $z(i, x)$ , we denote the set of its  $d + 1$  components by  $Z(i, x)$  and call it a *generator for*  $(i, x)$ .

For a lower semicontinuous (l.s.c.) function, equation (2.2) just defines its convex hull (cf. Lemma 5.5). This particular form of definition is proposed in [5].

PNEs and  $\epsilon$ -PNEs for the convexified game  $\tilde{\Gamma}$  are defined in the same way as for the game  $\Gamma$  in Definition 2.1.

The remainder of this subsection is dedicated to a preliminary analysis of the convexified game.

First let us introduce an assumption on the functions  $\theta_i$ 's characterizing the cost functions according to (2.1). It ensures the existence of generators for all  $(i, x) \in N \times \mathcal{X}$  (cf. Lemma 5.5 for a proof).

**Assumption 1.**

- (1) For any player  $i \in N$ , for any  $y \in \Omega \subset \mathbb{R}^d$ , the function  $x_i \mapsto \theta_i(x_i, y)$  is l.s.c. on  $\mathcal{X}_i$ .
- (2) There exist constants  $H > 0, \gamma > 0$  such that, for all  $i \in N$ , for all  $x_i \in \mathcal{X}_i$ , the function  $y \mapsto \theta_i(x_i, y)$  is  $(H, \gamma)$ -Hölder continuous on  $y \in \Omega$ , i.e.,

$$|\theta_i(x_i, y') - \theta_i(x_i, y)| \leq H \|y' - y\|^\gamma. \quad (2.3)$$

**Remark 2.1.** It is straightforward from Assumption 1 that  $f_i(\cdot, x_{-i})$  is l.s.c. in  $x_i \in \mathcal{X}_i$  for any fixed  $x_{-i} \in \tilde{\mathcal{X}}_{-i}$ .

According to Lemma 5.5,  $x_i \mapsto \tilde{f}_i(x_i, x_{-i})$  is convex and l.s.c on  $\tilde{\mathcal{X}}_i$ . Its subdifferential exists; let it be denoted by  $\partial_i \tilde{f}_i(\cdot, x_{-i})$ . Then, for each  $x_i \in \tilde{\mathcal{X}}_i$ ,  $\partial_i \tilde{f}_i(x_i, x_{-i})$  is a nonempty convex subset of  $\mathbb{R}^d$ .

**Proposition 2.3** (Existence of PNE in  $\tilde{\Gamma}$ ). *Under Assumption 1, the convexified game  $\tilde{\Gamma}$  admits a PNE.*

*Proof.* This results from Theorem 5.3 in Appendix A. □

**Remark 2.2.** Theorem 5.3 is a natural extension of Rosen's theorem on the existence of PNEs in games with convex continuous cost functions [36] to the case where the cost functions are only l.s.c. instead of being continuous with respect to the players' own actions.

The following example shows that even the continuity of  $f_i$  on  $\mathcal{X}_i$  cannot guarantee the continuity of  $\tilde{f}_i$  on  $\tilde{\mathcal{X}}_i$ , meaning that Rosen's theorem is not sufficient here.

Consider  $d = 3$ ,  $\mathcal{X}_i = T \cup B \cup S$  where  $T = \{(x^1, x^2, x^3) \in \mathbb{R}^3 | (x^1)^2 + (x^2)^2 = 1, x^3 = 1\}$ ,  $B = \{(x^1, x^2, x^3) \in \mathbb{R}^3 | (x^1)^2 + (x^2)^2 = 1, x^3 = -1\}$ , and  $S = \{(x^1, x^2, x^3) \in \mathbb{R}^3 | x^1 = 1, x^2 = 0, -1 \leq x^3 \leq 1\}$ ;  $f_i$  is independent of  $x_{-i}$ , and  $f_i(x) = 0$  for  $x \in T \cup B$ ,  $f_i(x) = |x^3| - 1$  for  $x \in S$ . Then, for all  $x \in \{(x^1, x^2, x^3) \in \mathbb{R}^3 | (x^1)^2 + (x^2)^2 = 1, x^3 = 0\} \subset \partial \tilde{\mathcal{X}}_i$ ,  $\tilde{f}_i(x) = 0$  except for  $x^* = (1, 0, 0)$ , but  $\tilde{f}_i(x^*) = f_i(x^*) = -1$ .

## 2.2 Existence and construction of an $\epsilon$ -PNE of the non-convex game

The following proposition shows the existence of an  $\epsilon$ -PNE in the non-convex game  $\Gamma$  and its construction from an exact PNE of the convexified game  $\tilde{\Gamma}$ . In particular, we observe that  $\epsilon$  is small when the number of players  $n$  is large with respect to  $q$ , the dimension of the space in which the aggregate  $\frac{1}{n} \sum_{i \in N} A_i x_i$  lies.

**Proposition 2.4** (Existence of  $\epsilon$ -PNE). *Under Assumption 1, the non-convex game  $\Gamma$  admits an  $\epsilon$ -PNE, where  $\epsilon = 2H \left( \frac{(\sqrt{q}+1)M\Delta}{n} \right)^\gamma$ .*

*In particular, suppose that  $\tilde{x} \in \tilde{\mathcal{X}}$  is a PNE in  $\tilde{\Gamma}$  (which exists according to Proposition 2.3), and  $Z(i, \tilde{x})$  is an arbitrary generator for each player  $i$ ; then,  $x^* \in \mathcal{X}$  such that*

$$x^* \in \operatorname{argmin}_{x_i \in Z(i, \tilde{x}), i \in N} \left\| \sum_{i \in N} A_i \tilde{x}_i - \sum_{i \in N} A_i x_i \right\|^2, \quad (2.4)$$

*is an  $\epsilon$ -PNE of the non-convex game  $\Gamma$ .*

*Sketch of the proof:* By the definition of the PNE in  $\tilde{\Gamma}$ ,  $\tilde{x}_i$  is a best response to  $\tilde{x}_{-i}$  in terms of  $\tilde{f}_i$ . By Lemma 5.5, all the points in  $Z(i, \tilde{x})$  are also best replies to  $\tilde{x}_{-i}$  in terms of  $\tilde{f}_i$ .

We then use the Shapley-Folkman Lemma (Lemma 5.4) to disaggregate  $\frac{1}{n} \sum_i A_i \tilde{x}_i$  over the sets  $Z(i, \tilde{x})$  to obtain a feasible profile  $x^*$ . Finally, we can show that  $\frac{1}{n} \sum_i A_i x_i^* \approx \frac{1}{n} \sum_i A_i \tilde{x}_i$  and that  $x_i^*$  is (almost) a best response to  $\frac{1}{n} \sum_i A_i x_i^*$ .

From an algorithmic point of view, a PNE is not always easy or fast to compute for the convexified game  $\tilde{\Gamma}$ , even though its existence is guaranteed. Even when we have a convergent algorithm, the outputs of the algorithm at each iteration provide only approximations of the exact PNE which may constitute  $\epsilon$ -PNEs but rarely exact PNEs. Then, the question that naturally arises is whether the idea above is still valid if  $\tilde{x}$  is only an  $\epsilon$ -PNE of  $\tilde{\Gamma}$ , i.e.,  $\tilde{x}_i$  is an  $\epsilon$ -best response to  $\tilde{x}_{-i}$  in terms of  $\tilde{f}_i$ . The answer is yes if the  $\epsilon$ -PNE  $\tilde{x}$  of the convexified game  $\tilde{\Gamma}$  satisfies a more demanding condition, introduced by the following definition.

**Definition 2.5** (Stability condition). In game  $\tilde{\Gamma}$ , for a given  $\eta \geq 0$ , a point  $\tilde{x} \in \tilde{\mathcal{X}}$  is said to satisfy the  $\eta$ -stability condition with respect to  $(Z(i, \tilde{x}))_i$  if, for each player  $i$ ,  $\tilde{f}_i(x_i, \tilde{x}_{-i}) \leq \tilde{f}_i(\tilde{x}_i, \tilde{x}_{-i}) + \eta$  for all  $x_i \in Z(i, \tilde{x})$ .

A point  $\tilde{x}$  is said to satisfy the  $\eta$ -stability condition if it satisfies the  $\eta$ -stability condition with respect to a certain generator profile  $(Z(i, \tilde{x}))_i$ .

A point  $\tilde{x}$  is said to satisfy the *full  $\eta$ -stability condition* if it satisfies the  $\eta$ -stability condition with respect to any generator profile  $(Z(i, \tilde{x}))_i$ .

The stability condition of  $\tilde{x}$  with respect to  $(Z(i, \tilde{x}))_i$  means that the cost for each player  $i$  is only slightly increased if the player's choice is unilaterally perturbed within the convex hull of the generator  $Z(i, \tilde{x})$ .

According to Lemma 5.5, a PNE of  $\tilde{\Gamma}$  satisfies the full 0-stability condition.

A sufficient condition for the  $\eta$ -stability of  $\tilde{x}$  is given by Lemma 2.6.

**Lemma 2.6.** *Under Assumption 1, for any action profile  $\tilde{x} \in \tilde{\mathcal{X}}$ , for any player  $i$ , if there is a generator  $Z(i, \tilde{x})$  and  $h \in \partial_i \tilde{f}_i(\tilde{x}_i, \tilde{x}_{-i})$  such that*

$$\langle h, x_i - \tilde{x}_i \rangle \geq -\eta \|x_i - \tilde{x}_i\|, \quad \forall x_i \in \text{conv} Z(i, \tilde{x}), \quad (2.5)$$

then,

$$|\tilde{f}_i(x_i, \tilde{x}_{-i}) - \tilde{f}_i(\tilde{x}_i, \tilde{x}_{-i})| \leq \eta \|x_i - \tilde{x}_i\|, \quad \text{for all } x_i \in Z(i, \tilde{x}).$$

In particular,  $\tilde{x}$  satisfies the  $\eta\Delta$ -stability condition with respect to  $(Z(i, \tilde{x}))_i$ .

The following theorem describes how to construct an  $\mathcal{O}(\frac{1}{n^\gamma})$ -PNE of the original non-convex game  $\Gamma$  when we know both an  $\epsilon$ -PNE of the convexified game  $\tilde{\Gamma}$  satisfying the  $\eta$ -stability condition and the associated generator profile.

**Theorem 2.7** (Construction of  $\epsilon$ -PNE of  $\Gamma$ ). *Under Assumption 1, suppose that  $\tilde{x} \in \tilde{\mathcal{X}}$  is an  $\epsilon$ -PNE in  $\tilde{\Gamma}$  that satisfies the  $\eta$ -stability condition with respect to a specific generator profile  $(Z(i, \tilde{x}))_i$ . Let  $x^* \in \mathcal{X}$  be such that*

$$x^* \in \underset{x_i \in Z(i, \tilde{x}), i \in N}{\text{argmin}} \left\| \sum_{i \in N} A_i \tilde{x}_i - \sum_{i \in N} A_i x_i \right\|^2. \quad (2.6)$$

Then,  $x^*$  is an  $\tilde{\epsilon}$ -PNE of the non-convex game  $\Gamma$ , where  $\tilde{\epsilon} = \epsilon + \eta + 2H(\frac{(\sqrt{q}+1)M\Delta}{n})^\gamma$ .

### 2.3 A distributed randomized ‘‘Shapley-Folkman disaggregation’’

Once an exact PNE or an  $\epsilon$ -PNE  $\tilde{x}$  satisfying the  $\eta$ -stability condition of the convexified game  $\tilde{\Gamma}$  is obtained, as well as the associated generator profile  $(Z(i, \tilde{x}))_i$ , we would like to find an  $\tilde{\epsilon}$ -PNE of the non-convex game  $\Gamma$  whose existence is shown by Proposition 2.4 and Theorem 2.7. However, solving (2.6) is generally hard (cf. Udell and Boyd [43] for such a ‘‘Shapley-Folkman disaggregation’’ in a particular optimization setting). In this section, we present a method for computing an  $\tilde{\epsilon}$ -mixed-strategy Nash equilibrium (MNE) in a distributed way, based on the known  $\epsilon$ -PNE of  $\tilde{\Gamma}$ , its associated generator profile, and its coefficients. The algorithm is called ‘‘distributed’’ because it computes the mixed-strategy  $\mu_i$  of each player  $i$  from the information of  $\tilde{x}_i$ , the generator  $Z(i, \tilde{x})$  and the corresponding coefficients only.

