



**HAL**  
open science

# Efficient Estimation of Equilibria of Large Congestion Games with Heterogeneous Players

Cheng Wan, Paulin Jacquot, Olivier Beauce, Nadia Oudjane

► **To cite this version:**

Cheng Wan, Paulin Jacquot, Olivier Beauce, Nadia Oudjane. Efficient Estimation of Equilibria of Large Congestion Games with Heterogeneous Players. 2018. hal-01904546v1

**HAL Id: hal-01904546**

**<https://hal.science/hal-01904546v1>**

Preprint submitted on 25 Oct 2018 (v1), last revised 26 Mar 2020 (v3)

**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.

# Efficient Estimation of Equilibria of Large Congestion Games with Heterogeneous Players

Cheng Wan, Paulin Jacquot, Olivier Beaude, Nadia Oudjane

**Abstract**—Computing an equilibrium in congestion games can be challenging when the number of players is large. Yet, it is a problem to be addressed in practice, for instance to forecast the state of the system and be able to control it. In this work, we analyze the case of generalized atomic congestion games, with coupling constraints, and with players that are heterogeneous through their action sets and their utility functions. We obtain an approximation of the variational Nash equilibria—a notion generalizing Nash equilibria in the presence of coupling constraints—of a large atomic congestion game by an equilibrium of an auxiliary population game, where each population corresponds to a group of atomic players of the initial game. Because the variational inequalities characterizing the equilibrium of the auxiliary game have smaller dimension than the original problem, this approach enables the fast computation of an estimation of equilibria in a large congestion game with thousands of heterogeneous players.

**Index Terms**—Atomic Congestion Game - Variational Nash Equilibrium - Variational Inequalities - Population Game

## I. INTRODUCTION

*a) Motivation:* Congestion games form a class of non-cooperative games [1]. In a congestion game, each player chooses a certain quantity of each of the available resources, and pays a cost for each resource obtained by the per-unit cost of that resource multiplied by the quantity she has chosen. A congestion game is said to be atomic if there is a finite number of players, and nonatomic if there is a continuum of infinitesimal players. The particularity of congestion games is that the per-unit cost of each resource depends only on its total demand.

Congestion games find practical applications in various fields such as traffic management [2], communications [3, 4] and more recently in electrical systems [5, 6].

The concept of Nash equilibrium (NE) [7] has emerged as the most credible outcome in the theory of noncooperative games. However, it is shown that computing a NE, when it exists, is a hard problem [8, 9]. NEs are often characterized by some variational inequalities. Therefore, the efficiency of the computation of NEs depends on the dimension of the variational inequalities in question, hence on the number of players and the number of constraints. The problem can be intractable at a large scale, when considering several thousands of heterogeneous agents, which is often the case

when describing real situations. The case of generalized Nash equilibria [10], when one considers coupling constraints—for instance capacity constraints—makes the problem even harder to solve. Meanwhile, coupling constraints commonly exist in real world. For example, in transportation, roads and communication channels have a limited capacity that should be considered. In the energy domain, production plants are also limited in the magnitude of variations of power, inducing some “ramp constraints” [11].

However, estimating the outcome situation—supposed to correspond to an equilibrium—is often a priority for the operator of the system. For instance, the operator controls some variables such as physical or managerial parameters, of a communication or transport network and wishes to optimize the performance of the system. The computation of equilibria or their approximation is also a key aspect in bi-level programming [12], where the lower level corresponds to a usually large scale game, and the upper level corresponds to a decision problem of an operator choosing optimal parameters. These parameters, such as prices or taxes, are to be applied in the low level game, with the aim of maximizing the revenue in various industrial sectors and public economics, such as highway management, urban traffic control, air industry, freight transport and radio network [13–17].

In this paper, we consider atomic congestion games with a finite but large number of players. We propose a method to compute an approximation of NEs, or variational Nash equilibria (VNEs) [18] in the presence of coupling constraints. The main idea is to reduce the dimension of the variational inequalities characterizing NEs or VNEs. The players are divided into groups with similar characteristics. Then, each group is replaced by a homogeneous population of nonatomic players. To provide an estimation of the equilibria of the original game, we compute a Wardrop equilibrium (WE) [19] in the approximating nonatomic population game, or a variational Wardrop equilibrium (VWE) in the case of coupling constraints. The quality of the estimation depends on how well the characteristics, such as action set and cost function, of each homogeneous population approximate those of the atomic players it replaces.

In addition to the reduction in dimension, another advantage of WE is that it is usually unique in congestion games, in contrast to NEs. In bi-level programming, the uniqueness of a low level equilibrium allows for clear-cut comparative statics and sensitivity analysis at the high level.

*b) Related works:* The relation between NEs in large games and WEs has been studied in Gentile et al. [20]. In their paper, the authors also consider atomic congestion games with coupling constraints and show, using variational inequalities approach, that the distance between a NE and a WE converges

Cheng Wan is with Université Paris-Sud, Inria Paris and RIIS, SUFE, Shanghai (cheng.wan.2005@polytechnique.org). Paulin Jacquot is with EDF R&D (OSIRIS), Inria and École polytechnique, CNRS, Palaiseau, France (paulin.jacquot@polytechnique.edu). Olivier Beaude and Nadia Oudjane are with EDF R&D (OSIRIS), Palaiseau, France ({olivier.beaude, nadia.oudjane}@edf.fr).

This work was partially supported by the PGMO-ICODE project “Jeu de pilotage de flexibilités de consommation électrique : dynamique et aspect composite”.

to zero when the number of players tends to infinity. Their WE corresponds to an equilibrium of the game where each atomic player is replaced by a population. The objective of our paper is different. We look for an approximation of NEs by reducing the dimension of the original game. To this end, we regroup many players into few homogeneous populations. Our results apply to the subdifferentiable case in contrast to the differential case considered in [20].

In [21], Jacquot and Wan show that, in congestion games with a continuum of heterogeneous players, the WE can be approximated by a NE of an approximating game with a finite number of players. In [22], those results are extended to aggregative games, a more general class of games including congestion games, furthermore with nonsmooth cost functions.

Different algorithms have been proposed to solve monotone variational inequalities corresponding to NE or WE, such as [23–27], and more recently [28–31] and the references therein.

The approach developed in the present paper is actually the inverse of the one taken in [21] and [22]: here, the WE in the auxiliary game serves as an approximation of an NE of the original large game.

c) *Main contributions*: The contributions of this paper are the following.

- We define an approximating population game (Sec. III-B). The idea is that the auxiliary game has smaller dimension but is close enough to the original large game—quantified through the Hausdorff distance between action sets and between subgradients of players’ objective functions.

- We show theoretically that a particular variational Wardrop equilibrium (VWE) of the approximating population game is close to any variational Nash equilibria (VNE) of the original game with or without coupling constraints, while the computation of the former is much faster than the later because of the dimension reduction. We provide an explicit expression of the error bound of the approximating VWE (Thm. 5).

- We give auxiliary results on variational equilibria: when the number of players is large, VNEs are close to each other (Thm. 2) and that VNEs are close to the approximating VWE (Thm. 6). This last theorem extends [20, Thm. 1] in the case of nondifferentiable cost functions, in the framework of congestion games.

- Last, we provide a numerical illustration of our results (Sec. IV) based on a practical application: the decentralized charging of electric vehicles through a demand response mechanism [32]. This example illustrates the nondifferentiable case through piece-wise linear electricity prices (“block rates tariffs”), with coupling constraints of capacities and limited variations on the aggregate load profile between time periods. This example shows that the proposed method is implementable and that it reduces the time needed to compute an equilibrium by computing its approximation (six times faster for an approximation with a relative error of less than 2%).

The remainder of this paper is organized as follows: Sec. II specifies the framework of congestion games with coupling constraints, and recalls the notions of variational equilibria and monotonicity for variational inequalities, as well as several results on the existence and uniqueness of equilibria. Sec. III formulates the main results: Sec. III-B shows that a VWE

approximates VNEs in large games and then, Sec. III-B formulates the approximating population game with the approximation measures, and gives an error bound on the VWE of the approximating game with respect to the original VNEs. Sec. IV presents a numerical illustration in the framework of demand response for electric vehicle smart charging.

## II. CONGESTION GAMES WITH COUPLING CONSTRAINTS

### A. Model and equilibria

The original game throughout this paper is an atomic splittable congestion game, a particular sort of aggregative games where a set of resources is shared among finitely many players, and each resource incurs a cost increasing with the aggregate demand for it. The formal definition is as follows.

**Definition 1.** An *atomic splittable congestion game*  $\mathcal{G}$  is defined by:

- a finite set of players:  $\mathcal{I} = \{1, \dots, i, \dots, I\}$ ,
- a finite set of resources:  $\mathcal{T} = \{1, \dots, t, \dots, T\}$ ,
- for each resource  $t$ , a cost function  $c_t : \mathbb{R}_+ \rightarrow \mathbb{R}$ ,
- for each player  $i$ , a set of feasible choices:  $\mathcal{X}_i \subset \mathbb{R}_+^T$ , an element  $\mathbf{x}_i = (x_{i,t})_{t \in \mathcal{T}} \in \mathcal{X}_i$  signifies that  $i$  has demand  $x_{i,t}$  for resource  $t$ ,
- for each player  $i$ , an individual utility function  $u_i : \mathcal{X}_i \rightarrow \mathbb{R}$ ,
- a coupling constraint set  $A \subset \mathbb{R}^T$ .

We denote by  $\tilde{\mathcal{X}} \triangleq \mathcal{X}_1 \times \dots \times \mathcal{X}_I$  the product set of action profiles. An action profile  $\mathbf{x} = (\mathbf{x}_i)_{i \in \mathcal{I}} \in \tilde{\mathcal{X}}$  induces a profile of aggregate demand for the resources, denoted by  $\mathbf{X} = (X_t)_{t \in \mathcal{T}} \triangleq (\sum_{i \in \mathcal{I}} x_{i,t})_{t \in \mathcal{T}}$ . We denote the set of feasible aggregate demand profiles by:

$$\bar{\mathcal{X}} \triangleq \{\mathbf{X} \in \mathbb{R}^T : \forall i \in \mathcal{I}, \exists \mathbf{x}_i \in \mathcal{X}_i \text{ s.t. } \sum_{i \in \mathcal{I}} \mathbf{x}_i = \mathbf{X}\}.$$

With coupling constraints, the set of feasible aggregate demand profiles with coupling constraints is  $\bar{\mathcal{X}} \cap A$ , and the set of feasible action profiles is denoted by  $\tilde{\mathcal{X}}(A) = \{\mathbf{x} \in \tilde{\mathcal{X}} : \sum_{i \in \mathcal{I}} \mathbf{x}_i \in \bar{\mathcal{X}} \cap A\}$ .

Let the vector of cost functions denoted by  $\mathbf{c}(\mathbf{X}) = (c_t(X_t))_{t \in \mathcal{T}}$ , where  $c_t(X_t)$  is the (per-unit of demand) cost of resource  $t$  when the aggregate demand for it is  $X_t$ .

Player  $i$ ’s cost function  $f_i : \mathcal{X}_i \times \bar{\mathcal{X}} \rightarrow \mathbb{R}$  is defined by:

$$f_i(\mathbf{x}_i, \mathbf{Y}) = \sum_{t \in \mathcal{T}} x_{i,t} c_t(Y_t) - u_i(\mathbf{x}_i), \quad \forall \mathbf{x}_i \in \mathcal{X}_i, \mathbf{Y} \in \bar{\mathcal{X}}. \quad (1)$$

Given  $\mathbf{x}_{-i} \in \prod_{j \neq i} \tilde{\mathcal{X}}_j$  and  $\mathbf{X}_{-i} \triangleq \sum_{j \neq i} \mathbf{x}_j$ , player  $i$ ’s cost is  $f_i(\mathbf{x}_i, \mathbf{x}_i + \mathbf{X}_{-i})$ , composed of the network costs and her individual utility.

This atomic congestion game with coupling constraints is defined as the tuple  $(\mathcal{I}, \mathcal{T}, \tilde{\mathcal{X}}, A, \mathbf{c}, (u_i)_{i \in \mathcal{I}})$ .

In an atomic game, there are finitely many players whose actions are not negligible on the aggregate profile and on the objectives of other players. The term “atomic” is opposed to “nonatomic” where players have an infinitesimal weight [1]. The term “splittable” refers to the infinite number of choices of pure actions  $\mathbf{x}_i \in \mathcal{X}_i$  for each player  $i$ , as opposed to the unsplittable case where each player can only choose one action in a finite subset of  $2^T$  [33]. Besides, atomic splittable

congestion games are particular cases of aggregative games [20]: each player's cost function depends on the actions of the others only through the aggregate profile  $\mathbf{X}$ .

The following standard assumptions are adopted in this paper.

**Assumption 1.**

- (1) For each player  $i \in \mathcal{I}$ , the set  $\mathcal{X}_i$  is a convex and compact subset of  $\mathbb{R}^T$  with nonempty relative interior.
- (2) The cost function  $c_t$  for each resource  $t \in \mathcal{T}$  is continuous, convex and non-decreasing on  $(-\eta, +\infty)$  for a positive  $\eta > 0$ .
- (3) For each player  $i \in \mathcal{I}$ , individual utility function  $u_i$  is continuous and concave in  $\mathbf{x}_i$  on  $\mathcal{X}_i$ .
- (4)  $A$  is a convex closed set of  $\mathbb{R}^T$ , and  $\bar{\mathcal{X}} \cap A$  is not empty.

An important class of atomic splittable congestion games corresponds to the case where resources constitute a *parallel-arc* network [34]. There, each player  $i$  has a total demand and specific bounds on the demand that she can have for each resource so that her strategy set is given by  $\mathcal{X}_i = \{ \mathbf{x}_i \in \mathbb{R}_+^T : \sum_t x_{i,t} = m_i \text{ and } \underline{x}_{i,t} \leq x_{i,t} \leq \bar{x}_{i,t} \}$ . Here,  $m_i$  can represent the mass of data to send over different canals, or the amount of energy to consume over several time periods [6]. In particular, the demand is continuous and splittable, as the mass  $m_i$  is split over the resources  $t \in \mathcal{T}$ .

