Nonatomic Aggregative Games with Infinitely Many Types ☆

Paulin Jacquot[1,*], Cheng Wan[2,*]

Abstract

We define and analyze the notion of variational Wardrop equilibrium for nonatomic aggregative games with an infinity of players types. These equilibria are characterized through an infinite-dimensional variational inequality. We show, under monotonicity conditions, a convergence theorem enables to approximate such an equilibrium with arbitrary precision. To this end, we introduce a sequence of nonatomic games with a finite number of players types, which approximates the initial game. We show the existence of a symmetric Wardrop equilibrium in each of these games. We prove that those symmetric equilibria converge to an equilibrium of the infinite game, and that they can be computed as solutions of finite-dimensional variational inequalities. The model is illustrated through an example from smart grids: the description of a large population of electricity consumers by a parametric distribution gives a nonatomic game with an infinity of different players types, with actions subject to coupling constraints.

Keywords: nonatomic aggregative game; coupling aggregative constraints; generalized variational inequality; monotone game; variational equilibrium

1. Introduction

We study the existence and uniqueness of variational Wardrop equilibrium (VWE) in nonatomic aggregative games with coupling aggregative constraints, where a continuum of players have heterogeneous compact convex pure-action sets and cost functions. We establish the convergence, to the VWE of such a game, of a sequence of symmetric VWE in auxiliary games with a finite number of types of players.

A motivating example. Consider the example of an energy operator studying the flexibility potential between peak and off-peak periods in a large population of energy consumers, for instance all households in France.

The operator considers that each household i has a certain quantity of energy E_i that can be balanced between consumption on peak period $x_{P,i}$ and consumption on off-peak periods $x_{O,i}$, such that $x_{O,i} + x_{P,i} = E_i$, depending on the cost (per unit of energy) $c_P(X_P)$ and $c_O(X_O)$ associated with the peak and off peak periods. The total on-peak consumption $X_P = \sum_i x_{P,i}$ and off-peak consumption $X_O = \sum_i x_{O,i}$ affect the prices on the energy market and, therefore, change the costs $c_P(X_P)$ and $c_O(X_O)$ set by the operator.

The operator wants to compute an equilibrium of this game (for instance to design tariffs). For practical and privacy reasons, it is impossible to have access to the flexibility potential E_i of the thirty millions of French households. However, the operator may have an easier access to a precise parametric, continuous distribution function of the flexibility potential among the French households.

Then, using the inverse transform sampling method, the game is replicated by modeling the population of households as a continuum $\Theta = [0, 1]$ and associating to each $\theta \in \Theta$ the flexible energy quantities $E_\theta = F_E^{-1}(\theta)$ from the inverse of the cumulative distribution function F_E. As the distribution is continuous, there is an infinity of different energy quantity E_θ i.e. an infinity of *players types* in the obtained game, where a *type* refers to the definition of a set of feasible actions and a payoff function. The operator has two questions: how to characterize an equilibrium of this nonatomic game with an infinity of players types and how to compute such an equilibrium. This paper provides answers to those two questions.

The game described above belongs to the class of aggregative games. In such a game, a player's payoff is determined by her own action and the aggregate of all the players' actions [11]. The setting of aggregative games is particularly relevant to the study of nonatomic games [37], i.e. games with a continuum of players. There, a player has an interaction with the other players only via an aggregate-level profile of their actions, while she has no interest or no way to know the behavior of any particular player or the identity of the player making a certain choice.

Nonatomic games are readily adapted to many situations in industrial engineering or public sectors where a huge number of users, such as traffic commuters and electricity consumers, are involved. These users have no direct interaction except through the aggregate congestion or consumption to which they are contributing collectively. These situations can often be modeled as a congestion game, a special class of aggregative games, both in nonatomic version and finite-player version. The latter, called atomic congestion game, was formally formulated by Rosenthal in 1973 [34], while related research work in transportation and traffic analysis, mostly in the nonatomic version, appeared much earlier [43, 8]. The theory of congestion games has also found numerous applications in telecommunications [31], distributed computing [2], energy management [3], and so on.

Nonatomic games are mathematical tools adapted to the modeling of interactions between a very large number of agents. Practical cases exist where a nonatomic model is intuitive and straightforward as when the modeler has an easier access to a description of the population through a parametric distribution of the types, as illustrated in the example.

As many distributions used in practice (e.g. normal distribution) are continuous, this implies that the nonatomic game obtained using these distributions will have an infinite number of players' types.

The concept of equilibrium in nonatomic games is captured by the so called Wardrop equilibrium (WE) [43]. A nonatomic player neglects the impact of her deviation on the aggregate profile of the whole population's actions, in contrast to a finite player.

For the computation of WE, existing results are limited to particular classes of nonatomic games, such as population games [29, 22, 36], where only a finite number of types of players are considered, each type sharing the same finite number of pure actions and the same payoff function.

The objective of this paper is to provide a model of nonatomic aggregative games with *infinitely many compact convex pure-action sets and infinitely many payoff functions*—in general a specific action set and a specific payoff function for each nonatomic player— then introduce a general form of coupling aggregative constraints into these games, define an appropriate notion of equilibrium, study the properties such as existence and uniqueness of these equilibria and, finally, their computation through an approximation.

Main results. After defining a pure-action profile in a nonatomic game where players have specific compact convex pure-action sets lying in $\mathbb{R}^T$, and specific cost functions, convex in their own action variable, Theorem 2.1 characterizes a WE as a solution to an infinite-dimensional variational inequality (IDVI).

Using the IDVI formulation, we extend this equilibrium notion to the case of a game with coupling aggregative constraints, by defining variational Wardrop equilibrium (VWE). Theorem 2.3 proves the existence of WE and VWE in monotone nonatomic games by showing the existence of solutions to the characteristic IDVI.

In Theorem 2.4, we establish the uniqueness of WE and VWE in case of strictly monotone or aggregatively strictly monotone games. The definition of monotone games is an extension of the stable games [21], also called dissipative games [40], in population games with a finite types of nonatomic players to the case with infinitely many types.

In the case where the nonatomic aggregative game has only a finite number of types of players, we define the notion of symmetric action profiles and symmetric VWE (SVWE), describing situations where all players of the same type play the same action. Proposition 2.2 shows that SVWEs are characterized as solutions of a *finite*-dimensional VI. Besides, Proposition 2.3 shows that, under monotonicity assumptions, there always exists an SVWE.

Theorem 3.1 is the main result of this paper. It shows that, for a sequence of finite-type approximating games, if the finite number of pure-action sets and cost functions converge to those of the players of a monotone nonatomic aggregative game, and if the aggregative constraint converges to the aggregative constraint of the infinite nonatomic game, then any sequence of SVWE associated to the sequence of approximating games converges in pure-action profile or in aggregate action profile to the VWE of the infinite-type game. We provide an upper bound on the distance between the approximating SVWE and the VWE, specified as a function of the parameters of the approximating finite-type games and the initial infinite nonatomic game.

This result allows the construction of a sequence of finite-type approximating games and associated SVWEs so as to approximate the infinite-dimensional VWE in the special class of strongly or aggregatively strongly monotone nonatomic aggregative games, with or without aggregative constraints. Since the resolution of finite-dimensional variational inequalities—characterizing SVWEs—is computationally tractable [14], it follows from our results that a VWE of a nonatomic game with infinitely many types can be approximated with arbitrary precision.

Section 3.3 shows how to construct an finite-type approximating games sequence for two general classes of nonatomic games. Appendix A gives the main ideas to extend all our results to the case where players have nonsmooth subdifferentiable cost functions: to get easily to the key ideas, we focus on the smooth case in the body of the paper. Appendix B explains how we can use the same arguments to show the convergence Nash equilibria of atomic finite-player games (instead of nonatomic finite-type games) to a VWE of a nonatomic game.

Related work. Extensive research has been conducted on WE in nonatomic congestion games via their formulation with variational inequalities [26]. In addition to their existence and uniqueness, the computational and dynamical aspects of equilibria as solutions to variational inequalities have also been studied [38, 46, 44, 10]. However, in most cases, the variational inequalities involved have finite dimensions, as opposed to the case of WE in this paper. Marcotte and Zhu [27] consider nonatomic players with continuous types (leading to a characterization of the WE as an IDVI) and studied the equilibrium in an aggregative game with nonatomic players differentiated through a linear parameter in their cost function.

Convergence of some dynamical systems describing the evolution of pure-action distribution in the population of a non atomic game has been established for some particular equilibria in some particular classes such as linear games [41], potential games [8, 35] and stable games [39, 21]. Algorithms corresponding to discretized versions of such dynamical systems for the computation of WE have been studied, in particular for congestion games [15, 44].

In engineering applications of nonatomic games such as the management of traffic flow or energy consumption, individual commuters or consumers often have specific choice sets due to individual constraints, and specific payoff functions due to personal preferences. Also, unlike for a transportation user who usually chooses a single path, an electricity consumer as modeled in the example above faces a resource allocation problem where she has to split the consumption of a certain quantity of energy over different time periods. Hence, her pure-action set is no longer a finite, discrete set as a commuter but a compact convex set in $\mathbb{R}^T$ where T is the total number of time periods. Few results exist for the computation of pure-action WE in the case where players have continuous action sets and in the case where there are infinitely many different types (i.e. action sets and payoff functions) of players. For example, [37] shows the existence of equilibrium in nonatomic games with finite action sets. Mas-Colell [28] and Carmona and Podzeck [9] consider compact strategy sets and show the existence of mixed strategy equilibria, and do not consider the case of aggregative games and pure-strategy equilibria. In their model, all players share the same actions set. Besides, most of the existing work assumes smooth cost functions of players which is somewhat restrictive in applications, as for instance electricity tariffs or tolls are usually not continuous.

Similarly, the subject of nonatomic games with (aggregative) coupling constraints has only been partially addressed. Coupling constraints at an aggregative level are to be considered in many of the above-mentioned applications, as also mentioned in [16]: for instance, when modeling the electricity consumption (see above example), some capacity constraints of the network or ramping constraints on the variation of total energy consumption between time periods are natural to consider from an engineering point of view. As seen in this paper, the presence of coupling constraints is not a simple artifact, as it adds non trivial difficulties in the analysis of WE and their computation. Indeed, an appropriate definition of equilibrium is already not obvious. An analog to the so-called generalized Nash equilibrium [18] for finite-player games does not exist for nonatomic games because a nonatomic player's behavior has no impact on the aggregative profile. Moreover, dynamical systems and algorithms used to compute WE in population games cannot be straightforwardly extended to this case. Indeed, in these dynamics and algorithms, players adapt their strategies unilaterally in their respective strategy spaces, which can well lead to a new strategy profile violating the coupling constraint.

Several works have quantified the relationship between Nash and Wardrop equilibria, a subject close to this paper, as shown in Appendix B. Haurie and Marcotte [20] show that in a sequence of atomic splittable games where atomic splittable players are replaced by smaller and smaller equal-size players with constant total weight, Nash equilibria converge to the WE of a nonatomic game. Their proof is based on the convergence of variational inequalities corresponding to the sequence of Nash equilibria,

a technique similar to the one used in this paper. Wan [42] generalizes this result to composite games where nonatomic players and atomic splittable players coexist, by allowing the atomic players to replace themselves by players with heterogeneous sizes.

Gentile et al. [32] consider a specific class of finite-player aggregative games with linear coupling constraints. They use the variational inequality formulations for the unique generalized Nash equilibrium and the unique generalized Wardrop-type equilibrium (which consists in letting each finite player act as if she was nonatomic) of the same finite-player game to show that, when the number of players grows, the former can be approximated by the latter. Their results are different from ours, as we consider nonatomic games with players of infinitely-many different types instead of finite-player games. Consequently, we consider VWE and symmetric VWE instead of generalized equilibria, which do not exist in nonatomic games. In contrast to generalized equilibria, a variational equilibrium is not characterized by a best reply condition for each of the finite or nonatomic players, as shown in Section 3. We also consider a general form of coupling constraints, and extend our results to nonsmooth cost functions, as shown in Appendix A (we focus on the differentiable case in the body of the paper).

Milchtaich [30] studies finite and nonatomic crowding games (similar to nonatomic aggregative games), where players have finitely many pure actions, and shows that, if each player in an n-person game is replaced by m identical replicas with constant total weight, pure Nash equilibria generically converge to the unique equilibrium of the limit nonatomic game as m goes to infinity. His proof is not based on a variational inequality formulation.

Structure. The remaining of the paper is organized as follows. Section 2 introduces the definitions of nonatomic aggregative games with and without aggregative constraints. After defining WE and VWE equilibrium in the case of an infinite number of players' types, we show, under monotonicity assumptions, the existence and uniqueness of equilibria there via generalized IDVIs. In the case of finite-type games, we define the notions of symmetric profiles and SVWE and show their characterization through finite-dimensional VIs and their existence. In Section 3, we give the definition of a sequence of finite-type approximating games for a nonatomic aggregative game with or without coupling constraints, and present the main theorem of the paper on the convergence of the sequence of S(V)WE, associated to the sequence of finite-type approximating games, to the (V)WE of the nonatomic game. In Section 3.3, the construction of sequences of finite-type approximating games is shown for two important classes of nonatomic aggregative games. In Section 4, we step back to the flexible energy example given above, and derive our results to the computation of an SVWE in this framework.

Last, in Appendix A, we show how our results extend to the case of nonsmooth cost functions and in Appendix B, we show how the results can be adapted to prove the convergence of Nash equilibria to a VWE of a nonatomic game.

Notation. Vectors are denoted by a bold font (e.g. $\boldsymbol{x}$) as opposed to scalars (e.g. x).

The transpose of vector $\boldsymbol{x}$ is denoted by $\boldsymbol{x}^\tau$.

The closed ball in a metric space, centered at $\boldsymbol{x}$ and of radius η, is denoted by $B_\eta(\boldsymbol{x})$.

For a nonempty convex set C in a Hilbert space $\mathcal{H}$ (over $\mathbb{R}$),

- $T_C(\boldsymbol{x}) = \{\boldsymbol{y} \in \mathcal{H} : \boldsymbol{y} = 0 \text{ or } \exists (\boldsymbol{x}_k)_k \text{ in } C \text{ s.t. } \boldsymbol{x}_k \not\equiv \boldsymbol{x}, \boldsymbol{x}_k \to \boldsymbol{x}, \frac{\boldsymbol{x}_k - \boldsymbol{x}}{\|\boldsymbol{x}_k - \boldsymbol{x}\|} \to \frac{\boldsymbol{y}}{\|\boldsymbol{y}\|}\}$ is the *tangent cone* of C at $\boldsymbol{x} \in C$;

- $\operatorname{span} C = \{\sum_{i=1}^k \alpha_i \boldsymbol{x}_i : k \in \mathbb{N}, \alpha_i \in \mathbb{R}, \boldsymbol{x}_i \in C\}$ is the *linear span* of C;

- $\operatorname{aff} C = \{\sum_{i=1}^k \alpha_i \boldsymbol{x}_i : k \in \mathbb{N}, \alpha_i \in \mathbb{R}, \sum_i \alpha_i = 1, \boldsymbol{x}_i \in C\}$ is the affine hull of C;

- $\operatorname{ri} C = \{\boldsymbol{x} \in C : \exists \eta > 0 \text{ s.t. } B_\eta(\boldsymbol{x}) \cap \operatorname{aff} C \subset C\}$ is the *relative interior* of C;

- $\operatorname{rbd} C$ is the *relative boundary* of C in $\mathcal{H}$, i.e. the boundary of C in $\operatorname{span} C$.

The *inner product* of two points $\boldsymbol{x}$ and $\boldsymbol{y}$ in any Euclidean space $\mathbb{R}^T$ is denoted by $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i=1}^T x_i y_i$. The l^2-*norm* of $\boldsymbol{x}$ is denoted by $\|\boldsymbol{x}\| \triangleq \langle \boldsymbol{x}, \boldsymbol{x} \rangle^{1/2}$.

We denote by $L^2([0,1], \mathbb{R}^T)$ the Hilbert space of measurable functions from $[0,1]$ (equipped with the Lebesgue measure μ) to $\mathbb{R}^T$ that are square integrable with respect to the Lebesgue measure. The *inner product* of two vector functions F and G is denoted by $\langle F, G \rangle_2 = \int_0^1 \langle F(\theta), G(\theta) \rangle \, d\theta$. The Hilbert space $L^2([0,1], \mathbb{R}^T)$ is endowed with L^2-*norm*: $\|F\|_2 = \langle F, F \rangle_2^{1/2}$.

The distance between a point $\boldsymbol{x}$ and a set A is denoted by $d_m(\boldsymbol{x}, A) \triangleq \inf_{\boldsymbol{y} \in A} \|\boldsymbol{x} - \boldsymbol{y}\|_m$, where m is omitted or is equal to 2, depending on whether we consider an Euclidean space or $L^2([0,1], \mathbb{R}^T)$.

Similarly, the *Hausdorff distance* between two sets A and B is denoted by $d_{H,m}(A, B)$, which is defined as $\max\{\sup_{\boldsymbol{x} \in A} d_m(\boldsymbol{x}, B), \sup_{\boldsymbol{y} \in B} d_m(\boldsymbol{y}, A)\}$.

For a function $(\boldsymbol{x}, \boldsymbol{X}) \mapsto f(\boldsymbol{x}, \boldsymbol{X})$ of two explicit variables, convex in $\boldsymbol{x}$, we denote by $\nabla_1 f(\boldsymbol{x}, \boldsymbol{X})$ the differential of function $f(\cdot, \boldsymbol{X})$ for any fixed $\boldsymbol{X}$, except in Appendix A in which ∂ (resp. ∂_1) is used to denote the (resp. partial) subdifferential.

2. Monotonicity, Coupling Constraints and Symmetric Equilibrium

2.1. Nonatomic aggregative games

In nonatomic aggregative games considered here, players have compact pure-action sets, and heterogeneous pure-action sets as well as heterogeneous cost function. This model is in line with Schmeidler's seminal paper [37]. It differs from most of the population games studied in game theory [22, 36], in which nonatomic players are grouped into several populations, with players in the same population having the same finite pure-action set and the same cost function.

Definition 2.1 (Nonatomic aggregative game). A *nonatomic aggregative game G* is defined by:

i) a continuum of players represented by the points on the real interval $\Theta = [0, 1]$ endowed with Lebesgue measure;

ii) a set of feasible pure actions $\mathcal{X}_\theta \subset \mathbb{R}^T$ for each player $\theta \in \Theta$, with $T \in \mathbb{N}^*$ a constant;

iii) a cost function $\mathcal{X}_\theta \times \mathbb{R}^T \to \mathbb{R} : (\boldsymbol{x}_\theta, \boldsymbol{X}) \mapsto f_\theta(\boldsymbol{x}_\theta, \boldsymbol{X})$ for each player θ, where $\boldsymbol{X} = (X_t)_{t=1}^T$ and $X_t \triangleq \int_0^1 \boldsymbol{x}_{\theta',t} \, d\theta'$ refers to an aggregate-action profile, given action profile $(\boldsymbol{x}_{\theta'})_{\theta' \in \Theta}$ for the population Θ.

The set of feasible pure-action profiles is defined by:

$$\boldsymbol{\mathcal{X}} \triangleq \left\{ \boldsymbol{x} \in L^2([0, 1], \mathbb{R}^T) \ : \ \forall \theta \in \Theta, \boldsymbol{x}_\theta \in \mathcal{X}_\theta \right\}.$$

Denote the game by $G = (\Theta, \boldsymbol{\mathcal{X}}, (f_\theta)_{\theta \in \Theta})$.

