Nonsmooth Aggregative Games with Coupling Constraints and Infinitely Many Classes of Players

Paulin Jacquot, Cheng Wan

▶ **To cite this version:**

Nonsmooth Aggregative Games with Coupling Constraints and Infinitely Many Classes of Players

Paulin Jacquot[*] and Cheng Wan [†]

Thursday 25th October, 2018

Abstract

After defining a pure-action profile in a nonatomic aggregative game, where players have specific compact convex pure-action sets and nonsmooth convex cost functions, as a square-integrable function, we characterize a Wardrop equilibrium as a solution to an infinite-dimensional generalized variational inequality. We show the existence of Wardrop equilibrium and variational Wardrop equilibrium, a concept of equilibrium adapted to the presence of coupling constraints, in monotone nonatomic aggregative games. The uniqueness of (variational) Wardrop equilibrium is proved for strictly or aggregatively strictly monotone nonatomic aggregative games. We then show that, for a sequence of finite-player aggregative games with aggregative constraints, if the players' pure-action sets converge to those of a strongly (resp. aggregatively strongly) monotone nonatomic aggregative game, and the aggregative constraints in the finite-player games converge to the aggregative constraint of the nonatomic game, then a sequence of so-called variational Nash equilibria in these finite-player games converge to the variational Wardrop equilibrium in pure-action profile (resp. aggregate-action profile). In particular, it allows the construction of an auxiliary sequence of games with finite-dimensional equilibria to approximate the infinite-dimensional equilibrium in such a nonatomic game. Finally, we show how to construct auxiliary finite-player games for two general classes of nonatomic games.

Keywords. nonatomic aggregative game, coupling aggregative constraints, generalized variational inequality, monotone game, nonsmooth cost function, variational equilibrium

1 Introduction.

This paper studies firstly the existence and uniqueness of variational Wardrop equilibrium in nonatomic aggregative games with coupling aggregative constraints, where a continuum of players have heterogeneous compact convex pure-action sets and cost functions. It then examines the convergence of a sequence of variational Nash equilibrium in auxiliary finite-player games to the variational Wardrop equilibrium.

Background. Aggregative games form a large class of non-cooperative games. In such a game, a player's payoff is determined by her own action and the aggregate of all the players' actions [13]. The setting of aggregative games is particularly relevant to the study of nonatomic games [45], 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, for example, the distribution of certain 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. Those users have no direct interaction with each other except through the aggregate congestion or consumption to which they are contributing simultaneously. 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 [42], while related research work in transportation and traffic analysis, mostly in the nonatomic version, appeared much earlier [9, 51]. The theory of congestion games has also found numerous applications in telecommunications [39], distributed computing [2], energy management [3], and so on.

The concept of equilibrium in nonatomic games is captured by the so called Wardrop equilibrium (WE) [51]. 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 [28, 36, 44], where finite types of players are considered, each type sharing the same finite number of pure actions and the same payoff function. Convergence of some dynamical systems describing the evolution of pure-action distribution in the population has been established for some particular equilibria in some particular classes such as linear games [49], potential games [9, 43] and stable games [27, 47]. Algorithms corresponding to discretized versions of such dynamical systems for the computation of WE have been studied, in particular for congestion games [20, 52].

Motivation. This paper is mainly motivated by two gaps in the literature on nonatomic games.

Firstly, 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 choose a single path, an electricity consumer faces to a resource allocation problem where she has to divide the consumption of a certain quantity of energy onto 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. Fewer results exist for the computation of WE for the case where players have infinitely many different types (i.e. action sets and payoff functions) or where they have continuous action sets. For example, most of the works in line with Schmeidler [10, 29, 35, 41] use fixed-point theorems to prove the existence of WE, though in fairly general settings. Besides, most of the existing work assumes smooth cost functions of players which is somewhat strong.

Secondly, in the above-mentioned applications of aggregative games, coupling constraints, especially those at aggregative level, commonly exist [22]. Examples are capacity constraints of the network or power-grid, and ramping constraints on the variation of total energy consumption between time periods. In this regard, there are even fewer studies in game-theoretical modeling of nonatomic game. However, the presence of coupling constraints 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 [24] 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 Wardrop equilibria 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.

In view of these two gaps, the main objective of this paper is to provide a model of nonatomic aggregative games with *infinitely many player-specific, compact convex pure-action sets and infinitely many player-specific nonsmooth payoff functions*, then introduce a general form of coupling aggregative constraints into these games, choose an appropriate equilibrium notion, study their properties such as existence and uniqueness and, finally, their computation.

Main results. After defining a pure-action profile in a nonatomic aggregative 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 but nonsmooth, Theorem 3.1 characterizes a Wardrop equilibrium (WE) as a solution to an infinite-dimensional generalized variational inequality (IDGVI).

Theorem 3.3 proves the existence of WE and variational Wardrop equilibrium (VWE), equilibrium notion in the presence of coupling constraints defined by a similar IDGVI, in monotone nonatomic games by showing the existence of solutions to the characteristic IDGVI. Then, Theorem 3.4 shows 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 [27], also called dissipative games [48], in population games with a finite types of nonatomic players to the case with infinitely many types.

Theorem 4.1 is the main result of this paper. It shows that, for a sequence of finite-player aggregative games, if the players' pure-action sets converge to those of a strongly monotone (resp. aggregatively strongly monotone) nonatomic aggregative game, and if the aggregative constraints in these finite-player games converge to the aggregative constraint in the nonatomic game, then a sequence of so-called variational Nash equilibria (VNE) in these finite-player games converge, in pure-action profile (resp. in aggregate-action profile), to the VWE. We provide an upper bound on the distance between the VNE and VWE, specified as a function of the parameters of the finite-player and nonatomic games.

This result allows the construction of an auxiliary sequence of finite-player games with finite-dimensional VNE 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 there are much more results [17] on the resolution of finite-dimensional variational inequalities characterizing VNE, we can therefore obtain an approximation of the VWE with arbitrary precision.

Finally, we show how to construct an AAS for two general classes of nonatomic games.

Related work. Extensive research has been conducted on Wardrop equilibria in nonatomic congestion games via their formulation with variational inequalities [33], while the similar characterization of Nash equilibria in atomic splittable games, where players have continuous action sets in contrast to unsplittable games where players have finite action sets, has been less studied [26, 39]. In addition to their existence and uniqueness, the computational and dynamical aspects of equilibria as solutions to variational inequalities have also been studied [11, 46, 52, 54]. However, in most cases, the variational inequalities involved have finite dimensions, as opposed to the case of Wardrop equilibrium in this paper. Marcotte and Zhu [34] consider nonatomic players with continuous types (leading to a characterization of the Wardrop equilibrium as an infinite-dimensional variational inequality) and studied the equilibrium in an aggregative game with nonatomic players differentiated through a linear parameter in their cost function.