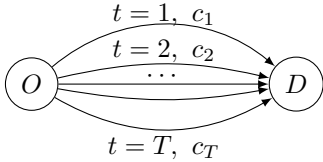


Fig. 1: A parallel-arc network with  $T$  arcs/resources.

Our model is more general than the one in [34], not only because the network topology can be arbitrary, but also because it allows for *elastic demands* from the players. For example, player  $i$ 's action set can be  $\mathcal{X}_i = \{ \mathbf{x}_i \in \mathbb{R}_+^T : \underline{m}_i \leq \sum_t x_{i,t} \leq \bar{m}_i \}$ . Indeed, the individual utility function counterbalances the network cost: a player may be willing to pay more congestion cost by increasing the demand, because she profits from a higher individual utility, and vice versa.

Let us cite two common forms of *individual utility function*. The first one measures the distance between a player's choice and her preference  $\mathbf{y}_i \in \mathcal{X}_i$ :  $u_i(\mathbf{x}_i) = -\omega_i \|\mathbf{x}_i - \mathbf{y}_i\|^2$ , where  $\omega_i > 0$  is the value that the player attaches to her preference. The second one is  $u_i(\mathbf{x}_i) = \omega_i \log(1 + \sum_t x_{i,t})$ , which is increasing in the player's total demand.

Finally, in congestion games, *aggregate constraints* are very common. For example, in routing games, there can be a capacity constraint linked to each arc. In energy consumption games, due to the operational constraints of the power grid, there can be both minimum and maximum consumption level for each time slot, and ramp constraints on the variation of energy consumption between time slots. This is why congestion games with aggregate constraints are of particular interest.

To separate player  $i$ 's choice from those of the other players in her cost function, define

$$\hat{f}_i(\mathbf{x}_i, \mathbf{Y}) \triangleq f_i(\mathbf{x}_i, \mathbf{Y} + \mathbf{x}_i)$$

for  $\mathbf{x}_i$  in  $\mathcal{X}_i$  and  $\mathbf{Y}$  in  $\bar{\mathcal{X}}_{-i} = \{ \sum_{j \in \mathcal{I} \setminus \{i\}} \mathbf{x}_j : \mathbf{x}_j \in \mathcal{X}_j \}$ .

Since  $c$  and  $u_i$ 's are not necessarily differentiable, we need to define the subdifferential of the players' utilities w.r.t. their actions for the characterization of equilibrium.

Let us define two correspondences,  $H$  and  $H'$ , from  $\bar{\mathcal{X}}$  to  $\mathbb{R}^{IT}$ : for any  $\mathbf{x} \in \bar{\mathcal{X}}$ ,

$$\begin{aligned} H(\mathbf{x}) &\triangleq \{ (\mathbf{h}_i)_{i \in \mathcal{I}} \in \mathbb{R}^{IT} : \mathbf{h}_i \in \partial_1 \hat{f}_i(\mathbf{x}_i, \mathbf{X}_{-i}), \forall i \in \mathcal{I} \} \\ &= \prod_{i \in \mathcal{I}} \partial_1 \hat{f}_i(\mathbf{x}_i, \mathbf{X}_{-i}) ; \\ H'(\mathbf{x}) &\triangleq \{ (\mathbf{h}_i)_{i \in \mathcal{I}} \in \mathbb{R}^{IT} : \mathbf{h}_i \in \partial_1 f_i(\mathbf{x}_i, \mathbf{X}), \forall i \in \mathcal{I} \} \\ &= \prod_{i \in \mathcal{I}} \partial_1 f_i(\mathbf{x}_i, \mathbf{X}) , \end{aligned}$$

where  $\partial_1$  signifies the partial differential w.r.t. the first variable of the function. The interpretation of  $H(\mathbf{x})$  is clear:  $\mathbf{h}_i$  is a subgradient of player  $i$ 's utility function  $\hat{f}_i$  w.r.t. her action  $\mathbf{x}_i$ . Let us leave the interpretation of  $H'(\mathbf{x})$  till Def. 3. For the moment, let us write the explicit expression of  $H$  and  $H'$ :

**Lemma 1.** For each  $\mathbf{x} \in \bar{\mathcal{X}}$ :

- $\mathbf{h} \in H(\mathbf{x})$  if and only if there are  $\mathbf{g}'_i \in \partial(-u_i)(\mathbf{x}_i)$  and  $\mathbf{a}_i \in \prod_{t \in \mathcal{T}} \partial c_t(X_t) \forall i$  s.t.:

$$\mathbf{h}_i = \mathbf{c}(\mathbf{X}) + (x_{i,t} \mathbf{a}_{i,t})_t + \mathbf{g}'_i, \quad \forall i \in \mathcal{I} ;$$

- $\mathbf{h}' \in H'(\mathbf{x})$  if and only if there is  $\mathbf{g}'_i \in \partial(-u_i)(\mathbf{x}_i) \forall i$  s.t.:

$$\mathbf{h}'_i = \mathbf{c}(\mathbf{X}) + \mathbf{g}'_i, \quad \forall i \in \mathcal{I} .$$

where  $\partial(-u_i)(\mathbf{x}_i)$  is the subdifferential of convex function  $-u_i$  at  $\mathbf{x}_i$  and  $\partial c_t(X_t)$  the subdifferential of  $c_t$  at  $X_t$ .

*Proof.* See Appendix A.  $\square$

In our framework with coupling constraints, the notion of Nash equilibrium (NE) [7] is replaced by that of Generalized Nash Equilibrium (GNE):  $\mathbf{x} \in \bar{\mathcal{X}}(A)$  is a GNE if, for each player  $i$ ,  $\hat{f}_i(\mathbf{x}_i, \mathbf{X}_{-i}) \leq \hat{f}_i(\mathbf{y}_i, \mathbf{X}_{-i})$  for all  $\mathbf{y}_i$  s.t.  $\mathbf{y}_i + \mathbf{X}_{-i} \in \bar{\mathcal{X}} \cap A$ . For atomic games, a special class of GNE is called Variational Nash Equilibria [10, 35], which enjoys some symmetric properties and can be easily characterized as the solution of the VI (2) below.

**Definition 2** (Variational Nash Equilibrium (VNE), [18]). A VNE is a solution  $\hat{\mathbf{x}} \in \bar{\mathcal{X}}(A)$  to the following GVI problem:

$$\exists \mathbf{g} \in H(\hat{\mathbf{x}}) \text{ s.t. } \langle \mathbf{g}, \mathbf{x} - \hat{\mathbf{x}} \rangle \geq 0, \quad \forall \mathbf{x} \in \bar{\mathcal{X}}(A). \quad (2)$$

In particular, if  $\bar{\mathcal{X}} \subset A$ , a VNE is a NE.

In this paper, we adopt VNE as the equilibrium notion in the presence of aggregate constraints.

As the first step of approximation, let us define a nonatomic congestion game  $\mathcal{G}(A)'$  associated to  $\mathcal{G}(A)$ . Let each player  $i$  be replaced by a continuum of identical nonatomic players, represented by interval  $[0, 1]$  with each point thereon corresponding to a nonatomic player. Each player in population  $i$  has action set  $\mathcal{X}_i$  and individual utility function  $u_i$ .

**Definition 3.** A symmetrical variational Wardrop equilibrium (SVWE) of  $\mathcal{G}(A)'$  is a solution to the following GVI:

$$\exists \mathbf{g} \in H'(\mathbf{x}^*) \text{ s.t. } \langle \mathbf{g}, \mathbf{x} - \mathbf{x}^* \rangle \geq 0, \forall \mathbf{x} \in \tilde{\mathcal{X}}(A). \quad (3)$$

For the definition of variational Wardrop equilibrium (VWE) and further discussion, we refer to [22]. In particular, a VWE is characterized by an infinite dimensional variational inequality. Here, we consider only those VWE where all the nonatomic players in population  $i$  take the same action  $\mathbf{x}_i$ . Such a SVWE exists because the players are identical in the same population.

The second interpretation of SVWE is the following: when the number of players is very large so that the individual contribution of each player on the aggregate action  $\mathbf{X}$  is almost negligible, the term  $x_{i,t} a_{i,t}$  in  $\mathbf{h} \in H(\mathbf{x})$  is so small that  $H(\mathbf{x})$  can be approximated by  $H'(\mathbf{x})$  (cf. Lemma 1). This is the interpretation adopted in [20]. However, note that a SVWE of  $\mathcal{G}(A)'$  is not an equilibrium of  $\mathcal{G}(A)$  in the sense of a ‘‘stable state’’ for the atomic congestion game.

The existence of equilibria defined in Defs. 2 and 3 are obtained without more conditions than Asm. 1:

**Proposition 1** (Existence of equilibria). Under Asm. 1,  $\mathcal{G}(A)$  (resp.  $\mathcal{G}(A)'$ ) admits a VNE (resp. SVWE).

*Proof.* : see Appendix B.

Before discussing the uniqueness of equilibria, let us recall some relevant monotonicity assumptions.

**Definition 4.** A correspondence  $\Gamma : \tilde{\mathcal{X}} \rightrightarrows \mathbb{R}^T$  is:

- *monotone* if for all  $\mathbf{x}, \mathbf{y} \in \tilde{\mathcal{X}}, \mathbf{g} \in \Gamma(\mathbf{x}), \mathbf{h} \in \Gamma(\mathbf{y})$ :

$$\sum_{i \in \mathcal{I}} \langle \mathbf{g}_i - \mathbf{h}_i, \mathbf{x}_i - \mathbf{y}_i \rangle \geq 0; \quad (4)$$

- *strictly monotone* if the equality in (4) holds iff  $\mathbf{x} = \mathbf{y}$ ;
- *aggregatively strictly monotone* if the equality in (4) holds iff  $\sum_i \mathbf{x}_i = \sum_i \mathbf{y}_i$ ;
- $\alpha$ -*strongly monotone* if  $\alpha > 0$  and, for all  $\mathbf{x}, \mathbf{y} \in \tilde{\mathcal{X}}$ :

$$\sum_{i \in \mathcal{I}} \langle \mathbf{g}_i - \mathbf{h}_i, \mathbf{x}_i - \mathbf{y}_i \rangle \geq \alpha \|\mathbf{x} - \mathbf{y}\|^2, \forall \mathbf{g} \in \Gamma(\mathbf{x}), \mathbf{h} \in \Gamma(\mathbf{y}); \quad (5)$$

- $\beta$ -*aggregatively strongly monotone* on  $\tilde{\mathcal{X}}$  if  $\beta > 0$  and, for all  $\mathbf{x}, \mathbf{y} \in \tilde{\mathcal{X}}$  with  $\mathbf{X} = \sum_i \mathbf{x}_i, \mathbf{Y} = \sum_i \mathbf{y}_i$ :

$$\sum_{i \in \mathcal{I}} \langle \mathbf{g}_i - \mathbf{h}_i, \mathbf{x}_i - \mathbf{y}_i \rangle \geq \beta \|\mathbf{X} - \mathbf{Y}\|^2, \forall \mathbf{g} \in \Gamma(\mathbf{x}), \mathbf{h} \in \Gamma(\mathbf{y}). \quad (6)$$

If  $T = 1$ , ‘‘monotone’’ corresponds to ‘‘increasing’’. Besides, (aggregatively) strict monotonicity implies monotonicity, while strong (resp. aggregatively strong) monotonicity implies strict (resp. aggregatively strict) monotonicity.

In Prop. 2 below, we recall some existing results concerning the uniqueness of VNE and SVWE, according to the monotonicity of  $H$  and  $H'$ :

**Proposition 2** (Uniqueness of equilibria). Under Asm. 1:

- (1) if  $H$  (resp.  $H'$ ) is strictly monotone, then  $\mathcal{G}(A)$  (resp.  $\mathcal{G}(A)'$ ) has a unique VNE (resp. SVWE);
- (2) if  $H$  (resp.  $H'$ ) is aggregatively strictly monotone, then all VNE (resp. SVWE) of  $\mathcal{G}(A)$  (resp.  $\mathcal{G}(A)'$ ) have the same aggregate profile;

(3) if  $H$  (resp.  $H'$ ) is only aggregatively strictly monotone but, in addition, for each  $i \in \mathcal{I}$ ,  $u_i(\mathbf{x})$  is strictly concave, then there is at most one NE (resp. WE) in the case without aggregative constraint.

*Proof.* see Appendix C.

Prop. 3 below gives sufficient conditions for the (strong) monotonicity to hold for  $H'$ .

**Proposition 3** (Monotonicity of  $H'$ ). Under Asm. 1,

- (1)  $H'$  is monotone.
- (2) If for each  $i \in \mathcal{I}$ ,  $u_i$  is  $\alpha_i$ -strongly concave, then  $H'$  is  $\alpha$ -strongly monotone with  $\alpha \triangleq \min_{i \in \mathcal{I}} \alpha_i$ .
- (3) If for each  $t \in \mathcal{T}$ ,  $c_t$  is  $\beta_t$ -strictly increasing, then  $H'$  is  $\beta$ -aggregatively strongly monotone with  $\beta \triangleq \min_{t \in \mathcal{T}} \beta_t$ .

*Proof.* See Appendix D.

As opposed to the monotonicity of  $H'$  shown in Prop. 3,  $H$  is rarely monotone (except in some particular cases, e.g. with  $c$  linear [34, 36]): even in the case where  $c$  is piece-wise linear, the Ex. 1 below shows that  $H$  can be non monotone.

**Example 1.** Let  $I = 2$  and  $T = 1$ ,  $\mathcal{X}_1 = \mathcal{X}_2 = [0, 4]$ . Consider the cost function  $c(X) = X$  for  $X \leq 4$  and  $c(X) = 3X - 8$  for  $X \geq 4$ . Asm. 1 holds. Consider the profiles  $x_1 = 3, x_2 = 1$  and  $y_1 = 4, y_2 = 0$ , then  $\mathbf{g} \triangleq (c(4) + 3x_1, c(4) + 3x_2) \in H(\mathbf{x})$  and  $\mathbf{h} \triangleq (c(4) + 1y_1, c(4) + 1y_2) \in H(\mathbf{y})$ , but:

$$\sum_{i \in \{1,2\}} \langle \mathbf{g}_i - \mathbf{h}_i, \mathbf{x}_i - \mathbf{y}_i \rangle = -2 < 0.$$

In view of Prop. 2, the absence of monotonicity of  $H$  can result in multiple VNEs [37].