Remark 2.1. The definition of a nonatomic aggregative game asks the pure-action profile $\boldsymbol{x}$ to be a measurable and integrable function on Θ instead of simply being a collection of $\boldsymbol{x}_\theta \in \mathcal{X}_\theta$ for $\theta \in \Theta$. In other words, a coupling constraint is inherent in the definition of nonatomic aggregative games and the notion of WE. This is in contrast with finite-player games.

The set of feasible aggregate actions is $\overline{\mathcal{X}} \triangleq \{\boldsymbol{X} \in \mathbb{R}^T : \exists \boldsymbol{x} \in \boldsymbol{\mathcal{X}} \text{ s.t. } \int_0^1 \boldsymbol{x}_\theta \, d\theta = \boldsymbol{X}\}$.

Further assumptions are needed for $\boldsymbol{\mathcal{X}}$ to be nonempty and for the existence of an equilibrium.

Assumption 1 (Nonatomic pure-action sets). *The correspondence* $\mathcal{X} : \Theta \rightrightarrows \mathbb{R}^T, \theta \mapsto \mathcal{X}_\theta$ *has nonempty, convex, compact values. Moreover, for all* $\theta \in \Theta$, $\mathcal{X}_\theta \subset B_R(\boldsymbol{0})$, *with* $R > 0$ *a constant.*

Under Assumption 1, a sufficient condition for $\boldsymbol{x}$ to be in $L^2([0, 1], \mathbb{R}^T)$ is that $\boldsymbol{x}$ is measurable.

Notations. We denote by $\mathcal{M} = [0, R + 1]^T$ the hypercube in $\mathbb{R}^T$ of edge $R + 1$.

Assumption 2 (Measurability). *The correspondence* $\mathcal{X} : \Theta \rightrightarrows \mathbb{R}^T, \theta \mapsto \mathcal{X}_\theta$ *has a measurable graph* $Gr_{\mathcal{X}} = \{(\theta, \boldsymbol{x}_\theta) \in \mathbb{R}^{T+1} : \theta \in \Theta, \boldsymbol{x}_\theta \in \mathcal{X}_\theta\}$, *i.e.* $Gr_{\mathcal{X}}$ *is a Borel subset of* $\mathbb{R}^{T+1}$.
The function $Gr_{\mathcal{X}} \to \mathbb{R}^T : (\theta, \boldsymbol{x}_\theta) \mapsto f_\theta(\boldsymbol{x}_\theta, \boldsymbol{Y})$ *is measurable for each* $\boldsymbol{Y} \in \mathbb{R}^T$.

Assumption 3 (Nonatomic convex cost functions). *For all* θ, f_θ *is defined on* $(\mathcal{M}')^2$, *where* $\mathcal{M}'$ *is a neighborhood of* $\mathcal{M}$, *and:*

i) *for each* $\theta \in \Theta$, *function* f_θ *is continuous. In particular,* f_θ *is bounded on* $\mathcal{M}^2$;

ii) *for each* $\theta \in \Theta$ *and each aggregate profile* $\boldsymbol{Y} \in \mathcal{M}$, $\boldsymbol{x} \mapsto f_\theta(\boldsymbol{x}, \boldsymbol{Y})$ *is differentiable and convex on* $\mathcal{M}'$;

iii) *there is* $L_{\mathbf{f}} > 0$ *such that* $\|\nabla_1 f_\theta(\boldsymbol{x}_\theta, \boldsymbol{Y})\| \leq L_{\mathbf{f}}$ *for each* $\boldsymbol{x}_\theta \in \mathcal{M}$, *each* $\boldsymbol{Y} \in \mathcal{M}$, *and each* $\theta \in \Theta$.

Remark 2.2. Assumption 3.iii) implies that $f_\theta(\cdot, \cdot)$'s are Lipschitz in the first variable with a uniform Lipschitz constant $L_{\mathbf{f}}$ on $\mathcal{M}^2$ for all θ.

We also need the continuity of the derivative of cost functions in the second (aggregate) variable:

Assumption 4. *For each $\theta \in \Theta$ and each $\boldsymbol{x}_\theta \in \mathcal{M}$, the function $\boldsymbol{Y} \mapsto \nabla_1 f_\theta(\boldsymbol{x}_\theta, \boldsymbol{Y})$ is continuous on $\mathcal{M}$.*

Wardrop equilibrium extends the notion of Nash equilibrium in the framework of nonatomic games, where a single player of measure zero has a negligible impact on the others.

Definition 2.2 (Wardrop Equilibrium (WE), [43])**.** A pure-action profile $\boldsymbol{x}^* \in \boldsymbol{\mathcal{X}}$ is a pure *Wardrop equilibrium* of nonatomic aggregative game G if we have, with $\boldsymbol{X}^* = \int_{\theta \in \Theta} \boldsymbol{x}_\theta^* \, \mathrm{d}\theta$:

$$f_\theta(\boldsymbol{x}_\theta^*, \boldsymbol{X}^*) \leq f_\theta(\boldsymbol{x}_\theta, \boldsymbol{X}^*), \quad \forall \boldsymbol{x}_\theta \in \mathcal{X}_\theta, \ \forall a.e.\, \theta \in \Theta \ .$$

All the actions and equilibria in this paper are pure, hence from now on, we no longer emphasize it.

Before characterizing WE by infinite-dimensional VI (IDVI), let us introduce some notions and a technical assumption ensuring that the IDVI is well-defined.

Lemma 2.1. *For all $\boldsymbol{x} \in L^2([0,1], \mathcal{M})$, the function $\mathbf{g}_{\boldsymbol{x}}$ defined from Θ to $\mathbb{R}^T$ by*

$$\mathbf{g}_{\boldsymbol{x}}(\theta) \triangleq \nabla_1 f_\theta(\boldsymbol{x}_\theta, \textstyle\int \boldsymbol{x}), \quad \forall \theta \in \Theta, \ \forall \boldsymbol{x} \in L^2([0,1], \mathcal{M}) \tag{1}$$

is measurable on Θ. Consequently, $\mathbf{g}_.$ is a mapping from $L^2([0,1], \mathcal{M})$ to $L^2([0,1], \mathbb{R}^T)$.

Proof of Lemma 2.1. For $n \in \mathbb{N}$ large enough, $\mathbf{g}_{\boldsymbol{x},n} \triangleq \theta \mapsto n\big(f_\theta(\boldsymbol{x}_\theta + \frac{1}{n}, \int \boldsymbol{x}) - f_\theta(\boldsymbol{x}_\theta, \int \boldsymbol{x})\big)$ is well defined. It is measurable according to Assumption 2. Thus $\mathbf{g}_{\boldsymbol{x}} = \lim_n \mathbf{g}_{\boldsymbol{x},n}$ is also measurable as a limit of measurable functions.

$\square$

Theorem 2.1 (IDVI formulation of WE)**.** *Under Assumptions 1 to 3, $\boldsymbol{x}^* \in \boldsymbol{\mathcal{X}}$ is a WE of nonatomic aggregative game G if and only if either of the following two equivalent conditions is true:*

$$\forall a.e.\, \theta \in \Theta, \quad \langle \nabla_1 f_\theta(\boldsymbol{x}_\theta^*, \boldsymbol{X}^*), \boldsymbol{x}_\theta - \boldsymbol{x}_\theta^* \rangle \geq 0, \quad \forall \boldsymbol{x}_\theta \in \mathcal{X}_\theta \ , \tag{2a}$$

$$\int_\Theta \langle \mathbf{g}_{\boldsymbol{x}^*}(\theta), \boldsymbol{x}_\theta - \boldsymbol{x}_\theta^* \rangle \, \mathrm{d}\theta \geq 0, \quad \forall \boldsymbol{x} \in \boldsymbol{\mathcal{X}} \ . \tag{2b}$$

Proof of Theorem 2.1. Given $\boldsymbol{X}^*$, (2a) is a necessary and sufficient condition for $\boldsymbol{x}_\theta^*$ to minimize the convex function $f_\theta(., \boldsymbol{X}^*)$ on $\mathcal{X}_\theta$. Condition (2a) implies condition (2b) because of Lemma 2.1.

For the converse, suppose that $\boldsymbol{x}^* \in \boldsymbol{\mathcal{X}}$ satisfies condition (2b) but not (2a). Then there must be a subset Θ' of Θ with strictly positive measure such that for each $\theta \in \Theta'$, $\boldsymbol{x}_\theta^* \notin \mathcal{Y}_\theta \triangleq \arg\min_{\mathcal{X}_\theta} f_\theta(\cdot, \boldsymbol{X}^*)$. In particular, for any $\boldsymbol{y}_\theta \in \mathcal{Y}_\theta$, $\langle \mathbf{g}_{\boldsymbol{x}^*}(\theta), \boldsymbol{y}_\theta - \boldsymbol{x}_\theta^* \rangle < f_\theta(\boldsymbol{y}_\theta, \boldsymbol{X}^*) - f_\theta(\boldsymbol{x}_\theta^*, \boldsymbol{X}^*) < 0$. A consequence of Assumptions 2 and 3 is that the function $\Theta \times \mathcal{M} \to \mathbb{R} : (\theta, \boldsymbol{z}) \mapsto f_\theta(\boldsymbol{z}, \boldsymbol{X}^*)$ is a Carathéodory function, that is, (i) $f_.(\boldsymbol{z}, \boldsymbol{X}^*)$ is measurable on Θ for each $\boldsymbol{z} \in \mathcal{M}$, and (ii) $f_\theta(\cdot, \boldsymbol{X}^*)$ is continuous on $\mathcal{M}$ for each $\theta \in \Theta$. Thus, according to the measurable maximum theorem [1, Thm. 18.19] applied to $f_.(\cdot, \boldsymbol{X}^*)$, there exists a selection $\boldsymbol{y}_\theta \in \arg\min_{\mathcal{X}_\theta} f_\theta(\cdot, \boldsymbol{X}^*)$ for $\theta \in \Theta'$ such that $\Theta' \to \mathbb{R}^T : \theta \mapsto \boldsymbol{y}_\theta$ is a measurable function. By defining $\boldsymbol{y}_\theta = \boldsymbol{x}_\theta^*$ for $\theta \notin \Theta'$, one has $\Theta \to \mathbb{R}^T : \theta \mapsto \boldsymbol{y}_\theta$ is measurable and hence belongs to $\boldsymbol{\mathcal{X}}$. However, $\int_\Theta \langle \mathbf{g}_{\boldsymbol{x}^*}(\theta), \boldsymbol{y}_\theta - \boldsymbol{x}_\theta^* \rangle \, \mathrm{d}\theta = \int_{\Theta'} \langle \mathbf{g}_{\boldsymbol{x}^*}(\theta), \boldsymbol{y}_\theta - \boldsymbol{x}_\theta^* \rangle \, \mathrm{d}\theta < 0$, contradicting (2b). $\square$

Remark 2.3. Condition (2a) is equivalent to $\langle \mathbf{g}_{\boldsymbol{x}^*}(\theta), \boldsymbol{y}_\theta \rangle \geq 0$ for all $\boldsymbol{y}_\theta \in T_{\mathcal{X}_\theta}(\boldsymbol{x}_\theta^*)$ for each θ. It means that no unilateral deviation is profitable. However, since each nonatomic player has measure zero, when considering a deviation in the action profile, one must let players in a set of strictly positive measure deviate: (2b) means that the collective deviation of players of any set of strictly positive measure increases their cost.

The existence of WE is obtained by an equilibrium existence theorem for nonatomic games.

Theorem 2.2 (Existence of a WE, [33])**.** *Under Assumption 1, Assumption 2 and Assumption 3.i), if for all θ and all $\boldsymbol{Y} \in \mathcal{M}$, $f_\theta(\cdot, \boldsymbol{Y})$ is continuous on $\mathcal{M}$, then the nonatomic aggregative game G admits a WE.*

Proof. The conditions required in Remark 8 in Rath's 1992 paper [33] on the existence of WE in aggregate games are satisfied. $\square$

Remark 2.4. No convexity of $f_\theta(\cdot, \boldsymbol{Y})$'s are needed for the existence.

2.2. Monotone Nonatomic Aggregative Games

For the uniqueness of WE and the existence of equilibrium notion to be introduced in the next subsection for the case with coupling constraints, let us consider the following notions of *monotonicity* in nonatomic aggregative games, also sometimes referred to as *stability*.

Definition 2.3. With notation $\mathbf{g}_{\boldsymbol{x}}(\theta) = \nabla_1 f_\theta(\boldsymbol{x}_\theta, \int \boldsymbol{x})$, for any $\theta \in \Theta$ and any $\boldsymbol{x}, \boldsymbol{y} \in L^2([0,1], \mathcal{M})$, we say that the nonatomic aggregative game G is

  i) *monotone* if

$$\int_\Theta \langle \mathbf{g}_{\boldsymbol{x}}(\theta) - \mathbf{g}_{\boldsymbol{y}}(\theta), \boldsymbol{x}_\theta - \boldsymbol{y}_\theta \rangle \, \mathrm{d}\theta \geq 0, \quad \forall \boldsymbol{x}, \boldsymbol{y} \in L^2([0,1], \mathcal{M}) \ . \tag{3}$$

  ii) *strictly monotone* if the equality in (3) holds if and only if $\boldsymbol{x} = \boldsymbol{y}$ almost everywhere.

  iii) *aggregatively strictly monotone* if the equality in (3) holds if and only if $\int \boldsymbol{x} = \int \boldsymbol{y}$.

  iv) *strongly monotone* with modulus α if

$$\int_\Theta \langle \mathbf{g}_{\boldsymbol{x}}(\theta) - \mathbf{g}_{\boldsymbol{y}}(\theta), \boldsymbol{x}_\theta - \boldsymbol{y}_\theta \rangle \, \mathrm{d}\theta \geq \alpha \|\boldsymbol{x} - \boldsymbol{y}\|_2^2, \ \forall \boldsymbol{x}, \boldsymbol{y} \in L^2([0,1], \mathcal{M}) \ . \tag{4}$$

  v) *aggregatively strongly monotone* with modulus β if

$$\int_\Theta \langle \mathbf{g}_{\boldsymbol{x}}(\theta) - \mathbf{g}_{\boldsymbol{y}}(\theta), \boldsymbol{x}_\theta - \boldsymbol{y}_\theta \rangle \, \mathrm{d}\theta \geq \beta \| \textstyle\int \boldsymbol{x} - \int \boldsymbol{y} \|^2, \ \forall \boldsymbol{x}, \boldsymbol{y} \in L^2([0,1], \mathcal{M}) \ . \tag{5}$$

Remark 2.5. A recent paper of Hadikhanloo [17] generalizes the notion of stable games in population games [21] to monotone games in anonymous games, an extension of population games with players having heterogeneous compact action sets but the same payoff function. He defines the notion of monotonicity directly on the distribution of actions among the players instead of action profile as we do. The two approaches are compatible.

Example of public products games.. An interesting example of aggregative games are given by cost functions of the form:

$$f_\theta(\boldsymbol{x}_\theta, \boldsymbol{X}) = \langle \boldsymbol{x}_\theta, \boldsymbol{c}(\boldsymbol{X}) \rangle - u_\theta(\boldsymbol{x}_\theta) \ , \tag{6}$$

where $\boldsymbol{c}(\boldsymbol{X})$ specifies the per-unit cost (or negative of per-unit utility) of each of the T "public products", which is a function of the aggregative contribution $\boldsymbol{X}$ to each of the "public products". Player θ's cost (resp. negative of utility) associated to these products is scaled by her own contribution $\boldsymbol{x}_\theta$. The function $u_\theta(\boldsymbol{x}_\theta)$ measures the private utility of player θ (resp. negative of private cost) for the contribution $\boldsymbol{x}_\theta$.

For instance, in a public goods game, $-c_t(X_t)$ is the common per-unit payoff for using public good t, determined by the total contribution X_t, while $-u_\theta(\boldsymbol{x}_\theta)$ is player θ's private cost of supplying $\boldsymbol{x}_\theta$ to the public goods; in a Cournot competition, $-c_t(X_t)$ is the common market price for product t, determined by its total supply X_t, while $-u_\theta(\boldsymbol{x}_\theta)$ is player θ's private cost of producing $x_{\theta,t}$ unit of product t for each product t; in a congestion game, $c_t(X_t)$ is the common per-unit cost for using arc t in a network, determined by the aggregate load X_t on arc t, while $u_\theta(\boldsymbol{x}_\theta)$ is player θ's private utility of her routing or energy consuming choice $\boldsymbol{x}_\theta$.

Proposition 2.1. *Under Assumptions 1 to 3, in a* public products *game G (i.e. with cost functions of form (6)), assume that $\boldsymbol{c}$ is monotone on $\mathcal{M}$ and, for each θ, u_θ is a concave function on $\mathcal{M}$. Then:*

  i) *G is a monotone game.*

  ii) *If u_θ is strictly concave on $\mathcal{M}$ for all $\theta \in \Theta$, then G is a strictly monotone game.*

  iii) *If $\boldsymbol{c}$ is strictly monotone on $\mathcal{M}$, then G is an aggregatively strictly monotone game.*

  iv) *If u_θ is strongly concave on $\mathcal{M}$ with modulus α_θ for each $\theta \in \Theta$ and $\inf_{\theta \in \Theta} \alpha_\theta = \alpha > 0$, then G is a strongly monotone game with modulus α.*

  v) *If $\boldsymbol{c}$ is strongly monotone on $\mathcal{M}$ with β, then G is an aggregatively strongly monotone game with modulus β.*

Proof. i) Let $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{\mathcal{X}}$ and $\boldsymbol{X} = \int \boldsymbol{x}$, $\boldsymbol{Y} = \int \boldsymbol{y}$. For each θ, $\nabla_1 f_\theta(\boldsymbol{x}_\theta, \boldsymbol{X}) = \boldsymbol{c}(\boldsymbol{X}) - \nabla u_\theta(\boldsymbol{x}_\theta)$. Then, given $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{\mathcal{X}}$, we have $\langle \nabla u_\theta(\boldsymbol{x}_\theta) - \nabla u_\theta(\boldsymbol{y}_\theta), \boldsymbol{x}_\theta - \boldsymbol{y}_\theta \rangle \leq 0$ because u_θ is concave. Then $\int_0^1 \langle \nabla_1 f_\theta(\boldsymbol{x}_\theta, \boldsymbol{X}) - \nabla_1 f_\theta(\boldsymbol{y}_\theta, \boldsymbol{Y}), \boldsymbol{x}_\theta - \boldsymbol{y}_\theta \rangle \mathrm{d}\theta = \langle \boldsymbol{c}(\boldsymbol{X}) - \boldsymbol{c}(\boldsymbol{Y}), \boldsymbol{X} - \boldsymbol{Y} \rangle - \int_0^1 \langle \nabla u_\theta(\boldsymbol{x}_\theta) - \nabla u_\theta(\boldsymbol{y}_\theta), \boldsymbol{x}_\theta - \boldsymbol{y}_\theta \rangle \mathrm{d}\theta \geq 0$ because $\boldsymbol{c}$ is monotone. Hence G is a monotone game.

The proof for ii)-v) is omitted. $\qquad\square$

In particular, if $\boldsymbol{c}(\boldsymbol{X}) = (c_t(X_t))_{t \in T}$, then $\boldsymbol{c}$ is monotone if c_t's are all non-decreasing, and $\boldsymbol{c}$ is strongly monotone if c_t's are all strictly increasing.

2.3. Aggregate constraints and VWE

Let us consider the aggregative constraint in nonatomic aggregative game G: $\boldsymbol{X} \in A$, where A is a convex compact subset of $\mathbb{R}^T$ such that $A \cap \overline{\mathcal{X}} \neq \emptyset$. Let $\boldsymbol{\mathcal{X}}(A)$ be a subset of $\boldsymbol{\mathcal{X}}$ defined by $\boldsymbol{\mathcal{X}}(A) \triangleq \{\boldsymbol{x} \in \boldsymbol{\mathcal{X}} : \boldsymbol{X} = \int \boldsymbol{x} \in A\}$. Let us denote the nonatomic aggregative game with aggregative constraint $\boldsymbol{X} \in A$ by $G(A)$.