Some results have already been given to quantify the relationship between Nash and Wardrop equilibria. Haurie and Marcotte [26] 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 Wardrop equilibrium 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 [50] 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. [40] 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. There are several differences between our model and theirs. Firstly, we consider nonatomic games with players of infinitely-many different types instead of finite-player games only. Secondly, we consider variational Nash and Wardrop equilibria instead of generalized equilibria (which does 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, which makes the study of its properties much more difficult, as shown in Section 4. Thirdly, we allow for nonsmooth cost functions and general form of coupling constraints while they consider differentiable cost functions and linear coupling constraints.

Milchtaich [37] studies finite and nonatomic crowding games (similar to 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 recalls the definition of finite-player aggregative games with and without aggregative constraints, the notion of equilibrium in these cases and their properties. Section 3 is dedicated to nonatomic aggregative games with and without aggregative constraints. After defining Wardrop equilibrium and variational Wardrop equilibrium, we concentrate on the special class of monotone games and show the existence and uniqueness of equilibria there via generalized infinite dimensional variational inequalities. In Section 4, we give the definition of an approximating sequence of finite-player 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 (variational) Nash equilibria of the approximating finite-player games to the (variational) Wardrop equilibrium of the nonatomic game. The construction of such a sequence of approximating finite-player games is shown for two important classes of nonatomic games.

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

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 $\boldsymbol{x} \in C$;

- 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;

- 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;

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

- rbd C is the *relative boundary* of C in $\mathcal{H}$, i.e. the boundary of C in 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 $\langle F, G \rangle_2 = \int_0^1 \langle F(\theta), G(\theta) \rangle \, \mathrm{d}\theta$. 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 $\max\{\sup_{\boldsymbol{x} \in A} d_m(\boldsymbol{x}, B), \sup_{\boldsymbol{y} \in B} d_m(\boldsymbol{y}, B)\}$. Later, we will define new metrics indexed by ν in Euclidean spaces. The point-set distances and Hausdorff distances are defined similarly and denoted with index ν.

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 if vector $\mathbf{g} \in \mathbb{R}^T$ is a subgradient of f at $\boldsymbol{x}$, then for all $\boldsymbol{z} \in C$, $f(\boldsymbol{z}) \geq f(\boldsymbol{x}) + \langle \mathbf{g}, \boldsymbol{z} - \boldsymbol{x} \rangle$.

For a function $(\boldsymbol{x}, \boldsymbol{X}) \mapsto f(\boldsymbol{x}, \boldsymbol{X})$ of two explicit variables, convex in $\boldsymbol{x}$, we denote by $\partial_1 f(\boldsymbol{x}, \boldsymbol{X})$ the (nonempty) subdifferential of function $f(\cdot, \boldsymbol{X})$ for any fixed $\boldsymbol{X}$.

2 Finite-player aggregative games.

This section recalls the definition of finite-player aggregative games with and without coupling aggregative constraints, and some notions of equilibrium in these games as well as their characterization by generalized variational inequalities.

Definition 2.1 (Finite-player aggregative game). A *finite-player aggregative game* is a non-cooperative game specified by:
(i) a finite set of players $\mathcal{I} = \{1, \dots, I\}$,
(ii) a set of feasible pure actions $\mathcal{X}_i \subset \mathbb{R}^T$ for each player i, where $T \in \mathbb{N}^*$ a constant, with a typical pure action $\boldsymbol{x}_i = (x_{i,t})_{t=1}^T \in \mathcal{X}_i$,
(iii) a cost function $\mathcal{X}_i \times \mathbb{R}^T \to \mathbb{R} : f_i(\boldsymbol{x}_i, \sum_{j \in I} \boldsymbol{x}_j)$ for each player i, so that a player's cost is determined by her own action and the aggregate action profile.

Denote $\widetilde{\mathcal{X}} \triangleq \mathcal{X}_1 \times \cdots \times \mathcal{X}_I$. The pure-action profile of the players $\boldsymbol{x} \triangleq (\boldsymbol{x}_i)_{i \in \mathcal{I}}$ induces an aggregate-action profile load attributed to arc $t \in \mathcal{T}$, which is denoted by $\boldsymbol{X} = (X_t)_{t=1}^T$ where $X_t = \sum_{i \in \mathcal{I}} x_{i,t}$. Denote the set of feasible aggregate-action profiles by $\overline{\mathcal{X}} \triangleq \{\boldsymbol{X} \in \mathbb{R}^T : \exists \boldsymbol{x} \in \widetilde{\mathcal{X}} \text{ s.t. } \sum_{i \in \mathcal{I}} \boldsymbol{x}_i = \boldsymbol{X}\}$.

Denote the game by $\mathcal{G} = (\mathcal{I}, \widetilde{\mathcal{X}}, (f_i)_{i \in \mathcal{I}})$.

The following assumptions and notations are adopted in this paper.

Assumption 1 (Convex costs). *For each $i \in \mathcal{I}$, the function $\hat{f}_i(\boldsymbol{x}_i, \boldsymbol{X}_{-i}) \triangleq f_i(\boldsymbol{x}_i, \boldsymbol{X})$, where $\boldsymbol{X}_{-i} = \sum_{j \in \mathcal{I}, j \neq i} \boldsymbol{x}_j$, is continuous in $\boldsymbol{x}_i$ and in $\boldsymbol{X}_{-i}$, and is convex in $\boldsymbol{x}_i$ for all $\boldsymbol{x}_{-i} = (\boldsymbol{x}_j)_{j \in \mathcal{I}, j \neq i} \in \prod_{j \in \mathcal{I}, j \neq i} \mathcal{X}_j$.*

Assumption 2 (Convex and compact strategy sets). *For each $i \in \mathcal{I}$, the set $\mathcal{X}_i$ is a nonempty, convex and compact subset of $\mathbb{R}^T$.*

Recall the definition of Nash Equilibrium in finite-player non-cooperative games.

Definition 2.2 (Nash Equilibrium (NE) [38]). A (pure) *Nash equilibrium* of $\mathcal{G}$ is a profile of pure actions $\hat{\boldsymbol{x}} \in \widetilde{\mathcal{X}}$ such that $\hat{f}_i(\hat{\boldsymbol{x}}_i, \boldsymbol{X}_{-i}) \leq \hat{f}_i(\boldsymbol{x}_i, \boldsymbol{X}_{-i})$ for all $\boldsymbol{x}_i \in \mathcal{X}_i$ and all $i \in \mathcal{I}$.