In [6], a particular case with parallel arc network is shown to have a unique NE. However, in the next section, we shall prove that, when the number of players is very large, all VNEs are close to each other and they can be well approximated by the unique SVWE.

### III. APPROXIMATING VNEs OF A LARGE GAME

#### A. Considering SVWE instead of VNE

The approximation of VNEs is done in two steps. The first step consists in replacing VNE by SVWE. According to the second interpretation of SVWE, the SVWE should be close to the VNEs in a large game. Now, let us formulate this idea and bound the distance between the two.

Denote by  $\mathcal{X}_0 \subset \mathbb{R}^T$  the convex closed hull of  $\bigcup_{i \in \mathcal{I}} \mathcal{X}_i$ , by  $m = \max_{\mathbf{x} \in \mathcal{X}_0} \|\mathbf{x}\|$  and  $M = Im \geq \max_{\mathbf{X} \in \bar{\mathcal{X}}} \|\mathbf{X}\|$ . Define compact set  $\mathcal{M} = [0, M + \delta I]^T$ , where  $\delta > 0$  is a constant to be specified later.

Denote by  $C \triangleq \sup\{b \in \mathbb{R} : b \in \partial c_t(X_t), \mathbf{X} \in \bar{\mathcal{X}}, t \in \mathcal{T}\}$  the upper bound on the subgradients of  $c$ .

Prop. 2 and Prop. 3 show that, in general, VNEs are not unique. However, when the set of players is large, VNEs are indeed close to each other:

**Theorem 2** (VNEs are close to each other). Under Asm. 1, let  $\mathbf{x}$  and  $\mathbf{y}$  in  $\tilde{\mathcal{X}}(A)$  be two distinctive VNEs of  $\mathcal{G}(A)$ . Then (1) if for each  $i \in \mathcal{I}$ ,  $u_i$  is  $\alpha_i$ -strongly concave, then:

$$\|\mathbf{x} - \mathbf{y}\| \leq 2M \sqrt{\frac{TC}{\alpha I}}, \quad (7)$$

with  $\alpha \triangleq \min_{i \in \mathcal{I}} \alpha_i$ ;

(2) if for each  $t \in \mathcal{T}$ ,  $c_t$  is  $\beta_t$ -strictly increasing, then:

$$\|\mathbf{X} - \mathbf{Y}\| \leq 2M \sqrt{\frac{TC}{\beta I}}, \quad (8)$$

with  $\beta \triangleq \min_{t \in \mathcal{T}} \beta_t$ .

*Proof.* : See Appendix E.

The first step of approximation is based upon the following theorem which gives an upper bound on the distance between a VNE and the unique SVWE.

**Theorem 3** (SVWE is close to VNE). Under Asm. 1, let  $\mathbf{x} \in \tilde{\mathcal{X}}(A)$  be a VNE of  $\mathcal{G}(A)$  and  $\mathbf{x}^* \in \tilde{\mathcal{X}}(A)$  a SVWE of  $\mathcal{G}(A)'$ , then:

(1) if for each  $i \in \mathcal{I}$ ,  $u_i$  is a  $\alpha_i$ -strongly concave, then  $\mathbf{x}^*$  is unique and:

$$\|\mathbf{x} - \mathbf{x}^*\| \leq M \sqrt{\frac{2TC}{\alpha I}}, \quad (9)$$

with  $\alpha \triangleq \min_i \alpha_i$ ;

(2) if for each  $t \in \mathcal{T}$ ,  $c_t$  is  $\beta_t$ -strictly increasing, then  $\mathbf{X}^*$  is unique and:

$$\|\mathbf{X} - \mathbf{X}^*\| \leq M \sqrt{\frac{2TC}{\beta I}}, \quad (10)$$

with  $\beta \triangleq \min_t \beta_t$ .

*Proof.* Similar to the proof of Thm. 2.  $\square$

An upper bound on the distance between two VNEs can also be derived from Thm. 3, applying the triangle inequality. However, Thm. 2 gives a tighter upper bound.

Thm. 3 shows that, if the number of players  $I$  is large, then the SVWE will provide a good approximation of a VNE of  $\mathcal{G}(A)$ . Similar results are obtained in [20]. However, this does not reduce the dimension of the GVI to resolve: the GVI characterizing the VNE and those characterizing the SVWE have the same dimension. For this reason, the second step of approximation consists in regrouping similar populations.

### B. Classification of populations

In this subsection, we shall regroup the populations in  $\mathcal{G}(A)'$  with similar strategy sets  $\mathcal{X}_i$  and utility subgradients  $\partial(-u_i)$  (w.r.t. the Hausdorff distance, denoted by  $d_H$ ) into larger populations, endow them with a common strategy set and a common utility function, so that the SVWE of this new population game approximates the SVWE of  $\mathcal{G}(A)$ .

At the SVWE of the new population game with a reduced dimension, all the nonatomic players in the same population play the same action, by the definition of SVWE. Therefore, in order for this new SVWE to well approximate the SVWE in  $\mathcal{G}(A)'$ , we must ensure that populations with similar characteristics in  $\mathcal{G}(A)'$  do play similar actions at the SVWE of  $\mathcal{G}(A)'$ . Prop. 4 formulates this results in the case without coupling constraint.

Without loss of generality, we assume that for each  $i \in \mathcal{I}$ ,  $u_i$  can be extended to a neighborhood of  $\frac{M}{I} = [0, m]^T$ , and is bounded on  $\frac{M}{I}$ . Denote  $B_{\mathbf{u}_i} = \sup\{\|\mathbf{g}'_i\| : \mathbf{g}'_i \in \partial(-u_i)(\mathbf{x}_i), \mathbf{x}_i \in \mathcal{X}_i\}$ , and  $B_c = \sup\{\|\mathbf{c}(\mathbf{X})\| : \mathbf{X} \in \tilde{\mathcal{X}}\}$ .

**Proposition 4.** Under Asm. 1, let  $\mathbf{x}^* \in \tilde{\mathcal{X}}$  be a SVWE of  $\mathcal{G}'$  (without coupling constraints). For two populations  $i$  and  $j$  in  $\mathcal{I}$ , if  $u_i$  is  $\alpha_i$ -strongly concave,  $d_H(\mathcal{X}_i, \mathcal{X}_j) \leq \delta$ , and  $\sup_{\mathbf{x}_j \in \mathcal{X}_j} \sup_{\mathbf{g}'_j \in \partial(-u_j)(\mathbf{x}_j)} d(\mathbf{g}'_j, \partial(-u_i)(\mathbf{x}_j)) \leq \lambda$ , then

$$\|\mathbf{x}_i^* - \mathbf{x}_j^*\|^2 \leq \frac{1}{\alpha_i} ((B_{\mathbf{u}_i} + B_{\mathbf{u}_j} + 2B_c)\lambda + 2\delta m).$$

*Proof.* : See Appendix F.

In the case with coupling constraints, the proof for a similar result is more complicated. Let us leave it to Cor. 1.

Let us now present the regrouping procedure.

Denote an auxiliary game  $\tilde{\mathcal{G}}^{\mathcal{N}}(A)$ , with a set  $\mathcal{N}$  of  $N$  populations. Each population  $n \in \mathcal{N}$  corresponds to a subset of populations in game  $\mathcal{G}(A)'$ , denoted by  $\mathcal{I}_n$ , and  $\bigcup_{n \in \mathcal{N}} \mathcal{I}_n = \mathcal{I}$ .

Denote  $I_n = |\mathcal{I}_n|$  the number of original populations now included in  $n$ . By abuse of notations, let  $n$  also denote the interval  $[0, I_n]$ , so that each nonatomic player in population  $n$  is represented by a point  $\theta \in [0, I_n]$ . The common action set of each nonatomic player in  $n$  is a compact convex subset of  $\mathbb{R}^T$ , denoted by  $\mathcal{X}_n$ .

Each player  $\theta$  in each population  $n$  having chosen action  $\mathbf{x}_\theta$ , let  $\mathbf{X} \triangleq \sum_{n \in \mathcal{N}} \int_{\theta \in n} \mathbf{x}_\theta d\theta$  denote the aggregate action profile. The aggregate action-profile set in  $\mathbb{R}^T$  is then:

$\tilde{\mathcal{X}}^{\mathcal{N}} = \{\sum_{n \in \mathcal{N}} \int_{\theta \in n} \mathbf{x}_\theta d\theta : \mathbf{x}_\theta \in \mathcal{X}_n, \forall \theta \in n, \forall n \in \mathcal{N}\}$ . The cost function of player  $\theta$  in population  $n \in \mathcal{N}$  is:

$$f_n(\mathbf{x}_\theta, \mathbf{X}) = \langle \mathbf{x}_\theta, \mathbf{c}(\mathbf{X}) \rangle - u_n(\mathbf{x}_\theta),$$

where the common individual utility function  $u_n$  for all the players in  $n$  is concave on a neighborhood of  $\mathcal{X}_n$ .

We are only interested in *symmetric* action profiles, i.e. where all the nonatomic players in the same population  $n$  play the same action. Denote the set of symmetric action profiles by  $\tilde{\mathcal{X}}^{\mathcal{N}} = \prod_{n \in \mathcal{N}} \mathcal{X}_n$ . Let us point out that a symmetric action profile happens as a specific case in the non-cooperative game, without any coordination between the players within a population. Besides, considering the coupling constraint that  $\mathbf{X} \in A$ , we define  $\tilde{\mathcal{X}}^{\mathcal{N}}(A) = \{\mathbf{x} \in \tilde{\mathcal{X}}^{\mathcal{N}} : \mathbf{X} = \sum_{n \in \mathcal{N}} I_n \mathbf{x}_n \in A\}$ .

Let us introduce two indicators to “measure” the quality of the clustering of  $\tilde{\mathcal{G}}^{\mathcal{N}}$ :

- $\bar{\delta} = \max_{n \in \mathcal{N}} \delta_n$ , where

$$\delta_n \triangleq \max_{i \in \mathcal{I}_n} d_H(\mathcal{X}_i, \mathcal{X}_n), \quad (11)$$

- $\bar{\lambda} = \max_{n \in \mathcal{N}} \lambda_n$ , where

$$\lambda_n \triangleq \max_{i \in \mathcal{I}_n} \sup_{\mathbf{x} \in \mathcal{X}_n} d_H(\partial(-u_n)(\mathbf{x}), \partial(-u_i)(\mathbf{x})). \quad (12)$$

The quantity  $\delta_n$  measures the heterogeneity in strategy sets of populations within the group  $\mathcal{I}_n$ , while  $\lambda_n$  measures the heterogeneity in the subgradients in the group  $\mathcal{I}_n$ .

Since the auxiliary game  $\tilde{\mathcal{G}}^{\mathcal{N}}$  is to be used to compute an approximation of an equilibrium of the large game  $\mathcal{G}$ , the indicators  $\delta_n$  and  $\lambda_n$  should be minimized when defining  $\tilde{\mathcal{G}}^{\mathcal{N}}$ . Thus, we assume that  $(\mathcal{X}_n)_n$  and  $(u_n)_n$  are chosen such that the following holds:

**Assumption 2.** For each  $n \in \mathcal{N}$ , we have:

- 1)  $\mathcal{X}_n$  is in the convex hull of  $\bigcup_{i \in \mathcal{I}_n} \mathcal{X}_i$ , so that  $\|\mathcal{X}_n\| \leq m$ . Moreover, for each  $i \in \mathcal{I}_n$ ,  $\text{aff } \mathcal{X}_i \subset \text{aff } \mathcal{X}_n$ , where  $\text{aff } S$  denotes the affine hull of set  $S$ ;

2) similarly,  $u_n$  is such that  $\partial u_n(\mathbf{x})$  is contained in the convex hull of  $\bigcup_{i \in \mathcal{I}_n} \partial(-u_i)(\mathbf{x})$  for all  $\mathbf{x} \in \mathcal{X}_n$ , so that  $\|\partial(-u_n)\|_\infty \leq \max_{i \in \mathcal{I}_n} B_{u_i}$ .

An interesting case in the perspective of minimizing the quantities  $\bar{\delta}$  and  $\bar{\lambda}$  is when  $\mathcal{I}$  can be divided into homogeneous populations, as in Ex. 2 below.

**Example 2.** The player set  $\mathcal{I}$  can be divided into a small number  $N$  of subsets  $(\mathcal{I}_n)_n$ , with homogeneous players inside each subset  $\mathcal{I}_n$  (i.e., for each  $n$  and  $i, j \in \mathcal{I}_n$ ,  $\mathcal{X}_i = \mathcal{X}_j$  and  $u_i = u_j$ ). In that case, consider an auxiliary game  $\tilde{\mathcal{G}}^{\mathcal{N}}$  with  $N$  populations and, for each  $n \in \mathcal{N}$  and  $i \in \mathcal{I}_n$ ,  $\mathcal{X}_i^{\mathcal{N}} \triangleq \mathcal{X}_i$  and  $u_n \triangleq u_i$ . Then,  $\bar{\delta} = \bar{\lambda} = 0$ .

In order to approximate the SVWE of  $\mathcal{G}(A)'$  by the SVWE of an auxiliary game  $\tilde{\mathcal{G}}^{\mathcal{N}}$ , let us first state the following result on the geometry of the action sets for technical use.

**Lemma 4.** Under Asm. 1, there exists a strictly positive constant  $\rho$  and an action profile  $\mathbf{z} \in \tilde{\mathcal{X}}(A)$  such that,  $d(\mathbf{z}_i, \text{rd} \mathcal{X}_i) \geq \rho$  for all  $i \in \mathcal{I}$ , where rbd stands for the relative boundary.

*Proof.* See Appendix G.  $\square$

Lemma 4 ensures the existence of a profile  $\mathbf{z}$  such that  $\mathbf{z}_i$  has uniform distance to the relative boundary of  $\mathcal{X}_i$  for all  $i$  and that  $\mathbf{z}$  satisfies the coupling constraint.