A notion of generalized WE similar to the one of generalized Nash equilibrium in finitely-many-player games—where each player does the best she can while not violating the coupling constraints given the choices of the others [18]—is not well-defined in a nonatomic game. Indeed, since the impact of a nonatomic player's choice on the aggregative profile is negligible, the feasible action set of a nonatomic player θ facing the choices of the others $\boldsymbol{x}_{-\theta}$ in a game with coupling constraint is not a well-established notion: either $\int \boldsymbol{x}_{-\theta} \in A$ then $\mathcal{X}_\theta = \mathcal{X}$, or $\int \boldsymbol{x}_{-\theta} \notin A$ then $\mathcal{X}_\theta = \emptyset$. Departing from an action profile in $\boldsymbol{\mathcal{X}}(A)$, simultaneous unilateral deviations by the players can lead to any profile in $\boldsymbol{\mathcal{X}}$. If only profiles in $\mathcal{X}(A)$ are allowed to be attained, then one lands on a notion similar to the so-called variational Nash equilibrium in finite-many-player games [18]. Indeed, the most natural notion of equilibrium with the presence of aggregative constraint is the notion of variational Wardrop equilibrium, where feasible deviations are defined on a collective basis.

Definition 2.4 (Variational Wardrop Equilibrium (VWE))**.** A solution to the following IDVI problem:

$$\text{Find } \boldsymbol{x}^* \in \boldsymbol{\mathcal{X}}(A) \text{ s.t. } \int_\Theta \langle \mathbf{g}_{\boldsymbol{x}^*}(\theta), \boldsymbol{x}_\theta - \boldsymbol{x}_\theta^* \rangle \mathrm{d}\theta \geq 0, \quad \forall \boldsymbol{x} \in \boldsymbol{\mathcal{X}}(A), \tag{7}$$

is called a *variational Wardrop equilibrium* of $G(A)$.

Remark 2.6 (VWE in the literature)**.** In the literature of congestion games, the equilibrium notion characterized by VI of form (7) but in finite dimension and with smooth cost functions has long been studied. For example, see [24, 25, 12, 45] and references therein.

The following facts are needed for later use.

Lemma 2.2. *Under Assumptions 1 and 2:*

i) $\boldsymbol{\mathcal{X}}$ *is a nonempty, convex, closed and bounded subset of* $L^2([0,1], \mathbb{R}^T)$*;*

ii) $\boldsymbol{\mathcal{X}}(A)$ *is a nonempty, convex and closed subset of* $\boldsymbol{\mathcal{X}}$*;*

iii) $\overline{\mathcal{X}}$ *and* $A \cap \overline{\mathcal{X}}$ *are nonempty, convex and compact subsets of* $\mathbb{R}^T$*.*

We omit the proof and only point out that $\boldsymbol{\mathcal{X}}$ and $\overline{\mathcal{X}}$ are nonempty because of Assumption 1 and the measurable selection theorem of Aumann [5], while aggregate-action set $\overline{\mathcal{X}}$ is compact by [4, Theorem 4].

Theorem 2.3 shows the existence of VWE via the VI approach.

Theorem 2.3 (Existence of VWE)**.** *Under Assumptions 1 to 4, if a nonatomic aggregative game with coupling constraint $G(A)$ is monotone on $\boldsymbol{\mathcal{X}}(A)$, then a VWE exists.*

Proof. We can apply [13, Corollary 2.1] which shows that Eq. (7) has a solution, as:

- $\boldsymbol{\mathcal{X}}(A)$ is bounded, closed and convex in $L^2([0,1], \mathbb{R}^T)$;

- $\mathbf{g}_. : L^2([0,1], \mathcal{M}) \rightrightarrows L^2([0,1], \mathbb{R}^T)$ is a monotone correspondence which is upper hemicontinuous from the line segments in $\boldsymbol{\mathcal{X}}(A)$ to the weak* topology of $L^2([0,1], \mathbb{R}^T)$. Notice that $\mathbf{g}_.$ has closed values. Let us do the proof in the general nonsmooth case.

Take $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\boldsymbol{\mathcal{X}}(A)$, consider sequence $(\boldsymbol{x}^{(k)})_k$ with $\boldsymbol{x}^{(k)} \in [\boldsymbol{x}, \boldsymbol{y}]$ with $\boldsymbol{x}^{(k)} \to \boldsymbol{x}$, and sequence $(\mathbf{g}^{(k)})_k$ such that $\mathbf{g}^{(k)} \in \mathbf{g}_{\boldsymbol{x}^{(k)}}$ and $\mathbf{g}^{(k)} \xrightarrow{*} \mathbf{g}^\infty$ with $\mathbf{g}^\infty \in L^2([0,1], \mathbb{R}^T)$. Let us show that $\mathbf{g}^\infty \in \mathbf{g}_{\boldsymbol{x}}$. We have $\boldsymbol{X} = \int \boldsymbol{x}$ converging to $\boldsymbol{X}^{(k)} = \int \boldsymbol{x}^{(k)}$ in l^2-norm.

By definition of $\mathbf{g}$ and convexity, for each $\boldsymbol{z} \in \mathcal{M}$, for each θ, $f_\theta(\boldsymbol{z}_\theta, \boldsymbol{X}_\theta^{(k)}) \geq f_\theta(\boldsymbol{x}_\theta^{(k)}, \boldsymbol{X}^{(k)}) + \langle \mathbf{g}_\theta^{(k)}, \boldsymbol{z}_\theta - \boldsymbol{x}_\theta^{(k)} \rangle$. Since f_θ is continuous in both variables, $f_\theta(\boldsymbol{z}_\theta, \boldsymbol{X}^{(k)}) \to f_\theta(\boldsymbol{z}_\theta, \boldsymbol{X})$ and $f_\theta(\boldsymbol{x}_\theta^{(k)}, \boldsymbol{X}^{(k)}) \to f_\theta(\boldsymbol{x}_\theta, \boldsymbol{X})$. Besides, $\langle \mathbf{g}_\theta^{(k)}, \boldsymbol{z}_\theta - \boldsymbol{x}_\theta^{(k)} \rangle = \langle \mathbf{g}_\theta^{(k)}, \boldsymbol{z}_\theta - \boldsymbol{x}_\theta \rangle + \langle \mathbf{g}_\theta^{(k)}, \boldsymbol{x}_\theta - \boldsymbol{x}_\theta^{(k)} \rangle$, and $\langle \mathbf{g}_\theta^{(k)}, \boldsymbol{z}_\theta - \boldsymbol{x}_\theta \rangle \to \langle \mathbf{g}_\theta^\infty, \boldsymbol{z}_\theta - \boldsymbol{x}_\theta \rangle$ because $\mathbf{g}^{(k)} \overset{*}{\rightharpoonup} \mathbf{g}^\infty$, while $\langle \mathbf{g}_\theta^{(k)}, \boldsymbol{x}_\theta - \boldsymbol{x}_\theta^{(k)} \rangle \to 0$ because $\mathbf{g}_\theta^{(k)}$'s are uniformly bounded by $L_{\mathbf{f}}$. Therefore, $f_\theta(\boldsymbol{z}_\theta, \boldsymbol{X}) \geq f_\theta(\boldsymbol{x}_\theta^{(k)}, \boldsymbol{X}) + \langle \mathbf{g}_\theta, \boldsymbol{z}_\theta - \boldsymbol{x}_\theta \rangle$ so that $\mathbf{g}_\theta^\infty \in \partial_1 f_\theta(\boldsymbol{x}_\theta, \boldsymbol{X})$. Since the limit of measurable functions is measurable, $\mathbf{g}$ is measurable. Hence $\mathbf{g}^\infty \in \mathbf{g}_{\boldsymbol{x}}$ (and $\mathbf{g}^\infty = \mathbf{g}_{\boldsymbol{x}}$ in the smooth case), which concludes.

$\square$

Theorem 2.4 (Uniqueness of VWE). *Under Assumptions 1 to 3:*

i) if $G(A)$ is strictly monotone on $\boldsymbol{\mathcal{X}}(A)$, then it has at most one VWE;

ii) if $G(A)$ is aggregatively strictly monotone on $\boldsymbol{\mathcal{X}}(A)$, then all VWE of $G(A)$ have the same aggregative profile;

iii) if G (without aggregative constraint) is only aggregatively strictly monotone but, for each $\theta \in \Theta$ and all $\boldsymbol{Y} \in \mathcal{M}$, $f_\theta(\boldsymbol{x}, \boldsymbol{Y})$ is strictly convex in $\boldsymbol{x}$, then there is at most one WE.

Proof. Suppose that $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{\mathcal{X}}(A)$ are both VWE. Let $\boldsymbol{X} = \int \boldsymbol{x}$ and $\boldsymbol{Y} = \int \boldsymbol{y}$. According to Theorem 2.1, we have $\int_\Theta \langle \mathbf{g}_{\boldsymbol{x}}(\theta), \boldsymbol{y}_\theta - \boldsymbol{x}_\theta \rangle \mathrm{d}\theta \geq 0$ and $\int_\Theta \langle \mathbf{g}_{\boldsymbol{y}}(\theta), \boldsymbol{x}_\theta - \boldsymbol{y}_\theta \rangle \mathrm{d}\theta \geq 0$. Adding up these two inequalities yields $\int_\Theta \langle \mathbf{g}_{\boldsymbol{x}}(\theta) - \mathbf{g}_{\boldsymbol{y}}(\theta), \boldsymbol{y}_\theta - \boldsymbol{x}_\theta \rangle \mathrm{d}\theta \geq 0$.

i) If $G(A)$ is a strictly monotone game, then $\int_\Theta \langle \mathbf{g}_{\boldsymbol{x}}(\theta) - \mathbf{g}_{\boldsymbol{y}}(\theta), \boldsymbol{x}_\theta - \boldsymbol{y}_\theta \rangle \mathrm{d}\theta = 0$ and thus $\boldsymbol{x} = \boldsymbol{y}$ almost everywhere.

ii)-iii) If $G(A)$ is an aggregatively strictly monotone game, then $\int_\Theta \langle \mathbf{g}_{\boldsymbol{x}}(\theta) - \mathbf{g}_{\boldsymbol{y}}(\theta), \boldsymbol{x}_\theta - \boldsymbol{y}_\theta \rangle \mathrm{d}\theta = 0$ and thus $\boldsymbol{X} = \boldsymbol{Y}$.

If there is no aggregative constraint and $f_\theta(\cdot, \boldsymbol{Z})$ is strictly convex for all $\boldsymbol{Z} \in \mathcal{M}$, then for all θ, $\boldsymbol{x}_\theta$ (resp. $\boldsymbol{y}_\theta$) is the unique minimizer of $f_\theta(\cdot, \boldsymbol{X})$ (resp. $f_\theta(\cdot, \boldsymbol{Y})$). Since $\boldsymbol{X} = \boldsymbol{Y}$, one has $\boldsymbol{x}_\theta = \boldsymbol{y}_\theta$. $\square$

2.4. Symmetric VWE with a finite number of types

A particular class of nonatomic aggregative games is those with only a finite number of types of players, that is, when the sets $\{\mathcal{X}_\theta\}_\theta$ and $\{f_\theta\}_\theta$ are both finite. Consider a nonatomic aggregative game with a set of I types $\mathcal{I} = \{1, \ldots, I\}$. The player set Θ is divided into I measurable subsets $\Theta_1, \ldots, \Theta_I$ such that each nonatomic player belonging to Θ_i is of type i. Let us denote the common action set of players in Θ_i by $\mathcal{X}_i$ and their common cost functions by f_i.

Let us consider a particular class of action profiles in these finite-type nonatomic aggregative games, called *symmetric action profiles*:

Definition 2.5 (Symmetric action profile and symmetric variational Wardrop equilibrium (SVWE)). The set of symmetric action profiles, denoted by $\boldsymbol{\mathcal{X}}_S$, is the set of action profiles where players of the same type play the same action:

$$\boldsymbol{\mathcal{X}}_S \triangleq \{x \in \boldsymbol{\mathcal{X}} : \boldsymbol{x}_\theta = \boldsymbol{x}_\xi, \forall \theta, \xi \in \Theta_i, \forall i \in \mathcal{I}\}$$

The set of symmetric action profiles satisfying the aggregative constraint is denoted by

$$\boldsymbol{\mathcal{X}}_S(A) \triangleq \boldsymbol{\mathcal{X}}_S \cap \boldsymbol{\mathcal{X}}(A). \tag{8}$$

A symmetric variational Wardrop equilibrium is a VWE that is symmetric.

For any symmetric action profile $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}_S$, let the common action of players of type $i \in \mathcal{I}$ be denoted by $\boldsymbol{x}_i$, so that the action profile can be specified by $(\boldsymbol{x}_i)_{i \in \mathcal{I}}$. Obviously, for $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}_S$, for each type $i \in \mathcal{I}$, $\mathbf{g}_{\boldsymbol{x}}(\theta) = \mathbf{g}_{\boldsymbol{x}}(\xi) = \nabla_1 f_i(\boldsymbol{x}_i, \int \boldsymbol{x})$ for all $\theta, \xi \in \mathcal{I}$. Let us abusively denote this quantity by $\mathbf{g}_{\boldsymbol{x}}(i)$.

Proposition 2.2. *In a finite-type nonatomic aggregative game $G(A)$ with an aggregative constraint, a VWE is a symmetric one if and only if it is a solution to the following VI:*

$$\text{Find } \hat{\boldsymbol{x}} \in \boldsymbol{\mathcal{X}}_S(A) \text{ s.t. } \sum_{i \in \mathcal{I}} \langle \mathbf{g}_{\hat{\boldsymbol{x}}}(i), \mu_i \boldsymbol{x}_i - \mu_i \hat{\boldsymbol{x}}_i \rangle \geq 0, \ \forall \boldsymbol{x} \in \boldsymbol{\mathcal{X}}_S(A), \tag{9}$$

where μ_i is the Lebesgue measure of Θ_i.

Proof. Since $\boldsymbol{\mathcal{X}}_S(A) \subset \boldsymbol{\mathcal{X}}(A)$, it is clear that a SVWE, characterized as a solution to the IDVI (7), is a solution to (9).

Conversely, suppose that $\hat{\boldsymbol{x}}$ is a solution to the VI problem (9), let us show that it also solves the IDVI (7). Indeed, for all $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}(A)$,

$$\int_\Theta \langle \mathbf{g}_{\hat{\boldsymbol{x}}}(\theta), \boldsymbol{x}_\theta - \hat{\boldsymbol{x}}_\theta \rangle \, \mathrm{d}\theta = \sum_{i \in \mathcal{I}} \int_{\Theta_i} \langle \mathbf{g}_{\hat{\boldsymbol{x}}}(\theta), \boldsymbol{x}_\theta - \hat{\boldsymbol{x}}_\theta \rangle \, \mathrm{d}\theta = \sum_{i \in \mathcal{I}} \langle \mathbf{g}_{\hat{\boldsymbol{x}}}(i), \int_{\Theta_i} \boldsymbol{x}_\theta - \mu_i \hat{\boldsymbol{x}}_i \rangle$$

$$= \sum_{i \in \mathcal{I}} \langle \mathbf{g}_{\hat{\boldsymbol{x}}}(i), \mu_i \frac{\int_{\Theta_i} \boldsymbol{x}_\theta}{\mu_i} - \mu_i \hat{\boldsymbol{x}}_i \rangle \geq 0$$

as for all $\theta \in \Theta_i$, $\boldsymbol{x}_\theta \in \mathcal{X}_i$ which is convex, hence $\frac{\int_{\Theta_i} \boldsymbol{x}_\theta}{\mu_i} \in \mathcal{X}_i$, so that (9) can be applied. $\square$

Proposition 2.3 (Existence of SVWE). *Under Assumptions 1, 3 and 4, a finite-type nonatomic aggregative game $G(A)$ admits a SVWE.*

Proof. First note that the VI problem (9) is equivalent to a finite dimension VI:

$$\text{Find } \hat{\boldsymbol{x}} \in \boldsymbol{\mathcal{X}}_S(A)' \text{ s.t. } \sum_{i \in \mathcal{I}} \langle \nabla_1 f_i(\hat{\boldsymbol{x}}_i, \sum_{j \in \mathcal{I}} \mu_j \hat{\boldsymbol{x}}_j), \mu_i \boldsymbol{x}_i - \mu_i \hat{\boldsymbol{x}}_i \rangle \geq 0, \; \forall \boldsymbol{x} \in \boldsymbol{\mathcal{X}}_S(A)' , \qquad (10)$$

where $\boldsymbol{\mathcal{X}}_S(A)'$ is the finite-dimensional set

$$\boldsymbol{\mathcal{X}}_S(A)' \triangleq \{\boldsymbol{x} \in \mathbb{R}^{IT} : \boldsymbol{x} \in \prod_{i \in \mathcal{I}} \mathcal{X}_i, \; \sum_{i \in \mathcal{I}} \mu_i \boldsymbol{x}_i \in A\} .$$

As the mapping $(\boldsymbol{x}_i)_{i \in \mathcal{I}} \mapsto \left(\nabla_1 f_i(\boldsymbol{x}_i, \sum_j \mu_j \hat{\boldsymbol{x}}_j)\right)_{i \in \mathcal{I}}$ is continuous from Assumption 4, then [19, Lemma 3.1] implies that the VI (10) has a solution on the finite dimensional convex compact $\boldsymbol{\mathcal{X}}_S(A)'$. $\square$

In this paper, only SVWE are considered for finite-type nonatomic aggregative games. For such equilibria, the distribution of different types on Θ is not relevant, since the equilibrium behavior of each player is only determined by the finite dimensional VI (10). Therefore, we shall specify a finite-type nonatomic aggregative game only by the tuple $\left((\mu_i)_{i \in \mathcal{I}}, (\mathcal{X}_i)_{i \in \mathcal{I}}, (f_i)_{i \in \mathcal{I}}, A\right)$. In particular, a symmetric action profile in such a game shall be specified by a vector $(x_i)_{i \in \mathcal{I}} \in \mathbb{R}^{IT}$, and the set of symmetric action profiles is nothing else but $\boldsymbol{\mathcal{X}}_S(A)'$.

3. Approximating an Infinite-type nonatomic aggregative game

3.1. Finite-type approximating game sequence

After introducing (V)WE in nonatomic aggregative games and SVWE in finite-type nonatomic aggregative games with coupling constraints, we study the relation between these notions. As WE is a particular case of VWE when the aggregate constraint set is any subset of $\mathbb{R}^T$ containing $\overline{\mathcal{X}}$, we can only consider the case of VWE and SVWE.

This section shows the following result: considering a sequence of equilibria of "approximating" finite-type nonatomic aggregative games $(G^\nu(A^\nu))_\nu$ of a nonatomic aggregative game $G(A)$, where each type of players in $G^\nu(A^\nu)$ corresponds to a collection of nonatomic players who are similar in their types, a sequence of SVWE in $(G^\nu(A^\nu))_\nu$ converges to the VWE of $G(A)$ when this one is (aggregatively) strongly monotone.

The particularity of SVWE is that it can be characterized by a finite dimensional VI. As opposed to the case of infinite dimensional ones, there is a large literature on algorithms for computing solutions of finite dimensional VI (e.g. [14] and references therein). Therefore, the result stated above can be practically used to compute a VWE, solution of an IDVI, with arbitrary precision.

In this section, we always suppose that Assumptions 1 to 4 hold.