Define a correspondence $H : \widetilde{\mathcal{X}} \rightrightarrows \mathbb{R}^{IT}$ by

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

Since the cost functions are convex in players' own strategies, NE can be characterized as solutions to generalized variational inequalities (GVI) [19].

Proposition 2.1 (GVI formulation of NE). *Under Assumptions 1 and 2, $\hat{\boldsymbol{x}} \in \widetilde{\mathcal{X}}$ is an NE of $\mathcal{G}$ if and only if either of the following two equivalent conditions holds:*

$$\forall i \in \mathcal{I}, \exists\, \mathbf{g}_i \in \partial_1 \hat{f}_i(\hat{\boldsymbol{x}}_i, \hat{\boldsymbol{X}}_{-i}) \ s.t. \ \langle \mathbf{g}_i, \boldsymbol{x}_i - \hat{\boldsymbol{x}}_i \rangle \geq 0, \ \forall \boldsymbol{x}_i \in \mathcal{X}_i, \tag{1a}$$

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

Proof. Eq. (1a) is a necessary and sufficient condition for $\hat{\boldsymbol{x}}_i$ to minimize the convex function $\hat{f}_i(.,\hat{\boldsymbol{X}}_{-i})$ on $\mathcal{X}_i$ ([8, Proposition 27.8]). The equivalence between (1a) and (1b) is obvious. □

Remark 2.1 (Generalized VI and Generalized NE are different things). The variational inequalities (VI) are of *generalized* type here because the subdifferentials of cost functions are not necessarily singled valued. In the case that cost functions are differentiable with respect to the players' own actions, the GVI are reduced to a VI. Do not confuse with *generalized NE* in *generalized games* (cf. Definition 2.3) which are characterized by *(generalized-)quasi-VI.*

The existence of an NE is obtained by a classical result in game theory for finite-player continuous game, since the players have convex continuous cost functions and convex compact pure-action sets. No differentiability condition is needed.

Proposition 2.2 (Existence of NE, [15, 18, 21]). *Under Assumptions 1 and 2, $\mathcal{G}$ admits an NE.*

Remark 2.2 (NE is a unilateral level stability condition). The NE condition ensures stability not only in terms of a single player's behavior but also in terms of their collective welfare. Indeed, on the one hand, condition (1a) is equivalent to $\langle \mathbf{g}_i, \boldsymbol{y}_i \rangle \geq 0$ for all $\boldsymbol{y}_i \in T_{\mathcal{X}_i}(\hat{\boldsymbol{x}}_i)$ for each i, i.e. a unilaterally feasible deviation of player i increases her cost; on the other hand, condition (1b) is equivalent to $\langle \mathbf{g}, \boldsymbol{y} \rangle \geq 0$ for all $\boldsymbol{y} \in T_{\widetilde{\mathcal{X}}}(\hat{\boldsymbol{x}})$, i.e. a collectively feasible deviation of all the players increases their total costs. The two conditions are equivalent because the players have independent pure-action spaces so that $T_{\widetilde{\mathcal{X}}}(\hat{\boldsymbol{x}}) = \prod_{i \in \mathcal{I}} T_{\mathcal{X}_i}(\hat{\boldsymbol{x}}_i)$, i.e. any collectively feasible deviation can be decomposed into unilaterally feasible deviations. This remark is important because it is no longer the case when one introduces a coupling constraint in the game.

The coupling aggregative constraint considered in this paper is of the following general form: There is a nonempty, convex and compact subset A of $\mathbb{R}^T$, whose intersection with $\overline{\mathcal{X}}$ is not empty, such that the aggregate-action profile $\boldsymbol{X} \in A$. An example is $A = \{\boldsymbol{X} \in \mathbb{R}_+^T : \underline{M}_t \leq X_t \leq \overline{M}_t, \forall t \in \mathcal{T}; a_t \leq X_{t+1} - X_t \leq b_t, \forall t \in \mathcal{T} \setminus \{T\}\}$.

Definition 2.3 (Finite-player aggregative game with aggregative constraints). Its only difference from the game defined in Definition 2.1 is that, for each player $i \in \mathcal{I}$, given the profile of pure actions of the others players $\boldsymbol{y}_{-i} \in \prod_{j \in \mathcal{I} \setminus \{i\}} \mathcal{X}_j$, her feasible pure-action set becomes $\mathcal{X}_i(\boldsymbol{y}_{-i}) \triangleq \{\boldsymbol{x}_i \in \mathcal{X}_i : (\boldsymbol{x}_i, \boldsymbol{y}_{-i}) \in \widetilde{\mathcal{X}}(A)\}$, where $\widetilde{\mathcal{X}}(A)$ is a subset of $\widetilde{\mathcal{X}}$ defined by

$$\widetilde{\mathcal{X}}(A) = \left\{ \boldsymbol{x} \in \widetilde{\mathcal{X}} : \boldsymbol{X} = \sum_{i \in \mathcal{I}} \boldsymbol{x}_i \in A \right\} \,.$$

This game is denoted by $\mathcal{G}(A) = (\mathcal{I}, \widetilde{\mathcal{X}}, A, (f_i)_{i \in \mathcal{I}})$ or simply $\mathcal{G}(A)$.

Finite-player non-cooperative games with coupling constraints are called generalized Nash games [24]. The extension from games to generalized games is not trivial. In the case with no coupling constraint, the pure-action spaces of the players are independent so that any collectively feasible deviation can be decomposed into unilaterally feasible deviation. This property does not always hold with a coupling constraint. To see this, we recall the following notion of generalized equilibrium in generalized games.

Definition 2.4 (Generalized Nash Equilibrium (GNE), [24]). A profile of pure actions $\hat{\boldsymbol{x}} \in \mathcal{X}$ is a *generalized Nash equilibrium* of $\mathcal{G}(A)$ if

$$\hat{\boldsymbol{x}}_i \in \mathcal{X}_i(\hat{\boldsymbol{x}}_{-i}) \ \text{ and } \ \hat{f}_i(\hat{\boldsymbol{x}}_i, \hat{\boldsymbol{X}}_{-i}) \leq \hat{f}_i(\boldsymbol{x}_i, \hat{\boldsymbol{X}}_{-i}), \quad \forall \boldsymbol{x}_i \in \mathcal{X}_i(\hat{\boldsymbol{x}}_{-i}), \quad \forall i \in \mathcal{I} \,.$$

Its characterization by generalized quasi-variational inequalities (GQVI) [12] can be proved as for Proposition 2.1 .