Recall that we are only interested in *symmetric* action profiles in population games  $\mathcal{G}(A)'$  and  $\tilde{\mathcal{G}}^{\mathcal{N}}$ . Given a symmetric action profile  $\mathbf{x}^{\mathcal{N}}$  in the auxiliary game  $\tilde{\mathcal{X}}^{\mathcal{N}}$  in  $\tilde{\mathcal{G}}^{\mathcal{N}}$ , we can define a corresponding symmetric action profile of  $\mathcal{G}(A)'$  such that all the nonatomic players in the populations regrouped in  $\mathcal{I}_n$  play the same action  $\mathbf{x}_n^{\mathcal{N}}$ . (It is allowed that  $\mathbf{x}_n^{\mathcal{N}}$  be not in  $\mathcal{X}_i$ . Recall that we can extend  $u_i$  to a neighborhood of  $\mathcal{M}/I$  such that  $u_i$  is bounded on  $\mathcal{M}/I$ ). Formally, define map  $\psi: \mathbb{R}^{N\mathcal{I}} \rightarrow \mathbb{R}^{I\mathcal{I}}$ :

$$\forall \mathbf{x}^{\mathcal{N}} \in \mathbb{R}^{N\mathcal{I}}, \psi(\mathbf{x}^{\mathcal{N}}) = (\mathbf{x}_i)_{i \in \mathcal{I}} \text{ where } \mathbf{x}_i = \mathbf{x}_n^{\mathcal{N}}, \forall i \in \mathcal{I}_n.$$

Conversely, for a symmetric action profile  $\mathbf{x}$  in  $\mathcal{G}(A)'$ , we define a corresponding symmetric action profile in the auxiliary game  $\tilde{\mathcal{G}}^{\mathcal{N}}$  by the following map  $\bar{\psi}: \mathbb{R}^{I\mathcal{I}} \rightarrow \mathbb{R}^{N\mathcal{I}}$ :

$$\forall \mathbf{x} \in \mathbb{R}^{I\mathcal{I}}, \bar{\psi}(\mathbf{x}) = (\mathbf{x}_n^{\mathcal{N}})_{n \in \mathcal{N}} \text{ where } \mathbf{x}_n^{\mathcal{N}} = \frac{1}{I_n} \sum_{i \in \mathcal{I}_n} \mathbf{x}_i.$$

Thm. 5 below is the main result of this subsection. It gives an upper bound on the distance between the SVWE of the population game  $\mathcal{G}(A)'$ , which has the same dimension as the original atomic game  $\mathcal{G}(A)$ , and that of an auxiliary game  $\tilde{\mathcal{G}}^{\mathcal{N}}(A)$ , which has a reduced dimension.

**Theorem 5** (SVWE of  $\tilde{\mathcal{G}}^{\mathcal{N}}(A)$  is close to SVWE of  $\mathcal{G}(A)'$ ). Under Asms. 1 and 2, in an auxiliary game  $\tilde{\mathcal{G}}^{\mathcal{N}}(A)$ ,  $\bar{\delta}$  and  $\bar{\lambda}$  are defined by Eqs. (11) and (12), with  $\bar{\delta} < \frac{\rho}{2}$ . Let  $\hat{\mathbf{x}}$  be a SVWE of  $\tilde{\mathcal{G}}^{\mathcal{N}}(A)$ , and  $\mathbf{x}^*$  a SVWE of  $\mathcal{G}(A)$ . Then:

(1) if  $H'$  is strongly monotone with modulus  $\alpha$ , then both  $\hat{\mathbf{x}}$  and  $\mathbf{x}^*$  are unique and

$$\|\psi(\hat{\mathbf{x}}) - \mathbf{x}^*\| \leq \frac{1}{\alpha} K(\bar{\delta}, \bar{\lambda}); \quad (13)$$

(2) if  $H'$  is aggregatively strongly monotone with modulus  $\beta$ , then both  $\hat{\mathbf{X}} = \sum_{n \in \mathcal{N}} \hat{\mathbf{x}}_n$  and  $\mathbf{X}^* = \sum_{i \in \mathcal{I}} \mathbf{x}_i^*$  are unique, and

$$\|\hat{\mathbf{X}} - \mathbf{X}^*\|^2 \leq \frac{1}{\beta} K(\bar{\delta}, \bar{\lambda}), \quad (14)$$

where  $K(\bar{\delta}, \bar{\lambda})$ , appearing in both inequalities, is:

$$K(\bar{\delta}, \bar{\lambda}) \triangleq 2M(3\frac{B_f}{\rho}\bar{\delta} + \bar{\lambda}), \quad (15)$$

with  $B_f \triangleq B_c + \max_{i \in \mathcal{I}} B_{u_i}$ . In particular,

$$K(\bar{\delta}, \bar{\lambda}) = \mathcal{O}(\bar{\delta} + \bar{\lambda}) \xrightarrow{\bar{\delta}, \bar{\lambda} \rightarrow 0} 0.$$

*Proof.* See Appendix H.  $\square$

We have pointed out that the approximation error depends on how the populations are clustered according to  $\mathcal{N}$ , and is related to the heterogeneity of players in  $\mathcal{I}$  rather than their number. In particular, in the case of Ex. 2, Thm. 5 states that the (aggregate) SVWE of the auxiliary game  $\tilde{\mathcal{G}}^{\mathcal{N}}(A)$  is exactly equal to the (aggregate) SVWE of the large game  $\mathcal{G}(A)'$ .

A direct corollary of Thm. 5-(1) is that two populations in  $\mathcal{G}(A)'$  with similar characteristics have similar behavior at a SVWE there. This is the extension of Prop. 4 in the presence of coupling constraints.

**Corollary 1.** Let  $\mathbf{x}^* \in \tilde{\mathcal{X}}$  be a SVWE of game  $\mathcal{G}(A)'$ . Under Asm. 1, for two populations  $i$  and  $j$  in  $\mathcal{I}$ , if  $d_H(\mathcal{X}_i, \mathcal{X}_j) \leq \delta$ ,  $\sup_{\mathbf{x} \in \mathcal{M}/I} d_H(\partial(-u_j)(\mathbf{x}), \partial(-u_i)(\mathbf{x})) \leq \lambda$ , and  $u_i$  (resp.  $u_j$ ) is  $\alpha_i$ - (resp.  $\alpha_j$ -)strongly concave, then

$$\|\mathbf{x}_i^* - \mathbf{x}_j^*\| \leq \left( \frac{1}{\sqrt{\alpha_i}} + \frac{1}{\sqrt{\alpha_j}} \right) K(\delta, \lambda)^{1/2}.$$

### C. Combining the two steps to approximate a VNE of $\mathcal{G}(A)$

The following theorem is the main result of the paper, which combines the two steps of approximation given in Thm. 3 and in Thm. 5, in the computation of a VNE of the original game  $\mathcal{G}(A)$ .

**Theorem 6** (SVWE of  $\tilde{\mathcal{G}}^{\mathcal{N}}(A)$  is close to VNEs of  $\mathcal{G}(A)$ ). Under Asms. 1 and 2, in an auxiliary game  $\tilde{\mathcal{G}}^{\mathcal{N}}(A)$ ,  $\bar{\delta}$  and  $\bar{\lambda}$  are defined by Eqs. (11) and (12), with  $\bar{\delta} < \frac{\rho}{2}$ . Let  $\hat{\mathbf{x}}$  be a SVWE of  $\tilde{\mathcal{G}}^{\mathcal{N}}(A)$ ,  $\mathbf{x} \in \tilde{\mathcal{X}}(A)$  be a VNE of  $\mathcal{G}(A)$ ,  $\hat{\mathbf{X}} = \sum_{n \in \mathcal{N}} \mathbf{x}_n$ ,  $\mathbf{X} = \sum_{i \in \mathcal{I}} \mathbf{x}_i$ , and  $K(\bar{\delta}, \bar{\lambda})$  the constant given by (15).

(1) if  $u_i$  is  $\alpha_i$ -strongly concave for each  $i \in \mathcal{I}$ , with  $\alpha \triangleq \min_i \alpha_i$ , then  $\hat{\mathbf{x}}$  is unique and

$$\|\psi(\hat{\mathbf{x}}) - \mathbf{x}\| \leq \sqrt{\frac{1}{\alpha}} K(\bar{\delta}, \bar{\lambda})^{1/2} + M \sqrt{\frac{2TC}{\alpha I}};$$

(2) if  $c_t$  is  $\beta_t$ -strictly increasing for each  $t \in \mathcal{T}$ , with  $\beta \triangleq \min_t \beta_t$ , then  $\hat{\mathbf{X}}$  is unique and

$$\|\hat{\mathbf{X}} - \mathbf{X}\| \leq \sqrt{\frac{1}{\beta}} K(\bar{\delta}, \bar{\lambda})^{1/2} + M \sqrt{\frac{2TC}{\beta I}}.$$

*Proof.* : This is an implication of the inequalities given in Thms. 3 and 5.

Given the large game  $\mathcal{G}(A)$  and a certain  $N \in \mathbb{N}^*$ , Thm. 6 suggests that we should find the auxiliary game  $\tilde{\mathcal{G}}^{\mathcal{N}}$  that minimizes  $K(\bar{\delta}, \bar{\lambda})$  in order to have the best possible

approximation of the equilibria. This would correspond to a “clustering problem” given as follows:

$$\min_{(\mathcal{I}_n)_{n \in \mathcal{P}(\mathcal{I})}} \min_{(\mathcal{X}_n)_n} \min_{(u_n)_n} K(\bar{\delta}, \bar{\lambda}), \quad (16)$$

where  $\mathcal{P}(\mathcal{I})$  denotes the set of partitions of  $\mathcal{I}$  of cardinal  $N$ , while  $(\mathcal{X}_n)_{n \in \mathcal{N}}$  and  $(u_n)_n$  are chosen according to Asm. 2.

The value of the optimal solutions of problem (16), and thus of the quality of the approximation in Thm. 6, depends on the homogeneity of the  $I$  players in  $\mathcal{I}$  in terms of action sets and utility functions. The “ideal” case is given in Ex. 2 where  $\mathcal{I}$  is composed of a small number  $N$  of homogeneous populations and thus  $K(\bar{\delta}, \bar{\lambda}) = 0$ .

In general, solving (16) is a hard problem in itself. It is indeed a generalization of the  $k$ -means clustering problem [38] (with  $k = N$  and considering a function of Hausdorff distances), which is itself NP-hard [39]. In Sec. IV, we illustrate how we use directly the  $k$ -means algorithm to compute efficiently an approximate solution  $(\mathcal{I}_n, \mathcal{X}_n, u_n)_{n \in \mathcal{N}}$  in the parametric case.

Finally, the number  $N$  in the definition of the auxiliary game should be chosen a priori as a trade-off between the minimization of  $K(\bar{\delta}, \bar{\lambda})$  and a sufficient minimization of the dimension. Indeed, with  $\mathcal{N} = \mathcal{I}$ ,  $\mathcal{X}_i = \mathcal{X}_n$  and  $u_n = u_i$ , we get  $\bar{\lambda} = \bar{\delta} = 0$ . However, the aim of Thm. 6 is to find an auxiliary game  $\tilde{\mathcal{G}}^{\mathcal{N}}$  with  $N \ll I$  so that the dimension of the GVIs characterizing the equilibria (and thus the time needed to compute their solutions) is significantly reduced, while ensuring a relatively small error, measured by  $\bar{\lambda}$  and  $\bar{\delta}$ .

#### IV. APPLICATION TO DEMAND RESPONSE FOR ELECTRIC VEHICLE SMART CHARGING

Demand response (DR) [40] refers to a set of techniques to influence, control or optimize the electric consumption of agents in order to provide some services to the grid, e.g. reduce production costs and CO<sub>2</sub> emissions or avoid congestion [6]. The increasing number of electric vehicles (EV) offers a new source of flexibility in the optimization of the production and demand, as electric vehicles require a huge amount of energy and enjoy a sufficiently flexible charging scheme (whenever the EV is parked). Because of the privacy of each consumer or EV owner’s information and the decentralized aspects of the DR problem, many relevant works adopt a game theoretical approach by considering consumers as players minimizing a cost function and a utility [41].

In this section, we consider the consumption associated to electric vehicle charging on a set of 24-hour time-periods  $\mathcal{T} = \{1, \dots, T\}$ , with  $T = 24$ , indexing the hours from 10 PM to 9PM the day after (including the night time periods where EVs are usually parked at home).

##### A. Price functions: block rates energy prices

As in the framework described in [6], we consider a centralized entity, called the aggregator, who manages the aggregate flexible consumption. The aggregator interacts with the electricity market and energy producers, with his own objectives such as minimizing his cost or achieving a target aggregate demand profile.

The aggregator imposes electricity prices on each time-period. We consider prices taking the specific form of inclining block-rates tariffs (IBR tariffs, [42]), i.e. a piece-wise affine function  $c(\cdot)$  which depends on the aggregate-demand  $X_t = \sum_{i \in \mathcal{I}} x_{i,t}$  for each time-period  $t$ , and is defined as follows:

$$\begin{aligned} c(X) &= 1 + 0.1X \text{ if } X \leq 500 \\ c(X) &= -49 + 0.2X \text{ if } 500 \leq X \leq 1000 \\ c(X) &= -349 + 0.5X \text{ if } 1000 \leq X. \end{aligned} \quad (17)$$

This function  $c$  is continuous and convex. Those price functions are transmitted by the aggregator to each consumer or EV owner. Thus, each consumer  $i$  minimizes an objective function of the form (1), with an energy cost determined by (17) and a utility function  $u_i$  defined below. An equilibrium gives a stable situation where each consumer minimizes her objective and has no interest to deviate from her current consumption profile.

##### B. Consumers’ constraints and parameters

We simulate the consumption of  $I = 2000$  consumers who have demand constraints of the form:

$$\mathcal{X}_i = \{ \mathbf{x}_i \in \mathbb{R}_+^T : \sum_t x_{i,t} = m_i \text{ and } \underline{x}_{i,t} \leq x_{i,t} \leq \bar{x}_{i,t} \} \quad (18)$$

where  $m_i$  is the total energy needed by  $i$ , and  $\underline{x}_{i,t}, \bar{x}_{i,t}$  the (physical) bounds on the power allowed to her at time  $t$ . The utility functions have form  $u_i(\mathbf{x}_i) = -\omega_i \|\mathbf{x}_i - \mathbf{y}_i\|^2$ .