Let us consider the following definition of an approximating game sequence:

Definition 3.1. Finite-type Approximating Games Sequence
A sequence of finite-type nonatomic aggregative games $\{G^\nu(A^\nu) = \left((\mu_i^\nu)_{i \in \mathcal{I}^\nu}, (\mathcal{X}_i^\nu)_{i \in \mathcal{I}^\nu}, (f_i^\nu)_{i \in \mathcal{I}^\nu}, A^\nu\right) : \nu \in \mathbb{N}^*\}$ with aggregative constraints is a *finite-type approximating game sequence* for the nonatomic aggregative game $G(A) = \left(\Theta, \boldsymbol{\mathcal{X}}, (f_\theta)_\theta, A\right)$ with an aggregative constraint if, for each $\nu \in \mathbb{N}^*$, there exists a partition $(\Theta_0^\nu, \Theta_1^\nu, \ldots, \Theta_{I^\nu}^\nu)$ of the set Θ, with $\mathcal{I}^\nu \triangleq \{1, \ldots, I^\nu\}$, such that the Lebesgue measure of Θ_0^ν is $\mu_0^\nu = 0$, and if, for each $i \in \mathcal{I}^\nu$, the Lebesgue measure of Θ_i^ν is μ_i^ν while the collection of nonatomic players in Θ_i^ν are getting close to the nonatomic players of type $i \in \mathcal{I}^\nu$ in the sense that, as $\nu \to +\infty$:

i) $\overline{\delta}^{\nu} \triangleq \max_{i \in \mathcal{I}^{\nu}} \delta_i^{\nu} \longrightarrow 0$, where δ_i^{ν} is the Hausdorff distance between the action sets of nonatomic players in Θ_i^{ν} in $G(A)$ and the action set of nonatomic players of type $i \in \mathcal{I}^{\nu}$ in $G^{\nu}(A^{\nu})$:

$$\delta_i^{\nu} \triangleq \sup_{\theta \in \Theta_i^{\nu}} d_H\left(\mathcal{X}_{\theta}, \mathcal{X}_i^{\nu}\right) , \tag{11}$$

and span $\mathcal{X}_i^{\nu} = \text{span } \mathcal{X}_{\theta}, \ \forall \theta \in \Theta_i^{\nu}$.

ii) $\overline{d}^{\nu} \triangleq \max_{i \in \mathcal{I}^{\nu}} d_i^{\nu} \longrightarrow 0$, where d_i^{ν} measures the distance between the differential of nonatomic players' cost functions in $G(A)$ and that of their corresponding players' cost functions in $G^{\nu}(A^{\nu})$:

$$d_i^{\nu} \triangleq \sup_{\theta \in \Theta_i} \sup_{(\boldsymbol{x}, \boldsymbol{Y}) \in \mathcal{M}^2} \|\nabla_1 f_i^{\nu}(\boldsymbol{x}_i, \boldsymbol{Y}) - \nabla_1 f_{\theta}(\boldsymbol{x}_{\theta}, \boldsymbol{Y})\| . \tag{12}$$

iii) $D^{\nu} \longrightarrow 0$, where $D^{\nu} \triangleq d_H\left(A^{\nu}, A\right)$ is the Hausdorff distance between the aggregative constraint set $A^{\nu} \subset \mathbb{R}^T$ and the aggregative constraint set $A \subset \mathbb{R}^T$. Besides, span $A = \text{span } A^{\nu}$ for all $\nu \in \mathbb{N}^*$.

Roughly speaking, along a sequence of finite-type approximating games, for each nonatomic player θ in Θ, the difference between her type and her corresponding type i in the approximating game $G^{\nu}(A)$ (in the sense that $\theta \in \Theta_i^{\nu}$) is disappearing as ν goes to infinity. Also, the aggregate-profile constraint sets of the sequence of approximating games converge to the one in $G(A)$.

Note that, except the last condition on D^{ν}, the other conditions are independent of the constraint sets $(A^{\nu})_{\nu}$ and A.

Remark 3.1. The assumption span $\mathcal{X}_i^{\nu} = \text{span } \mathcal{X}_{\theta}, \ \forall \theta \in \Theta_i^{\nu}$ is needed for our proofs because of the existence of coupling constraints. It implies in particular that the nonatomic infinite game considered is such that $\{\text{span } \mathcal{X}_{\theta}\}_{\theta \in \Theta}$ has a finite number of elements. This assumption is natural as, in many models, span $\mathcal{X}_{\theta}$ will be the same for all $\theta \in \Theta$ (see example of Section 4).

Remark 3.2. Without loss of generality, we assume $\text{ri}\left(A \cap \overline{\mathcal{X}}\right) \neq \emptyset$ in this section. Indeed, if the nonempty convex compact set $A \cap \overline{\mathcal{X}}$ has an empty relative interior, then it is reduced to a point hence the problem becomes trivial.

In Section 3.3, we will construct a sequence of finite-type approximating games for two fairly general cases of nonatomic aggregative games.

In order to compare symmetric action profiles in the approximating games and action profiles in the original game, we introduce the following linear mappings which define an equivalent action profile for a symmetric action profile in an approximating game, and vice versa.

First, define the set $L_{\text{S}}^{\mathcal{I}^{\nu}} \triangleq \{\boldsymbol{x} \in L^2([0,1], \mathcal{M}) : \boldsymbol{x}_{\theta} = \boldsymbol{x}_{\xi}, \forall \theta, \xi \in \Theta_i, \forall i \in \mathcal{I}^{\nu}\}$ and the mapping $\psi^{\nu} : L^2([0,1], \mathcal{M}) \to L_{\text{S}}^{\mathcal{I}^{\nu}}$ for each $\nu \in \mathbb{N}^*$ by

$$\forall \boldsymbol{x} \in L^2([0,1], \mathcal{M}^T), \ \psi^{\nu}(\boldsymbol{x}) = \left(\psi_{\theta}^{\nu}(\boldsymbol{x})\right)_{\theta \in \Theta}, \ \text{where } \forall i \in \mathcal{I}^{\nu}, \ \forall \theta \in \Theta_i^{\nu}, \ \psi_{\theta}^{\nu}(\boldsymbol{x}) = \frac{\int_{\Theta_i^{\nu}} \boldsymbol{x}_{\xi} \mathrm{d}\xi}{\mu_i^{\nu}} . \tag{13}$$

The interpretation of ψ^{ν} is that a nonatomic player $\theta \in \Theta_i$ (type i) adopts the average behavior of players in Θ_i^{ν}.

In the following, we assume that Assumptions 1, 3 and 4 also hold for each game G^{ν} of a sequence of finite-type approximating games: this appears naturally in many cases if $(G^{\nu})_{\nu}$ is built from G, as seen in Section 3.3. Finally, let us make the following additional assumption for this section.

Assumption 5. *There is a strictly positive constant η and an action profile $\bar{\boldsymbol{x}} \in \mathcal{X}$ such that, for almost all $\theta \in \Theta$, $d(\bar{\boldsymbol{x}}_{\theta}, \text{rbd } \mathcal{X}_{\theta}) > \eta$.*

It means that the action space of each player has an (aggregatively) nonempty relative interior and that the relative interior is not vanishing along any sequence of players.

3.2. Convergence of equilibrium profiles and aggregate equilibrium profiles

The following Theorem 3.1 gives the main result of this paper. It shows that a VWE in a strongly monotone nonatomic aggregative game can be approximated by SVWE of a finite-type approximating games sequence, both in the case with and without aggregative constraints.

Recall that, according to Theorem 2.4, a strongly monotone game is strictly monotone, hence the VWE is unique, while an aggregatively strongly monotone game is aggregatively strictly monotone, hence the aggregate-action profile at VWE is unique.

Theorem 3.1 (Convergence of SVWE to VWE). *Under Assumptions 1 to 5, let $(G^\nu(A^\nu))_\nu$ be a sequence of finite-type approximating games for the nonatomic aggregative game $G(A)$ with an aggregative constraint A. Let $\boldsymbol{x}^*$ be the VWE of $G(A)$, $\hat{\boldsymbol{x}}^\nu \in \boldsymbol{\mathcal{X}}^\nu(A^\nu)$ an SVWE of $G^\nu(A^\nu)$ for each $\nu \in \mathbb{N}^*$, and $\boldsymbol{X}^*$, $\hat{\boldsymbol{X}}^\nu$ their respective aggregate-action profiles. Then, there exists a constant $\rho > 0$ such that the following results hold with $K_A \triangleq \frac{R+1}{\rho}$:*

i) If G is aggregatively strongly monotone with modulus β, $(\hat{\boldsymbol{X}}^\nu)_\nu$ converges to $\boldsymbol{X}^*$: for all $\nu \in \mathbb{N}^*$ such that $\max(\overline{\delta}^\nu, D^\nu) < \rho$,

$$\|\hat{\boldsymbol{X}}^\nu - \boldsymbol{X}^*\|^2 \leq \frac{1}{\beta}\Big((4L_{\mathbf{f}} + 1)K_A \max(D^\nu, \overline{\delta}^\nu) + (2M+1)\overline{d}^\nu \Big). \tag{14}$$

ii) If G is strongly monotone with modulus α, then $(\hat{\boldsymbol{x}}^\nu)_\nu$, converges to $\boldsymbol{x}^*$ in L^2-norm: for all $\nu \in \mathbb{N}^*$ such that $\max(\overline{\delta}^\nu, D^\nu) < \rho$,

$$\|\hat{\boldsymbol{x}}^\nu - \boldsymbol{x}^*\|_2^2 \leq \frac{1}{\alpha}\Big((4L_{\mathbf{f}} + 1)K_A \max(D^\nu, \overline{\delta}^\nu) + (2M+1)\overline{d}^\nu \Big). \tag{15}$$

If there are no aggregate constraints, one can replace K_A (resp. D^ν) by $\frac{1}{2}$ (resp. 0) in (14) and (15), and Assumption 5 is no longer required.

Some notions and a series of lemmas are needed for the proof of Theorem 3.1.

Notations. Let $\Pi_i^\nu(\cdot)$ denote the (Euclidean) projection function onto $\mathcal{X}_i^\nu$ for $i \in \mathcal{I}^\nu$ and $\Pi_\theta(\cdot)$ the projection function onto $\mathcal{X}_\theta$ for $\theta \in \Theta$.

Let Π^ν denote the (Euclidean) projection function onto $\boldsymbol{\mathcal{X}}_S^\nu(A^\nu) \subset L_S^{I^\nu}$, and Π the projection function onto $\boldsymbol{\mathcal{X}}(A) \subset L^2([0,1], \mathbb{R}^T; \mu)$.

Since $\mathcal{X}_i^\nu$'s, $\mathcal{X}_\theta$'s, $\boldsymbol{\mathcal{X}}(A)$ and $\boldsymbol{\mathcal{X}}_S^\nu(A^\nu)$'s (as defined by (8)) are all convex and closed in their respective Hilbert spaces, the projection functions onto these sets are well defined.

The following Lemma 3.1 shows that the players become infinitesimal along a sequence of finite-type approximating games.

Lemma 3.1. *Under Assumption 1, for all $\nu \in \mathbb{N}^*$, $\|\boldsymbol{x}^\nu\|_2 \leq \overline{\delta}^\nu + R$ for all $\boldsymbol{x}^\nu \in \boldsymbol{\mathcal{X}}_S^\nu$.*

Proof. Let $\boldsymbol{x}_i^\nu \in \mathcal{X}_i^\nu$ and $\theta \in \Theta_i^\nu$. By definition of δ_i^ν, $\|\boldsymbol{x}_i^\nu - \Pi_\theta(\boldsymbol{x}_i^\nu)\| \leq \delta_i^\nu$ so that $\|\boldsymbol{x}_i^\nu\| \leq (\delta_i^\nu + \|\Pi_\theta(\boldsymbol{x}_i^\nu)\|) \leq (\delta_i^\nu + R)$. Then, $\|\boldsymbol{x}^\nu\|_2^2 = \sum_{i=1}^{I^\nu} \int_{\Theta_i} \|\boldsymbol{x}_i^\nu\|^2 \, d\theta = \sum_{i=1}^{I^\nu} \mu_i^\nu \|\boldsymbol{x}_i^\nu\|^2 \leq \sum_{i=1}^{I^\nu} \mu_i^\nu (\delta_i^\nu + R)^2 \leq (\overline{\delta}^\nu + R)^2$. $\qquad \square$

The following lemma shows that the convergence of each type of action set in finite-type game G^ν to that of her corresponding nonatomic player in G, assumed by Eq. (11), implies the convergence of the product action sets in $L^2([0,1], \mathcal{M})$.

Lemma 3.2 (Convergence of $\boldsymbol{\mathcal{X}}_S^\nu$ to $\boldsymbol{\mathcal{X}}$). *Under Assumption 1, for all $\nu \in \mathbb{N}^*$,*

i) for each $\boldsymbol{x}^\nu \in \boldsymbol{\mathcal{X}}_S^\nu$, $d_2(\boldsymbol{x}^\nu, \boldsymbol{\mathcal{X}}) \leq \overline{\delta}^\nu$;

ii) for each $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$, $d_2(\psi^\nu(\boldsymbol{x}), \boldsymbol{\mathcal{X}}_S^\nu) \leq \overline{\delta}^\nu$;

iii) for each $i \in \mathcal{I}^\nu$ and each $\boldsymbol{x}_i^\nu \in \mathcal{X}_i^\nu$, if $d(\boldsymbol{x}_i^\nu, \mathrm{rbd}\, \mathcal{X}_i^\nu) > \delta_i^\nu$, then $\boldsymbol{x}_i^\nu \in \mathcal{X}_\theta$ for all $\theta \in \Theta_i^\nu$;

iv) for each $i \in \mathcal{I}^\nu$, each $\theta \in \Theta_i^\nu$, and each $\boldsymbol{x}_\theta \in \mathcal{X}_\theta$, if $d(\boldsymbol{x}_\theta, \mathrm{rbd}\, \mathcal{X}_\theta) > \delta_i^\nu$, then $\boldsymbol{x}_\theta \in \mathcal{X}_i^\nu$.

Proof. i) Let $\boldsymbol{x}^\nu \in \boldsymbol{\mathcal{X}}_S^\nu$. For each $i \in \mathcal{I}^\nu$ and each $\theta \in \Theta_i^\nu$, define $\boldsymbol{y}_\theta = \Pi_\theta(\boldsymbol{x}^\nu) \in \mathcal{X}_\theta$, so that $\|\boldsymbol{y}_\theta - \boldsymbol{x}^\nu\| \leq \delta_i^\nu$. Let us show that $\boldsymbol{y}$ is measurable on each Θ_i^ν, hence measurable on Θ so that $\boldsymbol{y} \in \boldsymbol{\mathcal{X}}$. For that, define κ_i on $\Theta_i^\nu \times \mathcal{M}^T$ by $\kappa_i : (\theta, \boldsymbol{w}) \mapsto \|\boldsymbol{x}_i^\nu - \boldsymbol{w}\|$. Then, κ_i is a Carathéodory function. Since the correspondence $\Theta_i^\nu \ni \theta \mapsto \mathcal{X}_\theta$ is measurable, according to the measurable maximum theorem [1, Thm. 18.19], there is a measurable selection of $\boldsymbol{x}_\theta \in \arg\min_{\mathcal{X}_\theta} \kappa_i(\theta, \cdot)$. The minimum of $\kappa_i(\theta, \cdot)$ on $\mathcal{X}_\theta$ is unique and is just $\boldsymbol{y}_\theta$, hence $\boldsymbol{y}$ is measurable on Θ_i^ν.

Then, $\|\boldsymbol{x}^\nu - \boldsymbol{y}\|_2 \leq \overline{\delta}^\nu$, which shows that $d_2(\boldsymbol{x}^\nu, \boldsymbol{\mathcal{X}}) \leq \overline{\delta}^\nu$.

ii) Let $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$. For each $i \in \mathcal{I}^\nu$, $\theta \in \Theta_i^\nu$, $\|\boldsymbol{x}_\theta - \Pi_i^\nu(\boldsymbol{x}_\theta)\| \leq \delta_i^\nu$. Since $\mathcal{X}_i^\nu$ is a convex subset in $\mathbb{R}^T$, $\frac{1}{\mu_i^\nu} \int_{\Theta_i^\nu} \Pi_i^\nu(\boldsymbol{x}_\theta) \, \mathrm{d}\theta \in \mathcal{X}_i^\nu$. Define $\boldsymbol{y} \in \boldsymbol{\mathcal{X}}^\nu$ by $\boldsymbol{y}_\theta = \boldsymbol{y}_i \triangleq \frac{1}{\mu_i^\nu} \int_{\Theta_i^\nu} \Pi_i^\nu(\boldsymbol{x}_\theta) \, \mathrm{d}\theta \in \mathcal{X}_i^\nu$ for each $\theta \in \Theta_i$, for $i \in \mathcal{I}_i^\nu$. Then, from Cauchy-Schwartz inequality:

$$\|\psi^\nu(\boldsymbol{x}) - \boldsymbol{y}\|_2^2 = \sum_{i \in \mathcal{I}^\nu} \mu_i^\nu \|\psi_i^\nu(\boldsymbol{x}) - \frac{1}{\mu_i^\nu} \int_{\Theta_i^\nu} \Pi_i^\nu(\boldsymbol{x}_\theta) \, \mathrm{d}\theta\|^2 = \sum_{i \in \mathcal{I}^\nu} \frac{1}{\mu_i^\nu} \| \int_{\Theta_i^\nu} (\boldsymbol{x}_\theta - \Pi_i^\nu(\boldsymbol{x}_\theta)) \, \mathrm{d}\theta\|^2$$

$$\leq \sum_{i \in \mathcal{I}^\nu} \frac{1}{\mu_i^\nu} \mu_i^\nu \int_{\Theta_i} \|\boldsymbol{x}_\theta - \Pi_i^\nu(\boldsymbol{x}_\theta)\|^2 \, \mathrm{d}\theta = \sum_{i \in \mathcal{I}^\nu} \int_{\Theta_i} \|\boldsymbol{x}_\theta - \Pi_i^\nu(\boldsymbol{x}_\theta)\|^2 \, \mathrm{d}\theta \leq \sum_{i \in \mathcal{I}^\nu} \mu_i^\nu (\delta_i^\nu)^2 \leq (\overline{\delta}^\nu)^2,$$

so that $\|\psi^\nu(\boldsymbol{x}) - \boldsymbol{y}\|_2 \leq \overline{\delta}^\nu$. This concludes the proof.

iii) Fix $\nu \in \mathbb{N}^*$, $i \in \mathcal{I}^\nu$ and $\theta \in \Theta_i^\nu$. Consider $\boldsymbol{x}_i^\nu \in \mathcal{X}_i^\nu$ such that $d(\boldsymbol{x}_i^\nu, \mathrm{rbd}\, \mathcal{X}_i^\nu) > \eta$ for some $\eta > \delta_i^\nu$. Assume that $\boldsymbol{x}_i^\nu \notin \mathcal{X}_\theta$ i.e. $\left\|\boldsymbol{x}_i^\nu - \Pi_{\mathcal{X}_\theta}(\boldsymbol{x}_i^\nu)\right\| > 0$. Let $\boldsymbol{y}_i^\nu = \boldsymbol{x}_i^\nu + \eta \frac{\boldsymbol{x}_i^\nu - \Pi_{\mathcal{X}_\theta}(\boldsymbol{x}_i^\nu)}{\left\|\boldsymbol{x}_i^\nu - \Pi_{\mathcal{X}_\theta}(\boldsymbol{x}_i^\nu)\right\|} \in \mathcal{X}_i^\nu$ as span $\mathcal{X}_\theta \subset$ span $\mathcal{X}_i^\nu$.