Proposition 2.3 (GQVI formulation of GNE). *Under Assumptions 1 and 2, $\hat{\boldsymbol{x}} \in \widetilde{\mathcal{X}}$ is a GNE of $\mathcal{G}(A)$ if and only if one of the following two equivalent conditions holds:*

$$\forall i \in \mathcal{I} : \hat{\boldsymbol{x}}_i \in \widetilde{\mathcal{X}}(\hat{\boldsymbol{x}}_{-i}) \text{ and } \exists \mathbf{g}_i \in \partial_1 \hat{f}_i(\hat{\boldsymbol{x}}_i, \hat{\boldsymbol{X}}_{-i}) \text{ s.t. } \langle \mathbf{g}_i, \boldsymbol{x}_i - \hat{\boldsymbol{x}}_i \rangle \geq 0, \ \forall \boldsymbol{x}_i \in \mathcal{X}_i(\hat{\boldsymbol{x}}_{-i}) \tag{2a}$$

$$\hat{\boldsymbol{x}} \in \widetilde{\mathcal{X}}(\hat{\boldsymbol{x}}) \triangleq \prod_{i \in \mathcal{I}} \mathcal{X}_i(\boldsymbol{x}_{-i}) \text{ and } \exists \mathbf{g} \in H(\hat{\boldsymbol{x}}) \text{ s.t. } \langle \mathbf{g}, \boldsymbol{x} - \hat{\boldsymbol{x}} \rangle \geq 0, \ \forall \boldsymbol{x} \in \widetilde{\mathcal{X}}(\hat{\boldsymbol{x}}) \ . \tag{2b}$$

Remark 2.3 (Problematics of GNE). The notion of GNE can be problematic. Given a pure-action profile $\boldsymbol{x} \in \widetilde{\mathcal{X}}(A)$, firstly simultaneous unilateral deviations can lead to a new profile out of $\widetilde{\mathcal{X}}(A)$. Secondly, if $\widetilde{\mathcal{X}}(\boldsymbol{x}) \cap \widetilde{\mathcal{X}}(A)$ is a proper set of $\widetilde{\mathcal{X}}(A)$, profiles in $\widetilde{\mathcal{X}}(A) \setminus \widetilde{\mathcal{X}}(\boldsymbol{x})$ are not feasible. Indeed, the unilateral stability condition (2a) is equivalent to collective stability only among those collective deviations composed by unilaterally feasible deviations, i.e. condition (2b), because $T_{\mathcal{X}(\hat{\boldsymbol{x}})}(\hat{\boldsymbol{x}}) = \prod_{i \in \mathcal{I}} T_{\widetilde{\mathcal{X}}_i(\hat{\boldsymbol{x}}_{-i})}(\hat{\boldsymbol{x}}_i)$. Collective deviations towards $\widetilde{\mathcal{X}}(A) \setminus \widetilde{\mathcal{X}}(\boldsymbol{x})$ are not composed by unilaterally feasible deviations while they may effectively decrease the total cost or even each player's cost (cf. [24] for an example). A GNE can thus lose its stability when such collective deviations are allowed. To answer to this issue, we consider the stronger notion of equilibrium defined below:

Definition 2.5 (Variational Nash Equilibrium (VNE), [24]). A solution to the following GVI problem:

$$\text{Find } \hat{\boldsymbol{x}} \in \widetilde{\mathcal{X}}(A) \text{ s.t. } \exists \mathbf{g} \in H(\hat{\boldsymbol{x}}) \text{ s.t. } \langle \mathbf{g}, \boldsymbol{x} - \hat{\boldsymbol{x}} \rangle \geq 0, \ \forall \boldsymbol{x} \in \widetilde{\mathcal{X}}(A). \tag{3}$$

is called a variational Nash equilibrium of $\mathcal{G}(A)$. In particular, if $\overline{\mathcal{X}} \subset A$, a VNE is a NE.

A pure-action profile $\hat{\boldsymbol{x}} \in \widetilde{\mathcal{X}}(A)$ is a VNE if and only if a collective deviation to any profile in $\widetilde{\mathcal{X}}(A)$ is not collectively beneficial. Indeed, VNE is also unilaterally stable, because it is a GNE [24, Theorem 3].

Proposition 2.4 ([24]). *In $\mathcal{G}(A)$, under Assumption 2, any VNE is a GNE.*

Remark 2.4 (VNE refines GNE). VNE can be seen as a refinement of GNE [30]. With a small perturbation of A, a GNE which is not a VNE can no longer be an equilibrium. Harker [24] gives an example of a GNE which is not a VNE.

Hence, VNE is adopted in this paper as the equilibrium notion in the presence of aggregative constraints. Moreover, in Section 3.3 it is argued that the notion of generalized equilibrium cannot even be established in nonatomic games.

Proposition 2.5 (Existence of VNE). *Under Assumptions 1 and 2, $\mathcal{G}(A)$ admits a VNE.*

Proof. From the convexity of each $\hat{f}_i$, we deduce that H is a nonempty, convex, compact valued, upper hemicontinuous correspondence. Then [12, Corollary 3.1] shows that the GVI problem (3) admits a solution on the finite dimensional convex compact $\widetilde{\mathcal{X}}(A)$. (In the case that the $\hat{f}_i$ is partially differentiable with respect to $\boldsymbol{x}_i$, then the GVI is reduced to a VI, and Lemma 3.1 in [25] suffices to show the existence of a solution.) □

3 A continuum of players: nonatomic framework.

3.1 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 the Schmeidler's seminal paper [45], but in contrast to most of the population games studied in game theory [28, 44] where nonatomic players are grouped into several populations, with players in the same population have the same finite pure-action set and the same cost function.

Definition 3.1 (Nonatomic aggregative game). A *nonatomic aggregative game G* is defined by:
(i) a continuum of players represented by 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, and
(iii) a cost function $\mathcal{X}_\theta \times \mathbb{R}^T \to \mathbb{R} : f_\theta(\boldsymbol{x}_\theta, \boldsymbol{X})$ for each player θ, where $\boldsymbol{X} = (X_t)_{t=1}^T$ and $X_t \triangleq \int_0^1 \boldsymbol{x}_{\theta,t} \, \mathrm{d}\theta$ denotes the aggregate-action profile.

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

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

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

Remark 3.1. The definition of a nonatomic 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 games and the notion of WE. This is in contrast to finite-player games.

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

Further assumptions are necessary for $\widetilde{\mathcal{X}}$ to be nonempty and for the existence of equilibria to be discussed later.

Assumption 3 (Nonatomic pure-action sets)**.** *The correspondence $\mathcal{X} : \Theta \rightrightarrows \mathbb{R}^T, \theta \mapsto \mathcal{X}_\theta$ has nonempty, convex, compact values and 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 $\mathbb{R}^{T+1}$. Moreover, for all $\theta \in \Theta$, $\mathcal{X}_\theta \subset N_M(\boldsymbol{0})$, with $M > 0$ a constant.*

Under Assumption 3, a sufficient condition for $\boldsymbol{x}$ to be in $L^2([0,1], \mathbb{R}^T)$ is that $\boldsymbol{x}$ is measurable.