**Proposition 2.8.** *Under Assumption 1, suppose that  $\tilde{x} \in \tilde{\mathcal{X}}$  is an  $\epsilon$ -PNE of  $\tilde{\Gamma}$  satisfying the  $\eta$ -stability condition with respect to  $(Z(i, \tilde{x}))_i$ , and each player  $i$  plays a mixed strategy  $\tilde{\mu}_i$  independently, i.e., a random action  $X_i$  following the distribution  $\tilde{\mu}_i$  over  $\mathcal{X}_i$ , defined by  $\mathbb{P}(X_i = x_i^l) = \alpha_i^l$ , where  $x_i^l \in Z(i, \tilde{x})$  for  $l = 1, \dots, d+1$ , and  $(\alpha_i^l)_{l=1}^{d+1} = \alpha(i, \tilde{x})$  is the corresponding vector of coefficients. Then, for  $\gamma \leq 1$ ,  $\tilde{\mu} = (\tilde{\mu}_i)_i$  is an  $\tilde{\epsilon}$ -MNE of the non-convex game  $\Gamma$ , where  $\tilde{\epsilon} = \epsilon + \eta + 2H(\frac{(\sqrt{n}+1)M\Delta}{n})^\gamma$ , in the sense that*

$$\mathbb{E}[f_i(X_i, X_{-i})] \leq \mathbb{E}[f_i(x_i, X_{-i})] + \tilde{\epsilon}, \quad \forall x_i \in \mathcal{X}_i.$$

**Remark 2.3.** Note that this is not an ‘‘adaptive’’ algorithm that allows the players to attain an  $\epsilon$ -PNE/MNE in the non-convex game through a decentralized adaptation/learning process; instead, it is a distributed, randomized disaggregation algorithm to recover an  $\tilde{\epsilon}$ -MNE of  $\Gamma$  from a known  $\epsilon$ -PNE of  $\tilde{\Gamma}$  satisfying the  $\eta$ -stability condition and its generators.



Besides, when  $\tilde{x}$  is an exact PNE of  $\tilde{\Gamma}$ , all the generators of  $\tilde{x}$  can be used in this algorithm. On the contrary, when  $\tilde{x}$  is only an  $\epsilon$ -PNE of  $\tilde{\Gamma}$ , we need a specific profile of generators  $(Z(i, x))_i$ , with respect to which  $\tilde{x}$  satisfies the  $\eta$ -stability condition. How to find such a generator profile is not evident. However, in the next section, we provide an algorithm to compute, for a specific class of games, an  $\epsilon$ -PNE of  $\tilde{\Gamma}$  satisfying the full  $\eta$ -stability condition (i.e., with respect to any profile of generators of  $\tilde{x}$ ). In that case, any profile of generators can be used in the algorithm of Proposition 2.8.

Finally, note that the estimated error  $\tilde{\epsilon}$  for the distributed randomized approximate MNE in Proposition 2.8 is larger than the estimated error  $\tilde{\epsilon}$  for the approximate PNE  $x^*$  in Theorem 2.7.

### 3 Computing $\epsilon$ -equilibria for large non-convex congestion games

#### 3.1 Non-convex congestion games

Congestion games are an extensively studied class of sum-aggregative games. In this section, we present an iterative algorithm to compute an  $\omega(K, n)\Delta$ -PNE of the convexification of a specific congestion game, in which  $\omega(K, n)$  tends to zero when both the number of players  $n$  and the number of iterations  $K$  tend to  $+\infty$ , while  $\frac{n}{K}$  tends to zero. Note that any algorithm returning an approximate PNE of the convexified game will not necessarily ensure that it verifies the stability condition. The proposed algorithm is of particular interest because we can show that the iterates provide an  $\omega(K, n)\Delta$ -PNE of the convexified game that satisfies the full  $\omega(K, n)\Delta$ -stability condition (cf. Proposition 3.2). Then, taking  $K \sim \mathcal{O}(n^3)$ , Theorem 3.3 shows that one can recover an  $\mathcal{O}(\frac{1}{n})$ -PNE of the original non-convex congestion game from this  $\omega(K, n)\Delta$ -PNE of the convexified game.

Consider a congestion game in which each player  $i \in N$  has an action set  $\mathcal{X}_i \subset \mathbb{R}^d$  and a cost function of the following form:

$$\begin{aligned} f_i(x_i, x_{-i}) &= \left\langle g\left(\frac{1}{n} \sum_{j \in N} a_j x_j\right), x_i \right\rangle + h_i\left(\frac{1}{n} \sum_{j \in N} a_j x_j\right) + r_i(x_i) \\ &= \sum_{t=1}^d g_t\left(\frac{1}{n} \sum_{j \in N} a_j x_{j,t}\right) x_{i,t} + h_i\left(\frac{1}{n} \sum_{j \in N} a_j x_j\right) + r_i(x_i). \end{aligned} \tag{3.1}$$

Suppose that the following assumptions hold on  $\mathcal{X}_i$ ,  $(a_j)_{j \in N} \in \mathbb{R}^n$ ,  $g_t$ ,  $h_i$  and  $r_i$ .

#### Assumption 2.

- There exist constants  $m > 0$  and  $M > 0$  such that  $m \leq a_i \leq M$  for all  $i \in N$ .
- For  $t = 1, \dots, d$ , the function  $g_t : \mathbb{R} \rightarrow \mathbb{R}$  is  $L_{g_t}$ -Lipschitz continuous and nondecreasing on a neighborhood of  $[D_1, D_2]$ , where the constants  $D_1$  and  $D_2$  are such that  $D_1 \leq \min_{t=1, \dots, d; x \in \tilde{\mathcal{X}}} \frac{1}{n} \sum_{j \in N} a_j x_{j,t} \leq \max_{t=1, \dots, d; x \in \tilde{\mathcal{X}}} \frac{1}{n} \sum_{j \in N} a_j x_{j,t} \leq D_2$ .
- For each  $i \in N$ , the function  $h_i : \mathbb{R}^d \rightarrow \mathbb{R}$  is  $L_{h_i}$ -Lipschitz continuous on  $[D_1, D_2]^d$ .
- Players' local cost functions  $r_i : \mathbb{R}^d \rightarrow \mathbb{R}$  are uniformly bounded, i.e., there exists a constant  $B_r > 0$  such that, for all  $i \in N$  and all  $x_i \in \mathcal{X}_i$ ,  $|r_i(x_i)| \leq B_r$ .

**Notation.** Let the constant  $\Delta := \max\{\max_{i \in N} \max_{x_i \in \tilde{\mathcal{X}}_i} \|x_i\|, \max_{i \in N} |\tilde{\mathcal{X}}_i|\}$ . Let  $L_g := \max_{1 \leq t \leq d} L_{g_t}$ ,  $L_h := \max_{i \in N} L_{h_i}$ ,  $B_g := \max_{1 \leq t \leq d, D_1 \leq s \leq D_2} |g_t(s)|$ .

The convexification of  $\Gamma$  is rather complicated to compute in the general case. Let us first introduce an auxiliary game that is very close to  $\Gamma$  whose convexification is easier to obtain.

Fix arbitrarily  $x_i^+ \in \mathcal{X}_i$  for each player  $i \in N$ . The auxiliary game  $\bar{\Gamma}$  is defined as follows: the player set and each player's action set are the same as in  $\Gamma$ , but player  $i$ 's cost function is, for all  $x_i \in \mathcal{X}_i$  and all  $x_{-i} \in \tilde{\mathcal{X}}_{-i}$ ,

$$\bar{f}_i(x_i, x_{-i}) := \left\langle g\left(\frac{1}{n} \sum_{j \neq i} a_j x_j + \frac{1}{n} a_i x_i^+\right), x_i \right\rangle + r_i(x_i).$$

The original game  $\Gamma$  can be approximated by the auxiliary game  $\bar{\Gamma}$  because their approximate equilibria are very close to each other, as the following lemma shows (see its proof in Appendix B).

**Lemma 3.1.** *Under Assumption 2, for the auxiliary game  $\bar{\Gamma}$ ,*

- (1) *Assumption 1 is verified with  $H = L_g \Delta$  and  $\gamma = 1$ ;*
- (2) *an  $\epsilon$ -PNE of  $\bar{\Gamma}$  is an  $(\epsilon + \frac{L_h M \Delta}{n} + \frac{2L_g M \Delta^2}{n})$ -PNE of  $\Gamma$ .*

For any fixed  $x_{-i} \in \tilde{\mathcal{X}}_{-i}$ ,  $\bar{f}_i(\cdot, x_{-i})$  is composed of a linear function of  $x_i$  and a local function of  $x_i$ . By abuse of notation, let us still use  $\bar{f}_i$  to denote its convexification on  $\tilde{\mathcal{X}}_i$ . More explicitly,

$$\bar{f}_i(x_i, x_{-i}) := \left\langle g\left(\frac{1}{n} \sum_{j \neq i} a_j x_j + \frac{1}{n} a_i x_i^+\right), x_i \right\rangle + \tilde{r}_i(x_i), \quad (3.2)$$

where  $\tilde{r}_i$  is the convexification of  $r_i$  defined on  $\tilde{\mathcal{X}}_i$  in the same way as (2.2).

By abuse of notation, let  $\tilde{\Gamma}$  denote the convexification of  $\bar{\Gamma}$  on  $\tilde{\mathcal{X}}$ .

### 3.2 A gradient-proximal algorithm

This subsection presents a gradient-proximal algorithm based on the block coordination proximal algorithm introduced by Xu and Yin [47] that can be used to construct an  $\mathcal{O}(\frac{1}{n})$ -PNE of  $\tilde{\Gamma}$  that satisfies the full  $\mathcal{O}(\frac{1}{n})$ -stability condition.

---

**Algorithm 1:** Gradient-proximal algorithm for  $\tilde{\Gamma}$ 

---

**Initialization:** choose initial point  $x^0 = (x_1^0, x_2^0, \dots, x_n^0) \in \tilde{\mathcal{X}}$

**for**  $k = 1, 2, \dots$  **do**

**for**  $i = 1, 2, \dots, n$  **do**

$$x_i^k = \operatorname{argmin}_{x_i \in \tilde{\mathcal{X}}_i} \left\langle g \left( \frac{1}{n} \sum_{j < i} a_j x_j^k + \frac{1}{n} \sum_{j \geq i} a_j x_j^{k-1} \right), x_i - x_i^{k-1} \right\rangle + \frac{a_i L_g}{2n} \|x_i - x_i^{k-1}\|^2 + \tilde{r}_i(x_i) \quad (3.3)$$

**end**

**if** *stopping criterion is satisfied* **then**

        | return  $(x_1^k, x_2^k, \dots, x_n^k)$ . Break.

**end**

**end**

---

**Remark 3.1.** This is a decentralized-coordinated type of algorithm. The coordinator needs to know the current choices of the players and  $(a_i)_i$  to compute  $g(\frac{1}{n} \sum_{j < i} a_j x_j^k + \frac{1}{n} \sum_{j \geq i} a_j x_j^{k-1})$ . The value of the vector  $g(\frac{1}{n} \sum_{j < i} a_j x_j^k + \frac{1}{n} \sum_{j \geq i} a_j x_j^{k-1})$  is sent to player  $i$  by the coordinator in iteration  $k$ , when it is that player's turn to compute. No detailed information concerning the other players' choices is revealed. Receiving this value, player  $i$  uses her local information, i.e.,  $a_i$ ,  $\tilde{r}_i$  and  $\tilde{\mathcal{X}}_i$ , to update her choice according to (3.3) and then sends it to the coordinator.

**Proposition 3.2.** Under Assumption 2, for  $K \in \mathbb{N}^*$ , there is  $k^* \leq K$  such that  $x^{k^*}$  is an  $\omega(K, n)\Delta$ -PNE of game  $\tilde{\Gamma}$  that satisfies the full  $\omega(K, n)\Delta$ -stability condition, where

$$\omega(K, n) = \frac{\sqrt{2CL_g}M}{m} \sqrt{\frac{n}{K}} + \frac{2L_gM\Delta}{n}, \quad (3.4)$$

where  $C = (d\Delta L_g + 2B_r)M$ .

In particular, if the constant  $K \geq \frac{2C}{m^2 L_g} n^{1+2\delta} + 1$  for some constant  $\delta > 0$ , then there exists some  $k^* \leq K$  such that  $x^{k^*}$  is an  $L_g M \Delta (n^{-\delta} + 2\Delta n^{-1})$ -PNE of game  $\tilde{\Gamma}$  satisfying the full  $L_g M \Delta (n^{-\delta} + 2\Delta n^{-1})$ -stability condition.

**Theorem 3.3.** Under Assumption 2, for a constant  $\delta > 0$  and integer  $K \geq \frac{2C}{m^2 L_g} n^{1+2\delta} + 1$ , let  $x^* \in \mathcal{X}$  be the pure-strategy profile generated by (2.6), where  $\tilde{x}$  is replaced by  $x^{k^*}$  in Proposition 3.2. Then,  $x^*$  is a  $(2L_g M \Delta (n^{-\delta} + \frac{(\sqrt{q}+4)\Delta}{n}) + \frac{L_h M \Delta}{n})$ -PNE of the non-convex game  $\Gamma$ .

For the case in which a ‘‘Shapley-Folkman’’ disaggregation of  $x^{k^*}$  is not easy to obtain, one can use the distributed randomized disaggregation method introduced in Section 2.3 to immediately obtain an  $\tilde{\epsilon}$ -MNE, where  $\tilde{\epsilon}$  is given by the following corollary. However, the quality of approximation is worse than that of a ‘‘Shapley-Folkman’’ disaggregation.