Since $\mathcal{X}_i^\nu$ is convex, $d(\boldsymbol{y}_i^\nu, \mathcal{X}_\theta) = \left\|\boldsymbol{x}_i^\nu - \Pi_{\mathcal{X}_\theta}(\boldsymbol{x}_i^\nu)\right\| + \left\|\boldsymbol{y}_i^\nu - \boldsymbol{x}_i^\nu\right\| > \eta > \delta_i^\nu$, which contradicts the fact that $d(\mathcal{X}_\theta, \boldsymbol{x}_i^\nu) \leq \delta_i^\nu$. Hence $\boldsymbol{x}_i^\nu \in \mathcal{X}_\theta$.

iv) The proof is similar to that of iii).

$\square$

Note that, because of the convexity assumptions (Assumption 1), the sets of aggregate action profiles in the finite-type game $G^\nu(A^\nu)$, obtained by considering symmetric or by considering non symmetric profiles are the same, that is:

$$\overline{\mathcal{X}}^\nu \triangleq \left\{ \int_\Theta \boldsymbol{x}_\theta \, \mathrm{d}\theta \mid \boldsymbol{x} \in \boldsymbol{\mathcal{X}}^\nu \right\} = \left\{ \int_\Theta \boldsymbol{x}_\theta \, \mathrm{d}\theta \mid \boldsymbol{x} \in \boldsymbol{\mathcal{X}}_S^\nu \right\}, \tag{16}$$

and the same equality holds for $\overline{\mathcal{X}}^\nu \cap A^\nu$ when considering aggregate constraint A^ν.

The sequence of sets of aggregate action profiles $(\overline{\mathcal{X}}^\nu)_\nu$ in games $(G^\nu(A^\nu))_\nu$ *with aggregative constraints*, converges to the set of aggregate-action profiles of the nonatomic aggregative game $G(A)$ *with an aggregative constraint*, as the following lemma says.

Lemma 3.3. *Under Assumption 1, for $\nu \in \mathbb{N}^*$,*

i) $d_H(\overline{\mathcal{X}}^\nu, \overline{\mathcal{X}}) \leq \overline{\delta}^\nu$;

ii) for $\boldsymbol{X} \in \mathrm{ri}\, \overline{\mathcal{X}}$, if $d(\boldsymbol{X}, \mathrm{rbd}\, \overline{\mathcal{X}}) > \overline{\delta}^\nu$, then $\boldsymbol{X} \in \overline{\mathcal{X}}^\nu$;
for $\boldsymbol{X}^\nu \in \mathrm{ri}\, \overline{\mathcal{X}}^\nu$, if $d(\boldsymbol{X}, \mathrm{rbd}\, \overline{\mathcal{X}}^\nu) > \overline{\delta}^\nu$, then $\boldsymbol{X} \in \overline{\mathcal{X}}$;

iii) for $\boldsymbol{X} \in \mathrm{ri}\, A$, if $d(\boldsymbol{X}, \mathrm{rbd}\, A) > D^\nu$, then $\boldsymbol{X} \in A^\nu$;
for $\boldsymbol{X}^\nu \in \mathrm{ri}\, A^\nu$, if $d(\boldsymbol{X}^\nu, \mathrm{rbd}\, A^\nu) > D^\nu$, then $\boldsymbol{X}^\nu \in A$;

iv) for $\boldsymbol{X} \in \mathrm{ri}(\overline{\mathcal{X}} \cap A)$, if $d(\boldsymbol{X}, \mathrm{rbd}(\overline{\mathcal{X}} \cap A)) > \max(\overline{\delta}^\nu, D^\nu)$, then $\boldsymbol{X} \in \overline{\mathcal{X}}^\nu \cap A^\nu$;
for $\boldsymbol{X}^\nu \in \mathrm{ri}(\overline{\mathcal{X}}^\nu \cap A^\nu)$, if $d(\boldsymbol{X}^\nu, \mathrm{rbd}(\overline{\mathcal{X}}^\nu \cap A^\nu)) > \max(\overline{\delta}^\nu, D^\nu)$, then $\boldsymbol{X}^\nu \in \overline{\mathcal{X}} \cap A$.

Proof. i) Fix $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$. Consider $\boldsymbol{y} \in \boldsymbol{\mathcal{X}}_S^\nu$ such that $\|\psi^\nu(\boldsymbol{x}) - \boldsymbol{y}\|_2 \leq \overline{\delta}^\nu$ (cf. Lemma 3.2). Then, from Cauchy-Schwarz inequality, $\| \int \psi^\nu(\boldsymbol{x}) - \int \boldsymbol{y}\|^2 \leq \|\psi^\nu(\boldsymbol{x}) - \boldsymbol{y}\|_2^2 \leq (\overline{\delta}^\nu)^2$. Hence $d(\int \boldsymbol{x}, \overline{\mathcal{X}}^\nu) \leq \| \int \psi^\nu(\boldsymbol{x}) - \int \boldsymbol{y}\| \leq \overline{\delta}^\nu$.

On the other hand, let $\boldsymbol{x}^\nu \in \boldsymbol{\mathcal{X}}_S^\nu$, thus $\boldsymbol{X}^\nu \triangleq \int \boldsymbol{x}^\nu \in \overline{\mathcal{X}}^\nu$. For each $i \in \mathcal{I}^\nu$ and each $\theta \in \Theta_i^\nu$, define $\boldsymbol{y}_\theta = \Pi_\theta(\boldsymbol{x}_i^\nu) \in \mathcal{X}_\theta$, so that $\|\boldsymbol{x}_\theta^\nu - \boldsymbol{y}_\theta\| \leq \delta_i^\nu$. Then, $\| \int \boldsymbol{x}^\nu - \int \boldsymbol{y}\| \leq \int \|\boldsymbol{x}_\theta^\nu - \boldsymbol{y}_\theta\| \, \mathrm{d}\theta \leq \overline{\delta}^\nu$, which shows that $d(\boldsymbol{X}^\nu, \overline{\mathcal{X}}) \leq \overline{\delta}^\nu$ for all $\boldsymbol{X}^\nu \in \overline{\mathcal{X}}^\nu$.

ii-iii) The proof is similar to that for Lemma 3.2.iii).
iv) These are corollaries of ii) and iii).

$\square$

For the proof of the main theorem, we need to rely on nonatomic profiles that are away from the boundary of the feasible domain, giving us some margin. The existence of such profiles is ensured by the following lemma.

Lemma 3.4. *Under Assumptions 1 and 5, there is a strictly positive constant ρ^* and a nonatomic action profile $\boldsymbol{z} \in \boldsymbol{\mathcal{X}}$ such that $\int \boldsymbol{z} \in \mathrm{ri}(\overline{\mathcal{X}} \cap A)$ and, for almost all $\theta \in \Theta$, $d(\boldsymbol{z}_\theta, \mathrm{rbd}\, \mathcal{X}_\theta) > 3\rho^*$.*

Proof. Take $\bar{\boldsymbol{x}}$ the nonatomic action profile in Assumption 5 and an arbitrary $\boldsymbol{y} \in \boldsymbol{\mathcal{X}}(A)$ such that $\int \boldsymbol{y} \in \mathrm{ri}\,(\overline{\mathcal{X}} \cap A)$.

Denote $t = \frac{d(\int \boldsymbol{y}, \mathrm{rbd}\,(\overline{\mathcal{X}} \cap A))}{3M}$. Define profile $\boldsymbol{z} \in \boldsymbol{\mathcal{X}}$ by $\boldsymbol{z} = \boldsymbol{y} - t(\boldsymbol{y} - \bar{\boldsymbol{x}})$.

Firstly, $\| \int \boldsymbol{y} - \int \boldsymbol{z} \| = t \| \int \boldsymbol{y} - \int \bar{\boldsymbol{x}} \| \leq t 2M \leq \frac{2}{3} d(\int \boldsymbol{y}, \mathrm{rbd}\,(\overline{\mathcal{X}} \cap A))$, hence $\int \boldsymbol{z} \in \mathrm{ri}\,(\overline{\mathcal{X}} \cap A)$.

Besides, for any θ, $\boldsymbol{z}_\theta = \boldsymbol{y}_\theta - t(\boldsymbol{y}_\theta - \bar{\boldsymbol{x}}_\theta)$. Since $d(\bar{\boldsymbol{x}}_\theta, \mathrm{rbd}\,\mathcal{X}_\theta) > \eta$, $\boldsymbol{y}_\theta \in \mathcal{X}_\theta$, and $\mathcal{X}_\theta$ is convex, one has $d(\boldsymbol{z}_\theta, \mathrm{rbd}\,\mathcal{X}_\theta) > \eta t = \frac{\eta}{3M} d(\int \boldsymbol{y}, \mathrm{rbd}\,(\overline{\mathcal{X}} \cap A))$. One concludes by defining $\rho^* \triangleq \frac{\eta}{9M} d(\int \boldsymbol{y}, \mathrm{rbd}\,(\overline{\mathcal{X}} \cap A))$. $\quad\square$

Notations. Denote $\boldsymbol{Z} = \int \boldsymbol{z}$ where $\boldsymbol{z}$ is the one in Lemma 3.4 and define $\bar{\rho} \triangleq \frac{1}{3} d(\boldsymbol{Z}, \mathrm{rbd}\,\overline{\mathcal{X}} \cap A) > 0$ and the parameter $\rho \triangleq \min(\rho^*, \bar{\rho})$, appearing in the bounds of Theorem 3.1.
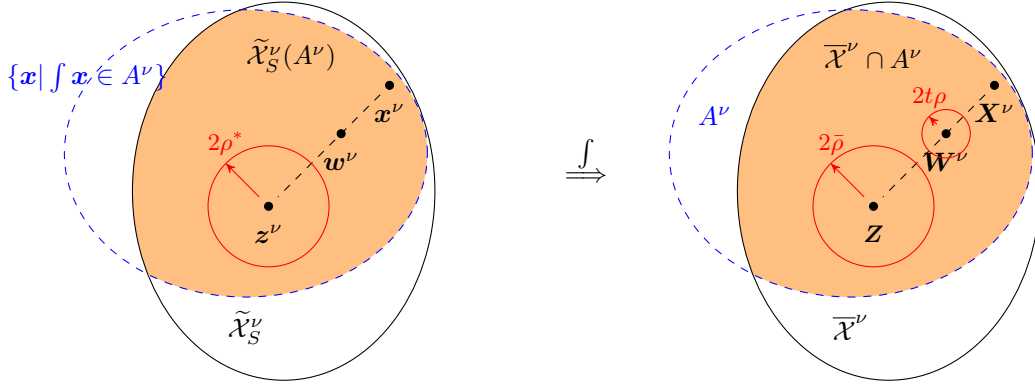

Figure 1: Illustration of the mapping $\int_\Theta(.)$ between $\boldsymbol{\mathcal{X}}^\nu(A^\nu)$ and $\overline{\mathcal{X}} \cap A$ used in Lemma 3.5

The following lemma shows that the space of symmetric action profiles in the finite-type game *with aggregative constraint*, $\boldsymbol{\mathcal{X}}^\nu_S(A^\nu)$, is converging to the space of action profiles in the nonatomic aggregative game *with aggregative constraint*, $\boldsymbol{\mathcal{X}}(A)$.

Lemma 3.5 (Convergence of $\boldsymbol{\mathcal{X}}^\nu_S(A^\nu)$ to $\boldsymbol{\mathcal{X}}(A)$). *Under Assumptions 1 and 5, let $K_A = \frac{R+1}{\rho}$. Then, for all $\nu \in \mathbb{N}^*$ such that $\max(\overline{\delta}^\nu, D^\nu) < \rho$,*

i) for each $\boldsymbol{x}^\nu \in \boldsymbol{\mathcal{X}}^\nu_S(A^\nu)$, $d_2(\boldsymbol{x}^\nu, \boldsymbol{\mathcal{X}}(A)) \leq 2K_A \max(D^\nu, \overline{\delta}^\nu)$;

ii) for each $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}(A)$, $d_2(\psi^\nu(\boldsymbol{x}), \boldsymbol{\mathcal{X}}^\nu_S(A^\nu)) \leq 2K_A \max(D^\nu, \overline{\delta}^\nu)$.

Proof. i) Consider $\boldsymbol{x}^\nu \in \boldsymbol{\mathcal{X}}^\nu_S(A^\nu)$ and $\boldsymbol{X}^\nu = \int \boldsymbol{x}^\nu$. Let $\boldsymbol{z}^\nu \triangleq \psi^\nu(\boldsymbol{z})$ where $\boldsymbol{z}$ is defined in Lemma 3.4. Since for each θ, $d(\boldsymbol{z}_\theta, \mathrm{rbd}\,\mathcal{X}_\theta) > 3\rho > \overline{\delta}^\nu$, one has $\boldsymbol{z}^\nu \in \boldsymbol{\mathcal{X}}^\nu_S$ according to Lemma 3.2.iv).

Define $\boldsymbol{w}^\nu \triangleq \boldsymbol{x}^\nu + t(\boldsymbol{z}^\nu - \boldsymbol{x}^\nu)$ with $t \triangleq \frac{\max(D^\nu, \overline{\delta}^\nu)}{\rho} < 1$, then $\boldsymbol{w}^\nu \in \boldsymbol{\mathcal{X}}^\nu_S$ from convexity and:

- we have $\|\boldsymbol{w}^\nu - \boldsymbol{x}^\nu\|_2 = \max(D^\nu, \overline{\delta}^\nu) \frac{\|\boldsymbol{z}^\nu - \boldsymbol{x}^\nu\|_2}{\rho} \leq \max(D^\nu, \overline{\delta}^\nu) \frac{2(R+1)}{\rho} \leq 2K_A \max(D^\nu, \overline{\delta}^\nu)$;

- let us show that $\boldsymbol{w}^\nu \in \boldsymbol{\mathcal{X}}$: from Lemma 3.2.iii), it is sufficient to show that $d(\boldsymbol{w}^\nu_i, \mathrm{rbd}\,\mathcal{X}^\nu_i) > \overline{\delta}^\nu_i$. For that, we show $d(\boldsymbol{z}^\nu_i, \mathrm{rbd}\,\mathcal{X}^\nu_i) \geq 2\rho$ which implies $d(\boldsymbol{w}^\nu_i, \mathrm{rbd}\,\mathcal{X}^\nu_i) \geq 2t\rho > \overline{\delta}^\nu$. For any arbitrary $\boldsymbol{y}_i \in B_{2\rho}(\boldsymbol{z}^\nu_i) \cap \mathrm{span}\,\mathcal{X}^\nu_i$, let $\boldsymbol{y}_\theta \triangleq \boldsymbol{z}_\theta + \boldsymbol{y}_i - \boldsymbol{z}^\nu_i$ for $\theta \in \Theta^\nu_i$. Then, $\boldsymbol{y}_\theta \in B_{2\rho}(\boldsymbol{z}_\theta) \cap \mathrm{span}\,\mathcal{X}_\theta$ as $\mathrm{span}\,\mathcal{X}_\theta = \mathrm{span}\,\mathcal{X}^\nu_i$, and $d(\boldsymbol{y}_\theta, \mathrm{rbd}\,\mathcal{X}_\theta) \geq d(\boldsymbol{z}_\theta, \mathrm{rbd}\,\mathcal{X}_\theta) - \|\boldsymbol{z}_\theta - \boldsymbol{y}_\theta\| > 3\rho - 2\rho = \rho > \overline{\delta}^\nu$. Hence, from Lemma 3.2.iv), one has $\boldsymbol{y}_\theta \in \mathcal{X}^\nu_i$ and, from convexity, $\boldsymbol{y}_i = \frac{1}{\mu^\nu_i} \int_{\Theta_i} \boldsymbol{y}_\theta \in \mathcal{X}^\nu_i$ which concludes;

- let us show that $\boldsymbol{W}^\nu = \int \boldsymbol{w}^\nu \in A$: from Lemma 3.3.iii), it is sufficient to show that $\boldsymbol{W}^\nu \in \mathrm{ri}\,A^\nu$ and $d(\boldsymbol{W}^\nu, \mathrm{rbd}\,A^\nu) > D^\nu$. First, since $\max(\overline{\delta}^\nu, D^\nu) < \bar{\rho}$, one has $d(\boldsymbol{Z}, \mathrm{rbd}\,(\overline{\mathcal{X}}^\nu \cap A^\nu)) \geq 3\bar{\rho} - \max(\overline{\delta}^\nu, D^\nu) \geq 2\bar{\rho}$. The linear mapping $\boldsymbol{x} \mapsto \int \boldsymbol{x}$ maps the segment linking $\boldsymbol{x}^\nu$ and $\boldsymbol{z}^\nu$ in $\boldsymbol{\mathcal{X}}^\nu(A^\nu)$ to a segment linking $\boldsymbol{X}^\nu$ and $\boldsymbol{Z} = \boldsymbol{Z}^\nu$ in $\overline{\mathcal{X}}^\nu \cap A^\nu$. Hence, by the definition of $\boldsymbol{w}^\nu$, $B_{2t\rho}(\boldsymbol{W}^\nu) \cap \mathrm{span}\,(\overline{\mathcal{X}}^\nu \cap A^\nu) \subset \overline{\mathcal{X}}^\nu \cap A^\nu$, because each point in $B_{2t\rho}(\boldsymbol{W}^\nu) \cap \mathrm{span}\,(\overline{\mathcal{X}}^\nu \cap A^\nu)$ is on the segment linking $\boldsymbol{X}^\nu$ and some point in $B_{2\rho}(\boldsymbol{Z}) \cap \mathrm{span}\,(\overline{\mathcal{X}}^\nu \cap A^\nu) \subset \overline{\mathcal{X}}^\nu \cap A^\nu$. We conclude with $d(\boldsymbol{W}^\nu, \mathrm{rbd}\,(\overline{\mathcal{X}}^\nu \cap A^\nu)) \geq 2t\rho = 2\max(D^\nu, \overline{\delta}^\nu) > D^\nu$.
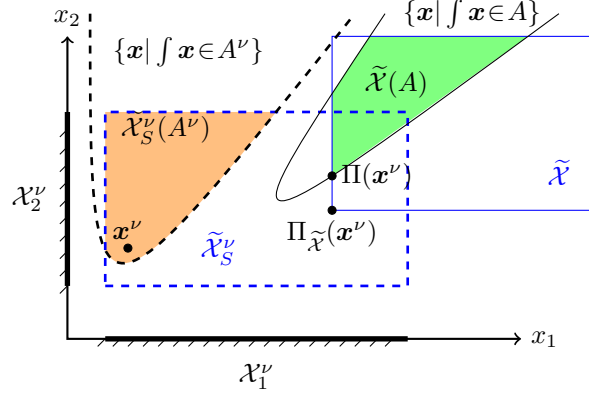
ii) The proof is similar and omitted.

$\quad\square$

Figure 2: Difference between projections on $\boldsymbol{\mathcal{X}}$ and on $\boldsymbol{\mathcal{X}}(A)$. (Since it is impossible to draw the graph of a L^2 action-profile space with a continuum of players, we illustrate the idea with two players.)

Remark 3.3 (Difference between unilateral projections of actions and collective projection of the action profile). Lemma 3.5 shows that $d_2(\boldsymbol{x}^\nu, \Pi(\boldsymbol{x}^\nu)) \leq 2K_A \max(D^\nu, \overline{\delta}^\nu)$ and $d_2(\psi^\nu(\boldsymbol{x}), \Pi^\nu(\psi^\nu(\boldsymbol{x}))) \leq 2K_A \max(D^\nu, \overline{\delta}^\nu)$, with Π and Π^ν the projection functions onto $\boldsymbol{\mathcal{X}}(A)$ and $\boldsymbol{\mathcal{X}}^\nu(A^\nu)$. Lemma 3.5 is of first importance for our proof of Theorem 3.1: without it, we only have the convergence of individual, i.e. unilateral action spaces in the sequence of approximating games, to the unilateral action spaces in the nonatomic aggregative game, as shown in Lemma 3.2. Without coupling constraints, this should be sufficient in the proof of the convergence of a sequence of SVWE. However, in the presence of coupling aggregative constraints, this convergence of unilateral action spaces is not enough. Given a profile in $\boldsymbol{\mathcal{X}}_S^\nu(A^\nu)$, unilateral projection of each type i player's action onto the corresponding nonatomic player's action space in $G(A)$, i.e. from $\mathcal{X}_i^\nu$ to $\mathcal{X}_\theta$, cannot guarantee that the resulting action profile is in $\boldsymbol{\mathcal{X}}(A)$, and vice versa. Lemma 3.5 shows that, for each action profile $\boldsymbol{x}^\nu \in \boldsymbol{\mathcal{X}}_S^\nu(A^\nu)$, its projection on the space of nonatomic action profiles in $\boldsymbol{\mathcal{X}}(A)$ is close to $\boldsymbol{x}^\nu$, and vice versa.