Notations. Denote $\mathcal{M} = [0, M+1]^T$.

Assumption 4 (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 is bounded on $\mathcal{M}^2$, and for each aggregate profile $\boldsymbol{Y} \in \mathcal{M}$,*
(i) function $Gr_{\mathcal{X}} \to \mathbb{R}^T : (\theta, \boldsymbol{x}_\theta) \mapsto f_\theta(\boldsymbol{x}_\theta, \boldsymbol{Y})$ is measurable.
(ii) for each $\theta \in \Theta$, function $\boldsymbol{x}_\theta \mapsto f_\theta(\boldsymbol{x}_\theta, \boldsymbol{Y})$ is continuous and convex on $\mathcal{M}'$;
(iii) There is $B_{\mathbf{f}} > 0$ such that $\|\mathbf{g}\| \leq B_{\mathbf{f}}$ for all subgradients $\mathbf{g} \in \partial_1 f_\theta(\boldsymbol{x}_\theta, \boldsymbol{Y})$ for each $\boldsymbol{x}_\theta \in \mathcal{M}$, each $\boldsymbol{Y} \in \mathcal{M}$, and each $\theta \in \Theta$.

Remark 3.2. Assumption 4.(iii) implies that $f_\theta(\cdot, \cdot)$'s are Lipschitz in the first variable with a uniform Lipschitz constant $B_{\mathbf{f}}$ on $\mathcal{M}^2$ for all θ. Besides, if $f_\theta(\cdot, \boldsymbol{y})$ is differentiable on $\mathcal{M}$, then $\partial_1 f_\theta(\boldsymbol{x}_\theta, \boldsymbol{Y})$ contains one element $\nabla_1 f_\theta(\boldsymbol{x}_\theta, \boldsymbol{Y})$, the gradient of $f_\theta(\cdot, \boldsymbol{Y})$ at $\boldsymbol{x}_\theta$.

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 3.2 (Wardrop Equilibrium (WE), [51])**.** A pure-action profile $\boldsymbol{x}^* \in \widetilde{\mathcal{X}}$ is a *Wardrop equilibrium* of nonatomic game G if

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

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

First, define a correspondence $H : L^2([0,1], \mathcal{M}) \rightrightarrows L^2([0,1], \mathbb{R}^T)$ as follows:

$$H(\boldsymbol{x}) \triangleq \{\mathbf{g} = (\mathbf{g}_\theta)_{\theta \in \Theta} \in L^2([0,1], \mathbb{R}^T) : \mathbf{g}_\theta \in \partial_1 f_\theta(\boldsymbol{x}_\theta, \textstyle\int \boldsymbol{x}), \forall \, a.e. \, \theta \in \Theta\}, \quad \forall \boldsymbol{x} \in L^2([0,1], \mathcal{M}). \quad (4)$$

In other words, $H(\boldsymbol{x})$ is the collection of measurable (and integrable because of Assumption 4.(iii)) selections of a subgradient for each $\boldsymbol{x}_\theta$.

Next, define a best-reply correspondence Br from the set of aggregate-action profiles $\overline{\mathcal{X}}$ to the set of pure-action profiles $\widetilde{\mathcal{X}}$:

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

Finally, fix $\boldsymbol{Y} \in \overline{\mathcal{X}}$ and $\boldsymbol{x} \in Br(\boldsymbol{Y})$, define a correspondence $\mathcal{D}(\boldsymbol{x}, \boldsymbol{Y})$ from Θ to $\mathbb{R}^T$ as follows:

$$\mathcal{D}(\boldsymbol{x}, \boldsymbol{Y})(\theta) \triangleq \{\mathbf{g}_\theta \in \partial_1 f_\theta(\boldsymbol{x}_\theta, \boldsymbol{Y}) : \langle \mathbf{g}_\theta, \boldsymbol{y}_\theta - \boldsymbol{x}_\theta \rangle \geq 0, \forall \boldsymbol{y}_\theta \in \mathcal{X}_\theta\}, \quad \forall \theta \in \Theta. \quad (5)$$

Clearly, this is a nonempty and closed-valued correspondence.

Assumption 5. *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.*

Theorem 3.1 (GVI formulation of WE). *Under Assumptions 3 to 5, $\boldsymbol{x}^* \in \widetilde{\mathcal{X}}$ is a WE of nonatomic game G if and only if either of the following two equivalent conditions is true:*

$$\forall\, a.e.\, \theta \in \Theta,\ \exists\, \mathbf{g}_\theta \in \partial_1 f_\theta(\boldsymbol{x}_\theta^*, \boldsymbol{X}^*)\ s.t\ \langle \mathbf{g}_\theta, \boldsymbol{x}_\theta - \boldsymbol{x}_\theta^* \rangle \geq 0,\ \forall \boldsymbol{x}_\theta \in \mathcal{X}_\theta\ , \tag{6a}$$

$$\exists\, \mathbf{g} \in H(\boldsymbol{x}^*)\ s.t\ \int_\Theta \langle \mathbf{g}_\theta, \boldsymbol{x}_\theta - \boldsymbol{x}_\theta^* \rangle\, \mathrm{d}\theta \geq 0,\ \forall \boldsymbol{x} \in \widetilde{\mathcal{X}}\ . \tag{6b}$$

We need the following lemma for the proof of Theorem 3.1.

Lemma 3.1.
(1) For all $\boldsymbol{x} \in L^2([0,1], \mathcal{M})$, $H(\boldsymbol{x})$ is nonempty.
(2) For all $\boldsymbol{Y} \in \overline{\mathcal{X}}$, $Br(\boldsymbol{Y})$ is nonempty.
(3) Under Assumption 5, for all $\boldsymbol{Y} \in \overline{\mathcal{X}}$ and all $\boldsymbol{x} \in Br(\boldsymbol{Y})$, there exists a measurable mapping $\theta \mapsto \mathbf{g}_\theta$ such that $\mathbf{g}_\theta \in \mathcal{D}(\boldsymbol{x}, \boldsymbol{Y})(\theta)$ for each $\theta \in \Theta$.

Proof. (1) For each θ, the subdifferential $\partial_1 f_\theta(\boldsymbol{x}_\theta, \int \boldsymbol{x})$ is nonempty and compact valued, so that a measurable selection exists according to the compact-valued selection theorem [6].