**Proposition 3.4.** Under Assumption 2, for a constant  $\delta > 0$  and integer  $K \geq \frac{2C}{m^2 L_g} n^{1+2\delta} + 1$ , let  $\tilde{\mu} = (\tilde{\mu}_i)_i$  be a profile of independent mixed strategies defined as in Lemma 5.7, where  $\tilde{x}$  is replaced by  $x^{k^*}$  in Proposition 3.2. Then,  $\tilde{\mu}$  is a  $(2L_g M \Delta (n^{-\delta} + \frac{(\sqrt{n}+4)\Delta}{n}) + \frac{L_h M \Delta}{n})$ -MNE of the non-convex game  $\Gamma$ .

## 4 Numerical example

In this section, we consider an example of flexible electric vehicle charging control; the convex version of this problem was studied by Jacquot et al. [20]. Each player must charge the battery of her electric vehicle when she arrives home after work. A player's cost is defined in the form of (2.1) so that it depends on both her own consumption and the aggregate consumption of all the players. Such a design is intended to ensure that the Nash equilibria or  $\epsilon$ -Nash equilibria attain the goal of decreasing the peak demand and smoothing the load curve of the power grid. In this context, the technical constraints of the battery of an electric vehicle limit the number of feasible consumption profiles. They also generally imply non-convex action sets; for example, they only allow for discrete power consumption profiles.

More specifically, one day is divided into peak hours (e.g., 6 am–10 pm) and off-peak hours. The electricity production cost function for total flexible loads of  $\ell^P$  and  $\ell^{OP}$  at peak and off-peak hours are, respectively,  $C^P(\ell^P) = \alpha_0^P \ell^P + \beta_0 (\ell^P)^2$  and  $C^{OP}(\ell^{OP}) = \alpha_0^{OP} \ell^{OP} + \beta_0 (\ell^{OP})^2$ , where  $\alpha_0^P > \alpha_0^{OP} > 0$  and  $\beta_0 > 0$ . Player  $i$ 's action is denoted by  $\ell_i = (\ell_i^P, \ell_i^{OP})$ , where  $\ell_i^P$  (resp.  $\ell_i^{OP}$ ) is the peak (resp. off-peak) consumption of player  $i$ . Player  $i$ 's electricity bill is then defined by

$$b_i(\ell_i, \ell_{-i}) := \frac{C^P(\ell^P)}{\ell^P} \ell_i^P + \frac{C^{OP}(\ell^{OP})}{\ell^{OP}} \ell_i^{OP},$$

where  $\ell^P = \sum_i \ell_i^P$ , and  $\ell^{OP} = \sum_i \ell_i^{OP}$ . Player  $i$ 's cost is then defined by

$$\phi_i(\ell_i, \ell_{-i}) = b_i(\ell_i, \ell_{-i}) + \gamma_i \|\ell_i - \ell_i^{ref}\|^2 \quad (4.1)$$

where  $\gamma_i$  indicates the player's sensitivity to the deviation from her preference  $\ell_i^{ref}$ . In [20], the action set of player  $i$  is the convex compact set  $S_i = \{\ell_i = (\ell_i^P, \ell_i^{OP}) \mid \ell_i^P + \ell_i^{OP} = e_i, \underline{\ell}_i^P \leq \ell_i^P \leq \overline{\ell}_i^P, \underline{\ell}_i^{OP} \leq \ell_i^{OP} \leq \overline{\ell}_i^{OP}\}$ , where  $e_i$  stands for the energy required by player  $i$  to charge an electric vehicle battery and  $\underline{\ell}_i^P$  and  $\overline{\ell}_i^P$  (resp.  $\underline{\ell}_i^{OP}$  and  $\overline{\ell}_i^{OP}$ ) are the minimum and maximum power consumption for player  $i$  during peak (resp. off-peak) hours. However, for various reasons, such as finite choices for charging power or battery protection guidelines that indicate that the charging must be interrupted as infrequently as possible, the players' action sets can be non-convex. For example, in this paper a particular case with the non-convex action set  $S_i^{NC} = \{\ell_i = (\ell_i^P, \ell_i^{OP}) \mid \ell_i^P + \ell_i^{OP} = e_i, \ell_i^P \in \{\underline{\ell}_i^P, \overline{\ell}_i^P\}\}$  is adopted for numerical simulation.

Let us apply Algorithm 1 to this game. The asymptotic performance of the algorithm for large  $n$  is illustrated.

First, game (4.1) is reformulated with uni-dimensional actions. For simplification, suppose that all the players have the same type of electric vehicle (EV), a 2018 Nissan Leaf, with a battery capacity  $e$ , and two charging rate levels  $p_{\min}$  and  $p_{\max}$ . The total consumption of player  $i$  is denoted by  $e_i$  and determined by a parameter  $\tau_i$  as follows:  $e_i = (1 - \tau_i)e = \ell_i^P + \ell_i^{OP}$ , where  $\tau_i \in [0, 1]$  signifies the remaining proportion of energy in the player's battery when she arrives at home. Let  $x_i := \frac{\ell_i^P}{e}$  denote player  $i$ 's strategy in the following reformulation of game (4.1):

$$\tilde{f}_i^{(n)}(x_i, x_{-i}) = \tilde{b}_i^{(n)}(x_i, x_{-i}) + \tilde{\gamma}_i \|x_i - x_i^{ref}\|^2, \quad (4.2)$$

where  $\tilde{\gamma}_i$  indicates how much player  $i$  cares about deviating from her preferred consumption profile and is uniformly set to be  $ne$  for simplification, and

$$\begin{aligned}\tilde{b}_i^{(n)}(x_i, x_{-i}) &= (\alpha_0^P + \beta_0 ne \frac{1}{n} \sum_j (1 - \tau_j) x_j) \ell_i^P + (\alpha_0^{OP} + \beta_0 ne \frac{1}{n} \sum_j (1 - \tau_j) (1 - x_j)) \ell_i^{OP} \\ &= e(1 - \tau_i) \left[ (\alpha_0^P - \alpha_0^{OP} - \beta_0 ne + 2\beta_0 ne \frac{1}{n} \sum_j (1 - \tau_j) x_j) x_i \right. \\ &\quad \left. + \alpha_0^{OP} + \beta_0 ne - \beta_0 ne \frac{1}{n} \sum_j (1 - \tau_j) x_j \right].\end{aligned}$$

The non-convex action set of player  $i$ , introduced in Section 1 as  $S_i^{NC} = \{\ell_i = (\ell_i^P, \ell_i^{OP}) \mid \ell_i^P + \ell_i^{OP} = e_i, \ell_i^P \in \{\underline{\ell}_i^P, \overline{\ell}_i^P\}\}$ , is now translated into  $\mathcal{X}_i = \{\underline{x}_i, \bar{x}_i\} \subset [0, 1]$ , where  $\underline{x}_i$  and  $\bar{x}_i$  correspond, respectively, to charging at  $p_{min}$  and  $p_{max}$ .

By extracting the common factor  $ne(1 - \tau_i)$ , player  $i$ 's cost function becomes

$$f_i^{(n)}(x_i, x_{-i}) := \left\langle g^{(n)}\left(\frac{1}{n} \sum_{j=1}^n (1 - \tau_j) x_j\right), x_i \right\rangle + h^{(n)}\left(\frac{1}{n} \sum_{j=1}^n (1 - \tau_j) x_j\right) + \frac{r_i^{(n)}(x_i)}{1 - \tau_i}, \quad (4.3)$$

where  $g^{(n)}(y) := \frac{\alpha_0^P - \alpha_0^{OP}}{n} + \beta_0 e(2y - 1)$ ,  $h^{(n)}(y) := \frac{\alpha_0^{OP}}{n} + \beta_0 e(1 - y)$ , and  $r_i^{(n)}(y) := \|y - x_i^{ref}\|^2$  for  $y \in \mathbb{R}$ , where  $\alpha_0^P = -4.17 + 0.59 \times 12n$  (€/kWh),  $\alpha_0^{OP} = -4.17 + 0.59 \times 8n$  (€/kWh), and  $\beta_0 = 0.295$  (€/kWh<sup>2</sup>) according to Jacquot et al. [20].

**Simulation parameters** The peak hours are between 6 am and 10 pm, while the remaining hours of the day are off-peak hours. The battery capacity of a 2018 Nissan Leaf is  $e = 40$  kWh. The discrete action set of player  $i$  is determined as follows. The players' arrival times at home are independently generated according to a *Von Mises* distribution with  $\kappa = 1$  between 5 pm and 7 pm. Their departure times are independently generated according to a *Von Mises* distribution with  $\kappa = 1$  between 7 am and 9 am. The proportion  $\tau_i$  of energy in the battery when a player arrives at home is independently generated according to a *Beta* distribution with the parameter  $\beta(2, 5)$ . Once a player arrives at home, she starts charging at one of the two available levels,  $p_{min} = 3.7$  kW or  $p_{max} = 7$  kW. This power level is maintained until the energy requirement  $e_i$  is reached. The arrival and departure time parameters are defined such that the problem is always feasible, i.e., the energy requirement  $e_i$  can always be reached during the charging period by choosing the power level  $p_{max}$ . Players are all assumed to prefer to charge their vehicle as fast as possible, so that  $x_i^{ref} = \bar{x}_i$  for all  $i$ . Fifty instances of the problem are considered for the numerical test. They are obtained by independent simulations of the aforementioned parameters (players' arrival and departure times and remaining energy when they arrive at home).

Algorithm 1 is applied to the EV charging game  $\Gamma^{(n)}$  (4.3) for  $n = 2^s$ ,  $s = 1, \dots, 15$ . For each game  $\Gamma^{(n)}$ , for each iteration  $k$  of the algorithm, let  $x^{(n),k}$  denote the  $k^{th}$  iterate of Algorithm 1 applied to game  $\Gamma^{(n)}$ . Then, the relative error  $\epsilon^{(n),k}$  of  $x^{(n),k}$  is given by

$$\begin{aligned}\epsilon^{(n),k} &:= \min \left\{ \epsilon \geq 0 \mid f_i^{(n)}(x_i^{(n),k}, x_{-i}^{(n),k}) - \inf_{x_i \in \mathcal{X}_i} f_i^{(n)}(x_i, x_{-i}^{(n),k}) \right. \\ &\quad \left. \leq \epsilon \left( \sup_{x_i \in \mathcal{X}_i} f_i^{(n)}(x_i, x_{-i}^{(n),k}) - \inf_{x_i \in \mathcal{X}_i} f_i^{(n)}(x_i, x_{-i}^{(n),k}) \right) \right\}.\end{aligned}$$

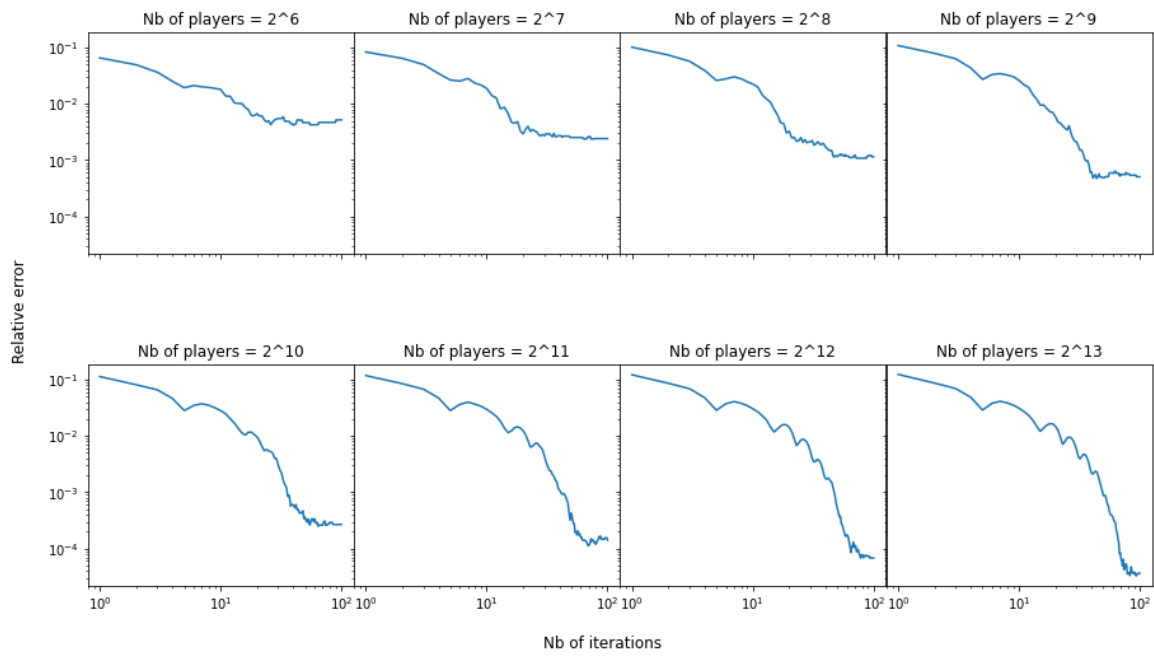


Figure 1: Log-log chart of relative error  $\epsilon^{(n),k}$  (averaged over fifty instances of the problem) as a function of the number of iterations  $k$  (for a fixed number of players  $n = 2^6, 2^7, \dots, 2^{13}$ ).