We are finally ready to prove Theorem 3.1.

Proof of Theorem 3.1. Fix $\nu \in \mathbb{N}^*$, define $\hat{\boldsymbol{z}}^\nu \triangleq \Pi(\hat{\boldsymbol{x}}^\nu) \in L^2([0,1], \mathbb{R}^T)$. Then $\hat{\boldsymbol{z}}^\nu \in \boldsymbol{\mathcal{X}}(A)$ is an action profile in nonatomic aggregative game $G(A)$. By the definition of VWE (Definition 2.4), we have $\int_0^1 \langle \mathbf{g}_{\boldsymbol{x}^*}(\theta), \ \boldsymbol{x}_\theta^* - \hat{\boldsymbol{z}}_\theta^\nu \rangle \, \mathrm{d}\theta \leq 0$.

Secondly, by the definition of SVWE, we have $\int_0^1 \langle \mathbf{h}_{\hat{\boldsymbol{x}}^\nu}(\theta), \ \hat{\boldsymbol{x}}_\theta^\nu - \boldsymbol{z}_\theta^\nu \rangle \, \mathrm{d}\theta \leq 0$ for any $\boldsymbol{z}^\nu \in \boldsymbol{\mathcal{X}}^\nu(A)$, where $\mathbf{h}_{\hat{\boldsymbol{x}}^\nu}(\theta) = \nabla_1 f_\theta^\nu(\hat{\boldsymbol{x}}_\theta^\nu, \boldsymbol{X}^\nu) = \nabla_1 f_i^\nu(\hat{\boldsymbol{x}}_i^\nu, \boldsymbol{X}^\nu) \triangleq \mathbf{h}_{\hat{\boldsymbol{x}}^\nu}(i)$, for each $\theta \in \Theta_i^\nu$ and each type $i \in \mathcal{I}^\nu$.

For all $i \in \mathcal{I}^\nu$ and $\theta \in \Theta_i^\nu$, by the definition of d_i^ν (cf. Eq. (12)), we have $\|\mathbf{h}_{\hat{\boldsymbol{x}}^\nu} - \mathbf{g}_{\hat{\boldsymbol{x}}^\nu}\|_2 \leq d_i^\nu$.

Thirdly, $\|\hat{\boldsymbol{x}}^\nu - \hat{\boldsymbol{z}}^\nu\|_2 \leq 2K_A \max(D^\nu, \overline{\delta}^\nu)$ by Lemma 3.5.

With these two results, while noticing that $\hat{\boldsymbol{x}}_\theta^\nu \leq R + \overline{\delta}^\nu$ for all θ by Lemma 3.1, one has:

$$
\int_\Theta \left\langle \mathbf{g}_{\boldsymbol{x}^*}(\theta) - \mathbf{g}_{\hat{\boldsymbol{x}}^\nu}(\theta), \ \boldsymbol{x}_\theta^* - \hat{\boldsymbol{x}}_\theta^\nu \right\rangle \mathrm{d}\theta
$$

$$
= \int_\Theta \langle \mathbf{g}_{\boldsymbol{x}^*}(\theta), \ \boldsymbol{x}_\theta^* - \hat{\boldsymbol{z}}_\theta^\nu \rangle \, \mathrm{d}\theta + \int_\Theta \langle \mathbf{g}_{\boldsymbol{x}^*}(\theta), \ \hat{\boldsymbol{z}}_\theta^\nu - \hat{\boldsymbol{x}}_\theta^\nu \rangle \, \mathrm{d}\theta
$$

$$
+ \int_\Theta \left\langle \mathbf{g}_{\hat{\boldsymbol{x}}^\nu}(\theta) - \mathbf{h}_{\hat{\boldsymbol{x}}^\nu}(\theta), \ \hat{\boldsymbol{x}}_\theta^\nu - \boldsymbol{x}_\theta^* \right\rangle \mathrm{d}\theta + \int_\Theta \left\langle \mathbf{h}_{\hat{\boldsymbol{x}}^\nu}(\theta), \ \hat{\boldsymbol{x}}_\theta^\nu - \boldsymbol{x}_\theta^* \right\rangle \mathrm{d}\theta
$$

$$
\leq 0 + \|\mathbf{g}_{\boldsymbol{x}^*}\|_2 \, \|\hat{\boldsymbol{z}} - \hat{\boldsymbol{x}}^\nu\|_2 + \|\mathbf{g}_{\hat{\boldsymbol{x}}^\nu} - \mathbf{h}_{\hat{\boldsymbol{x}}^\nu}\|_2 \, \|\hat{\boldsymbol{x}}^\nu - \boldsymbol{x}^*\|_2 + J^\nu
$$

$$
\leq 2L_{\mathbf{f}} \, K_A \max(D^\nu, \overline{\delta}^\nu) + (2M + \overline{\delta}^\nu)\overline{d}^\nu + J^\nu \tag{17}
$$

where $J^\nu \triangleq \int_\Theta \left\langle \mathbf{h}_{\hat{\boldsymbol{x}}^\nu}(\theta), \ \hat{\boldsymbol{x}}_\theta^\nu - \boldsymbol{x}_\theta^* \right\rangle \mathrm{d}\theta = \sum_{i \in \mathcal{I}^\nu} \int_{\Theta_i^\nu} \left\langle \mathbf{h}_{\hat{\boldsymbol{x}}^\nu}(i), \ \hat{\boldsymbol{x}}_\theta^\nu - \boldsymbol{x}_\theta^* \right\rangle \mathrm{d}\theta$.

Next, for the VWE $\boldsymbol{x}^* \in \boldsymbol{\mathcal{X}}(A)$, let $\boldsymbol{y}^{*\nu} = \psi(\boldsymbol{x}^*) \in L^2([0,1], \mathcal{M})$ and $\boldsymbol{z}^{*\nu} \triangleq \Pi^\nu(\boldsymbol{y}^{*\nu}) \in \boldsymbol{\mathcal{X}}^\nu(A^\nu)$:

$$J^\nu = \sum_{i \in \mathcal{I}^\nu} \left\langle \mathbf{h}_{\hat{\boldsymbol{x}}^\nu}(i), \int_{\Theta_i^\nu} \hat{\boldsymbol{x}}_\theta^\nu - \boldsymbol{x}_\theta^* \,\mathrm{d}\theta \right\rangle = \sum_{i \in \mathcal{I}^\nu} \left\langle \mathbf{h}_{\hat{\boldsymbol{x}}^\nu}(i), \mu_i^\nu(\hat{\boldsymbol{x}}_i^\nu - \boldsymbol{y}_i^{*\nu}) \right\rangle \tag{18}$$

$$= \sum_{i \in \mathcal{I}^\nu} \left\langle \mathbf{h}_{\hat{\boldsymbol{x}}^\nu}(i), \mu_i^\nu(\hat{\boldsymbol{x}}_i^\nu - \boldsymbol{z}_i^{*\nu}) \right\rangle + \sum_{i \in \mathcal{I}^\nu} \left\langle \mathbf{h}_{\hat{\boldsymbol{x}}^\nu}(i), \mu_i^\nu(\boldsymbol{z}_i^{*\nu} - \boldsymbol{y}_i^{*\nu}) \right\rangle \tag{19}$$

$$\leq 0 + (L_{\mathbf{f}} + \overline{d}^\nu) \left\| \boldsymbol{z}^{*\nu} - \boldsymbol{y}^{*\nu} \right\|_2 \leq (L_{\mathbf{f}} + \overline{d}^\nu) 2 K_A \max(D^\nu, \overline{\delta}^\nu), \tag{20}$$

because of the definition of SVWE $\hat{\boldsymbol{x}}^\nu$ and Proposition 2.2, the definition of $\overline{d}^\nu$ and Lemma 3.5.(i).

Let us summarize by combining (17) and (20), and considering ν large enough such that $\overline{d}^\nu, \overline{\delta}^\nu \leq 1$:

$$\int_\Theta \left\langle \mathbf{g}_{\boldsymbol{x}^*}(\theta) - \mathbf{h}_{\hat{\boldsymbol{x}}^\nu}(\theta), \boldsymbol{x}_\theta^* - \hat{\boldsymbol{x}}_\theta^\nu \right\rangle \mathrm{d}\theta \leq \Omega^\nu \quad \text{with} \quad \Omega^\nu \triangleq (4L_{\mathbf{f}} + 1)K_A \max(D^\nu, \overline{\delta}^\nu) + (2M + 1)\overline{d}^\nu. \tag{21}$$

Last, using the monotonicity definitions (Definition 2.3):

- if G is strongly monotone with modulus α, then $\alpha \left\| \hat{\boldsymbol{x}}^\nu - \boldsymbol{x}^* \right\|_2^2 \leq \Omega^\nu$;

- if G is aggregatively strongly monotone with modulus β, then $\beta \| \hat{\boldsymbol{X}}^\nu - \boldsymbol{X}^* \|^2 \leq \Omega^\nu$,

which lead to the results announced in Theorem 3.1. $\qquad \square$

Remark 3.4. The strong monotonicity of the nonatomic aggregative game G, either with respect to action profile or with respect to aggregate-action profile, is essential in this result. In contrast to finite-player games (cf. [42]), strict monotonicity is not enough to obtain such results using the same techniques. Indeed, since $L^2([0,1], \mathcal{M}^T)$ is only weakly compact, one cannot ensure that $\int_\Theta \langle \mathbf{g}_{\boldsymbol{x}^*}(\theta) - \mathbf{h}_{\hat{\boldsymbol{x}}^\nu}(\theta), \boldsymbol{x}_\theta^* - \hat{\boldsymbol{x}}_\theta^\nu \rangle \,\mathrm{d}\theta$ tends to $\int_\Theta \langle \mathbf{g}_{\boldsymbol{x}^*}(\theta) - \mathbf{h}_{\hat{\boldsymbol{z}}}(\theta), \boldsymbol{x}_\theta^* - \hat{\boldsymbol{z}}_\theta \rangle \,\mathrm{d}\theta$ in (21), where $\hat{\boldsymbol{z}}$ is an accumulation point of $(\hat{\boldsymbol{x}}^\nu)_\nu$ in the weak topology.

3.3. Construction of a sequence of finite-type approximating games

As seen in our previous results, a nonatomic player θ is characterized by two elements: her action set $\mathcal{X}_\theta$, and her gradient $\nabla_1 f_\theta$ defined from $\mathcal{M}^2$ to $\mathbb{R}^T$: $(\boldsymbol{x}, \boldsymbol{Y}) \mapsto \nabla_1 f_\theta(\boldsymbol{x}, \boldsymbol{Y})$.

Note that it is the gradient of the cost function $\nabla_1 f_\theta$, instead of the cost function f_θ itself, that characterizes a nonatomic player's type. For example, two players θ and ξ with $\mathcal{X}_\theta = \mathcal{X}_\xi$ and $f_\theta(\boldsymbol{x}, \boldsymbol{Y}) \equiv f_\xi(\boldsymbol{x}, \boldsymbol{Y}) + C$ where C is a strictly positive constant can be seen as identical in their behavior.

This section presents the construction of a sequence of finite-type approximating games for a given nonatomic aggregative game G in two particular cases: 1) the player characteristic profile $\theta \mapsto (\mathcal{X}_\theta, \nabla_1 f_\theta)$ is piecewise continuous (cf. Definition 3.2) and, 2) $\{\mathcal{X}_\theta, \theta \in \Theta\}$ and $\{f_\theta, \theta \in \Theta\}$ are respectively polytopes and functions parameterized by a finite number of real parameters. Those two cases are fairly general. As illustrated in Section 4, they emerge naturally when the nonatomic game comes from the modeling of a population described by parametric probability distributions, the main motivation for considering infinite nonatomic games.

Case 1: Piecewise Continuous Characteristics – Uniform Splitting.

Definition 3.2 (Continuity of nonatomic player characteristic profile)**.** The player characteristic profile $\theta \mapsto (\mathcal{X}_\theta, \nabla_1 f_\theta)$ in nonatomic aggregative game G is *continuous* at $\theta \in \Theta$ if, for all $\varepsilon > 0$, there exists $\eta > 0$ such that: for each $\theta' \in \Theta$

$$|\theta - \theta'| \leq \eta \ \Rightarrow \begin{cases} d_H(\mathcal{X}_\theta, \mathcal{X}_{\theta'}) \leq \varepsilon \\ \sup_{(\boldsymbol{x}, \boldsymbol{Y}) \in \mathcal{M} \times \mathcal{M}} \|\nabla_1 f_\theta(\boldsymbol{x}, \boldsymbol{Y}) - \nabla_1 f_{\theta'}(\boldsymbol{x}, \boldsymbol{Y})\| \leq \varepsilon \end{cases}. \tag{22}$$

If (22) is true for all θ and θ' on an interval $\Theta' \subset \Theta$, then the player characteristic profile is *uniformly continuous* on Θ'.

Assume that the player characteristic profile $\theta \mapsto (\mathcal{X}_\theta, \nabla_1 f_\theta)$ of nonatomic aggregative game G is piecewise continuous, with a finite number K of discontinuity points $\sigma_0 = 0 \leq \sigma_1 < \sigma_2 < \cdots < \sigma_K \leq \sigma_K = 1$, and that it is uniformly continuous on (σ_k, σ_{k+1}), for each $k \in \{0, \ldots, K-1\}$.

For $\nu \in \mathbb{N}^*$, define an ordered set of I_ν cutting points by $\{v_i^\nu, i = 0, \ldots, I^\nu\} := \left\{\frac{k}{\nu}\right\}_{0 \leq k \leq \nu} \cup \{\sigma_k\}_{1 \leq k \leq K}$ and the corresponding partition $(\Theta_i^\nu)_{i \in \mathcal{I}^\nu}$ of Θ by:

$$\Theta_i^\nu = [v_{i-1}^\nu, v_i^\nu) \text{ for } i \in \{1, \ldots, I^\nu - 1\}; \quad \Theta_{I^\nu}^\nu = [v_{I_\nu - 1}^\nu, 1].$$

Hence, $\mu_i^\nu = v_i^\nu - v_{i-1}^\nu$. Denote $\bar{v}_i^\nu = \frac{v_{i-1}^\nu + v_i^\nu}{2}$.

Proposition 3.1. *Let Assumptions 1 to 4 hold, and assume that* {span $\mathcal{X}_\theta$}$_{\theta \in \Theta}$ *has a finite number of elements. For $\nu \in \mathbb{N}^*$, consider the finite-type game $G^\nu(A^\nu)$ with aggregative constraint $A^\nu \triangleq A$, set of types $\mathcal{I}^\nu \triangleq \{1 \dots I^\nu\}$, where for each type $i \in \mathcal{I}^\nu$:*

$$\mathcal{X}_i^\nu \triangleq \mathcal{X}_{\bar{v}_i^\nu} \quad and \quad f_i^\nu(\boldsymbol{x}, \boldsymbol{Y}) \triangleq f_{\bar{v}_i^\nu}\Big(\boldsymbol{x}, \boldsymbol{Y}\Big), \quad \forall (\boldsymbol{x}, \boldsymbol{Y}) \in \mathcal{M} \times \mathcal{M}.$$

Then $\big(G^\nu(A)\big)_\nu = \big(\mathcal{I}^\nu, \boldsymbol{\mathcal{X}}^\nu, A, (f_i^\nu)_{i \in \mathcal{I}^\nu}\big)_\nu$ is a sequence of finite-type approximating games of nonatomic aggregative game $G(A)$.

Proof. Let us show the three points required by Definition 3.1 as follows.

i) Given an arbitrary $\varepsilon > 0$, there is a common modulus of uniform continuity η such that (22) holds for all the intervals (σ_k, σ_{k+1}). For ν large enough, one has, for each $i \in \mathcal{I}^\nu$, $\mu_i^\nu < \eta$ so that for all $\theta \in \Theta_i^\nu$, $|\bar{v}_i^\nu - \theta| < \eta$; hence $d_H\big(\mathcal{X}_\theta, \mathcal{X}_i^\nu\big) = d_H(\mathcal{X}_\theta, \mathcal{X}_{\bar{v}_i^\nu}) < \varepsilon$.

ii) According to the continuity property, for all $(\boldsymbol{x}, \boldsymbol{Y}) \in \mathcal{M}^2$:

$$\|\nabla_1 f_i^\nu(\mu_i^\nu \boldsymbol{x}, \boldsymbol{Y}) - \nabla_1 f_\theta(\boldsymbol{x}, \boldsymbol{Y})\| = \left\|\nabla_1 f_{\bar{v}_i^\nu}\Big(\boldsymbol{x}, \boldsymbol{Y}\Big) - \nabla_1 f_\theta(\boldsymbol{x}, \boldsymbol{Y})\right\| < \varepsilon.$$

To be rigorous, one would we need to ensurespan $\mathcal{X}_\theta$ to be the same for all $\theta \in \Theta_i^\nu$: if not, one can further divide Θ_i^ν into a finite number groups so that players in each group have the same span $\mathcal{X}_\theta$. This is possible because {span $\mathcal{X}_\theta$}$_\theta$ is finite.

iii) By definition, $D^\nu = 0$. $\qquad\square$

Case 2: Finite-dimensions Parameterized Characteristics – Meshgrid Approximation.

Assume that the nonatomic aggregative game G satisfy two conditions:

(i) The feasible action sets are K-dimensional polytopes: there exists a constant real-valued $K \times T$ matrix $\boldsymbol{P}$, and a bounded mapping $\boldsymbol{b} : \Theta \to \mathbb{R}^K$, such that for any θ, $\mathcal{X}_\theta = \{\boldsymbol{x} \in \mathbb{R}^T : \boldsymbol{P}\boldsymbol{x} \leq \boldsymbol{b}_\theta\}$, which is a nonempty, bounded, closed and convex polytope in $\mathbb{R}^T$.

(ii) There is a bounded mapping $\boldsymbol{s} : \Theta \to \mathbb{R}^l$ such that for any $\theta \in \Theta$, $f_\theta(\cdot, \cdot) = f(\cdot, \cdot; \boldsymbol{s}_\theta)$. Furthermore, for all $(\boldsymbol{x}, \boldsymbol{Y}) \in \mathcal{M}^2$, $\nabla_1 f(\boldsymbol{x}, \boldsymbol{Y}; \cdot)$ is Lipschitz-continuous in $\boldsymbol{s}$ and with a Lipschitz constant L_3, independent of $\boldsymbol{x}$ and $\boldsymbol{Y}$.

Denote $\underline{b}_k = \min_\theta b_{\theta,k}$, $\bar{b}_k = \max_\theta b_{\theta,k}$ for $k \in \{1 \dots K\}$ and $\underline{s}_k = \min_\theta s_{\theta,k}$, $\bar{s}_k = \max_\theta s_{\theta,k}$ for $k \in \{1 \dots l\}$. The characteristics of player θ are parameterized by point $(\boldsymbol{b}_\theta, \boldsymbol{s}_\theta)$ in $\prod_{k=1}^K [\underline{b}_k, \bar{b}_k] \times \prod_{k=1}^l [\underline{s}_k, \bar{s}_k]$, a compact subset of $\mathbb{R}^{K+l}$.