(2) Fix $\boldsymbol{Y} \in \overline{\mathcal{X}}$. A consequence of Assumption 4.(i-ii) is that the function $\Theta \times \mathcal{M} \to \mathbb{R} : (\theta, \boldsymbol{z}) \mapsto f_\theta(\boldsymbol{z}, \boldsymbol{Y})$ is a Carathéodory function, that is, (i) $f_\cdot(\boldsymbol{z}, \boldsymbol{Y})$ is measurable on Θ for each $\boldsymbol{z} \in \mathcal{M}$, and (ii) $f_\theta(\cdot, \boldsymbol{Y})$ is continuous on $\mathcal{M}$ for each $\theta \in \Theta$. Thus, according to the measurable maximum theorem [1, Thm. 18.19] applied to $f_\cdot(\cdot, \boldsymbol{Y})$, there exists a selection $\boldsymbol{x}_\theta \in \arg\min_{\mathcal{X}_\theta} f_\theta(\cdot, \boldsymbol{Y})$ such that $\boldsymbol{x}$ is a measurable function on Θ.

(3) Because of Assumption 5, one can apply the compact-valued selection theorem [6]. $\square$

Proof of Theorem 3.1. Given $\boldsymbol{X}^*$, (6a) is a necessary and sufficient condition for $\boldsymbol{x}_\theta^*$ to minimize the convex function $f_\theta(., \boldsymbol{X}^*)$ on $\mathcal{X}_\theta$. Condition (6a) implies condition (6b) because of Assumption 5.

For the converse, suppose that $\boldsymbol{x}^* \in \widetilde{\mathcal{X}}$ satisfies condition (6b) but not (6a). 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 $\boldsymbol{y}_\theta \in \mathcal{Y}_\theta$, $\langle \mathbf{g}_\theta, \boldsymbol{y}_\theta - \boldsymbol{x}_\theta^* \rangle < f_\theta(\boldsymbol{y}_\theta, \boldsymbol{X}^*) - f_\theta(\boldsymbol{x}_\theta^*, \boldsymbol{X}^*) < 0$. By the same argument as in the proof of Lemma 3.1, one can select a $\boldsymbol{y}_\theta \in \arg\min_{\mathcal{X}_\theta} f_\theta(\cdot, \boldsymbol{X}^*)$ for $\theta \in \Theta'$ such that $\Theta' \to \mathbb{R}^T : \theta \mapsto \boldsymbol{y}_\theta$ is measurable. 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 $\widetilde{\mathcal{X}}$. However, $\int_\Theta \langle \mathbf{g}_\theta, \boldsymbol{y}_\theta - \boldsymbol{x}_\theta^* \rangle\, \mathrm{d}\theta = \int_{\Theta'} \langle \mathbf{g}_\theta, \boldsymbol{y}_\theta - \boldsymbol{x}_\theta^* \rangle\, \mathrm{d}\theta < 0$, contradicting (6b). $\square$

Remark 3.3. Condition (6a) is equivalent to $\langle \mathbf{g}_\theta(\boldsymbol{x}_\theta^*, \boldsymbol{X}^*), \boldsymbol{y}_\theta \rangle \geq 0$ for all $\boldsymbol{y}_\theta \in T_{\mathcal{X}_\theta}(\boldsymbol{x}_\theta^*)$ for each θ. The interpretation is the same as for atomic players at NE: no unilateral deviation is profitable. However, since each nonatomic player has measure zero, when considering a deviation in the profile of pure actions, one must let players in a set of strictly positive measure deviate: (6b) means that the collective deviation of players of any set of strictly positive measure increases their cost. Note that the GVI problem (6b) has infinite dimensions.

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

Theorem 3.2 (Existence of a WE, [41]). *Under Assumptions 3 and 4.(1), 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 [41] on the existence of WE in aggregate games are satisfied. $\square$

Remark 3.4. No convexity of $f_\theta(\cdot, \boldsymbol{Y})$'s are needed.

3.2 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 introduce the following notions of monotone nonatomic games.

Definition 3.3. The nonatomic aggregative game G is *monotone* if

$$\int_\Theta \langle \mathbf{g}_\theta - \mathbf{h}_\theta, \boldsymbol{x}_\theta - \boldsymbol{y}_\theta \rangle\, \mathrm{d}\theta \geq 0, \quad \forall \boldsymbol{x}, \boldsymbol{y} \in L^2([0,1], \mathcal{M}) \text{ and } \mathbf{g} \in H(\boldsymbol{x}), \mathbf{h} \in H(\boldsymbol{y}). \tag{7}$$

It is *strictly monotone* if the equality in (7) holds if and only if $\boldsymbol{x} = \boldsymbol{y}$ almost everywhere.

8

It is *aggregatively strictly monotone* if the equality in (7) holds if and only if $\int \boldsymbol{x} = \int \boldsymbol{y}$.

It is *strongly monotone* with modulus α if

$$\int_\Theta \langle \mathbf{g}_\theta - \mathbf{h}_\theta, \boldsymbol{x}_\theta - \boldsymbol{y}_\theta \rangle \, \mathrm{d}\theta \geq \alpha \|\boldsymbol{x} - \boldsymbol{y}\|_2^2, \ \forall \boldsymbol{x}, \boldsymbol{y} \in L^2([0,1], \mathcal{M}) \text{ and } \mathbf{g} \in H(\boldsymbol{x}), \mathbf{h} \in H(\boldsymbol{y}). \tag{8}$$

It is *aggregatively strongly monotone* with modulus β if

$$\int_\Theta \langle \mathbf{g}_\theta - \mathbf{h}_\theta, \boldsymbol{x}_\theta - \boldsymbol{y}_\theta \rangle \, \mathrm{d}\theta \geq \beta \| \int \boldsymbol{x} - \int \boldsymbol{y} \|^2, \ \forall \boldsymbol{x}, \boldsymbol{y} \in L^2([0,1], \mathcal{M}) \text{ and } \mathbf{g} \in H(\boldsymbol{x}), \mathbf{h} \in H(\boldsymbol{y}). \tag{9}$$

Remark 3.5. Eq. (7) means nothing else but H is a monotone correspondence on $L^2([0,1], \mathcal{M})$ (cf. [8, Definition 20.1] for the definition of monotone correspondence in Hilbert spaces).

Remark 3.6. A recent paper of Hadikhanloo [23] generalizes the notion of stable games in population games [27] to monotone games in anonymous games, an extension of population games with players having heterogeneous compact pure-action sets but the same payoff function. He defines the notion of monotonicity directly on the distribution of pure-actions among the players instead of pure-action profile as we do. The two approaches are compatible.