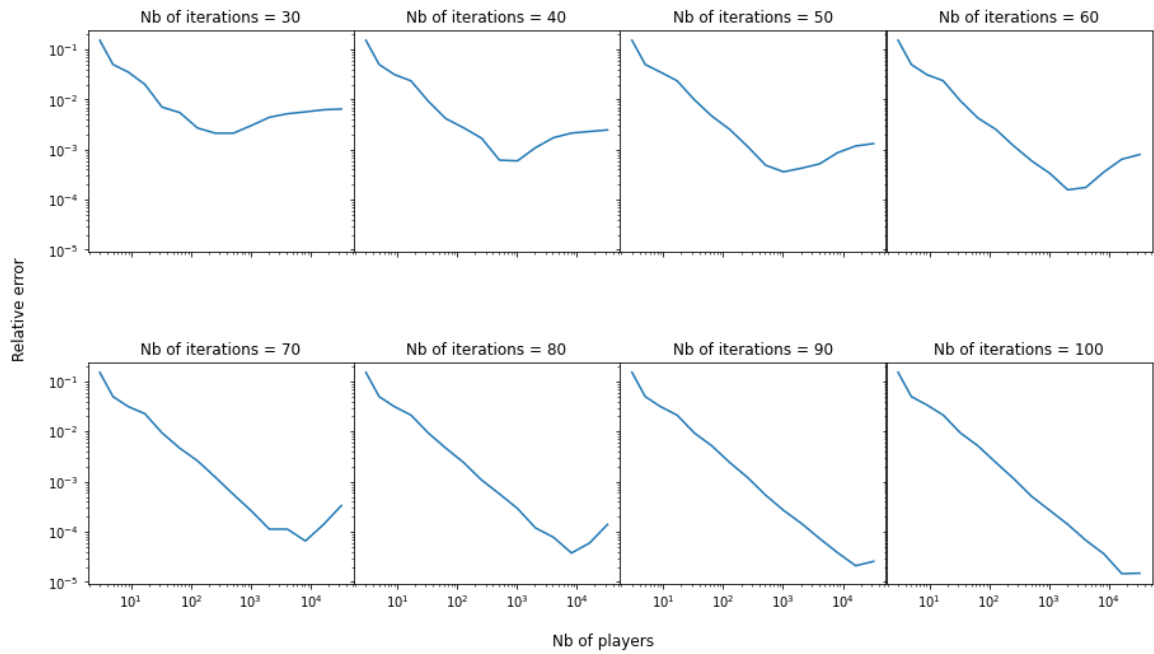


Figure 2: Log-log chart of relative error  $\epsilon^{(n),k}$  (averaged over fifty instances of the problem) as a function of the number of players  $n$  (for a fixed number of iterations  $k = 30, 40, \dots, 90, 100$ ).

As Figure 1 shows, the relative error decreases with the number of iterations to a certain limit. This limiting relative error decreases with the number of players  $n$ . This observation is consistent with equation (3.4) in Proposition 3.2. For Figure 2, according to Proposition 3.2, when the iteration number  $k$  is fixed, due to the domination of the term  $\frac{2L_g M \Delta}{n}$  in equation (3.4) when  $n$  is small,  $\epsilon^{(n),k}$  first decreases linearly with  $n$  before reaching a certain threshold. After that,  $\frac{\sqrt{2CL_g M}}{m} \sqrt{\frac{n}{k}}$  dominates the relative error value so that  $\epsilon^{(n),k}$  may increase with  $n$ . The threshold itself increases with the iteration number  $k$ . This is exactly what Figure 2 shows.

## 5 Conclusion and perspectives

This paper developed an original approach for the study of non-convex games. Non-convexities are widely present in real applications, and they are known to add nontrivial difficulties in the analysis of existence and computation of equilibria. Our approach is restricted to large aggregative games because it is based on the Shapley-Folkman Lemma which essentially exploits the aggregative form with a large number of players. This category covers nevertheless a broad class of games with practical interest, including congestion games. In particular, we illustrated the relevance of this approach with an industrial application to the coordination of electric vehicle charging.

**Distributed and randomized “Shapley-Folkman disaggregation”.** In Section 2.3, a distributed disaggregating method is introduced to obtain a randomized “Shapley-Folkman disaggregation” for the case  $\gamma \leq 1$ . It is extremely fast and easy to carry out: once an  $\epsilon$ -PNE  $\tilde{x}$  is obtained for the convexified game, as well as the profile of generators  $(Z(i, \tilde{x}))_i$ , each player  $i$  randomly chooses one feasible action that is in  $Z(i, \tilde{x})$ , according to the distribution law  $\alpha(i, x)$ . This procedure returns an  $\mathcal{O}(\frac{1}{\sqrt{n^\gamma}})$ -MNE, with the error vanishing when the number of players goes to infinity. However, even if an  $\mathcal{O}(\frac{1}{n^\gamma})$ -PNE can be difficult to obtain using an exact “Shapley-Folkman disaggregation” especially if a large centralized program is involved, for example, to solve (2.6), it would be desirable to find other algorithms that can find better approximations of the Nash equilibria of a non-convex game. Distributed and randomized algorithms are appealing because they can be faster to carry out, they require less coordination and hence are more tractable, and they take advantage of the law of large numbers when  $n$  is large.

**Aggregation and disaggregation of clusters.** In a power grid management setting, flexible agents can be regrouped into clusters, and each cluster is commanded by a so-called aggregator. The EV charging game considered in this paper then takes place between the relatively few aggregators instead of the individuals. This “aggregate game” is different from the EV charging game described in this paper, as the individuals are no longer autonomous but are commanded by their respective aggregators rather than choosing their own charging behaviors. One can build an aggregate model for each aggregator by defining his action set as the set of the aggregate actions of the individuals in his cluster and his cost function as an aggregate of the individuals’ costs. When the clusters are large, it is possible to show, with the help of the Shapley-Folkman Lemma, that the aggregators’ action sets and cost functions

are almost convex. Then, the game admits an  $\epsilon$ -PNE (via Rosen’s existence theorem), and its computation could be relatively easier owing to the small number of players. However, each aggregator then has to reconstruct for each individual under his control a feasible action consistent with their aggregate action at the equilibrium of this “aggregate game.” When the constraints of each flexible individual are non-convex, this aggregation/disaggregation approach can be rather difficult to implement. An original technique based on the Shapley-Folkman Lemma is proposed in Hreinsson et al. [19] within the optimization framework, with applications to the management of consumption flexibilities in power systems.

**Acknowledgments.** We are grateful to J. Frédéric Bonnans and Rebecca Jeffers for stimulating discussions and comments. We particularly thank the reviewers and the associated editor for their very relevant remarks, which have helped greatly to improve the paper.

## Appendix A: PNE in l.s.c. convex games

Since we have not found a specific reference of the extension of Rosen’s theorem to the l.s.c. case, we prefer to provide our own proof for the sake of completeness.

**Lemma 5.1.** *Let  $R$  be a nonempty convex compact set in  $\mathbb{R}^n$ . If the real-valued function  $\rho(x, y)$  defined on  $R \times R$  is continuous in  $x$  on  $R$  for any fixed  $y$  in  $R$ , l.s.c. in  $(x, y)$  on  $R \times R$ , and convex in  $y$  on  $R$  for any fixed  $x$  in  $R$ , then the set-valued map  $\zeta : R \rightarrow R$ ,  $x \mapsto \zeta(x) = \operatorname{argmin}_{z \in R} \rho(x, z)$  has a fixed point.*

*Proof.* Kakutani’s fixed-point theorem [23] will be applied for the proof. First, let us show that  $\zeta$  is a Kakutani map, i.e., (i)  $\Gamma$  is upper semicontinuous (u.s.c.) in the set map sense and (ii) for all  $x \in R$ ,  $\zeta(x)$  is non-empty, compact and convex.

(i) Fix  $x \in R$ . On the one hand, since  $\rho(x, y)$  is convex w.r.t  $y$ ,  $\zeta(x)$  is convex. On the other hand,  $\rho(x, y)$  is l.s.c in  $y$ , while  $R$  is compact; hence  $\rho(x, y)$  can attain its minimum w.r.t  $y$  and  $\zeta(x)$  is thus nonempty. Besides, since  $\rho$  is l.s.c.,  $\zeta(x) = \{y | \rho(x, y) \leq \min_{z \in R} \rho(x, z)\}$  is a closed subset of compact set  $R$ ; hence it is compact.

(ii) Recall that the set-valued map  $\zeta$  is u.s.c. if, for any open set  $w \subset R$ , set  $\{x \in R | \zeta(x) \subset w\}$  is open.

Let us first show by contradiction that, for arbitrary  $x_0 \in R$ , for any  $\epsilon > 0$ , there exists  $\delta > 0$  such that for all  $z \in B(x_0, \delta)$ ,  $\zeta(z) \subset \zeta(x_0) + B(0, \epsilon)$ . If this is not true, then there exists  $\epsilon_0 > 0$  and, for all  $n \in \mathbb{N}^*$ , the point  $z_n \in B(x_0, \frac{1}{n})$  such that there exists  $y_n \in \zeta(z_n)$  with  $d(y_n, \zeta(x_0)) > \epsilon_0$ . Since the sequence  $\{y_n\}$  is in the compact set  $R$ , it has a subsequence  $y_{\phi(n)}$  converging to some  $\bar{y}$  in  $R$ , and  $d(\bar{y}, \zeta(x_0)) \geq \epsilon_0$ . Then, for all  $y \in R$ ,

$$\rho(x_0, \bar{y}) \leq \liminf_{n \rightarrow \infty} \rho(z_{\phi(n)}, y_{\phi(n)}) \leq \liminf_{n \rightarrow \infty} \rho(z_{\phi(n)}, y) = \rho(x_0, y),$$

where the first inequality is due to the lower semicontinuity of  $\rho$  in  $(x, y)$ , the second inequality is due to the definition of  $\zeta(z_{\phi(n)})$ , and the third equality is due to the continuity of  $\rho$  in  $x$ . This shows that  $\bar{y} \in \zeta(x_0)$ , in contradiction with the fact that  $d(\bar{y}, \zeta(x_0)) \geq \epsilon_0$ .

Now fix arbitrarily an open set  $w \subset R$  and some  $x_0 \in R$  such that  $\zeta(x_0) \subset w$ . Since  $\zeta(x_0)$  is compact while  $w$  is open, there exists  $\epsilon > 0$  such that  $\zeta(x_0) + B(0, \epsilon) \subset w$ . According to the result of the previous paragraph, for this particular  $\epsilon$ , there exists  $\delta > 0$  such that



$\zeta(z) \subset \zeta(x_0) + B(0, \epsilon) \subset w$  for all  $z \in B(x_0, \delta)$ . This means that  $B(x_0, \delta) \subset \{x \in R \mid \zeta(x) \subset w\}$ . As a result, the set  $\{x \in R \mid \zeta(x) \subset w\}$  is open.

Finally, according to Kakutani's fixed-point theorem, there exists  $\tilde{x} \in R$  such that  $\tilde{x} \in \zeta(\tilde{x})$ .  $\square$

**Definition 5.2.** A family of real-valued functions  $\{f(\cdot, y) : \mathcal{X} \rightarrow \mathbb{R} \mid y \in \mathcal{Y}\}$  indexed by  $y$ , with  $\mathcal{X} \subset \mathbb{R}^{d_1}$  and  $\mathcal{Y} \subset \mathbb{R}^{d_2}$ , is *uniformly equicontinuous* if, for all  $\epsilon > 0$ , there exists  $\delta$  such that, for all  $y \in \mathcal{Y}$ ,  $\|f(x_1, y) - f(x_2, y)\| \leq \epsilon$  whenever  $\|x_1 - x_2\| \leq \delta$ .

**Theorem 5.3** (Existence of PNE in l.s.c. convex games). *In an  $n$ -player game  $\Gamma$ , if, for each player  $i \in \{1, \dots, n\}$ ,*

- (1) *the action set  $\mathcal{X}_i$  is a convex compact subset of  $\mathbb{R}^d$ ,*
  - (2) *the cost function  $f_i(x_i, x_{-i}) : \mathcal{X}_i \times \prod_{j \neq i} \mathcal{X}_j \rightarrow \mathbb{R}$  is convex and l.s.c. in  $x_i \in \mathcal{X}_i$  for any fixed  $x_{-i} \in \prod_{j \neq i} \mathcal{X}_j$ , and*
  - (3) *the family of functions  $\{f_i(x_i, \cdot) : \prod_{j \neq i} \mathcal{X}_j \rightarrow \mathbb{R} \mid x_i \in \mathcal{X}_i\}$  are uniformly equicontinuous,*
- then  $\Gamma$  admits a PNE.*

*Proof.* Define the function  $\rho(x, y) : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$  by  $\rho(x, y) = \sum_{i=1}^n f_i(y_i, x_{-i})$ , where  $\mathcal{X} = \prod_i \mathcal{X}_i$ . It is easy to see that a fixed point of the set-valued map  $\zeta : \mathcal{X} \rightarrow \mathcal{X}$ ,  $x \mapsto \zeta(x) = \arg \min_{z \in R} \rho(x, z)$  is a Nash equilibrium of game  $\Gamma$ .