Fix $\nu \in \mathbb{N}^*$, consider a uniform partition of the compact set $\prod_{k=1}^K [\underline{b}_k, \bar{b}_k] \times \prod_{k=1}^l [\underline{s}_k, \bar{s}_k]$, obtained by dividing each dimension of this compact set into ν equal parts. Hence, the partition is composed of $I^\nu \triangleq \nu^{K+l}$ equal-sized subsets of $\prod_{k=1}^K [\underline{b}_k, \bar{b}_k] \times \prod_{k=1}^l [\underline{s}_k, \bar{s}_k]$. The cutting points of the partition are $\underline{b}_{k,n_k} \triangleq \underline{b}_k + \frac{n_k}{\nu}(\bar{b}_k - \underline{b}_k)$ for $k \in \{1, \dots, K\}$, and $\underline{s}_{k,n_k} \triangleq \underline{s}_k + \frac{n_k}{\nu}(\bar{s}_k - \underline{s}_k)$ for $k \in \{1, \dots, l\}$, with $n_k \in \{0, \dots, \nu\}$. Let the set of *vectorial* indices, indexing the partition, be denoted by:

$$\Gamma^\nu \triangleq \{\boldsymbol{n} = (n_k)_{k=1}^{K+l} \in \mathbb{N}^{K+l} \,|\, n_k \in \{1, \dots, \nu\}\}.$$

Define the corresponding partition of the interval Θ: $\Theta = \dot{\bigcup}_{\boldsymbol{n} \in \Gamma^\nu} \Theta_{\boldsymbol{n}}^\nu$, where:

$$\Theta_{\boldsymbol{n}}^\nu \triangleq \Big\{\theta \in \Theta : b_{\theta,k} \in [\underline{b}_{k,n_k-1}, \underline{b}_{k,n_k}) \text{ for } 1 \leq k \leq K; \, s_{\theta,k} \in [\underline{s}_{k,n_k-1}, \underline{s}_{k,n_k}) \text{ for } 1 \leq k \leq l\Big\}.$$

To be rigorous, when $\underline{b}_{k,n_k} = \bar{b}_k$ or $\underline{s}_{k,n_k} = \bar{s}_k$, the parameter interval is closed at the right.

Finally, define the set of players $\mathcal{I}^\nu$ as the elements $\boldsymbol{n}$ in Γ^ν such that $\mu(\Theta_{\boldsymbol{n}}^\nu) > 0$.

Proposition 3.2. *For $\nu \in \mathbb{N}^*$, let the nonatomic finite-type game $G^\nu(A^\nu)$ with an aggregative constraint $A^\nu \triangleq A$, set of types $\mathcal{I}^\nu \triangleq \{\boldsymbol{n} \in \Gamma^\nu : \mu(\Theta_{\boldsymbol{n}}^\nu) > 0\}$ and, for each type $\boldsymbol{n} \in \mathcal{I}^\nu$,*

$$\mathcal{X}_{\boldsymbol{n}}^\nu \triangleq \{\boldsymbol{x} \in \mathbb{R}^T | \boldsymbol{P}\boldsymbol{x} \leq \textstyle\int_{\Theta_{\boldsymbol{n}}^\nu} \boldsymbol{b}_\theta \, d\theta\},$$

$$f_{\boldsymbol{n}}^\nu(\boldsymbol{x}, \boldsymbol{Y}) \triangleq \mu_{\boldsymbol{n}}^\nu f\big(\tfrac{1}{\mu_{\boldsymbol{n}}^\nu} \boldsymbol{x}, \boldsymbol{Y}; \tfrac{1}{\mu_{\boldsymbol{n}}^\nu} \textstyle\int_{\Theta_{\boldsymbol{n}}^\nu} \boldsymbol{s}_\theta \, d\theta\big), \quad \forall (\boldsymbol{x}, \boldsymbol{Y}) \in \mu_i^\nu \mathcal{M} \times \mathcal{M}.$$

Then, under Assumptions 1 to 4, $(G^\nu(A))_\nu = \big(\mathcal{I}^\nu, \boldsymbol{\mathcal{X}}^\nu, A, (f_i^\nu)_{i \in \mathcal{I}^\nu}\big)_\nu$ is a sequence of finite-type approximating games of the nonatomic aggregative game $G(A)$ with an aggregative constraint.

Proof. Let us show the three properties required by Definition 3.1 as follows.

i) For each $\boldsymbol{n} \in \mathcal{I}^{\nu}$, $\mathcal{X}_{\boldsymbol{n}}^{\nu} = \left\{ \boldsymbol{x} \in \mathbb{R}^T : \boldsymbol{P}\boldsymbol{x} \leq \frac{1}{\mu_{\boldsymbol{n}}^{\nu}} \int_{\Theta_{\boldsymbol{n}}^{\nu}} \boldsymbol{b}_{\theta} \, \mathrm{d}\theta \right\}$. Then, by a result generalized from [7, Thm. 4.1], there is a constant C_0 such that, for each $\theta' \in \Theta_{\boldsymbol{n}}^{\nu}$: $d_H\left(\mathcal{X}_{\theta'}, \mathcal{X}_{\boldsymbol{n}}^{\nu}\right) \leq C_0 \left\| \boldsymbol{b}_{\theta'} - \frac{1}{\mu_{\boldsymbol{n}}^{\nu}} \int_{\Theta_{\boldsymbol{n}}^{\nu}} \boldsymbol{b}_{\theta} \, \mathrm{d}\theta \right\| \leq \frac{C_0}{\nu} \left\| \overline{\boldsymbol{b}} - \underline{\boldsymbol{b}} \right\|$. Hence, $\overline{\delta}^{\nu}$ tends to 0. Note that in this case, the assumption of $\{\operatorname{span} \mathcal{X}_{\theta}\}_{\theta \in \Theta}$ being finite is naturally satisfied.

ii) For each $\boldsymbol{n} \in \mathcal{I}^{\nu}$ and each $\theta' \in \Theta_{\boldsymbol{n}}^{\nu}$, for all $(\boldsymbol{x}, \boldsymbol{Y}) \in \mathcal{M}^2$, one has:

$$\|\nabla_1 f_{\boldsymbol{n}}^{\nu}(\boldsymbol{x}, \boldsymbol{Y}) - \nabla_1 f_{\theta'}(\boldsymbol{x}, \boldsymbol{Y})\| = \left\| \nabla_1 f\left(\boldsymbol{x}, \boldsymbol{Y}; \tfrac{1}{\mu_{\boldsymbol{n}}^{\nu}} \int_{\Theta_{\boldsymbol{n}}^{\nu}} \boldsymbol{s}_{\theta} \, \mathrm{d}\theta\right) - \nabla_1 f(\boldsymbol{x}, \boldsymbol{Y}; \boldsymbol{s}_{\theta'}) \right\|$$

$$\leq L_3 \| \frac{1}{\mu_{\boldsymbol{n}}^{\nu}} \int_{\Theta_{\boldsymbol{n}}^{\nu}} \boldsymbol{s}_{\theta} \, \mathrm{d}\theta - \boldsymbol{s}_{\theta'} \| \leq \frac{L_3}{\nu} \| \overline{\boldsymbol{s}} - \underline{\boldsymbol{s}} \| \ ,$$

by the Lipschitz continuity of $\nabla_1 f(\boldsymbol{x}, \boldsymbol{Y}; \cdot)$. Hence, $\overline{d}^{\nu}$ tends to 0.

iii) By definition, $D^{\nu} = 0$. □

Remark 3.5. In Proposition 3.2, instead of the average value of the characteristics of nonatomic players on $\Theta_{\boldsymbol{n}}^{\nu}$, one can use the characteristic value of any nonatomic player in $\Theta_{\boldsymbol{n}}^{\nu}$.

Remark 3.6. By construction, in both sequences above, the compacity and convexity of the feasibility sets $(\mathcal{X}_i)_i$ and the convexity and continuity of cost functions $(f_i)_i$ are naturally inherited from the properties assumed on $(\mathcal{X}_{\theta})_{\theta}$ and $(f_{\theta})_{\theta}$. This should often be the case when building a sequence of approximating games from a nonatomic game with an infinity of types.

4. Illustration on a Smart Grid Example

In this section the results are derived on a simple example for illustration, in the framework stated in introduction. In this example, we will be able to compute explicitly the aggregate equilibrium of the infinite-type nonatomic game.

We suppose that the energy operator has access to the probability distribution of the amount of flexible energy in the $N = 30$ millions French households: let us assume that this distribution is uniform on $[0, E_{\max}]$ with $E_{\max} = 20$kWh (kiloWatthour), that is $\phi_E(E) = \frac{1}{E_{\max}}$ for $E \in [0, E_{\max}]$.

Then the quantile function (or inverse cumulative distribution function), scaled by the population size, is given by $E_{\theta} = F_E^{-1}(\theta) = \theta E_{\max} N$, for each $\theta \in \Theta = [0, 1]$. In this case, the action set mapping $\mathcal{X} : \Theta \rightrightarrows \mathbb{R}^2$ is given by:

$$\forall \theta \in \Theta, \ \mathcal{X}_{\theta} = \left\{ \boldsymbol{x}_{\theta} = (x_{\theta, O}, x_{\theta, P}) \in \mathbb{R}_+^2 \mid x_{\theta, O} + x_{\theta, P} = E_{\theta} \right\} \ ,$$

which gives an infinity of different action sets. Let us consider, as said in the introduction, that there are two prices:

$$c_O(\boldsymbol{X}) = \frac{a_O}{N} X_O \text{ and } c_P(\boldsymbol{X}) = \frac{a_P}{N} X_P$$

for off peak and on peak periods, with $a_P > a_O$, that depend only the aggregate energy on off peak period X_O and on peak period X_P (or rather on the average energy that consumers ask on these periods). Thus, the cost function of each player θ is given, as in the example of public products game given by (6):

$$\forall \boldsymbol{x}_{\theta} \in \mathcal{X}_{\theta}, \ f_{\theta}(\boldsymbol{x}_{\theta}) = x_{\theta, O} \times c_O(\boldsymbol{X}) + x_{\theta, P} \times c_P(\boldsymbol{X}) = \langle \boldsymbol{x}_{\theta}, \boldsymbol{c}(\boldsymbol{X}) \rangle,$$

where $\boldsymbol{c} = (c_0, c_P)$. Hence, all players have the same cost function: the infinite number of types is only due to the infinite number of different action sets.

Owing to to Proposition 2.1, the nonatomic game G obtained is aggregatively strongly monotone with modulus $\beta = \frac{a_O}{N}$. However, the game is not strongly monotone.

It turns out that on this toy example, one can directly compute the aggregate profile of the VWE, as the IDVI of Definition 2.4 asks to find $\boldsymbol{x}^* \in \boldsymbol{\mathcal{X}}$ such that:

$$\int_{\Theta} \langle \mathbf{g}_{\boldsymbol{x}^*}(\theta), \boldsymbol{x}_{\theta} - \boldsymbol{x}_{\theta}^* \rangle \, \mathrm{d}\theta \geq 0, \quad \forall \boldsymbol{x} \in \boldsymbol{\mathcal{X}}$$

$$\iff \int_{\Theta} \langle \boldsymbol{c}(\boldsymbol{X}^*), \boldsymbol{x}_{\theta} - \boldsymbol{x}_{\theta}^* \rangle \, \mathrm{d}\theta \geq 0, \quad \forall \boldsymbol{x} \in \boldsymbol{\mathcal{X}}$$

$$\iff \langle \boldsymbol{c}(\boldsymbol{X}^*), \boldsymbol{X} - \boldsymbol{X}^* \rangle \geq 0, \quad \forall \boldsymbol{X} \in \overline{\boldsymbol{\mathcal{X}}} \ . \tag{23}$$

This simplification holds because, for each θ, $\mathbf{g}_{\boldsymbol{x}^*}(\theta)$ depends only on the aggregate $\boldsymbol{X}^*$ (which would not be the case for general nonlinear cost functions). As a result the VI obtained is of finite dimension.

In this example, the aggregate action set $\overline{\mathcal{X}}$ can also be characterized easily, although this would not be the case for arbitrary sets $(\mathcal{X}_\theta)_\theta$. In fact, as the aggregate flexible energy available is $E_{\text{tot}} \triangleq \int_\Theta E_\theta \, d\theta = \frac{1}{2} N E_{\max}$, we obtain:

$$\overline{\mathcal{X}} = \left\{ (X_O, X_P) \in \mathbb{R}_+^2 \mid X_O + X_P = E_{\text{tot}} \right\} . \tag{24}$$

Indeed, if $X_O + X_P = E_{\text{tot}}$, then taking $\boldsymbol{x}_\theta = \boldsymbol{X} \frac{E_\theta}{E_{\text{tot}}}$, we have $\boldsymbol{x} \in \mathcal{X}$ and $\int_\Theta \boldsymbol{x}_\theta \, d\theta = \boldsymbol{X}$. The converse inclusion is clear. Consequently, we obtain from (23) that $\boldsymbol{X}^*$ is the solution of the quadratic program:

$$\min_{\boldsymbol{X}} \frac{a_O}{N} \times \frac{1}{2} X_O^2 + \frac{a_P}{N} \times \frac{1}{2} X_P^2$$
$$X_O + X_P = E_{\text{tot}}$$
$$0 \leq X_O, X_P$$

that is: $\boldsymbol{X}^* = (X_O^*, X_P^*) = \left(\frac{a_P}{a_O + a_P} E_{\text{tot}}, \frac{a_O}{a_O + a_P} E_{\text{tot}} \right)$.

Now, let us define a sequence of finite-type approximating games G^ν to approximate G, with for each $\nu \in \mathbb{N}^*$, $I^\nu = \nu$. Let us drop the index ν for simplicity in the remaining. Let us split up the population uniformly with $\Theta_i^\nu = [\frac{i-1}{I}, \frac{i}{I}]$, for each $i \in \mathcal{I} = \{1, \dots, I\}$.

Because of the linearity of $\theta \mapsto E_\theta$, considering the uniform approximation detailed in Section 3.3 case 1, one will obtain directly $\boldsymbol{X}^*$. For the example, let us rather consider the approximating games defined with, for each $i \in \mathcal{I}$:

$$\mathcal{X}_i \triangleq \left\{ \boldsymbol{x}_i \in \mathbb{R}_+^2 \mid x_{i,O} + x_{i,P} = E_i \triangleq \frac{i}{I} N E_{\max} \right\} . \tag{25}$$

Besides, we naturally take $f_i \triangleq f_\theta$ for each i (as the cost function is the same for each player).

One can observe that we get for each i, $\delta_i = \frac{N E_{\max}}{I} = \frac{2 E_{\text{tot}}}{I} \to 0$, and of course $d_i = 0$.

On the other hand, computing the aggregate approximate equilibrium, similarly to (23), one obtains:

$$\hat{\boldsymbol{X}}^I = \left(\frac{a_P}{a_O + a_P} E_{\text{tot}} (1 + \frac{1}{I}), \frac{a_O}{a_O + a_P} E_{\text{tot}} (1 + \frac{1}{I}) \right) = (1 + \frac{1}{I}) \boldsymbol{X}^* ,$$

and thus we have:

$$\| \hat{\boldsymbol{X}}^I - \boldsymbol{X}^* \| = \frac{\| \boldsymbol{X}^* \|}{I} = \frac{\sqrt{a_O^2 + a_P^2}}{a_O + a_P} E_{\text{tot}} \times \frac{1}{I}. \tag{26}$$

However, from Theorem 3.1, as we can compute $L_{\mathbf{f}} = \max_{\boldsymbol{X} \in \overline{\mathcal{X}}} \| \boldsymbol{c}(\boldsymbol{X}) \| = \frac{a_P}{N} E_{\text{tot}}$, we obtain the more conservative convergence bound (we can replace $(4 L_{\mathbf{f}} + 1)$ by $2 L_{\mathbf{f}}$ since $D^\nu = \overline{d}^\nu = 0$):

$$\| \hat{\boldsymbol{X}}^I - \boldsymbol{X}^* \|^2 \leq \frac{1}{\beta} 2 L_{\mathbf{f}} \overline{\delta}^I = \frac{N}{a_O} 2 \frac{a_P}{N} E_{\text{tot}} \times \frac{2 E_{\text{tot}}}{I}$$
$$\iff \| \hat{\boldsymbol{X}}^I - \boldsymbol{X}^* \| \leq 2 E_{\text{tot}} \sqrt{\frac{a_P}{a_O}} \times \frac{1}{\sqrt{I}} .$$

5. Conclusion

Theorem 3.1 provides a precise theoretical result for the use of symmetric, finite-dimensional, (variational) Wardrop equilibria (S(V)WE) as an approximation of the (V)WE in a strongly monotone or aggregatively strongly monotone nonatomic aggregative game with an infinity of players types, with or without aggregative constraints. There are numerous research themes related to this result and our topic in general.

First, one needs to find efficient algorithms for the computation of finite dimensional variational inequalities arising as the characterization of SVWE. An extensive literature exists in this regard but our particular case of aggregative game with aggregative constraints may lead to special methods or improvements on existing results [16].

Then, the extension of evolutionary dynamics for population games and the related algorithms, to the framework of nonatomic games with infinitely many classes of players can be non trivial. A recent work [17] proposes online learning methods for population games with heterogeneous convex action sets. The presence of aggregate constraints adds two additional difficulties for considering evolutionary dynamics in population games, as those dynamics are based on unilateral adaptations from players. On the one

hand, in the presence of coupling constraints, unilateral deviations by players may well lead to an action profile violating the coupling constraint. On the other hand, a feasible deviation in the action profile cannot always be decomposed into unilateral deviations of players.

Last, our results are limited to monotone games and the convergence result is limited to strongly monotone games. The study of nonatomic aggregative games that are not monotone needs probably other approaches. Indeed, even for population games where there are only finitely many types of players, there exist much fewer results for games that are not linear, potential or monotone. The question of whether or not it is possible to obtain similar convergence results as those stated in Theorem 3.1 without monotonicity assumptions constitutes an interesting path for future work.

Acknowledgments

Appendix A. Extension of results to the subdifferentiable case

In this section, we explain briefly how our results extend to the case of convex nonsmooth cost functions, considering subdifferential instead of gradients of convex costs.

The essence of the proofs are roughly the same as in the smooth case, but considering subdifferentials requires some additional technical arguments. The full proofs can be found in [23]. The authors decided to formulate the results in the smooth case so that the key arguments and ideas appear clearly.

Recall that the *subdifferential*, i.e. set of *subgradients* of a convex function f at $\boldsymbol{x} \in \mathbb{R}^T$ in its domain C, which is a convex set in $\mathbb{R}^T$, is denoted by $\partial f(\boldsymbol{x})$. Recall that $\mathbf{g} \in \mathbb{R}^T$ is a subgradient of f at $\boldsymbol{x}$, denoted $\mathbf{g} \in \partial f(\boldsymbol{x})$, iff for all $\boldsymbol{z} \in C$, $f(\boldsymbol{z}) \geq f(\boldsymbol{x}) + \langle \mathbf{g}, \boldsymbol{z} - \boldsymbol{x} \rangle$.