The parameters are chosen as follows:

- $m_i$  is drawn uniformly between 1 and 30 kWh, which corresponds to a typical charge of a residential electric vehicle.
  - $\underline{x}_i, \bar{x}_i$ : First, we generate, in two steps, a continual set of charging time-periods  $\mathcal{T}_i = \{h_i - \frac{\tau_i}{2}, \dots, h_i + \frac{\tau_i}{2}\}$ :
    - the duration  $\tau_i$  is uniformly drawn from  $\{4, \dots, T\}$ ;
    - $h_i$  is then uniformly drawn from  $\{1 + \frac{\tau_i}{2}, \dots, T - \frac{\tau_i}{2}\}$ .
- Next, for  $t \notin \mathcal{T}_i$ , let  $\underline{x}_{i,t} = \bar{x}_{i,t} = 0$ . Finally, for  $t \in \mathcal{T}_i$ ,  $\underline{x}_{i,t}$  (resp.  $\bar{x}_{i,t}$ ) is drawn uniformly from  $[0, \frac{m_i}{\tau_i}]$  (resp.  $[\frac{m_i}{\tau_i}, m_i]$ ).
- $\omega_i$  is drawn uniformly from  $[1, 10]$ .
  - $\mathbf{y}_i$  is taken equal to  $\bar{x}_{i,t}$  on the first time periods of  $\mathcal{T}_i$  (first available time periods) until reaching  $m_i$  (which corresponds to a profile “the sooner the better” or “plug and charge”).

##### C. Coupling constraints on capacities and limited variations

We consider the following *coupling* constraints on the aggregate demand  $\mathbf{X}$  which are often encountered in energy applications:

$$\begin{aligned} -50 &\leq X_T - X_1 \leq 50 & (19) \\ X_t &\leq 1400, \quad \forall t \in \mathcal{T} & (20) \end{aligned}$$

Here, Constraint (19) imposes that the demand  $X_T$  at the very end of the time horizon is relatively close to the first aggregate  $X_1$ , so that the demand response profiles computed for the finite time set  $\mathcal{T}$  can be applied on a day-to-day, periodical basis.



Constraint (20) is a capacity constraint, induced by the maximal capacity of the electrical lines or by the generation capacities of electricity producers.

These linear coupling constraints can be written in the closed form:

$$A\mathbf{X} \leq b, \quad (21)$$

where  $A \in \mathcal{M}_{T+2,T}(\mathbb{R})$ ,  $b \in \mathbb{R}^{T+2}$ .

#### D. Computing populations with $k$ -means

Since  $I$  is very large, determining an exact VNE is computationally demanding. Thus, we apply the clustering procedure described in Sec. III-B to regroup the players.

We use the  $k$ -means algorithm [38], where “ $k$ ” =  $N$  is the number of populations (groups) to replace the large set of  $I$  players. For each player  $i \in \mathcal{I}$ , we define her parametric description vector:

$$\mathbf{p}_i = [\omega_i, \mathbf{y}_i, m_i, \underline{\mathbf{x}}_i, \bar{\mathbf{x}}_i] \in \mathbb{R}^{3T+2}. \quad (22)$$

Then, the  $k$ -means algorithm finds an approximate solution of finding a partition  $(S_n)_{1 \leq n \leq N}$  of  $\mathcal{I}$  into  $N$  clusters. The algorithm solves the combinatorial minimization problem:

$$\min_{S_1, \dots, S_N} \sum_{1 \leq n \leq N} \sum_{\mathbf{p} \in S_n} \|\mathbb{E}_{S_n}(\mathbf{p}) - \mathbf{p}\|^2 = \min_{S_1, \dots, S_N} \sum_{1 \leq n \leq N} |S_n| \text{Var}(S_n),$$

where  $\mathbb{E}_{S_n}(\mathbf{p}) = \frac{1}{|S_n|} \sum_{i \in S_n} \mathbf{p}_i$  denotes the average value of  $\mathbf{p}$  over the set  $S_n$ . These average values are taken to be  $w_n$ ,  $\mathbf{y}_n$ ,  $m_n$ ,  $\underline{\mathbf{x}}_n$  and  $\bar{\mathbf{x}}_n$ .

The simulations are run with different population numbers, with  $N \ll I$  chosen among  $\{5, 10, 20, 50, 100\}$ .

Since the  $k$ -means algorithm minimizes the squared distance of the average vector of parameters in  $S_n$  to the vectors of parameters of the points in  $S_n$ , the clustered populations obtained can be sub-optimal in terms of  $K(\bar{\delta}, \bar{\lambda})$ . As explained above, choosing the optimal populations  $\mathcal{N}$ , as formulated in problem (16), is a complex problem in itself which deserves further research. Our example shows that the  $k$ -means algorithm gives a practical and efficient way to compute a heuristic solution in the case where  $u_i$  and  $\mathcal{X}_i$  are parameterized.

#### E. Computation methods

We compute a VNE (Def. 2) with the original set of  $I$  players and the approximating SVWE (Def. 3) as solutions of the associated GVI (2).

We employ a standard projected descent algorithm, as described in [20, Algo. 2] and recalled below in Algorithm 1. It is adapted to the subdifferentiable case that we consider in this work. In particular, the fixed step  $\tau$  used in [20] is replaced by a variable step  $\tau^{(k)} = 1/k$ . The coupling constraint (21) is relaxed and the Lagrangian multipliers  $\boldsymbol{\lambda} \in \mathbb{R}_+^{T+2}$  associated to these constraints are considered as extra variables. Thus, we can perform the projections on the sets  $\mathcal{X}_i$  and on  $\mathbb{R}_+^{T+2}$ .

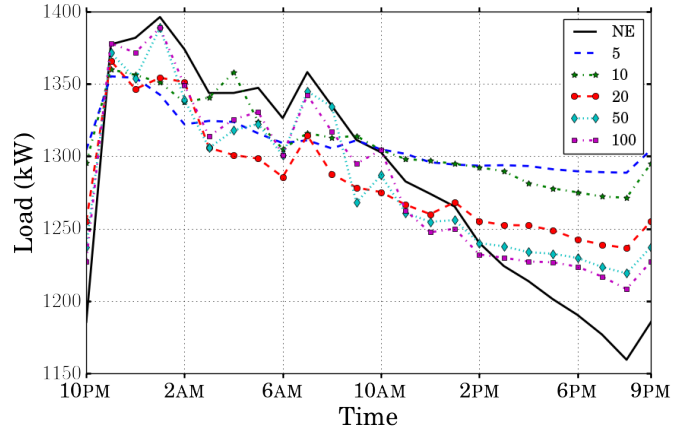


Fig. 2: Convergence of the aggregate SVWE profile of auxiliary games to a VNE profile of the original game.

---

#### Algo. 1 Projected Descent Algorithm

---

**Require:**  $\mathbf{x}^{(0)}, \boldsymbol{\lambda}^{(0)}$ , stopping criterion

- 1:  $k \leftarrow 0$
  - 2: **while** stopping criterion not true **do**
  - 3:   **for**  $n = 1$  to  $N$  **do**
  - 4:     take  $g_n^{(k)} \in \partial_1 f_n(\mathbf{x}_n^{(k)}, \mathbf{X}^{(k)})$
  - 5:      $\mathbf{x}_n^{(k+1)} \leftarrow \Pi_{\mathcal{X}_n} \left( \mathbf{x}_n^{(k)} - \tau^{(k)}(g_n^{(k)} + \boldsymbol{\lambda}^{(k)T} A) \right)$
  - 6:   **end for**
  - 7:    $\boldsymbol{\lambda}^{(k+1)} \leftarrow (\boldsymbol{\lambda}^{(k)} - \tau^{(k)}(b - 2A\mathbf{X}^{(k+1)} + A\mathbf{X}^{(k)}))^+$
  - 8:    $k \leftarrow k + 1$
  - 9: **end while**
- 

The convergence of Algorithm 1 is shown in [23, Thm.3.1]. The stopping criterion that we adopt here is the distance between two iterates: the algorithm stops when  $\|(\boldsymbol{\lambda}^{(k+1)}, \mathbf{x}^{(k+1)}) - (\boldsymbol{\lambda}^{(k)}, \mathbf{x}^{(k)})\|_2 \leq 10^{-3}$ .

Due to the form of the strategy sets considered (18), the projection steps (Line 5) can be computed efficiently and exactly in  $\mathcal{O}(T)$  with the Brucker algorithm [43]. However, if we consider more general strategy sets (arbitrary convex sets), this projection step can be costly: in that case, other algorithms such as [24] would be more efficient.

#### F. A trade-off between precision and computation time

Simulations were run using Python on a single core Intel Xeon @3.4Ghz and 16GB of RAM.

Fig. 2 shows the different aggregate SVWE profiles  $\hat{\mathbf{X}}^{\mathcal{N}}$  obtained for sets  $\mathcal{N}$  of different sizes, as well as a VNE of the original game for comparison. Thanks to the specific form of the strategy sets (18)—which enables a fast projection—we are able to compute a VNE  $\mathbf{X}$  of the original game with  $I = 2000$  players.

Fig. 3a and Fig. 3b show the two main metrics to consider to choose a relevant number of populations  $N$ : the precision of the SVWE approximating the equilibrium (measured by the distance of the aggregate SVWE profile to the aggregate profile of the VNE computed along), and the CPU time needed to compute the SVWE.

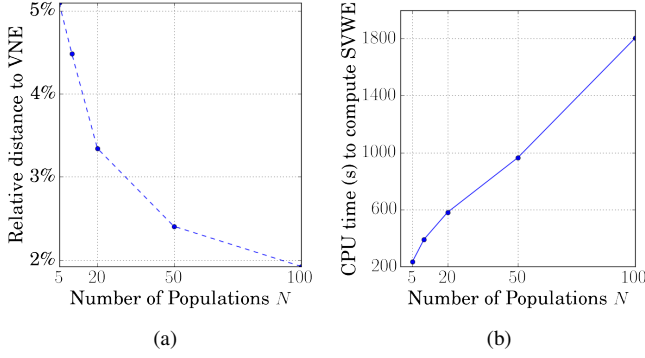


Fig. 3: (a) Relative error to actual VNE ; (b) Time to compute SVWE. *The time needed to compute SVWE (b) increases roughly linearly with  $N$ , at a faster rate than the error on the approximation of the VNE decreases (a).*

First notice on Fig. 3a that the distance between the aggregate equilibrium profile and its estimation decreases with  $N$  at a sublinear rate. This is partially explained in light of Thm. 5 and in addition with the following remarks:

- the Hausdorff distance of two parameterized polyhedral sets is Lipschitz continuous w.r.t their parameter vectors (generalization of [44]), which ensures that there is  $K > 0$  s.t. for all  $n$ :

$$\delta_n^{\mathcal{N}} = \max_{i \in \mathcal{I}_n^{\mathcal{N}}} d_H(\mathcal{X}_i, \mathcal{X}_n^{\mathcal{N}}) \leq K \left\| \begin{pmatrix} m_n \\ \underline{x}_n \\ \bar{x}_n \end{pmatrix} - \begin{pmatrix} m_i \\ \underline{x}_i \\ \bar{x}_i \end{pmatrix} \right\| ;$$

- similarly, as subgradients of utility functions are reduced to a point, one has, for all  $n$ :

$$\begin{aligned} \lambda_n &= \max_{i \in \mathcal{I}_n} \max_{\mathbf{x} \in \mathcal{M}} 2 \|\omega_n(\mathbf{x} - \mathbf{y}_n) - \omega_i(\mathbf{x} - \mathbf{y}_i)\| \\ &= \mathcal{O}(|\omega_n - \omega_i| + \|\mathbf{y}_n - \mathbf{y}_i\|) . \end{aligned}$$

Fig. 3b shows the CPU time needed to compute the WE with a stopping criterion of a maximum improvement between iterates of  $\|(\lambda^{(k+1)}, \mathbf{x}^{(k+1)}) - (\lambda^{(k)}, \mathbf{x}^{(k)})\|_2 \leq 10^{-3}$ . Computing a solution of the clustering problem with the  $k$ -means algorithm takes, for each value of  $N$ , less than ten seconds. This time is negligible in comparison to the time needed for convergence of Algorithm 1.

As a reference time, to compute a VNE of the original game (observed on Fig. 2) with the same stopping criterion and the same CPU configuration, we needed 3 hours 26 minutes. This is more than six times longer than the CPU time to compute the SVWE with one hundred populations.

On this figure, we see that the CPU time evolves linearly with the number of populations  $N$ . This is explained by the structure of Algorithm 1, as each iteration  $k$  is executed in a time proportional to  $N$  due to the **for** loop.

Last, one observes from Fig. 3a that, in our example, the error between the aggregate demand profile at equilibrium and its approximation is between 2% and 5%, which remains significant. However, as pointed out in Sec. III, the quality of the approximation depends on the heterogeneity of the set of players  $\mathcal{I}$ . In the example of this section, as the parameters are drawn uniformly (see Sec. IV-B), the set of players  $\mathcal{I}$  presents

a large variance so that it is a “worst” case as opposed to the case of Ex. 2 which is “optimal”.

## V. CONCLUSION

This paper shows that equilibria in splittable congestion games with a very large number of atomic players can be approximately computed with a Wardrop equilibrium of an auxiliary population game of smaller dimension. Our results give explicit bounds on the distance of this approximating equilibrium to the equilibria of the original large game. These theoretical results can be used in practice to solve, by an iterative method, complex nonconvex bi-level programs where the lower level is the equilibrium of a large congestion game, for instance, to optimize tariffs or tolls for the operator of a network. A detailed analysis of such a procedure would be an extension of the present work.

## VI. ACKNOWLEDGMENTS

We thank the PGMFO foundation for the financial support to our project “*Jeux de pilotage de flexibilités de consommation électrique : dynamique et aspect composite*”.