In order to apply Lemma 5.1, one needs to show the following: (i)  $\rho(x, y)$  is continuous in  $x$  for each fixed  $y$ ; (ii)  $\rho(x, y)$  is l.s.c. in  $(x, y)$ ; (iii)  $\rho(x, y)$  is convex in  $y$  for each fixed  $x$ .

Results (i) and (iii) follow straightforwardly from the definition of  $\rho$ .

For (ii), first note that, by the uniform equicontinuity of  $\{f_i(x_i, \cdot) : \prod_{j \neq i} \mathcal{X}_j \rightarrow \mathbb{R} \mid x_i \in \mathcal{X}_i\}$  for each  $i$  and the fact that  $n$  is finite,  $\{\rho(\cdot, y), y \in R\}$  is uniformly equicontinuous. Let  $(x^k, y^k)$  be a sequence in  $\mathcal{X} \times \mathcal{X}$  indexed by  $k$  that converges to  $(x, y) \in \mathcal{X} \times \mathcal{X}$ . Then,

$$\begin{aligned} \underline{\lim}_{k \rightarrow \infty} (\rho(x^k, y^k) - \rho(x, y)) &= \underline{\lim}_{k \rightarrow \infty} (\rho(x^k, y^k) - \rho(x, y^k) + \rho(x, y^k) - \rho(x, y)) \\ &= \underline{\lim}_{k \rightarrow \infty} (\rho(x, y^k) - \rho(x, y)) \\ &\geq 0, \end{aligned}$$

where the second equality is due to the uniform equicontinuity of  $\{\rho(\cdot, y), y \in \mathcal{X}\}$  and the last inequality is because  $\rho(x, y)$  is l.s.c. in  $y$  for any fixed  $x$ .  $\square$

**Remark 5.1.** The property (3) is weaker than the condition that  $f_i$  is continuous on  $\mathcal{X}$ . Indeed, since  $\mathcal{X}$  is compact,  $f_i(x_i, x_{-i})$  is uniformly continuous on  $\mathcal{X}_i \times \prod_{j \neq i} \mathcal{X}_j$  which implies the equicontinuity of  $\{f_i(x_i, \cdot) : \prod_{j \neq i} \mathcal{X}_j \rightarrow \mathbb{R} \mid x_i \in \mathcal{X}_i\}$ . In other words, Rosen's theorem on the existence of convex continuous games with compact convex actions sets is a corollary of Theorem 5.3.

## Appendix B: Other proofs and lemmata

**Lemma 5.4** (Shapley-Folkman Lemma [41]). *For  $n$  compact subsets  $S_1, \dots, S_n$  of  $\mathbb{R}^q$ , let  $x \in \text{conv} \sum_{i=1}^n S_i = \sum_{i=1}^n \text{conv} S_i$ , where  $\text{conv}$  signifies the convex hull, and the sum over sets is to be understood as a Minkowski sum. Then,*

- *there is a point  $x_i \in \text{conv} S_i$  for each  $i$  such that  $x = \sum_{i=1}^n x_i$ , and  $x_i \in S_i$  except for at most  $q$  values of  $i$ ; and*
- *there is a point  $y_i \in S_i$  for each  $i$  such that  $\|x - \sum_{i=1}^n y_i\|_{\mathbb{R}^q} \leq \sqrt{\min\{q, n\}}d$ , where  $d$  denotes the maximal diameter of  $S_i$ .*

In the proofs of Lemmata 5.5, 5.6, and 2.6, in order to simplify the notation,  $i$  and  $x_{-i} \in \tilde{\mathcal{X}}_{-i}$  are arbitrarily fixed. Index  $i$  and the parameter  $x_{-i}$  are thus omitted in  $f_i$ ,  $\tilde{f}_i$ ,  $\mathcal{X}_i$ ,  $\tilde{\mathcal{X}}_i$  and  $Z(i, \cdot)$ .

**Lemma 5.5.** *Under Assumption 1, for each  $x_{-i} \in \tilde{\mathcal{X}}_{-i}$ ,*

- (1)  $\tilde{f}_i(x_i, x_{-i}) \leq f_i(x_i, x_{-i})$  for all  $x_i \in \mathcal{X}_i$ ;
- (2) the infimum in (2.2) can be attained, i.e., it is in fact a minimum for all  $x_i \in \mathcal{X}_i$ ;
- (3) the function  $\tilde{f}_i(\cdot, x_{-i})$  is l.s.c. and convex on  $\tilde{\mathcal{X}}_i$ , and  $\text{conv}(\text{epi } f_i(\cdot, x_{-i})) = \text{epi } \tilde{f}_i(\cdot, x_{-i}) = \overline{\text{conv}}(\text{epi } f_i(\cdot, x_{-i}))$ ;
- (4) both  $\tilde{f}_i(\cdot, x_{-i})$  and  $f_i(\cdot, x_{-i})$  attain their minima on  $\tilde{\mathcal{X}}_i$  and  $\mathcal{X}_i$  respectively; and

$$\min_{\tilde{x}_i \in \tilde{\mathcal{X}}_i} \tilde{f}_i(\tilde{x}_i, x_{-i}) = \min_{x_i \in \mathcal{X}_i} f_i(x_i, x_{-i}). \quad (5.1)$$

*In particular, if  $\tilde{x}_i \in \arg \min_{y_i \in \tilde{\mathcal{X}}_i} \tilde{f}_i(y_i, \tilde{x}_{-i})$ , then  $Z(i, \tilde{x}) \subset \arg \min_{y_i \in \mathcal{X}_i} f_i(y_i, \tilde{x}_{-i})$ , where  $Z(i, \tilde{x})$  is an arbitrary generator for  $(i, \tilde{x})$  defined in Definition 2.2.*

*Proof of Lemma 5.5.*

The lemma is a particular case of more general results well-known in the field of convex analysis that have been shown in various works, such as [16, Lemma X.1.5.3]. We will provide a proof for this particular case for the sake of completeness.

(1) For  $x \in X$ , in the definition of  $\tilde{f}(x)$ , take  $x^k = x$ ,  $\alpha^k = \frac{1}{d+1}$  for all  $k$ . By definition,  $\tilde{f}(x) \leq \sum_{k=1}^{d+1} \alpha^k f(x^k) = f(x)$ .

(2) Suppose that  $((\alpha^{k,n})_k, (x^{k,n})_k)_{n \in \mathbb{N}}$  is a minimizing sequence for  $\tilde{f}(\tilde{x})$ , i.e.,  $\tilde{f}(\tilde{x}) = \lim_{n \rightarrow \infty} \sum_{k=1}^{d+1} \alpha^{k,n} f(x^{k,n})$ , with  $((\alpha^{k,n})_k, (x^{k,n})_k)_{n \in \mathbb{N}}$  satisfying the conditions in (2.2). Since  $(\alpha^{1,n}) \in [0, 1]$  for all  $n$ , it has a convergent subsequence  $\alpha^{1, \phi_1(n)}$ , which converges to some  $\alpha^1$ . Consider sequence  $\alpha^{2, \phi_2(n)}$  which has a subsequence  $\alpha^{2, \phi_2(n)}$  converging to some  $\alpha^2$ . Note that  $\phi_2(n)$  is a subsequence of  $\phi_1(n)$ . Repeat this operation  $d+1$  times and obtain the subsequences  $\phi_1(n), \dots, \phi_{d+1}(n)$  such that  $\alpha^{k, \phi_k(n)}$  converges to  $\alpha^k$ , for  $k = 1, \dots, d+1$ . Consider  $x^{1, \phi_{d+1}(n)}$ , which is in the compact set  $X$ . It has a convergent subsequence  $x^{1, \phi_{d+2}(n)}$  converging to  $x^1 \in X$ . Again, take a subsequence  $\phi_{d+3}(n)$  such that

$x^{2, \phi_{d+3}(n)}$  converges to  $x^k$ , and so on. Finally, one obtains a subsequence  $\phi_{2d+2}(n)$  of  $\mathbb{N}$  such that

$$\tilde{f}(\tilde{x}) = \lim_{n \rightarrow \infty} \sum_{k=1}^{d+1} \alpha^{k, \phi_{2d+2}(n)} f(x^{k, \phi_{2d+2}(n)}), \quad (5.2)$$

$$\alpha^k = \lim_{n \rightarrow \infty} \alpha^{k, \phi_{2d+2}(n)}, \quad \alpha^k \in [0, 1], \quad k = 1, 2, \dots, d+1, \quad (5.3)$$

$$\sum_{k=1}^{d+1} \alpha^k = \lim_{n \rightarrow \infty} \sum_{k=1}^{d+1} \alpha^{k, \phi_{2d+2}(n)} = 1, \quad (5.4)$$

$$x^k = \lim_{n \rightarrow \infty} x^{k, \phi_{2d+2}(n)}, \quad x^k \in \mathcal{X}, \quad k = 1, 2, \dots, d+1, \quad (5.5)$$

$$\sum_{k=1}^{d+1} \alpha^k x^k = \lim_{n \rightarrow \infty} \sum_{k=1}^{d+1} \alpha^{k, \phi_{2d+2}(n)} x^{k, \phi_{2d+2}(n)} = \lim_{n \rightarrow \infty} \tilde{x} = \tilde{x}. \quad (5.6)$$

Then,

$$\begin{aligned} \sum_{k=1}^{d+1} \alpha^k f(x^k) &\leq \lim_{n \rightarrow \infty} \sum_{k=1}^{d+1} \alpha^{k, \phi_{2d+2}(n)} f(x^{k, \phi_{2d+2}(n)}) = \lim_{n \rightarrow \infty} \sum_{k=1}^{d+1} \alpha^{k, \phi_{2d+2}(n)} f(x^{k, \phi_{2d+2}(n)}) \\ &= \tilde{f}(\tilde{x}) \leq \sum_{k=1}^{d+1} \alpha^k f(x^k). \end{aligned}$$

where the first inequality is due to (5.5), the second equality is due to (5.3), the third equality is due to (5.2) and the fourth inequality due to (5.4), (5.6) and (2.2). This shows that  $\tilde{f}(\tilde{x}) = \sum_{k=1}^{d+1} \alpha^k f(x^k)$ , i.e.,  $(\alpha^k, x^k)_{k=1}^{d+1}$ , is a minimizer.

(3) On the one hand, for all  $(x, y) \in \text{conv}(\text{epi } f)$ , by the Caratheodory theorem [6, Proposition 1.2.1], there exist  $(x^k, y^k) \in \text{epi } f$ ,  $k = 1, \dots, d+1$  such that  $(x, y) = \sum_{k=1}^{d+1} \alpha^k (x^k, y^k)$ , with  $\alpha \in \mathcal{S}_d$ . Hence,  $y^k \geq f(x^k)$ , and  $y = \sum_{k=1}^{d+1} \alpha^k y^k \geq \sum_{k=1}^{d+1} \alpha^k f(x^k) \geq \tilde{f}(x)$ . This shows that  $(x, y) \in \text{epi } \tilde{f}$ . Therefore,  $\text{conv}(\text{epi } f) \subset \text{epi } \tilde{f}$ . Recall that  $f$  is l.s.c.; hence,  $\text{epi } f$  is a closed set and thus so is  $\text{conv}(\text{epi } f)$ . Thus,  $\overline{\text{conv}}(\text{epi } f) \subset \text{epi } \tilde{f}$ .

On the other hand, for all  $(x, y) \in \text{epi } \tilde{f}$ ,  $y \geq \tilde{f}(x)$ . Let  $((\alpha^{k,n})_k, (x^{k,n})_k)_{n \in \mathbb{N}}$  be the minimizing sequence for  $\tilde{f}(x)$ , i.e.,  $\tilde{f}(x) = \lim_{n \rightarrow \infty} \sum_{k=1}^{d+1} \alpha^{k,n} f(x^{k,n})$ , with  $\alpha^{k,n}, x^{k,n}$  satisfying the conditions in (2.2). Then,  $y = \lim_{n \rightarrow \infty} \sum_{k=1}^{d+1} \alpha^{k,n} (f(x^{k,n}) + \frac{\delta}{d+1})$ , where  $\delta = y - \tilde{f}(x) \geq 0$ . Denote  $y^n = \sum_{k=1}^{d+1} \alpha^{k,n} (f(x^{k,n}) + \frac{\delta}{d+1})$ . Then,  $(x, y^n) \in \text{conv}(\text{epi } f)$ , and  $\lim_{n \rightarrow \infty} (x, y^n) = (x, y)$ . This means that  $(x, y) \in \overline{\text{conv}}(\text{epi } f)$  and, therefore,  $\text{epi } \tilde{f} \subset \overline{\text{conv}}(\text{epi } f)$ .

In conclusion,  $\text{epi } \tilde{f}(\cdot) = \overline{\text{conv}}(\text{epi } f(\cdot))$ , which implies that the epigraph of  $\tilde{f}$  is closed and convex. Thus,  $\tilde{f}$  is l.s.c. and convex on  $\tilde{\mathcal{X}}$ .