One has to consider the correspondence of subdifferential $\mathcal{H} : L^2([0,1], \mathcal{M}) \rightrightarrows L^2([0,1], \mathbb{R}^T)$, which associate to each profile $\boldsymbol{x} \in L^2([0,1], \mathcal{M})$ and each player θ the set of subgradients of her cost functions:

$$\mathcal{H}(\boldsymbol{x}) \triangleq \left\{ \mathbf{g} = (\mathbf{g}_\theta)_{\theta \in \Theta} \mid \mathbf{g}_\theta \in \partial_1 f_\theta(\boldsymbol{x}_\theta, \textstyle\int \boldsymbol{x}), \ \forall a.e. \theta \in \Theta \right\}, \quad \forall \boldsymbol{x} \in L^2([0,1], \mathcal{M}). \tag{A.1}$$

In other words, $\mathcal{H}(\boldsymbol{x})$ is the collection of measurable (and integrable because of Assumption 3.iii adapted to assume uniform boundedness of $\partial_1 f_\theta$) selections of a subgradient for each $\boldsymbol{x}_\theta$. Most of the paper can be interpreted in the nonsmooth framework by

- replacing $\mathbf{g}_{\boldsymbol{x}}$ by an element of $\mathcal{H}(\boldsymbol{x})$ in the equations,

- considering the Hausdorff distance d_H between subdifferentials instead of the Euclidean distance between two gradients (e.g. for d in Definition 3.1 of a sequence of finite-type approximating games),

- a direct implication is that we have to consider *generalized* variational inequalities (GVI), finite or infinite-dimensional, instead of VIs.

For instance, $\langle \mathbf{g}_{\boldsymbol{x}^*}, \boldsymbol{x}^* - \boldsymbol{x} \rangle \leq 0, \forall \boldsymbol{x} \in \boldsymbol{\mathcal{X}}$ becomes $\exists \mathbf{g} \in \mathcal{H}(\boldsymbol{x}^*), \ \langle \mathbf{g}, \boldsymbol{x}^* - \boldsymbol{x} \rangle \leq 0, \forall \boldsymbol{x} \in \boldsymbol{\mathcal{X}}$.

It is useful to introduce the best-reply correspondence $Br : \overline{\mathcal{X}} \rightrightarrows \boldsymbol{\mathcal{X}}$:

$$Br(\boldsymbol{Y}) \triangleq \left\{ \boldsymbol{x} \in \boldsymbol{\mathcal{X}} : \boldsymbol{x}_\theta \in \arg\min_{\mathcal{X}_\theta} f_\theta(\cdot, \boldsymbol{Y}), \forall \theta \in \Theta \right\}, \quad \forall \boldsymbol{Y} \in \overline{\mathcal{X}},$$

and, for $\boldsymbol{Y} \in \overline{\mathcal{X}}$ and $\boldsymbol{x} \in Br(\boldsymbol{Y})$, the correspondence $\mathcal{D}(\boldsymbol{x}, \boldsymbol{Y})$: $\Theta \to \mathbb{R}^T$ defined by:

$$\mathcal{D}(\boldsymbol{x}, \boldsymbol{Y})(\theta) \triangleq \left\{ \mathbf{g}_\theta \in \partial_1 f_\theta(\boldsymbol{x}_\theta, \boldsymbol{Y}) \mid \langle \mathbf{g}_\theta, \boldsymbol{z}_\theta - \boldsymbol{x}_\theta \rangle \geq 0, \quad \forall \boldsymbol{z}_\theta \in \mathcal{X}_\theta \right\} \quad \forall \theta \in \Theta$$

which is nonempty (by first order conditions) and closed-valued. To get similar results as in the smooth case, we need to make the following additional assumption:

Assumption 6. *For all $\boldsymbol{Y} \in \overline{\mathcal{X}}$ and all $\boldsymbol{x} \in Br(\boldsymbol{Y})$, $\mathcal{D}(\boldsymbol{x}, \boldsymbol{Y})$ is a measurable correspondence.*

One can show that $\mathcal{H}(.)$ and Br have nonempty values. Then, instead of Lemma 2.1, we use the compact-valued selection theorem [6] to obtain for each $\boldsymbol{x} \in Br(\boldsymbol{Y})$ the existence of a measurable mapping $\theta \mapsto \mathbf{g}_{\boldsymbol{x}}(\theta)$ such that $\forall \theta \in \Theta, \mathbf{g}_{\boldsymbol{x}}(\theta) \in \mathcal{D}(\boldsymbol{x}, \boldsymbol{Y})(\theta)$.

We can then obtain a characterization of WE similar to Theorem 2.1, where the differential $\mathbf{g}_{\boldsymbol{x}^*}$ is replaced by the existence of an element in $\mathcal{H}(\boldsymbol{x}^*)$. The existence result in Theorem 2.2 is also valid in the subdifferentiable case.

The monotonicity of G is defined as in Definition 2.3, where the inequalities on $\mathbf{g}_x$ and $\mathbf{g}_y$ now have to hold for each pair of elements of the correspondences $(\mathbf{g}_{\boldsymbol{x}}, \mathbf{g}_{\boldsymbol{y}}) \in \mathcal{H}(\boldsymbol{x}) \times \mathcal{H}(\boldsymbol{y})$. Properties characterizing monotonicity given in Proposition 2.1 follow with essentially the same proof, having in mind that for any θ, $\partial_1 f_\theta(\boldsymbol{x}_\theta, \boldsymbol{Y}) = \{\boldsymbol{c}(\boldsymbol{Y}) + \mathbf{g} : \mathbf{g} \in \partial(-u_\theta)(\boldsymbol{x}_\theta)\}$.

In presence of coupling constraints, a VWE (Definition 2.4) is also defined by $\boldsymbol{x}^* \in \boldsymbol{\mathcal{X}}(A)$ and the existence of an element $\mathbf{g} \in \mathcal{H}(x^*)$ satisfying the infinite dimensional GVI (7).

The existence of a VWE (similar to Theorem 2.3) can also be obtained in the nonsmooth case, where the continuity of the gradient is replaced by upper-hemicontinuity of the correspondence $\mathcal{H}$ and applying results of [13, Corollary 2.1]. The uniqueness conditions associated to monotonicity detailed in Theorem 2.4 follow as well with essentially the same proof as for the smooth case.

The main result, Theorem 3.1, is obtained for the nonsmooth case with the same bounds on the convergence rate.

Appendix B. On the Relationship between Nash and Wardrop Equilibria

The objective of this paper is to approximate the equilibrium of a nonatomic game with an infinity of players types, by considering approximating games with a finite number of players types.

A natural idea would also be to consider approximating games with a finite number of *players*, this number of players growing to infinity to approximate the nonatomic population game.

Indeed, we briefly explain in this appendix how we can obtain similar convergence results by adopting this approach considering finite-player atomic games. The approach is fully developed in [23].

The main difference and difficulty under this approach is that the equilibrium concept to consider for finite-player games is no longer Wardrop Equilibrium, but Nash Equilibrium (NE). As the number of players is finite, an individual action $\boldsymbol{x}_i$ of a player i does have an impact on the aggregate action $\boldsymbol{X} = \sum_i \boldsymbol{x}_i$.

As a result, the modified cost function $\hat{f}_i : (\boldsymbol{x}_i, \boldsymbol{X}_{-i}) \mapsto f_i(\boldsymbol{x}_i, \boldsymbol{X}_{-i} + \boldsymbol{x}_i)$ naturally appears, where the action of i is taken into account in the aggregate action.

This modified cost function, and the impact of individual actions in general, have to be considered both in the assumptions and in the definitions of the different concepts used in this paper.

Nash equilibrium are naturally characterized by finite-dimensional variational inequalities under convexity hypotheses: in the atomic case, we need the additional following assumption:

Assumption 7. *For an atomic game $\mathcal{G}(A) = (\mathcal{I}, (f_i)_i, (\mathcal{X}_i)_i, A)$ with a finite set of players $\mathcal{I}$ and cost functions $(f_i)_i$, the associated functions $\left(\hat{f}_i(., \boldsymbol{X})\right)_i$ are convex.*

Note that this convexity is not necessarily implied by the convexity of $f_i(., \boldsymbol{X})$. Under this additional assumption, we obtain a GVI (where G stands for *generalized* in the nonsmooth case, see Appendix A) characterization of NE, similar to the one for SVWE given in Proposition 2.2, and an existence result:

Definition Appendix B.1 (Variational Nash Equilibrium (VNE), [18])**.** A (variational) Nash equilibrium of atomic game $\mathcal{G}(A)$ is a solution to the following GVI problem:

$$\text{Find } \hat{\boldsymbol{x}} \in \boldsymbol{\mathcal{X}}(A) \text{ s.t. } \exists \mathbf{g} \in H(\hat{\boldsymbol{x}}) \text{ s.t. } \langle \mathbf{g}, \boldsymbol{x} - \hat{\boldsymbol{x}} \rangle \geq 0, \ \forall \boldsymbol{x} \in \boldsymbol{\mathcal{X}}(A). \tag{B.1}$$

where the subgradients correspondence $H : \boldsymbol{\mathcal{X}} \rightrightarrows \mathbb{R}^{IT}$ is given as:

$$\forall \boldsymbol{x} \in \boldsymbol{\mathcal{X}}, \ H(\boldsymbol{x}) \triangleq \{(\mathbf{g}_i)_{i \in \mathcal{I}} \in \mathbb{R}^{IT} : \mathbf{g}_i \in \partial_1 \hat{f}_i(\boldsymbol{x}_i, \boldsymbol{X}_{-i}), \ \forall i \in \mathcal{I}\} = \prod_{i \in \mathcal{I}} \partial_1 \hat{f}_i(\boldsymbol{x}_i, \boldsymbol{X}_{-i}) \ .$$

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

Proposition Appendix B.1 (Existence of VNE)**.** *Under Assumptions 1 and 3 (compacity and convexity) on $(\mathcal{X}_i)_i$ and Assumption 7, the atomic game $\mathcal{G}(A) = (\mathcal{I}, (f_i)_i, (\mathcal{X}_i)_i, A)$ admits a VNE.*

To obtain a convergence result of the (V)NEs in a sequence of atomic games, we need some stronger properties than for the sequence of finite-type approximating games. In addition to Definition 3.1, we assume that:

1. the number of players tends to infinity: $I^\nu \underset{\nu \to \infty}{\longrightarrow} \infty$;

2. each player becomes infinitesimal: $\mu_i = \mu(\Theta_i) \underset{\nu \to \infty}{\longrightarrow} 0$;

3. in the gradient (or subdifferential) of a player, the impact of her own action on the aggregate profile vanishes along the sequence, by considering the additional parameter (given in the subdifferential case):
$$\lambda_i^\nu \triangleq \sup_{(\boldsymbol{x}, \boldsymbol{Y}) \in \mathcal{M}^2} \sup_{\mathbf{g} \in \partial_1 \hat{f}_i^\nu(\mu_i^\nu \boldsymbol{x}, \boldsymbol{Y} - \mu_i^\nu \boldsymbol{x})} d\left(\mathbf{g}, \partial_1 f_i^\nu(\mu_i^\nu \boldsymbol{x}, \boldsymbol{Y})\right) \underset{\nu \to \infty}{\longrightarrow} 0 \tag{B.2}$$

Then one obtains similar convergence result as Theorem 3.1, with the only difference being that:

- in the upper bound, $\overline{d}^\nu$ is replaced by $(\overline{d}^\nu + \overline{\lambda}^\nu)$, where $\overline{\lambda}^\nu \triangleq \max_i \lambda_i^\nu$,

- for the convergence in $\Lambda_2([0, 1], \mathcal{M})$, one has to consider a projection of the (V)NE on $\Lambda_2([0, 1], \mathcal{M})$, where each $\theta \in \Theta_i$ is associated to the action of $i \in \mathcal{I}$.

Note that this approach has also another interest, as the convergence theorem that we obtain in this case can be interpreted in the reverse way: under the right assumptions, in a sequence of atomic games converging (in the sense given by the definition of finite-type approximating games) to a nonatomic aggregative game, the sequence of Nash equilibria converge to an equilibrium (the SVWE) of this limit nonatomic game.

In some particular cases where the continuous SVWE can be computed explicitly as a function $\boldsymbol{x}$: $\Theta \to \mathcal{M}$ (for instance in the example derived in Section 4), this would give an approximation of a Nash equilibrium in the atomic (reality) aggregative game, as the finite number of atomic players is very large.

[1] ALIPRANTIS, C. D. and BORDER, K. (2006). *Infinite Dimensional Analysis – A Hitchhiker's Guide.* 3rd ed. Springer-Verlag Berlin Heidelberg.

[2] ALTMAN, E., KAMEDA, H. and HOSOKAWA, Y. (2002). Nash equilibria in load balancing in distributed computer systems. *International Game Theory Review*, **4** 91–100.

[3] ATZENI, I., ORDÓÑEZ, L. G., SCUTARI, G., PALOMAR, D. P. and FONOLLOSA, J. R. (2013). Demand-side management via distributed energy generation and storage optimization. *IEEE Transactions on Smart Grid*, **4** 866–876.

[4] AUMANN, R. J. (1965). Integrals of set-valued functions. *Journal of Mathematical Analysis and Applications*, **12** 1–12.

[5] AUMANN, R. J. (1969). Measurable utility and the measurable choice problem. In *La Décision*. Paris. Colloque Internationaux du C. N. R. S., 15–26.

[6] AUMANN, R. J. (1976). An elementary proof that integration preserves uppersemicontinuity. *Journal of Mathematical Economics*, **3** 15–18.

[7] BATSON, R. G. (1987). Combinatorial behavior of extreme points of perturbed polyhedra. *Journal of Mathematical Analysis and Applications*, **127** 130–139.

[8] BECKMANN, M., MCGUIRE, C. B. and WINSTON, C. B. (1956). *Studies in the Economics of Transportation.* Yale University Press.

[9] CARMONA, G. and PODCZECK, K. (2009). On the existence of pure-strategy equilibria in large games. *Journal of Economic Theory*, **144** 1300–1319.

[10] CAVAZZUTI, E., PAPPALARDO, M. and PASSACANTANDO, M. (2002). Nash equilibria, variational inequalities, and dynamical systems. *Journal of Optimization Theory and Applications*, **114** 491–506.

[11] CORCHÓN, L. C. (1994). Comparative statics for aggregative games the strong concavity case. *Mathematical Social Sciences*, **28** 151–165.

[12] CORREA, J. R., SCHULZ, A. S. and STIER-MOSES, N. E. (2004). Selfish routing in capacitated networks. *Mathematics of Operations Research*, **29** 961–976.

[13] DING, X. P. and TARAFDAR, E. (1996). Monotone generalized variational inequalities and generalized complementarity problems. *Journal of Optimization Theory and Applications*, **88** 107–122.

[14] FACCHINEI, F. and PANG, J.-S. (2007). *Finite-Dimensional Variational Inequalities and Complementarity Problems.* Springer.

[15] FRIESZ, T. L., BERNSTEIN, D., MEHTA, N. J., TOBIN, R. L. and GANJALIZADEH, S. (1994). Day-to-day dynamic network disequilibria and idealized traveler information systems. *Operations Research*, **42** 1120–1136.

[16] GRAMMATICO, S. (2017). Dynamic control of agents playing aggregative games with coupling constraints. *IEEE Transactions on Automatic Control*, **62** 4537–4548.

[17] HADIKHANLOO, S. (2017). Learning in anonymous nonatomic games with applications to first-order mean field games. *arXiv:1704.00378*.

[18] HARKER, P. T. (1991). Generalized Nash games and quasi-variational inequalities. *European Journal of Operational Research*, **54** 81–94.

[19] HARTMAN, P. and STAMPACCHIA, G. (1966). On some non-linear elliptic differential-functional equations. *Acta Mathematica*, **115** 271–310.

[20] HAURIE, A. and MARCOTTE, P. (1985). On the relationship between Nash-Cournot and Wardrop equilibria. *Networks*, **15** 295–308.

[21] HOFBAUER, J. and SANDHOLM, W. H. (2009). Stable games and their dynamics. *Journal of Economic Theory*, **144** 1665–1693, 1693.e4.

[22] HOFBAUER, J. and SIGMUND, K. (1998). *Evolutionary Games and Population Dynamics*. Cambridge University Press.

[23] JACQUOT, P. and WAN, C. (2018). Nonsmooth aggregative games with coupling constraints and infinitely many classes of players. *arXiv preprint arXiv:1806.06230*.

[24] LARSSON, T. and PATRIKSSON, M. (1999). Side constrained traffic equilibrium models – analysis, computation and applications. *Transportation Research Part B: Methodological*, **33** 233–264.

[25] MARCOTTE, P., NGUYEN, S. and SCHOEB, A. (2004). A strategic flow model of traffic assignment in static capacitated networks. *Operations Research*, **52** 191–212.

[26] MARCOTTE, P. and PATRIKSSON, M. (2007). Traffic equilibrium. In *Transportation* (C. Barnhart and G. Laporte, eds.), vol. 14 of *Handbooks in Operations Research and Management Science*. Elsevier, 623–713.

[27] MARCOTTE, P. and ZHU, D.-L. (1997). Equilibria with infinitely many differentiated classes of customers. In *Complementarity and Variational Problems, State of Art*. SIAM, 234–258.

[28] MAS-COLELL, A. (1984). On a theorem of Schmeidler. *Journal of Mathematical Economics*, **13** 201–206.

[29] MAYNARD SMITH, J. (1982). *Evolution and the Theory of Games*. Cambridge University Press.

[30] MILCHTAICH, I. (2000). Generic uniqueness of equilibrium in large crowding games. *Mathematics of Operations Research*, **25** 349–364.

[31] ORDA, A., ROM, R. and SHIMKIN, N. (1993). Competitive routing in multiuser communication networks. *IEEE/ACM Transactions on Networking*, **1** 510–521.

[32] PACCAGNAN, D., GENTILE, B., PARISE, F., KAMGARPOUR, M. and LYGEROS, J. (2018). Nash and Wardrop equilibria in aggregative games with coupling constraints. *IEEE Transactions on Automatic Control* 1–1.

[33] RATH, K. P. (1992). A direct proof of the existence of pure strategy equilibria in games with a continuum of players. *Economic Theory*, **2** 427–433.

[34] ROSENTHAL, R. W. (1973). The network equilibrium problem in integers. *Networks*, **3** 53–59.

[35] SANDHOLM, W. H. (2001). Potential games with continuous player sets. *Journal of Economic Theory*, **97** 81–108.

[36] SANDHOLM, W. H. (2011). *Population Games and Evolutionary Dynamics*. MIT Press.

[37] SCHMEIDLER, D. (1973). Equilibrium points of nonatomic games. *Journal of statistical Physics*, **7** 295–300.

[38] SMITH, M. (1984). A descent algorithm for solving monotone variational inequalities and monotone complementarity problems. *Journal of Optimization Theory and Applications*, **44** 485–496.

[39] SMITH, M. J. (1984). The stability of a dynamic model of traffic assignment—an application of a method of Lyapunov. *Transportation Science*, **18** 245–252.

[40] SORIN, S. and WAN, C. (2016). Finite composite games: Equilibria and dynamics. *Journal of Dynamics & Games*, **3** 101–120.

[41] Taylor, P. D. and Jonker, L. B. (1978). Evolutionary stable strategies and game dynamics. *Mathematical Biosciences*, **40** 145–156.

[42] Wan, C. (2012). Coalitions in nonatomic network congestion games. *Mathematics of Operations Research*, **37** 654–669.

[43] Wardrop, J. G. (1952). Some theoretical aspects of road traffic research. In *Proc. of the Inst. of Civil Eng., Part II, 1.* 325–378.

[44] Zhang, D. and Nagurney, A. (1997). Formulation, stability, and computation of traffic network equilibria as projected dynamical systems. *Journal of Optimization Theory and Applications*, **93** 417–444.

[45] Zhong, R., Sumalee, A., Friesz, T. and Lam, W. H. (2011). Dynamic user equilibrium with side constraints for a traffic network: Theoretical development and numerical solution algorithm. *Transportation Research Part B: Methodological*, **45** 1035–1061.

[46] Zhu, D. and Marcotte, P. (1994). An extended descent framework for variational inequalities. *Journal of Optimization Theory and Applications*, **80** 349–366.