Examples of aggregative games are given by cost functions of the form:

$$f_\theta(\boldsymbol{x}_\theta, \boldsymbol{X}) = \langle \boldsymbol{x}_\theta, \boldsymbol{c}(\boldsymbol{X}) \rangle - u_\theta(\boldsymbol{x}_\theta) \ . \tag{10}$$

Here $\boldsymbol{c}(\boldsymbol{X})$ 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 3.1. *Under Assumptions 3 to 5, in a nonatomic aggregative game G with cost functions of form (10), assume that $\boldsymbol{c}$ is monotone on $\mathcal{M}$ and, for each θ, u_θ is a concave function on $\mathcal{M}$. Then:*
(1) G is a monotone game.
(2) If u_θ is strictly concave on $\mathcal{M}$ for all $\theta \in \Theta$, then G is a strictly monotone game.
(3) If $\boldsymbol{c}$ is strictly monotone on $\mathcal{M}$, then G is an aggregatively strictly monotone game.
(4) 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.
(5) If $\boldsymbol{c}$ is strongly monotone on $\mathcal{M}$ with β, then G is an aggregatively strongly monotone game with modulus β.

Proof. (1) Let $\boldsymbol{x}, \boldsymbol{y} \in \widetilde{\mathcal{X}}$ and $\boldsymbol{X} = \int \boldsymbol{x}$, $\boldsymbol{Y} = \int \boldsymbol{y}$. For each θ, $\partial_1 f_\theta(\boldsymbol{x}_\theta, \boldsymbol{Y}) = \{\boldsymbol{c}(\boldsymbol{Y}) + \mathbf{g} : \mathbf{g} \in \partial(-u_\theta)(\boldsymbol{x}_\theta)\}$. Then, given $\boldsymbol{x}, \boldsymbol{y} \in \widetilde{\mathcal{X}}$, with $\mathbf{g}_\theta \in \partial(-u_\theta)(\boldsymbol{x}_\theta)$ and $\mathbf{h}_\theta \in \partial(-u_\theta)(\boldsymbol{y}_\theta)$ for each θ, one has $\langle \mathbf{g}_\theta(\boldsymbol{x}_\theta) - \mathbf{h}_\theta(\boldsymbol{y}_\theta), \boldsymbol{x}_\theta - \boldsymbol{y}_\theta \rangle \geq 0$ because u_θ is concave so that $\partial(-u_\theta)$ is a monotone correspondence on $\mathcal{X}_\theta$.

Then $\int_0^1 \langle (\boldsymbol{c}(\boldsymbol{X}) + \mathbf{g}_\theta(\boldsymbol{x}_\theta)) - (\boldsymbol{c}(\boldsymbol{Y}) - \mathbf{h}_\theta(\boldsymbol{y}_\theta)), \boldsymbol{x}_\theta - \boldsymbol{y}_\theta \rangle \, \mathrm{d}\theta = \langle \boldsymbol{c}(\boldsymbol{X}) - \boldsymbol{c}(\boldsymbol{Y}), \boldsymbol{X} - \boldsymbol{Y} \rangle + \int_0^1 \langle \mathbf{g}_\theta(\boldsymbol{x}_\theta) - \mathbf{h}_\theta(\boldsymbol{y}_\theta), \boldsymbol{x}_\theta - \boldsymbol{y}_\theta \rangle \, \mathrm{d}\theta \geq 0$ because $\boldsymbol{c}$ is monotone. Hence G is a monotone game.

The proof for (2)-(5) 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.

3.3 Nonatomic aggregative games with aggregate constraints.

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 $\widetilde{\mathcal{X}}(A)$ be a subset of $\widetilde{\mathcal{X}}$ defined by $\widetilde{\mathcal{X}}(A) \triangleq \{\boldsymbol{x} \in \widetilde{\mathcal{X}} : \boldsymbol{X} = \int \boldsymbol{x} \in A\}$. Denote the nonatomic game with aggregative constraint $G(A)$.

In contrast to games with finitely many players, a generalized equilibrium in the style of Definition 2.4 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 pure-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 a pure-action profile in $\widetilde{\mathcal{X}}(A)$, simultaneous unilateral deviations by the players can lead to any profile in $\widetilde{\mathcal{X}}$. If only profiles in $\mathcal{X}(A)$ are allowed to be attained, then one lands on a notion similar to VNE. 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 3.4 (Variational Wardrop Equilibrium (VWE)). A solution to the following infinite dimensional GVI problem:

$$\text{Find } \boldsymbol{x}^* \in \widetilde{\mathcal{X}}(A) \text{ and } \mathbf{g} \in H(\boldsymbol{x}^*) \text{ s.t. } \int_{\Theta} \langle \mathbf{g}_{\theta}, \boldsymbol{x}_{\theta} - \boldsymbol{x}_{\theta}^* \rangle \, \mathrm{d}\theta \geq 0, \quad \forall \boldsymbol{x} \in \widetilde{\mathcal{X}}(A), \tag{11}$$

is called a *variational Wardrop equilibrium* of $G(A)$, where the correspondence H is as defined by Eq. (4).

Remark 3.7 (Justification for VWE). From a game theoretical point of view, the notion of VWE can be problematic as well. Each nonatomic player can deviate unilaterally without having any impact on the aggregate profile. The unilateral stability as for NE and VNE (cf. Remarks 2.2 and 2.4) is lost. However, our main theorem, Theorem 4.1, shows that VWE can be seen as the limit of a sequence of VNE which are unilaterally stable. Finally, in the literation of congestion games, the equilibrium notion characterized by VI of form (11) but in finite dimension and with smooth cost functions has long been studied. For example, see [14, 31, 32, 53] and references therein.

The following facts are needed for later use. Under Assumption 3:

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

- $\widetilde{\mathcal{X}}(A)$ is a nonempty, convex and closed subset of $\widetilde{\mathcal{X}}$;

- $\overline{\mathcal{X}}$ 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 $\widetilde{\mathcal{X}}$ and $\overline{\mathcal{X}}$ are nonempty because of Assumption 3 and the measurable selection theorem of Aumann [5], while aggregate-action set $\overline{\mathcal{X}}$ is compact by [4, Theorem 4]. Theorem 3.3 shows the existence of VWE via the VI approach. Compared with Theorem 3.2, much stronger conditions are required on cost functions.