## APPENDIX A

### PROOF OF LEMMA 1: EXPRESSIONS OF SUBGRADIENTS

Recall that  $\hat{f}_i(\mathbf{x}_i, \mathbf{X}_{-i}) \triangleq f_i(\mathbf{x}_i, \mathbf{X}_{-i} + \mathbf{x}_i)$ . According to [45, Proposition 16.6],  $\partial_1 \hat{f}_i(\mathbf{x}_i, \mathbf{X}_{-i}) \subset \{(I_T, I_T) \mathbf{g} : \mathbf{g} \in \partial \psi_i(\mathbf{x}_i)\}$ , where  $\partial f_i(\mathbf{x}_i, \mathbf{X}_{-i} + \mathbf{x}_i)$  is the subdifferential of  $\psi_i(\cdot) \triangleq f_i(\cdot, \mathbf{X}_{-i} + \cdot)$ , at  $\mathbf{x}_i$ . On the other hand, according to [45, Proposition 16.7],  $\partial \psi_i(\mathbf{x}_i)$  is a subset of:

$$\left\{ (\mathbf{g}_{i,1}, \mathbf{g}_{i,2}) : \mathbf{g}_{i,1} \in \partial_1 f_i(\mathbf{w}_i, \mathbf{Y})|_{\mathbf{w}_i=\mathbf{x}_i, \mathbf{Y}=\mathbf{X}_{-i}+\mathbf{x}_i}, \right. \\ \left. \mathbf{g}_{i,2} \in \partial_2 f_i(\mathbf{w}_i, \mathbf{Y})|_{\mathbf{w}_i=\mathbf{x}_i, \mathbf{Y}=\mathbf{X}_{-i}+\mathbf{x}_i} \right\} .$$

Therefore,  $\partial_1 \hat{f}_i(\mathbf{x}_i, \mathbf{X}_{-i})$  is a subset of:

$$\begin{aligned} &\{\mathbf{c}(\mathbf{X}) + \mathbf{g}'_{i,1} + \mathbf{g}_{i,2} : \mathbf{g}'_{i,1} \in \partial(-u_i)(\mathbf{x}_i), \mathbf{g}_{i,2} \in \partial_2 f_i(\mathbf{x}_i, \mathbf{X})\} \\ &= \{\mathbf{c}(\mathbf{X}) + \mathbf{g}'_{i,1} + (a_{i,t} x_{i,t})_t : \\ &\quad \mathbf{g}'_{i,1} \in \partial(-u_i)(\mathbf{x}_i), a_{i,t} \in \partial c_t(X_t) \forall t \in T\} . \end{aligned}$$

By the definition of subdifferential, it is easy to show that  $\{\mathbf{c}(\mathbf{X}) + \mathbf{g}'_{i,1} + (a_{i,t} x_{i,t})_t : \mathbf{g}'_{i,1} \in \partial(-u_i)(\mathbf{x}_i), a_{i,t} \in \partial c_t(X_t) \forall t \in T\}$  is a subset of  $\partial_1 \hat{f}_i(\mathbf{x}_i, \mathbf{X}_{-i})$ .

The proof for  $\partial_1 f_i(\mathbf{x}_i, \mathbf{X})$  is similar.

## APPENDIX B

### PROOF OF PROP. 1: EXISTENCE OF EQUILIBRIA

It is easy to see that  $\hat{f}_i(\cdot, \mathbf{X}_{-i})$  is convex on  $\mathcal{X}_i$  for all  $\mathbf{X}_{-i}$  in  $\sum_{j \in \mathcal{I} \setminus \{i\}} \mathcal{X}_j$ ,  $f_i(\cdot, \mathbf{X})$  is convex on  $\mathcal{X}_i$  for all  $\mathbf{X} \in \bar{\mathcal{X}}$ , and  $H$  and  $H'$  are nonempty, convex, compact valued, upper hemicontinuous correspondences. Then [46, Corollary 3.1] shows that the GVI problems (2) and (3) both admit a solution on the finite dimensional convex compact  $\tilde{\mathcal{X}}(A)$ .

## APPENDIX C

## PROOF OF PROP. 2: UNIQUENESS OF EQUILIBRIA

We prove for SVWE only and the proof for VNE is the same. Suppose that  $\mathbf{x}, \mathbf{y} \in \tilde{\mathcal{X}}(A)$  are both SVWE, with  $\mathbf{X} = \sum_i \mathbf{x}_i$  and  $\mathbf{Y} = \sum_i \mathbf{y}_i$ . According to the definition of SVWE, there is  $\mathbf{g} \in H'(\mathbf{x})$  and  $\mathbf{h} \in H'(\mathbf{y})$  such that  $\sum_i \langle \mathbf{g}_i, \mathbf{y}_i - \mathbf{x}_i \rangle \geq 0$  and  $\sum_i \langle \mathbf{h}_i, \mathbf{x}_i - \mathbf{y}_i \rangle \geq 0$ . Adding up these two inequalities yields  $\sum_i \langle \mathbf{g}_i - \mathbf{h}_i, \mathbf{y}_i - \mathbf{x}_i \rangle \geq 0$ .

(1) If  $H'$  is a strictly monotone, then  $\sum_i \langle \mathbf{g}_i - \mathbf{h}_i, \mathbf{x}_i - \mathbf{y}_i \rangle = 0$  and thus  $\mathbf{x} = \mathbf{y}$ .

(2-3) If  $H'$  is an aggregatively strictly monotone, then  $\sum_i \langle \mathbf{g}_i - \mathbf{h}_i, \mathbf{x}_i - \mathbf{y}_i \rangle = 0$  and thus  $\mathbf{X} = \mathbf{Y}$ . If there is no aggregative constraint and  $u_i$  is strictly concave, then  $\mathbf{x}_i$  (resp.  $\mathbf{y}_i$ ) is the unique minimizer of  $f_i(\cdot, \mathbf{X})$  (resp.  $f_i(\cdot, \mathbf{Y})$ ). Since  $\mathbf{X} = \mathbf{Y}$ , one has  $\mathbf{x}_i = \mathbf{y}_i$ .

## APPENDIX D

PROOF OF PROP. 3: MONOTONICITY OF  $H'$ 

(1) Let  $\mathbf{x}, \mathbf{y} \in \tilde{\mathcal{X}}$  and  $\mathbf{X} = \sum_i \mathbf{x}_i$ ,  $\mathbf{Y} = \sum_i \mathbf{y}_i$ . Recall that

$$\begin{aligned} \partial_1 f_i(\mathbf{x}_i, \mathbf{X}) &= \{ \mathbf{c}(\mathbf{X}) + \mathbf{g} : \mathbf{g} \in \partial(-u_i(\mathbf{x}_i)) \}, \\ \partial_1 f_i(\mathbf{y}_i, \mathbf{Y}) &= \{ \mathbf{c}(\mathbf{Y}) + \mathbf{h} : \mathbf{h} \in \partial(-u_i(\mathbf{y}_i)) \}. \end{aligned}$$

Let  $\mathbf{g}_i \in \partial(-u_i(\mathbf{x}_i))$  and  $\mathbf{h}_i \in \partial(-u_i(\mathbf{y}_i))$ . One has  $\langle \mathbf{g}_i(\mathbf{x}_i) - \mathbf{h}_i(\mathbf{y}_i), \mathbf{x}_i - \mathbf{y}_i \rangle \geq 0$  because  $u_i$  is concave so that  $\partial(-u_i)$  is monotone on  $\mathcal{X}_i$ . Then we get:

$$\begin{aligned} &\sum_i \langle (\mathbf{c}(\mathbf{X}) + \mathbf{g}_i(\mathbf{x}_i)) - (\mathbf{c}(\mathbf{Y}) + \mathbf{h}_i(\mathbf{y}_i)), \mathbf{x}_i - \mathbf{y}_i \rangle \\ &= \langle \mathbf{c}(\mathbf{X}) - \mathbf{c}(\mathbf{Y}), \mathbf{X} - \mathbf{Y} \rangle + \sum_i \langle \mathbf{g}_i(\mathbf{x}_i) - \mathbf{h}_i(\mathbf{y}_i), \mathbf{x}_i - \mathbf{y}_i \rangle \end{aligned}$$

$\geq 0$  because  $\mathbf{c}$  is monotone. Hence  $H'$  is monotone.

(2) By the definition of  $\alpha_i$ -strong concavity:

$$\sum_i \langle \mathbf{g}_i(\mathbf{x}_i) - \mathbf{h}_i(\mathbf{y}_i), \mathbf{x}_i - \mathbf{y}_i \rangle \geq \sum_i \alpha_i \|\mathbf{x}_i - \mathbf{y}_i\|^2 \geq \alpha \|\mathbf{x} - \mathbf{y}\|^2.$$

(3) By the definition of  $\beta_t$ -strong monotonicity:

$$\begin{aligned} \langle \mathbf{c}(\mathbf{X}) - \mathbf{c}(\mathbf{Y}), \mathbf{X} - \mathbf{Y} \rangle &= \sum_t \langle c_t(X_t) - c_t(Y_t), X_t - Y_t \rangle \\ &\geq \sum_t \beta_t \|X_t - Y_t\|^2 \geq \beta \|\mathbf{X} - \mathbf{Y}\|^2. \end{aligned}$$

## APPENDIX E

## PROOF OF THM. 2 : VNEs ARE CLOSE TO EACH OTHER

(1) Let  $\mathbf{x}, \mathbf{y} \in \tilde{\mathcal{X}}$  be two VNEs. Then, by (2), there are  $\mathbf{g}_i \in \partial_1 \hat{f}_i(\mathbf{x}_i, \mathbf{X}_{-i})$  and  $\mathbf{h}_i \in \partial_1 \hat{f}_i(\mathbf{y}_i, \mathbf{Y}_{-i})$  for each  $i$  with  $\mathbf{g}_i = \mathbf{c}(\mathbf{X}) + \mathbf{g}'_i + (a_{i,t} x_{i,t})_t$ ,  $\mathbf{h}_i = \mathbf{c}(\mathbf{Y}) + \mathbf{h}'_i + (b_{i,t} y_{i,t})_t$ , where  $\mathbf{g}'_i \in \partial(-u_i)(\mathbf{x}_i)$ ,  $\mathbf{h}'_i \in \partial(-u_i)(\mathbf{y}_i)$ ,  $a_{i,t} \in \partial c_t(X_t)$  and  $b_{i,t} \in \partial c_t(Y_t)$  for all  $t$ , such that  $\sum_i \langle \mathbf{g}_i, \mathbf{y}_i - \mathbf{x}_i \rangle \geq 0$  and  $\sum_i \langle \mathbf{h}_i, \mathbf{x}_i - \mathbf{y}_i \rangle \geq 0$ . Summing up these two inequalities:

$$\begin{aligned} 0 &\leq \sum_i \langle \mathbf{g}_i - \mathbf{h}_i, \mathbf{y}_i - \mathbf{x}_i \rangle \\ &= \sum_i \langle \mathbf{c}(\mathbf{X}) + \mathbf{g}'_i + (a_{i,t} x_{i,t})_t - \mathbf{c}(\mathbf{Y}) - \mathbf{h}'_i - (b_{i,t} y_{i,t})_t, \mathbf{y}_i - \mathbf{x}_i \rangle \\ &= \sum_i \langle \mathbf{c}(\mathbf{X}) - \mathbf{c}(\mathbf{Y}), \mathbf{y}_i - \mathbf{x}_i \rangle + \sum_i \langle \mathbf{g}'_i - \mathbf{h}'_i, \mathbf{y}_i - \mathbf{x}_i \rangle \\ &\quad + \sum_i \langle (a_{i,t} x_{i,t})_t - (b_{i,t} y_{i,t})_t, \mathbf{y}_i - \mathbf{x}_i \rangle \\ &= \langle \mathbf{c}(\mathbf{X}) - \mathbf{c}(\mathbf{Y}), \mathbf{Y} - \mathbf{X} \rangle + \sum_i \langle \mathbf{g}'_i - \mathbf{h}'_i, \mathbf{y}_i - \mathbf{x}_i \rangle \\ &\quad + \sum_i \langle (a_{i,t} x_{i,t})_t - (b_{i,t} y_{i,t})_t, \mathbf{y}_i - \mathbf{x}_i \rangle. \end{aligned}$$

$$\begin{aligned} &\implies \langle \mathbf{c}(\mathbf{X}) - \mathbf{c}(\mathbf{Y}), \mathbf{X} - \mathbf{Y} \rangle + \sum_i \langle \mathbf{g}'_i - \mathbf{h}'_i, \mathbf{x}_i - \mathbf{y}_i \rangle \\ &\leq - \sum_i \sum_t (a_{i,t} x_{i,t} - b_{i,t} y_{i,t})(x_{i,t} - y_{i,t}) \\ &\leq \sum_{i,t} (2C \frac{M}{T})(2 \frac{M}{T}) = \frac{4TCM^2}{T}. \end{aligned}$$

Since  $\mathbf{c}$  is monotone and so are  $\partial u_i$ 's because  $u_i$ 's are concave,  $\langle \mathbf{c}(\mathbf{X}) - \mathbf{c}(\mathbf{Y}), \mathbf{X} - \mathbf{Y} \rangle \geq 0$ ,  $\sum_i \langle \mathbf{g}'_i - \mathbf{h}'_i, \mathbf{x}_i - \mathbf{y}_i \rangle \geq 0$ .

If for each  $i$ ,  $u_i$ 's are  $\alpha_i$ -strongly concave, then  $\alpha \sum_i \|\mathbf{x}_i - \mathbf{y}_i\|^2 \leq \sum_i \alpha_i \|\mathbf{x}_i - \mathbf{y}_i\|^2 \leq \sum_i \langle \mathbf{g}'_i - \mathbf{h}'_i, \mathbf{x}_i - \mathbf{y}_i \rangle \leq \frac{4TCM^2}{I}$  so that  $\|\mathbf{x} - \mathbf{y}\| \leq 2M \sqrt{\frac{TC}{\alpha I}}$ .

If for each  $t$ ,  $c_t$  is  $\beta_t$ -strictly increasing, then  $\beta \|\mathbf{X} - \mathbf{Y}\|^2 \leq \langle \mathbf{c}(\mathbf{X}) - \mathbf{c}(\mathbf{Y}), \mathbf{X} - \mathbf{Y} \rangle \leq \frac{4TCM^2}{T}$  thus  $\|\mathbf{X} - \mathbf{Y}\| \leq 2M \sqrt{\frac{TC}{\beta T}}$ .