(4) By the lower semicontinuity of  $\tilde{f}$  and  $f$  on the compact sets  $\tilde{\mathcal{X}}$  and  $X$ , their minima can be attained. The equality (5.1) is thus clear by the definition in (2.2).  $\square$

**Remark 5.2.** If  $f_i(\cdot, x_{-i})$  is not l.s.c, the inclusion relationship in Lemma 5.5(2) can be strict, as shown, respectively, by the following two examples of dimension 1.

- $\mathcal{X} = \{0\} \cup \{\pm \frac{1}{z}\}_{z \in \mathbb{N}^*}$ ,  $f(x) = |x|$  for  $x \in \mathcal{X} \setminus \{0\}$ , and  $f(0) = 1$ . Then,  $\tilde{f}(x) = |x|$ , for all  $x \in \tilde{\mathcal{X}} = [-1, 1]$ , and  $\text{conv}(\text{epi } f) \subsetneq \text{epi } \tilde{f}$ .

- $\mathcal{X} = [0, 1]$ ,  $f(x) = 0$  for  $x \neq 0$ , and  $f(0) = 1$ . Then,  $\tilde{f}(x) = f(x)$  for all  $x \in [0, 1]$ , and  $\text{epi } \tilde{f} \subsetneq \overline{\text{conv}}(\text{epi } f)$ .

**Lemma 5.6.** *Under Assumption 1, for any profile  $\tilde{x} \in \tilde{\mathcal{X}}$ , for any player  $i$ , for all  $x_i \in Z(i, \tilde{x})$ ,*

$$(1) f_i(x_i, \tilde{x}_{-i}) = \tilde{f}_i(x_i, \tilde{x}_{-i});$$

$$(2) \text{ for any } h \in \partial_i \tilde{f}_i(\tilde{x}_i, \tilde{x}_{-i}),$$

$$f_i(x_i, \tilde{x}_{-i}) = \tilde{f}_i(x_i, \tilde{x}_{-i}) = \tilde{f}_i(\tilde{x}_i, \tilde{x}_{-i}) + \langle h, x_i - \tilde{x}_i \rangle. \quad (5.7)$$

*Proof of Lemma 5.6.* Let  $\{x^1, \dots, x^{d+1}\} \subset \mathcal{X}$  be a generator of  $(\tilde{x}, \tilde{f}(\tilde{x}))$  and let  $\alpha \in \mathcal{S}_d$  be their corresponding weights.

(1) Suppose that there is a  $k$  such that  $f(x^k) > \tilde{f}(x^k)$ . Then, there exists  $(y^l)_l$  in  $\mathcal{X}$  and  $\beta \in \mathcal{S}_d$  such that  $x^k = \sum_l \beta^l y^l$  and  $\tilde{f}(x^k) = \sum_l \beta^l f(y^l) < f(x^k)$ . In consequence,  $\tilde{f}(\tilde{x}) = \sum_m \alpha^m f(x^m) > \sum_{m \neq k} \alpha^m f(x^m) + \sum_l \alpha^k \beta^l f(y^l)$ , while  $\sum_{m \neq k} \alpha^m x^m + \sum_l \alpha^k \beta^l y^l = \tilde{x}$  and  $\sum_{m \neq k} \alpha^m + \sum_l \alpha^k \beta^l = 1$ , contradicting the definition of  $\tilde{f}(\tilde{x})$ .

(2) By the definition of the subdifferential, one has

$$\tilde{f}(x^k) \geq \tilde{f}(\tilde{x}) + \langle h, x^k - \tilde{x} \rangle, \quad \forall k = 1, \dots, d+1. \quad (5.8)$$

Multiplying (5.8) by  $\alpha^k$  for each  $k$  and adding the  $d+1$  inequalities yield

$$\sum_{k=1}^{d+1} \tilde{f}(x^k) \geq \tilde{f}(\tilde{x}) + \langle h, \sum_k \alpha^k x^k - \tilde{x} \rangle \Leftrightarrow \tilde{f}(\tilde{x}) \geq \tilde{f}(\tilde{x}). \quad (5.9)$$

If, for at least one  $k$ , the inequality in (5.8) is strict, then the inequalities in (5.9) are strict as well, which is absurd. Therefore, for each  $k$ ,  $\tilde{f}(x^k) = \tilde{f}(\tilde{x}) + \langle h, x^k - \tilde{x} \rangle$ .  $\square$

*Proof of Lemma 2.6.* First note that  $\tilde{x}$  is in  $\text{ri}(\text{conv } Z(\tilde{x}))$ , the relative interior of  $\text{conv } Z(\tilde{x})$ . Hence, for  $t > 0$  small enough,  $\tilde{x} \pm t(x - \tilde{x})$  is in  $\text{ri}(\text{conv } Z(\tilde{x})) \subset \tilde{\mathcal{X}}$ . By (2.5),  $\langle h, \tilde{x} \pm t(x - \tilde{x}) - \tilde{x} \rangle \geq -\eta \|\tilde{x} \pm t(x - \tilde{x}) - \tilde{x}\|$ , which yields  $|\langle h, x - \tilde{x} \rangle| \leq \eta \|x - \tilde{x}\|$ . Then, by Lemma 5.6,  $|\tilde{f}(x) - \tilde{f}(\tilde{x})| = |\langle h, x - \tilde{x} \rangle| \leq \eta \|x - \tilde{x}\|$ .  $\square$

*Proof of Theorem 2.7.* For each  $i \in N$ , define a set  $E_i(\tilde{x}) := A_i Z(i, \tilde{x})$  in  $\mathbb{R}^q$ . Since  $\tilde{x}_i \in \text{conv}(Z(i, \tilde{x}))$ , one has  $\sum_{i \in N} A_i \tilde{x}_i \in \sum_{i \in N} \text{conv}(E_i(\tilde{x})) = \text{conv}\left(\sum_{i \in N} E_i(\tilde{x})\right)$  by the linearity of the  $A_i$ 's. According to the Shapley-Folkman Lemma, there exists  $e_i \in \text{conv}(E_i(\tilde{x}))$  for each  $i \in N$ , and a subset  $I \subset N$  with  $|I| \leq q$ , such that (i)  $\sum_{i \in N} A_i \tilde{x}_i = \sum_{i \in N} e_i$ , and (ii)  $e_i \in E_i(\tilde{x})$  for all  $i \notin I$ . Thus, for all  $i \notin I$ , there exists  $\bar{x}_i \in Z(i, \tilde{x})$ , such that  $e_i = A_i \bar{x}_i$ . For all  $i \in I$ , take arbitrarily  $\bar{x}_i \in Z(i, \tilde{x})$ . Then,

$$\begin{aligned} \left\| \sum_{i \in N} A_i \tilde{x}_i - \sum_{i \in N} A_i \tilde{x}_i^* \right\| &\leq \left\| \sum_{i \in N} A_i \tilde{x}_i - \sum_{i \in N} A_i \bar{x}_i \right\| = \left\| \sum_{i \in N} e_i - \sum_{i \in N} A_i \bar{x}_i \right\| = \left\| \sum_{i \in I} A_i (\tilde{x}_i - \bar{x}_i) \right\| \\ &\leq \sqrt{q} M \Delta. \end{aligned} \quad (5.10)$$

Now, for all  $i$ ,  $x_i^* \in Z(i, \tilde{x})$ , so that it satisfies

$$f_i(x_i^*, \tilde{x}_{-i}) \leq \tilde{f}_i(\tilde{x}_i, \tilde{x}_{-i}) \leq \tilde{f}_i(x_i, \tilde{x}_{-i}) + \epsilon + \eta \leq f_i(x_i, \tilde{x}_{-i}) + \epsilon + \eta, \quad \text{for all } x_i \in \mathcal{X}_i, \quad (5.11)$$

according to Lemma 5.6.(1), Lemma 2.6 and Lemma 5.5.(1).

Recall that  $f_i(x) = \theta_i(x_i, \frac{1}{n} \sum_{j \in N} A_j x_j)$ . Hence, for any  $x_i \in \mathcal{X}_i$

$$\begin{aligned} f_i(x_i, \tilde{x}_{-i}) &= \theta_i\left(x_i, \frac{1}{n} A_i x_i + \frac{1}{n} \sum_{j \in N_{-i}} A_j \tilde{x}_j\right) \\ &= \theta_i\left(x_i, \frac{1}{n} A_i x_i + \frac{1}{n} \sum_{j \in N_{-i}} A_j x_j^* + \frac{1}{n} A_i (x_i^* - \tilde{x}_i) + \frac{1}{n} \sum_{j \in N} A_j (\tilde{x}_j - x_j^*)\right) \\ &= \theta_i\left(x_i, \frac{1}{n} A_i x_i + \frac{1}{n} \sum_{j \in N_{-i}} A_j x_j^* + \delta_i\right) - \theta_i\left(x_i, \frac{1}{n} A_i x_i + \frac{1}{n} \sum_{j \in N_{-i}} A_j x_j^*\right) + f_i(x_i, x_{-i}^*), \end{aligned}$$

where  $\delta_i := \frac{1}{n} A_i (x_i^* - \tilde{x}_i) + \frac{1}{n} \sum_{j \in N} A_j (\tilde{x}_j - x_j^*)$ .

By (5.10),  $\|\delta_i\| \leq \frac{(\sqrt{q}+1)M\Delta}{n}$ . Using Assumption 1, we can show that, for any  $x_i \in \mathcal{X}_i$ ,

$$|f_i(x_i, x_{-i}^*) - f_i(x_i, \tilde{x}_{-i})| \leq H \left( \frac{(\sqrt{q}+1)M\Delta}{n} \right)^\gamma.$$

Injecting this result into (5.11) yields

$$f_i(x_i^*, x_{-i}^*) \leq f_i(x_i, x_{-i}^*) + \epsilon + \eta + 2H \left( \frac{(\sqrt{q}+1)M\Delta}{n} \right)^\gamma, \quad \forall x_i \in \mathcal{X}_i, \forall i \in N. \quad (5.12)$$

□

**Lemma 5.7.** *Under Assumption 1, suppose that  $\tilde{x} \in \tilde{\mathcal{X}}$  is an  $\epsilon$ -PNE in  $\tilde{\Gamma}$  satisfying the  $\eta$ -stability condition with respect to  $(Z(i, \tilde{x}))_i$ , where  $Z(i, \tilde{x}) = \{x_i^1, x_i^2, \dots, x_i^{l_i}\}$  with  $1 \leq l_i \leq d+1$  and  $\tilde{x}_i = \sum_{l=1}^{l_i} \alpha_i^l x_i^l$ , where  $\alpha \in \mathcal{S}_{l_i-1}$ . Each player  $i$  plays a mixed strategy independently, i.e., a random action  $X_i$  following the distribution  $\tilde{\mu}_i$  over  $\mathcal{X}_i$  defined by  $\mathbb{P}(X_i = x_i^l) = \alpha_i^l$ . In other words,*

$$\tilde{\mu}_i = \sum_{l=1}^{l_i} \alpha_i^l \delta_{x_i^l}, \quad (5.13)$$

where  $\delta_{x_i^l}$  stands for the Dirac distribution on  $x_i^l$ . Then,

$$\mathbb{E} \left\| \sum_{i \in N} A_i \tilde{x}_i - \sum_{i \in N} A_i X_i \right\| \leq \sqrt{n} M \Delta.$$

*Proof.* By the independence of  $X_i$ , the  $A_i X_i$  are independent of each other. From the definition of  $\tilde{\mu}_i$ ,  $\mathbb{E}(A_i X_i) = A_i \tilde{x}_i$ . Therefore,

$$\left( \mathbb{E} \left\| \sum_{i \in N} A_i \tilde{x}_i - \sum_{i \in N} A_i X_i \right\| \right)^2 \leq \mathbb{E} \left[ \left\| \sum_{i \in N} A_i \tilde{x}_i - \sum_{i \in N} A_i X_i \right\|^2 \right] = \sum_{i \in N} \text{Var}(A_i X_i) \leq n M^2 \Delta^2,$$

where the first inequality is due to Jensen's inequality. □

*Proof of Proposition 2.8.* By the same arguments used in the proof of Theorem 2.7, one has

$$|f_i(x_i, \tilde{x}_{-i}) - f_i(x_i, X_{-i})| \leq H \|\delta_i(X)\|^\gamma,$$

where  $\delta_i(X) := \frac{1}{n}A_i(X_i - \tilde{x}_i) + \frac{1}{n}\sum_{j \in N} A_j(\tilde{x}_j - X_j)$ . By Lemma 5.7,

$$\mathbb{E}\|\delta_i(X)\| \leq \frac{1 + \sqrt{n}}{n}M\Delta.$$

Besides, since  $X_i$  takes values in  $Z(i, \tilde{x})$ ,

$$\begin{aligned} f_i(X_i, X_{-i}) &= f_i(X_i, X_{-i}) - f_i(X_i, \tilde{x}_{-i}) + f_i(X_i, \tilde{x}_{-i}) \\ &\leq f_i(X_i, X_{-i}) - f_i(X_i, \tilde{x}_{-i}) + f_i(x_i, \tilde{x}_{-i}) + \epsilon + \eta \\ &= f_i(X_i, X_{-i}) - f_i(X_i, \tilde{x}_{-i}) + f_i(x_i, \tilde{x}_{-i}) - f_i(x_i, X_{-i}) + f_i(x_i, X_{-i}) + \epsilon + \eta, \end{aligned}$$

so that

$$\begin{aligned} f_i(X_i, X_{-i}) - f_i(x_i, X_{-i}) &\leq |f_i(X_i, X_{-i}) - f_i(X_i, \tilde{x}_{-i})| + |f_i(x_i, \tilde{x}_{-i}) - f_i(x_i, X_{-i})| + \epsilon + \eta \\ &\leq 2H(\delta_i(X))^\gamma + \epsilon + \eta. \end{aligned}$$