## APPENDIX F

## PROOF OF PROP. 4: SWE BEHAVIOR FOR SIMILAR PLAYERS

Let  $\mathbf{g}'_i(\mathbf{x}_i^*) \in \partial(-u_i)(\mathbf{x}_i^*)$  be s.t., for all  $\mathbf{x}_i \in \mathcal{X}_i$ ,  $\langle \mathbf{c}(\mathbf{X}^*) + \mathbf{g}'_i(\mathbf{x}_i^*), \mathbf{x}_i^* - \mathbf{x}_i \rangle \leq 0$ . Let  $\mathbf{h}'_i(\mathbf{x}_i^*) \in \partial(-u_i)(\mathbf{x}_i^*)$  be such that  $\|\mathbf{h}'_i(\mathbf{x}_i^*) - \mathbf{g}'_j(\mathbf{x}_j^*)\| \leq \lambda$ . Then, by the strong concavity of  $u_i$ :

$$\begin{aligned} \alpha_i \|\mathbf{x}_i^* - \mathbf{x}_j^*\|^2 &\leq \langle \mathbf{g}'_i(\mathbf{x}_i^*) - \mathbf{h}'_i(\mathbf{x}_i^*), \mathbf{x}_i^* - \mathbf{x}_j^* \rangle \\ &= \langle \mathbf{g}'_i(\mathbf{x}_i^*) - \mathbf{g}'_j(\mathbf{x}_j^*) + \mathbf{g}'_j(\mathbf{x}_j^*) - \mathbf{h}'_i(\mathbf{x}_i^*), \mathbf{x}_i^* - \mathbf{x}_j^* \rangle \\ &\leq \langle \mathbf{g}'_i(\mathbf{x}_i^*) - \mathbf{g}'_j(\mathbf{x}_j^*), \mathbf{x}_i^* - \mathbf{x}_j^* \rangle + \lambda 2m \\ &= \langle \mathbf{g}'_i(\mathbf{x}_i^*) + \mathbf{c}(\mathbf{X}^*), \mathbf{x}_i^* - \mathbf{x}_j^* \rangle \\ &\quad + \langle \mathbf{g}'_j(\mathbf{x}_j^*) + \mathbf{c}(\mathbf{X}^*), \mathbf{x}_j^* - \mathbf{x}_i^* \rangle + 2\lambda m \\ &= \langle \mathbf{g}'_i(\mathbf{x}_i^*) + \mathbf{c}(\mathbf{X}^*), \mathbf{x}_i^* - \Pi_i(\mathbf{x}_j^*) + \Pi_i(\mathbf{x}_j^*) - \mathbf{x}_j^* \rangle \\ &\quad + \langle \mathbf{g}'_j(\mathbf{x}_j^*) + \mathbf{c}(\mathbf{X}^*), \mathbf{x}_j^* - \Pi_j(\mathbf{x}_i^*) + \Pi_j(\mathbf{x}_i^*) - \mathbf{x}_i^* \rangle + 2\lambda m \\ &\leq \langle \mathbf{g}'_i(\mathbf{x}_i^*) + \mathbf{c}(\mathbf{X}^*), \Pi_i(\mathbf{x}_j^*) - \mathbf{x}_j^* \rangle \\ &\quad + \langle \mathbf{g}'_j(\mathbf{x}_j^*) + \mathbf{c}(\mathbf{X}^*), \Pi_j(\mathbf{x}_i^*) - \mathbf{x}_i^* \rangle + 2\lambda m \\ &\leq (B_{u_i} + B_{u_j} + 2B_c)\delta + 2\lambda m. \end{aligned}$$

where  $\Pi_i$  (resp.  $\Pi_j$ ) is the projector on  $\mathcal{X}_i$  (resp.  $\mathcal{X}_j$ ).

## APPENDIX G

## PROOF OF LEMMA 4: EXISTENCE OF INTERIOR PROFILE

Let  $\bar{\mathbf{x}} \in \tilde{\mathcal{X}}$  be s.t.  $d(\bar{\mathbf{x}}_i, \text{rbd } \mathcal{X}_i) = \max_{\mathbf{x} \in \mathcal{X}_i} d(\mathbf{x}, \text{rbd } \mathcal{X}_i) \triangleq \eta_i$ , for all  $i$ . Denote  $\bar{\mathbf{X}} = \sum_i \bar{\mathbf{x}}_i$  and  $\eta = \min_i \eta_i > 0$ .

Let  $\mathbf{y} \in \tilde{\mathcal{X}}(A)$  and  $\mathbf{Y} = \sum_i \mathbf{y}_i$  be s.t.  $d(\mathbf{Y}, \text{rbd } A) = \max_{\mathbf{X} \in \bar{\mathcal{X}} \cap A} d(\mathbf{X}, \text{rbd } A)$ . Denote  $t = \frac{d(\mathbf{Y}, \text{rbd } A)}{3M}$ .

Define  $\mathbf{z} = \mathbf{y} - t(\mathbf{y} - \bar{\mathbf{x}}) \in \tilde{\mathcal{X}}$ . Let  $\mathbf{Z} = \sum_i \mathbf{z}_i$ .

Firstly,  $\|\mathbf{Y} - \mathbf{Z}\| = t\|\mathbf{Y} - \bar{\mathbf{X}}\| \leq t2M \leq \frac{2}{3}d(\mathbf{Y}, \text{rbd } A)$ , hence  $\mathbf{Z} \in \bar{\mathcal{X}} \cap \text{ri } A$ , where  $\text{ri}$  means the relative interior. Besides, for any  $i$ ,  $\mathbf{z}_i = \mathbf{y}_i - t(\mathbf{y}_i - \bar{\mathbf{x}}_i)$ . Since  $d(\bar{\mathbf{x}}_i, \text{rbd } \mathcal{X}_i) \geq \eta$ ,  $\mathbf{y}_i \in \mathcal{X}_i$ , and  $\mathcal{X}_i$  is convex, one has  $d(\mathbf{z}_i, \text{rbd } \mathcal{X}_i) \geq \eta t = \frac{\eta}{3M}d(\mathbf{Y}, \text{rbd } A)$ . Finally, define  $\rho \triangleq \frac{\eta}{3M}d(\mathbf{Y}, \text{rbd } A)$ .

## APPENDIX H

## PROOF OF THM. 5: APPROXIMATION OF SVWE

**Lemma 7.**

(1) For each  $n \in \mathcal{N}$  and  $\mathbf{x} \in \mathcal{X}_n$ , if  $d(\mathbf{x}, \text{rbd } \mathcal{X}_n) > \delta_n$ , then  $\mathbf{x} \in \mathcal{X}_i$  for each  $i \in \mathcal{I}_n$ .

(2) For each  $n \in \mathcal{N}$ ,  $i \in \mathcal{I}_n$  and  $\mathbf{x} \in \mathcal{X}_i$ , if  $d(\mathbf{x}, \text{rbd } \mathcal{X}_i) > \delta_n$ , then  $\mathbf{x} \in \mathcal{X}_n$ .

*Proof of Lemma 7.* (1) Suppose  $\mathbf{x} \notin \mathcal{X}_i$ . Let  $\mathbf{y} \triangleq \Pi_{\mathcal{X}_i}(\mathbf{x}) \neq \mathbf{x}$ . As  $\mathbf{y} \in \text{aff } \mathcal{X}_i \subset \text{aff } \mathcal{X}_n$ , then  $\mathbf{x} - \mathbf{y} \in \text{aff } \mathcal{X}_n$ . Let  $\mathbf{z} \triangleq \mathbf{x} + \delta_n \frac{\mathbf{x} - \mathbf{y}}{\|\mathbf{x} - \mathbf{y}\|}$ . Then,  $\mathbf{z} \in \mathcal{X}_n$  because  $\|\mathbf{z} - \mathbf{x}\| \leq \delta_n$ . By the convexity of  $\mathcal{X}_i$  and the definition of  $\mathbf{y}$ , we have  $d(\mathbf{z}, \mathcal{X}_i) = d(\mathbf{x}, \mathcal{X}_i) + \delta_n > \delta_n$ , contradicting the fact that  $\delta_n \geq d_H(\mathcal{X}_i, \mathcal{X}_n)$ . (2) Symmetric proof.  $\square$

**Lemma 8.** Under Asm. 1, if  $\bar{\delta} < \frac{\rho}{2}$ , then

(1) for each  $\mathbf{x} \in \tilde{\mathcal{X}}^{\mathcal{N}}(A)$ , there is  $\mathbf{w} \in \tilde{\mathcal{X}}(A)$  such that  $\|\mathbf{w}_i - \psi_i(\mathbf{x})\| \leq 4m\frac{\bar{\delta}}{\rho}$  for each  $i \in \mathcal{I}$ ;

(2) for each  $\mathbf{x} \in \tilde{\mathcal{X}}(A)$ , there is  $\mathbf{w} \in \tilde{\mathcal{X}}^{\mathcal{N}}(A)$  such that  $\|\mathbf{w}_n - \bar{\psi}_n(\mathbf{x})\| \leq 2mI_n\frac{\bar{\delta}}{\rho}$  for each  $n \in \mathcal{N}$ .

*Proof of Lemma 8.*

(1) For  $\mathbf{x} \in \tilde{\mathcal{X}}^{\mathcal{N}}(A)$ , define  $\mathbf{w} \in \tilde{\mathcal{X}}$  as follows:  $\forall n \in \mathcal{N}$ ,  $\forall i \in \mathcal{I}_n$ , let  $\mathbf{w}_i \triangleq \mathbf{x}_n + t(\mathbf{z}_i - \mathbf{x}_n)$  where  $\mathbf{z}$  is defined in Lemma 4, with  $t \triangleq 2\bar{\delta}/\rho < 1$ .

On the one hand,  $\forall n \in \mathcal{N}$ ,  $\forall i \in \mathcal{I}_n$ ,  $d(\mathbf{z}_i, \text{rbd } \mathcal{X}_n) \geq \rho - \bar{\delta}$  implies that  $d(\mathbf{w}_i, \text{rbd } \mathcal{X}_n) \geq t(\rho - \bar{\delta}) > t\rho/2 = \bar{\delta}$ . (This is because each point in the ball with radius  $t(\rho - \bar{\delta})$  centered at  $\mathbf{w}_i$  is on the segment linking  $\mathbf{x}_n$  and some point in the ball with radius  $\rho - \bar{\delta}$  centered at  $\mathbf{z}_i$  which is contained in  $\mathcal{X}_i$ .) Thus,  $\mathbf{w}_i \in \mathcal{X}_i \forall i \in \mathcal{I}_n$  according to Lemma 7.(1). On the other hand, the linear mapping  $S : \mathbb{R}^{IT} \ni \mathbf{v} \mapsto \sum_{i \in \mathcal{I}} \mathbf{v}_i$  maps the segment linking  $\psi(\mathbf{x})$  and  $\mathbf{z}$  in  $\tilde{\mathcal{X}}(A)$  to a segment linking  $\mathbf{X} = \sum_n I_n \mathbf{x}_n$  and  $\mathbf{Z}$  in the convex  $A$ . Hence  $\sum_{i \in \mathcal{I}} \mathbf{w}_i = t\mathbf{Z} + (1-t)\mathbf{X}$  is in  $A$  as well. Therefore,  $\mathbf{w} \in \mathcal{X}(A)$ .

Finally,  $\|\mathbf{w}_i - \psi_i(\mathbf{x})\| = t\|\mathbf{z}_i - \psi_i(\mathbf{x})\| \leq t2m = 4m\frac{\bar{\delta}}{\rho}$ .

(2) For  $\mathbf{x} \in \tilde{\mathcal{X}}(A)$ , let  $\mathbf{y} \triangleq \mathbf{x} + t(\mathbf{z} - \mathbf{x})$  with  $t \triangleq \frac{\bar{\delta}}{\rho}$ . Then, by similar arguments as above,  $d(\mathbf{y}_i, \text{rbd } \mathcal{X}_i) \geq \bar{\delta}$  hence  $\mathbf{y}_i \in \mathcal{X}_n$  and  $\bar{\psi}(\mathbf{y}) \in \tilde{\mathcal{X}}^{\mathcal{N}}$ . Besides,  $\sum_i \mathbf{y}_i = t\mathbf{Z} + (1-t)(\sum_i \mathbf{x}_i)$  so that  $\sum_i \mathbf{y}_i$  is in the convex  $A$ . Hence  $\mathbf{w} \triangleq \bar{\psi}(\mathbf{y}) \in \tilde{\mathcal{X}}^{\mathcal{N}}(A)$ . Finally,  $\|\mathbf{w}_n - \psi_n(\mathbf{x})\| = t\|\sum_{i \in \mathcal{I}_n} (\mathbf{z}_i - \mathbf{x}_i)\| \leq 2mI_n\frac{\bar{\delta}}{\rho}$ .  $\square$

Let  $\mathbf{w} \in \tilde{\mathcal{X}}(A)$  be s.t.  $\forall i \in \mathcal{I}$ ,  $\|\mathbf{w}_i - \psi_i(\hat{\mathbf{x}})\| \leq 4m\bar{\delta}/\rho$  (cf. Lemma 8). Since  $\mathbf{x}^*$  is a SVWE in  $\mathcal{G}(A)'$ , there is  $\mathbf{g}'_i \in \partial(-u_i)(\mathbf{x}^*_i), \forall i \in \mathcal{I}$  s.t.  $\sum_i \langle \mathbf{c}(\mathbf{X}^*) + \mathbf{g}'_i, \mathbf{x}^*_i - \mathbf{w}_i \rangle \leq 0$ . Secondly, since  $\hat{\mathbf{x}}$  is a SVWE in  $\tilde{\mathcal{G}}(A)$ , there is  $\mathbf{h}'_n \in \partial(-u_n)(\hat{\mathbf{x}}_n), \forall n \in \mathcal{N}$  s.t.  $\sum_n I_n \langle \mathbf{c}(\hat{\mathbf{X}}) + \mathbf{h}'_n, \hat{\mathbf{x}}_n - \mathbf{y}_n \rangle \leq 0$  for all  $\mathbf{y} \in \tilde{\mathcal{X}}^{\mathcal{N}}(A)$ . Thirdly,  $\forall n, \forall i \in \mathcal{I}_n$ , by the definition of  $\lambda_i$ , there is  $\mathbf{r}'_i \in \partial(-u_i)(\hat{\mathbf{x}}_n)$  such that  $\|\mathbf{r}'_i - \mathbf{h}'_n\| \leq \lambda_i$ .