Therefore,

$$\mathbb{E}[f_i(X_i, X_{-i}) - f_i(x_i, X_{-i})] \leq 2H \mathbb{E}[\|\delta_i(X)\|^\gamma] + \epsilon + \eta \leq 2H \left( \frac{(\sqrt{n} + 1)M\Delta}{n} \right)^\gamma + \epsilon + \eta.$$

□

*Proof of Lemma 3.1.* (1) First show that, for any fixed  $x_i \in \mathcal{X}_i$ , the function  $\theta_i(x_i, y) := \langle g(y + \frac{a_i}{n}(x_i^0 - x_i)), x_i \rangle + \ell_i(x_i)$  is  $L_g\Delta$ -Lipschitz in  $y$  on  $\Omega$ . For this, fix  $x_i \in \mathcal{X}_i$ . For any  $y$  and  $y'$  in  $\Omega$ ,

$$\begin{aligned} |\theta_i(x_i, y') - \theta_i(x_i, y)|^2 &= \left| \left\langle g\left(y' + \frac{a_i}{n}(x_i^0 - x_i)\right) - g\left(y + \frac{a_i}{n}(x_i^0 - x_i)\right), x_i \right\rangle \right|^2 \\ &\leq \left\| g\left(y' + \frac{a_i}{n}(x_i^0 - x_i)\right) - g\left(y + \frac{a_i}{n}(x_i^0 - x_i)\right) \right\|^2 \Delta^2 \\ &= \sum_{t=1}^d \left( g_t\left(y'_t + \frac{a_i}{n}(x_{i,t}^0 - x_{i,t})\right) - g_t\left(y_t + \frac{a_i}{n}(x_{i,t}^0 - x_{i,t})\right) \right)^2 \Delta^2 \\ &\leq L_g^2 \Delta^2 \sum_{t=1}^d (y'_t - y_t)^2 \\ &= L_g^2 \Delta^2 \|y' - y\|^2, \end{aligned}$$

where the first inequality results from the Cauchy-Schwarz inequality, while the second inequality is true because  $g_t$  is  $L_{g_t}$ -Lipschitz.

(2) It is easy to see that  $|f_i(x_i, x_{-i}) - h_i(\frac{1}{n}\sum_{j \in N} a_j x_j) - \bar{f}_i(x_i, x_{-i})| \leq \frac{L_g M \Delta^2}{n}$  for all  $x_i \in \mathcal{X}_i$  and all  $x_{-i} \in \tilde{\mathcal{X}}_{-i}$ . Hence, if  $\bar{x} \in X$  is an  $\epsilon$ -PNE of  $\bar{\Gamma}$ , then, for each  $i$ , for any

$x_i \in \mathcal{X}_i$ ,

$$\begin{aligned}
f_i(\bar{x}_i, \bar{x}_{-i}) &\leq \bar{f}_i(\bar{x}_i, \bar{x}_{-i}) + h_i \left( \frac{1}{n} \sum_{j \in N} a_j \bar{x}_j \right) + \frac{L_g M \Delta^2}{n} \\
&\leq \bar{f}_i(x_i, \bar{x}_{-i}) + \epsilon + h_i \left( \frac{1}{n} a_i x_i + \frac{1}{n} \sum_{j \neq i}^n a_j \bar{x}_j \right) + \frac{L_h M \Delta}{n} + \frac{L_g M \Delta^2}{n} \\
&\leq f_i(x_i, \bar{x}_{-i}) + \epsilon + \frac{L_h M \Delta}{n} + \frac{2L_g M \Delta^2}{n},
\end{aligned}$$

where the second inequality is due to the definition of  $\epsilon$ -PNE and the Lipschitz continuity of  $h_i$ .  $\square$

**Lemma 5.8.** *Under Assumption 2, let  $(x^k)_{k \in \mathbb{N}}$  be the sequence generated by Algorithm 1 with some initial point  $x^0 \in \tilde{\mathcal{X}}$ . Then,*

(1)  $\sum_{k=1}^{\infty} \|x^{k-1} - x^k\|^2 \leq \frac{2n^2}{m^2 L_g} C$ , where  $C = (d\Delta L_g + 2B_r)M$ ;

(2) for any  $K \in \mathbb{N}^*$ , there exists some  $k^* \leq K$ , such that  $\|x^{k^*-1} - x^{k^*}\| \leq \frac{\sqrt{2C}n}{m\sqrt{L_g K}}$ .

*Proof of Lemma 5.8.* Consider the following two real-valued functions defined on  $\tilde{\mathcal{X}}$ :

$$G_0(x) := \sum_{t=1}^d G_t \left( \frac{1}{n} \sum_{j \in N} a_j x_{j,t} \right), \quad G(x) := G_0(x) + \sum_{j \in N} \frac{a_j}{n} \tilde{r}_j(x_j), \quad (5.14)$$

where  $G_t$  is a primitive function of  $g_t$ , which exists thanks to Assumption 2.

Note that the function  $G_0$  is convex and differentiable on a neighborhood of  $\tilde{\mathcal{X}}$ , and the convex function  $\tilde{r}_j$  is uniformly bounded on  $\tilde{\mathcal{X}}_j$  for all  $j \in N$  with the same bound  $B_\ell$ , according to Assumption 2.

Besides, it is easy to see that, for any  $i$  and fixed  $x_{-i} \in \tilde{\mathcal{X}}_{-i}$ ,  $\nabla_i G_0(x_i, x_{-i}) := \frac{\partial G_0(x_i, x_{-i})}{\partial x_i} = \frac{a_i}{n} g \left( \frac{1}{n} a_i x_i + \frac{1}{n} \sum_{j \neq i} a_j x_j \right)$  is  $\frac{a_i^2 L_g}{n^2}$ -Lipschitz continuous on  $\tilde{\mathcal{X}}_i$ .

Therefore, Assumptions 1 and 2 in [47] are verified. One can thus apply Lemma 2.2 from [47] and obtain

$$\sum_{i \in N} \frac{a_i^2 L_g}{2n^2} \|x_i^k - x_i^{k+1}\|^2 \leq G(x^k) - G(x^{k+1}),$$

so that

$$\|x^k - x^{k+1}\|^2 \leq \frac{2n^2}{m^2 L_g} (G(x^k) - G(x^{k+1})).$$

In consequence,

$$\sum_{k=0}^{\infty} \|x^k - x^{k+1}\|^2 \leq \frac{2n^2}{m^2 L_g} (G(x^0) - G_{\min}), \quad (5.15)$$

where  $G_{min}$ , defined as  $\inf_{\{x \in \tilde{\mathcal{X}}\}} G(x)$ , exists and is finite, because  $G$  is l.s.c. on the compact set  $\tilde{\mathcal{X}}$ . Suppose that  $G_{min}$  is attained at  $\underline{x} \in \tilde{\mathcal{X}}$ ; then,

$$\begin{aligned} G(x^0) - G_{min} &= G(x^0) - G(\underline{x}) \\ &= \sum_{t=1}^d \int_{\frac{1}{n} \sum_{j \in N} a_j \underline{x}_{j,t}}^{\frac{1}{n} \sum_{j \in N} a_j x_{j,t}^0} g_t(s) ds + \sum_{j \in N} \frac{a_j}{n} (\tilde{r}_j(x_j^0) - \tilde{r}_j(\underline{x}_j)) \\ &\leq dM\Delta B_g + 2MB, \end{aligned} \quad (5.16)$$

where the last inequality is due to the mean value theorem and Assumption 2. Combining (5.15) and (5.16) yields  $\sum_{k=0}^{\infty} \|x^k - x^{k+1}\|^2 \leq \frac{2n^2}{m^2 L_g} C$ . This immediately implies

$$\sum_{k=1}^K \|x^{k-1} - x^k\|^2 \leq \frac{2n^2}{m^2 L_g} C.$$

The second result of the lemma is then straightforward.  $\square$

*Proof of Proposition 3.2.* First, notice that the vector function  $\zeta : \tilde{\mathcal{X}} \rightarrow \mathbb{R}^d, x \mapsto \zeta(x) = g(\frac{1}{n} \sum_{j \in N} a_j x_j)$  is  $\frac{L_g M}{\sqrt{n}}$ -Lipschitz continuous, i.e.,  $\|\zeta(x) - \zeta(y)\| \leq \frac{L_g M}{\sqrt{n}} \|x - y\|$ , for all  $x, y \in \tilde{\mathcal{X}}$ . Indeed,  $\|\zeta(x) - \zeta(y)\|^2 = \sum_{t=1}^d |g_t(\frac{1}{n} \sum_{j \in N} a_j x_{j,t}) - g_t(\frac{1}{n} \sum_{j \in N} a_j y_{j,t})|^2 \leq \sum_{t=1}^d |\frac{L_{g_t}}{n} \sum_{j \in N} a_j (x_{j,t} - y_{j,t})|^2 \leq \sum_{t=1}^d (\frac{L_{g_t}^2}{n^2} \sum_{j=1}^n a_j^2 \sum_{j \in N} (x_{j,t} - y_{j,t})^2) \leq \frac{L_g^2 M^2}{n} \|x - y\|^2$ , where the first inequality is because  $g_t$  is  $L_{g_t}$ -Lipschitz, while the second results from the Cauchy-Schwarz inequality.

Next, suppose that the sequence  $(x^k)_{k \in \mathbb{N}}$  is generated by Algorithm 1 with some initial point  $x^0 \in \tilde{\mathcal{X}}$ . Let us show that, if  $\|x^{k-1} - x^k\| \leq u_k$ , then,  $x^k$  satisfies the full  $\eta(u_k)\Delta$ -stability condition and, furthermore, it is an  $\eta(u_k)\Delta$ -PNE of game  $\tilde{\Gamma}$ , where  $\eta(u_k) = \frac{L_g M u_k}{\sqrt{n}} + \frac{2L_g M \Delta}{n}$ .

Since  $\|x^k - x^{k-1}\| \leq u_k$ , one has  $\|(x_1^k, \dots, x_{i-1}^k, x_i^k, x_{i+1}^k, \dots, x_n^k) - (x_1^{k-1}, \dots, x_{i-1}^{k-1}, x_i^{k-1}, \dots, x_n^{k-1})\| \leq u_k$ . Thus, the Lipschitz continuity of  $\zeta$  on  $\tilde{\mathcal{X}}$  and the Lipschitz continuity of  $g$  in  $x_i$  imply that

$$\begin{aligned} &\left\| g\left(\frac{1}{n} \sum_{j \neq i} a_j x_j^k + \frac{1}{n} a_i x_i^0\right) - g\left(\frac{1}{n} \sum_{j < i} a_j x_j^k + \frac{1}{n} a_i x_i^{k-1} + \frac{1}{n} \sum_{j > i} a_j x_j^{k-1}\right) \right\| \\ &\leq \left\| g\left(\frac{1}{n} \sum_{j \neq i} a_j x_j^k + \frac{1}{n} a_i x_i^0\right) - g\left(\frac{1}{n} \sum_{j \in N} a_j x_j^k\right) \right\| \\ &\quad + \left\| g\left(\frac{1}{n} \sum_{j \in N} a_j x_j^k\right) - g\left(\frac{1}{n} \sum_{j < i} a_j x_j^k + \frac{1}{n} a_i x_i^{k-1} + \frac{1}{n} \sum_{j > i} a_j x_j^{k-1}\right) \right\| \\ &\leq \frac{L_g M \Delta}{n} + \frac{L_g M u_k}{\sqrt{n}}. \end{aligned} \quad (5.17)$$

The first order condition of optimality of the optimization problem (3.3) is the following: there exists some  $p_i$  in the subdifferential of  $\tilde{r}_i(x_i^k)$  at  $x_i^k$ , denoted by  $\partial \tilde{r}_i(x_i^k)$ , such that for



all  $x_i \in \tilde{\mathcal{X}}_i$ ,

$$\left\langle g\left(\frac{1}{n} \sum_{j<i} a_j x_j^k + \frac{1}{n} \sum_{j \geq i} a_j x_j^{k-1}\right) + \frac{a_i L_g}{n} (x_i^k - x_i^{k-1}) + p_i, x_i - x_i^k \right\rangle \geq 0. \quad (5.18)$$