The above results and  $\hat{\mathbf{x}}_n \leq m, \forall n$  imply:

$$\begin{aligned} & \langle \mathbf{c}(\mathbf{X}^*) - \mathbf{c}(\hat{\mathbf{X}}), \mathbf{X}^* - \hat{\mathbf{X}} \rangle + \sum_i \langle \mathbf{g}'_i - \mathbf{r}'_i, \mathbf{x}^*_i - \hat{\mathbf{x}}_n \rangle \\ &= \langle \mathbf{c}(\mathbf{X}^*) - \mathbf{c}(\hat{\mathbf{X}}), \mathbf{X}^* - \hat{\mathbf{X}} \rangle + \sum_{n,i \in \mathcal{I}_n} \langle \mathbf{g}'_i - \mathbf{r}'_i, \mathbf{x}^*_i - \hat{\mathbf{x}}_n \rangle \\ &= \sum_{n,i \in \mathcal{I}_n} [\langle \mathbf{c}(\mathbf{X}^*) + \mathbf{g}'_i, \mathbf{x}^*_i - \mathbf{w}_i \rangle + \langle \mathbf{c}(\mathbf{X}^*) + \mathbf{g}'_i, \mathbf{w}_i - \hat{\mathbf{x}}_n \rangle] \\ & \quad + \sum_{n,i \in \mathcal{I}_n} [\langle \mathbf{r}'_i - \mathbf{h}'_n, \hat{\mathbf{x}}_n - \mathbf{x}^*_i \rangle + \langle \mathbf{c}(\hat{\mathbf{X}}) + \mathbf{h}'_n, \hat{\mathbf{x}}_n - \mathbf{x}^*_i \rangle] \\ & \leq 0 + \sum_{n,i \in \mathcal{I}_n} \|\mathbf{c}(\mathbf{X}^*) + \mathbf{g}'_i\| \|\mathbf{w}_i - \hat{\mathbf{x}}_n\| \\ & \quad + \sum_{n,i \in \mathcal{I}_n} \|\mathbf{r}'_i - \mathbf{h}'_n\| \|\hat{\mathbf{x}}_n - \mathbf{x}^*_i\| + J \\ & \leq B_{\mathbf{f}} 4M\frac{\bar{\delta}}{\rho} + 2M\bar{\lambda} + J \end{aligned} \quad (23)$$

where  $J \triangleq \sum_{n,i \in \mathcal{I}_n} \langle \mathbf{c}(\hat{\mathbf{X}}) + \mathbf{h}'_n, \hat{\mathbf{x}}_n - \mathbf{x}^*_i \rangle$ .

Next, for SVWE  $\mathbf{x}^* \in \tilde{\mathcal{X}}(A)$ , let  $\mathbf{y} \in \tilde{\mathcal{X}}^{\mathcal{N}}(A)$  be s.t.  $\forall n$ ,  $\|\mathbf{y}_n - \psi_n(\mathbf{x})\|_{\mathcal{N}} \leq 2mI_n\bar{\delta}/\rho$  (cf. Lemma 8). Then we get:

$$\begin{aligned} J &= \sum_{n \in \mathcal{N}} \langle \mathbf{c}(\hat{\mathbf{X}}) + \mathbf{h}'_n, \hat{\mathbf{x}}_n - \bar{\psi}_n(\mathbf{x}^*) \rangle \\ &= \sum_{n \in \mathcal{N}} \langle \mathbf{c}(\hat{\mathbf{X}}) + \mathbf{h}'_n, \hat{\mathbf{x}}_n - \mathbf{y}_n \rangle \\ & \quad + \sum_{n \in \mathcal{N}} \langle \mathbf{c}(\hat{\mathbf{X}}) + \mathbf{h}'_n, \mathbf{y}_n - \bar{\psi}_n(\mathbf{x}^*) \rangle \\ & \leq 0 + \sum_{n \in \mathcal{N}} B_{\mathbf{f}} \|\bar{\psi}_n(\mathbf{x}^*) - \mathbf{y}_n\| \leq B_{\mathbf{f}} 2M\frac{\bar{\delta}}{\rho}, \end{aligned} \quad (24)$$

We conclude by combining (23) and (24) and using the  $\alpha$ -strong monotonicity or  $\beta$ -aggregative strong monotonicity of  $H'$ .

## REFERENCES

- [1] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, *Algorithmic Game Theory*. Cambridge University Press Cambridge, 2007, vol. 1.
- [2] A. Ziegelmeyer, F. Koessler, K. B. My, and L. Denant-Boèmont, "Road traffic congestion and public information: an experimental investigation," *JTEP*, vol. 42, no. 1, pp. 43–82, 2008.
- [3] G. Scutari, D. P. Palomar, F. Facchinei, and J.-S. Pang, "Monotone games for cognitive radio systems," in *Distributed Decision Making and Control*. Springer, 2012, pp. 83–112.
- [4] E. Altman, T. Boulogne, R. El-Azouzi, T. Jiménez, and L. Wynter, "A survey on networking games in telecommunications," *Computers & Operations Research*, vol. 33, no. 2, pp. 286–311, 2006.
- [5] A.-H. Mohsenian-Rad, V. W. Wong, J. Jatskevich, R. Schober, and A. Leon-Garcia, "Autonomous demand-side management based on game-theoretic energy consumption scheduling for the future smart grid," *IEEE Trans. Smart Grid*, vol. 1, pp. 320–331, 2010.
- [6] P. Jacquot, O. Beaudé, S. Gaubert, and N. Oudjane, "Analysis and implementation of an hourly billing mechanism for demand response management," *IEEE Trans. Smart Grid*, pp. 1–14, 2018.
- [7] J. F. Nash, "Equilibrium points in  $n$ -person games," *Proc. of the Nat. Acad. of Sci. of the U.S.A.*, vol. 36, no. 1, pp. 48–49, 1950.
- [8] H. Ackermann, H. Röglin, and B. Vöcking, "On the impact of combinatorial structure on congestion games," *JACM*, vol. 55, no. 6, p. 25, 2008.
- [9] A. Fabrikant, C. Papadimitriou, and K. Talwar, "The complexity of pure Nash equilibria," in *ACM S. Theory Comput.* ACM, 2004, pp. 604–612.
- [10] P. T. Harker, "Generalized Nash games and quasi-variational inequalities," *European journal of Operational research*, vol. 54, no. 1, pp. 81–94, 1991.
- [11] M. Carrión and J. M. Arroyo, "A computationally efficient mixed-integer linear formulation for the thermal unit commitment problem," *IEEE Trans. Power Syst.*, vol. 21, no. 3, pp. 1371–1378, 2006.
- [12] B. Colson, P. Marcotte, and G. Savard, "An overview of bilevel optimization," *Annals of Operations Research*, vol. 153, no. 1, pp. 235–256, 2007.
- [13] M. Labbé, P. Marcotte, and G. Savard, "A bilevel model of taxation and its application to optimal highway pricing," *Management Science*, vol. 44, no. 12-part-1, pp. 1608–1622, 1998.
- [14] L. Brotcorne, M. Labbé, P. Marcotte, and G. Savard, "A bilevel model and solution algorithm for a freight tariff-setting problem," *Transportation Science*, vol. 34, no. 3, pp. 289–302, 2000.
- [15] —, "A bilevel model for toll optimization on a multicommodity transportation network," *Transportation Science*, vol. 35, no. 4, pp. 345–358, 2001.
- [16] J.-P. Côté, P. Marcotte, and G. Savard, "A bilevel modelling approach to pricing and fare optimisation in the airline industry,"

- Journal of Revenue and Pricing Management*, vol. 2, no. 1, pp. 23–36, 2003.
- [17] J. Elias, F. Martignon, L. Chen, and E. Altman, “Joint operator pricing and network selection game in cognitive radio networks: Equilibrium, system dynamics and price of anarchy,” *IEEE Trans. Veh. Technol.*, vol. 62, no. 9, pp. 4576–4589, 2013.
- [18] P. T. Harker, “Generalized Nash games and quasi-variational inequalities,” *European Journal of Operational Research*, vol. 54, no. 1, pp. 81–94, 1991.
- [19] J. G. Wardrop, “Some theoretical aspects of road traffic research,” in *Proc. of the Inst. of Civil Eng., Part II, 1*, 1952, pp. 325–378.
- [20] D. Paccagnan, B. Gentile, F. Parise, M. Kamgarpour, and J. Lygeros, “Nash and Wardrop equilibria in aggregative games with coupling constraints,” *IEEE Trans. Autom. Control*, pp. 1–1, 2018.
- [21] P. Jacquot and C. Wan, “Routing game on parallel networks: the convergence of atomic to nonatomic,” in *Proc. of the 57th IEEE Conference on Decision and Control (CDC)*. IEEE, 2018.
- [22] —, “Nonsmooth aggregative games with coupling constraints and infinitely many classes of players,” *arXiv:1806.06230*, 2018.
- [23] G. Cohen, “Auxiliary problem principle extended to variational inequalities,” *J. Optim. Theory Appl.*, vol. 59, no. 2, pp. 325–333, 1988.
- [24] M. Fukushima, “A relaxed projection method for variational inequalities,” *Math. Program.*, vol. 35, no. 1, pp. 58–70, 1986.
- [25] D. Zhu and P. Marcotte, “Modified descent methods for solving the monotone variational inequality problem,” *Oper. Res. Lett.*, vol. 14, no. 2, pp. 111–120, 1993.
- [26] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*. Springer, 2007.
- [27] F. Facchinei and C. Kanzow, “Generalized Nash equilibrium problems,” *Annals of Operations Research*, vol. 175, no. 1, pp. 177–211, 2010.
- [28] P. Yi and L. Pavel, “Asynchronous distributed algorithm for seeking generalized Nash equilibria,” *arXiv preprint arXiv:1801.02967*, 2018.
- [29] —, “A distributed primal-dual algorithm for computation of generalized Nash equilibria with shared affine coupling constraints via operator splitting methods,” *arXiv preprint arXiv:1703.05388*, 2017.
- [30] F. Parise, B. Gentile, and J. Lygeros, “A distributed algorithm for average aggregative games with coupling constraints,” *arXiv preprint arXiv:1706.04634*, 2017.
- [31] T. Tatarenko and M. Kamgarpour, “Learning generalized Nash equilibria in a class of convex games,” *IEEE Trans. Autom. Control*, 2018.
- [32] P. Palensky and D. Dietrich, “Demand side management: Demand response, intelligent energy systems, and smart loads,” *IEEE Trans. Ind. Inform.*, vol. 7, no. 3, pp. 381–388, 2011.
- [33] R. W. Rosenthal, “A class of games possessing pure-strategy Nash equilibria,” *International Journal of Game Theory*, vol. 2, no. 1, pp. 65–67, 1973.
- [34] A. Orda, R. Rom, and N. Shimkin, “Competitive routing in multiuser communication networks,” *IEEE/ACM Trans. Networking*, vol. 1, no. 5, pp. 510–521, 1993.
- [35] A. A. Kulkarni and U. V. Shanbhag, “On the variational equilibrium as a refinement of the generalized Nash equilibrium,” *Automatica*, vol. 48, no. 1, pp. 45–55, 2012.
- [36] O. Richman and N. Shimkin, “Topological uniqueness of the Nash equilibrium for selfish routing with atomic users,” *Math. Oper. Reas.*, vol. 32, no. 1, pp. 215–232, 2007.
- [37] U. Bhaskar, L. Fleischer, D. Hoy, and C.-C. Huang, “Equilibria of atomic flow games are not unique,” in *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms*, 2009, pp. 748–757.
- [38] S. Lloyd, “Least squares quantization in PCM,” *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 129–137, 1982.
- [39] M. Garey, D. Johnson, and H. Witsenhausen, “The complexity of the generalized Lloyd-max problem (corresp.),” *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 255–256, 1982.
- [40] A. Ipakchi and F. Albuyeh, “Grid of the future,” *IEEE power and energy magazine*, vol. 7, no. 2, pp. 52–62, 2009.
- [41] W. Saad, Z. Han, H. V. Poor, and T. Basar, “Game-theoretic methods for the smart grid: An overview of microgrid systems, demand-side management, and smart grid communications,” *IEEE Signal Process Mag.*, vol. 29, no. 5, pp. 86–105, 2012.
- [42] Z. Wang and R. Paranjape, “Optimal residential demand response for multiple heterogeneous homes with real-time price prediction in a multiagent framework,” *IEEE Trans. Smart Grid*, vol. 8, no. 3, pp. 1173–1184, 2017.
- [43] P. Brucker, “An  $o(n)$  algorithm for quadratic knapsack problems,” *Oper. Res. Lett.*, vol. 3, no. 3, pp. 163–166, 1984.
- [44] R. G. Batson, “Combinatorial behavior of extreme points of perturbed polyhedra,” *Journal of Mathematical Analysis and Applications*, vol. 127, no. 1, pp. 130–139, 1987.
- [45] H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, 2nd ed. Springer International Publishing, 2011.
- [46] D. Chan and J. S. Pang, “The generalized quasi-variational inequality problem,” *Math. Oper. Reas.*, vol. 7, no. 2, pp. 211–222, 1982.



**Cheng Wan** received a B.Sc in Mathematics from Fudan University, a M.Sc. in Engineering from École polytechnique with a major in Applied Mathematics and Economics, and a M.Sc. in optimization, game theory and economic modeling as well as a PhD in Mathematics specialized in Game Theory from Université Paris VI. She has been a research fellow in Economics at the University of Oxford and is now a postdoctoral researcher in distributed control and machine learning at Université Paris-Sud, in collaboration with EDF R&D.



**Paulin Jacquot** owns an M.Sc. in engineering from École polytechnique with a major in applied mathematics, and an M.Sc in Operations Research from Université Paris-Saclay. He is currently pursuing a Ph.D. in applied mathematics at EDF R&D and at the Inria center of Saclay, located at the École polytechnique. His research is related to decentralized electric systems and the application of optimization and game theory.



**Olivier Beude** received the M.Sc. degree with specialization in applied mathematics and economics from École polytechnique, the M.Sc. degree in optimization, game theory, and economic modeling from Université Paris VI, in 2011, and the Ph.D. degree with the Laboratory of Signals and Systems, Centrale Suplec, and the R&D Center of Renault. He is currently research engineer at EDF R&D where his research interests include controlling the impact of EV charging on the grid, and the development of coordination mechanisms and incentives for local

electricity systems.



**Nadia Oudjane** is currently a research expert at EDF R&D in probability and statistic. She holds an engineering degree in signal processing from SupAero, a PhD in Applied Mathematics and an HDR in sciences. She has been working for more than 15 years on various topics involving numerical methods for energy management in a stochastic environment.