Then,

$$\begin{aligned} & \left\langle g\left(\frac{1}{n} \sum_{j \neq i} a_j x_j^k + \frac{1}{n} a_i x_i^+\right) + p_i, x_i - x_i^k \right\rangle \\ &= \left\langle g\left(\frac{1}{n} \sum_{j \neq i} a_j x_j^k + \frac{1}{n} a_i x_i^+\right) - g\left(\frac{1}{n} \sum_{j<i} a_j x_j^k + \frac{1}{n} \sum_{j \geq i} a_j x_j^{k-1}\right), x_i - x_i^k \right\rangle \\ &+ \left\langle g\left(\frac{1}{n} \sum_{j<i} a_j x_j^k + \frac{1}{n} \sum_{j \geq i} a_j x_j^{k-1}\right) + \frac{a_i L_g}{n} (x_i^k - x_i^{k-1}) + p_i, x_i - x_i^k \right\rangle \\ &- \left\langle \frac{a_i L_g}{n} (x_i^k - x_i^{k-1}), x_i - x_i^k \right\rangle \\ &\geq \left\langle g\left(\frac{1}{n} \sum_{j \neq i} a_j x_j^k + \frac{1}{n} a_i x_i^+\right) - g\left(\frac{1}{n} \sum_{j<i} a_j x_j^k + \frac{1}{n} a_i x_i^k + \frac{1}{n} \sum_{j>i} a_j x_j^{k-1}\right), x_i - x_i^k \right\rangle \\ &- \left\langle \frac{a_i L_g}{n} (x_i^k - x_i^{k-1}), x_i - x_i^k \right\rangle \\ &\geq -\left(\frac{L_g M u_k}{\sqrt{n}} + \frac{L_g M \Delta}{n} + \frac{a_i L_g \Delta}{n}\right) \|x_i - x_i^k\| \\ &\geq -\left(\frac{L_g M u_k}{\sqrt{n}} + \frac{2L_g M \Delta}{n}\right) \|x_i - x_i^k\| = -\eta(u_k) \|x_i - x_i^k\|, \end{aligned}$$

where the first inequality is due to (5.18), while the second inequality is due to (5.17) and the Cauchy-Schwarz inequality. Then, according to Lemma 2.6,  $x^k$  satisfies the full  $\eta(u_k)\Delta$ -stability condition for game  $\tilde{\Gamma}$ , where  $\eta(u_k) = \frac{L_g M u_k}{\sqrt{n}} + \frac{2L_g M \Delta}{n}$ .

Furthermore, since  $\tilde{r}_i$  is convex on  $\tilde{\mathcal{X}}_i$ ,

$$\begin{aligned} \bar{f}_i(x_i, x_{-i}^k) - \bar{f}_i(x^k) &= \left\langle g\left(\frac{1}{n} \sum_{j \neq i} a_j x_j^k + \frac{1}{n} a_i x_i^+\right), x_i - x_i^k \right\rangle + \tilde{r}_i(x_i) - \tilde{r}_i(x_i^k) \\ &\geq \left\langle g\left(\frac{1}{n} \sum_{j \neq i} a_j x_j^k + \frac{1}{n} a_i x_i^+\right), x_i - x_i^k \right\rangle + \langle p_i, x_i - x_i^k \rangle \\ &\geq -\left(\frac{L_g M u_k}{\sqrt{n}} + \frac{2L_g M \Delta}{n}\right) \|x_i - x_i^k\|. \end{aligned}$$

Thus,  $x^k$  is an  $\eta(u_k)\Delta$ -PNE of game  $\tilde{\Gamma}$ .

For any  $K \in \mathbb{N}^*$ , there exists some  $k^* \leq K$  such that  $\|x^{k^*-1} - x^{k^*}\| \leq \frac{\sqrt{2Cn}}{m\sqrt{L_g K}}$  according to Lemma 5.8(2). The conclusion is immediately obtained by taking  $\omega(K, n) = \eta\left(\frac{\sqrt{2Cn}}{m\sqrt{L_g K}}\right)$ .  $\square$

*Proof of Theorem 3.3.* Proposition 3.2 shows that  $x^{k^*}$  is an approximate PNE of game  $\tilde{\Gamma}$  (obtained through the convexification of the non-convex auxiliary game  $\tilde{\Gamma}$ ). Then, Theorem

2.7 is applied to show that (the “Shapley-Folkman disaggregation” of  $x^{k^*}$ )  $x^*$  is an approximate PNE of the non-convex auxiliary game  $\bar{\Gamma}$ . The use of Theorem 2.7 is justified by Lemma 3.1(1). Finally, Lemma 3.1(2) is evoked to show that  $x^*$  is an approximate PNE of the original non-convex game  $\Gamma$ .  $\square$

## References

- [1] J.P. Aubin and I. Ekeland, *Estimates of the duality gap in nonconvex optimization*, Mathematics of Operations Research **1** (1976), no. 3, 225–245.
- [2] A. Basile, M.G. Graziano, and M. Pesce, *Oligopoly and cost sharing in economics with public goods*, International Economic Review **57** (2016), no. 2, 487–505.
- [3] D. Bertsekas, G. Lauer, N. Sandell, and T. Posbergh, *Optimal short-term scheduling of large-scale power systems*, IEEE Transactions on Automatic Control **28** (1983), no. 1, 1–11.
- [4] D.P. Bertsekas, *Convexification procedures and decomposition methods for nonconvex optimization problems*, Journal of Optimization Theory and Applications **29** (1979), 169–197.
- [5] ———, *Constrained-Optimization and Lagrangian Multiplier Methods*, Athena Scientific, 1996.
- [6] ———, *Convex Optimization Theory*, Athena Scientific, 2009.
- [7] D.P. Bertsekas and N.R. Sandell, *Estimates of the duality gap for large-scale separable nonconvex optimization problems*, 21st IEEE Conference on Decision and Control, 1982, pp. 782–785.
- [8] Y. Bi and A. Tang, *Duality gap estimation via a refined Shapley–Folkman lemma*, SIAM Journal on Optimization **30** (2020), no. 2, 1094–1118.
- [9] L.C. Corchón, *Comparative statics for aggregative games the strong concavity case*, Mathematical Social Sciences **28** (1994), no. 3, 151–165.
- [10] S. Dafermos, *Traffic equilibrium and variational inequalities*, Transportation Science **14** (1980), no. 1, 42–54.
- [11] J. David, P. Castrillo, and T. Verdier, *A general analysis of rent-seeking games*, Public Choice **73** (1992), no. 3, 335–350.
- [12] I. Ekeland and R. Témam, *Convex Analysis and Variational Problems*, Society for Industrial and Applied Mathematics, 1999.
- [13] F. Facchinei and J. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Springer-Verlag New York, 2003.
- [14] E.X. Fang, H. Liu, and M. Wang, *Blessing of massive scale: spatial graphical model estimation with a total cardinality constraint approach*, Mathematical Programming **176** (2019), no. 1–2, 175–205.

- [15] R. Foucart and C. Wan, *Strategic decentralization and the provision of global public goods*, Journal of Environmental Economics and Management **92** (2018), 537–558.
- [16] J.B. Hiriart-Urruty and C. Lemarechal, *Convex Analysis and Minimization Algorithms II: Advanced Theory and Bundle Methods*, Springer-Verlag Berlin Heidelberg, 1993.
- [17] J. Hofbauer and K. Sigmund, *Evolutionary Games and Population Dynamics*, Cambridge University Press, 1998.
- [18] J. Horta, E. Altman, M. Caujolle, D. Kofman, and D. Menga, *Real-time enforcement of local energy market transactions respecting distribution grid constraints*, 2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm), 2018, pp. 1–7.
- [19] K. Hreinsson, A. Scaglione, M. Alizadeh, and Y. Chen, *New insights from the Shapley-Folkman lemma on dispatchable demand in energy markets*, IEEE Transactions on Power Systems **36** (2021), no. 5, 4028–4041.
- [20] P. Jacquot, O. Beauce, S. Gaubert, and N. Oudjane, *Demand response in the smart grid: The impact of consumers temporal preferences*, 2017 IEEE International Conference on Smart Grid Communications (SmartGridComm), 2017, pp. 540–545.
- [21] P. Jacquot, C. Wan, O. Beauce, and N. Oudjane, *Efficient estimation of equilibria in large aggregative games with coupling constraints*, IEEE Transactions on Automatic Control **66** (2021), 2762–2769.
- [22] M.K. Jensen, *Aggregative games and best-reply potentials*, Economic theory **43** (2010), no. 1, 45–66.
- [23] S. Kakutani, *A generalization of Brouwer’s fixed point theorem*, Duke Mathematical Journal **8** (1941), no. 3, 457–459.
- [24] T. Kerdreux, I. Colin, and A. d’Aspremont, *An approximate Shapley-Folkman theorem*, arXiv:1712.08559 (2019).
- [25] G.S. Lauer, N.R. Sandell, D.P. Bertsekas, and T.A. Posbergh, *Solution of large-scale optimal unit commitment problems*, IEEE Transactions on Power Apparatus and Systems **PAS-101** (1982), no. 1, 79–86.
- [26] L. Libman and A. Orda, *Atomic resource sharing in noncooperative networks*, Proceedings of INFOCOM ’97, vol. 3, 1997, pp. 1006–1013.
- [27] P. Marcotte and M. Patriksson, *Chapter 10. Traffic Equilibrium*, Transportation, vol. 14, Elsevier, 2007, pp. 623–713.
- [28] C. Meyers, *Network Flow Problems and Congestion Games: Complexity and Approximation Results*, PhD dissertation, MIT, 2006.
- [29] F.H. Murphy, H.D. Sherali, and A.L. Soyster, *A mathematical programming approach for determining oligopolistic market equilibrium*, Mathematical Programming **24** (1982), 92–106.

- [30] R.B. Myerson and R.J. Weber, *A theory of voting equilibria*, The American Political Science Review **87** (1993), no. 1, 102–114.
- [31] A. Orda, R. Rom, and N. Shimkin, *Competitive routing in multiuser communication networks*, IEEE/ACM Transactions on Networking **1** (1993), no. 5, 510–521.
- [32] D. Paccagnan, B. Gentile, F. Parise, M. Kamgarpour, and J. Lygeros, *Nash and Wardrop equilibria in aggregative games with coupling constraints*, IEEE Transactions on Automatic Control **64** (2019), 1373–1388.
- [33] D. Paccagnan, M. Kamgarpour, and J. Lygeros, *On aggregative and mean field games with applications to electricity markets*, 2016 European Control Conference (ECC) (2016), 196–201.
- [34] T.R. Palfrey and H. Rosenthal, *A strategic calculus of voting*, Public Choice **41** (1983), no. 1, 7–53.
- [35] M. Pappalardo, *On the duality gap in nonconvex optimization*, Mathematics of Operations Research **11** (1986), no. 1, 30–35.
- [36] J.B. Rosen, *Existence and uniqueness of equilibrium points for concave  $N$ -person games*, Econometrica **33** (1965), no. 3, 520–534.
- [37] R.W. Rosenthal, *A class of games possessing pure-strategy Nash equilibria*, International Journal of Game Theory **2** (1973), 65–67.
- [38] S. Sagratella, *Computing all solutions of Nash equilibrium problems with discrete strategy sets*, SIAM Journal on Optimization **26** (2016), no. 4, 2190–2218.
- [39] G. Scutari, F. Facchinei, J. Pang, and D. Palomar, *Real and complex monotone communication games*, IEEE Transactions on Information Theory **60** (2014), 4197–4231.
- [40] R. Selten, *Preispolitik der Mehrproduktenunternehmung in der Statischen Theorie*, Springer Verlag Berlin, 1970.
- [41] R.M. Starr, *Quasi-equilibria in markets with non-convex preferences*, Econometrica **37** (1969), no. 1, 25–38.
- [42] L. Tran-Thanh, M. Polukarov, A. Chapman, A. Rogers, and N.R. Jennings, *On the existence of pure strategy Nash equilibria in integer-splittable weighted congestion games*, Algorithmic Game Theory, Springer Berlin Heidelberg, 2011, pp. 236–253.
- [43] M. Udell and S. Boyd, *Bounding duality gap for separable problems with linear constraints*, Computational Optimization and Applications **64** (2016), 355–378.
- [44] R. Vujanic, M.E. Peyman, P. Goulart, and M. Morari, *Large scale mixed-integer optimization: A solution method with supply chain applications*, 22nd Mediterranean Conference on Control and Automation (2014), 804–809.
- [45] R. Vujanic, M.E. Peyman, P.J. Goulart, M. Sebastien, and M. Manfred, *A decomposition method for large scale MILPs, with performance guarantees and a power system application*, Automatica **67** (2016), 144–156.

- [46] M. Wang, *Vanishing price of decentralization in large coordinative nonconvex optimization*, SIAM Journal on Optimization **27** (2017), no. 3, 1977–2009.
- [47] Y. Xu and W. Yin, *A block coordinate descent method for regularized multiconvex optimization with applications to nonnegative tensor factorization and completion*, SIAM Journal of Imaging Sciences **6** (2013), no. 3, 1758–1789.
- [48] W. Yu and R. Lui, *Dual methods for nonconvex spectrum optimization of multicarrier systems*, IEEE Transactions on Communications **54** (2006), no. 7, 1310–1322.