Assumption 6 (Continuity of cost function in aggregate action). *For each $\theta \in \Theta$ and $\boldsymbol{x}_{\theta} \in \mathcal{M}$, $f_{\theta}(\boldsymbol{x}_{\theta}, \cdot)$ is continuous on $\mathcal{M}$.*

Theorem 3.3 (Existence of VWE). *Under Assumptions 3, 4 and 6, if nonatomic game with coupling constraint $G(A)$ is monotone on $\widetilde{\mathcal{X}}(A)$, then a VWE exists.*

Proof. Let us apply [16, Corollary 2.1] to show that Eq. (11) has a solution. This theorem states that if $\widetilde{\mathcal{X}}(A)$ is bounded, closed and convex in $L^2([0, 1], \mathbb{R}^T)$, and if $H : L^2([0, 1], \mathcal{M}) \rightrightarrows L^2([0, 1], \mathbb{R}^T)$ is a monotone correspondence which is upper hemicontinuous from the line segments in $\widetilde{\mathcal{X}}(A)$ to the weak* topology of $L^2([0, 1], \mathbb{R}^T)$, then (11) admits a solution. We only need to show the upper hemicontinuity property. First notice that $H(\cdot)$ has closed values. Take $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\widetilde{\mathcal{X}}(A)$, consider sequence $(\boldsymbol{x}^{(k)})_k$ with $\boldsymbol{x}^{(k)} = \boldsymbol{x} + \frac{1}{k}(\boldsymbol{y} - \boldsymbol{x})$, and sequence $(\mathbf{g}^{(k)})_k$ such that $\mathbf{g}^{(k)} \in H(\boldsymbol{x}^{(k)})$ and $\mathbf{g}^{(k)} \overset{*}{\rightharpoonup} \mathbf{g}$ with $\mathbf{g} \in L^2([0, 1], \mathbb{R}^T)$. Let us show that $\mathbf{g} \in H(\boldsymbol{x})$.

Denote $\boldsymbol{X} = \int \boldsymbol{x}$ and $\boldsymbol{X}^{(k)} = \int \boldsymbol{x}^{(k)}$. Then $\boldsymbol{X}^{(k)}$ converges to $\boldsymbol{X}$ in l^2-norm.

By definition of H, 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 $\langle \mathbf{g}_{\theta}^{(k)}, \boldsymbol{z}_{\theta} - \boldsymbol{x}_{\theta} \rangle \to \langle \mathbf{g}_{\theta}, \boldsymbol{z}_{\theta} - \boldsymbol{x}_{\theta} \rangle$ because $\mathbf{g} \overset{*}{\rightharpoonup} \mathbf{g}$, 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 $B_{\mathbf{f}}$. 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 $\mathbf{g}_{\theta} \in \partial_1 f_{\theta}(\boldsymbol{x}_{\theta}, \boldsymbol{X})$. Since the limit of measurable functions is measurable, $\mathbf{g}$ is measurable. Hence $\mathbf{g} \in H(\boldsymbol{x})$, which concludes the proof. $\qquad \square$

Theorem 3.4 (Uniqueness of VWE). *Under Assumptions 3 and 4:*
(1) if $G(A)$ is strictly monotone on $\widetilde{\mathcal{X}}(A)$, then it has at most one VWE;
(2) if $G(A)$ is aggregatively strictly monotone on $\widetilde{\mathcal{X}}(A)$, then all VWE of $G(A)$ have the same aggregative profile;
(3) 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 $\boldsymbol{x}, \boldsymbol{y} \in \widetilde{\mathcal{X}}(A)$ are both VWE. Let $\boldsymbol{X} = \int \boldsymbol{x}$ and $\boldsymbol{Y} = \int \boldsymbol{y}$. According to Theorem 3.1, there exist $\mathbf{g} \in H(\boldsymbol{x})$ an $\mathbf{h} \in H(\boldsymbol{y})$ such that $\int_\Theta \langle \mathbf{g}_\theta, \boldsymbol{y}_\theta - \boldsymbol{x}_\theta \rangle \, \mathrm{d}\theta \geq 0$ and $\int_\Theta \langle \mathbf{h}_\theta, \boldsymbol{x}_\theta - \boldsymbol{y}_\theta \rangle \, \mathrm{d}\theta \geq 0$. Adding up these two inequalities yields $\int_\Theta \langle \mathbf{g}_\theta - \mathbf{h}_\theta, \boldsymbol{y}_\theta - \boldsymbol{x}_\theta \rangle \, \mathrm{d}\theta \geq 0$.

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

(2-3) If $G(A)$ is an aggregatively strictly monotone game, then $\int_\Theta \langle \mathbf{g}_\theta - \mathbf{h}_\theta, \boldsymbol{x}_\theta - \boldsymbol{y}_\theta \rangle \, \mathrm{d}\theta = 0$ 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 $\boldsymbol{x}_\theta = \boldsymbol{y}_\theta$. $\qquad\square$

4 Approximating a nonatomic aggregative game.

4.1 Atomic approximating sequence.

After discussing the existence and uniqueness properties of NE/VNE and WE/VWE in finite-player and nonatomic aggregative games, we shall study the relationship between these notions. First note that NE (resp. WE) is a particular case of VNE (resp. VWE) when the aggregate constraint set is any subset of $\mathbb{R}^T$ containing $\overline{\mathcal{X}}$. On the one hand, one may naturally expect that, when the number of players grows very large in a finite-player aggregative game, the game gets "close" to a nonatomic aggregative game. On the other hand, since there is a larger literature on algorithms for finite dimensional VI than for infinite dimensional ones, it can be helpful for the computation of WE/VWE to find a NE/VNE that approximates the former with arbitrary precision.

This section shows the following result: Considering a sequence of equilibria of "approximating" finite-player aggregative games $(\mathcal{G}^\nu(A^\nu))_\nu$ of a nonatomic game $G(A)$, where each player in $\mathcal{G}^\nu(A^\nu)$ represents a collection of nonatomic players who are similar in their action sets and cost functions, a sequence of VNE in $(\mathcal{G}^\nu(A^\nu))_\nu$ converges to the VWE of $G(A)$ when this one is (aggregatively) strongly monotone.

In this section, we always consider that Assumptions 1 to 6 hold.

Let us consider the following definition of an approximating sequence:

Definition 4.1. Atomic Approximating Sequence (AAS)
A sequence of finite-player aggregative games $\{\mathcal{G}^\nu(A^\nu) = \left(\mathcal{I}^\nu, \widetilde{\mathcal{X}}^\nu, (f_i^\nu)_i, A^\nu\right) : \nu \in \mathbb{N}^*\}$ with aggregative constraints is an *atomic approximating sequence* (AAS) for the nonatomic aggregative game $G(A) = \left(\Theta, \widetilde{\mathcal{X}}, (f_\theta)_\theta, A\right)$ 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, where $I^\nu \triangleq |\mathcal{I}^\nu|$, 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 $\mu_i^\nu > 0$ while the collection of nonatomic players in Θ_i^ν corresponding to finite player $i \in \mathcal{I}^\nu$ satisfies that, as $\nu \to +\infty$:

i) $\overline{\delta}^\nu \triangleq \max_{i \in \mathcal{I}^\nu} \delta_i^\nu \longrightarrow 0$, where δ_i is the Hausdorff distance between the feasible pure-action sets of nonatomic players in Θ_i^ν and the scaled feasible pure-action set of finite player $i \in \mathcal{I}^\nu$:

$$\delta_i^\nu \triangleq \sup_{\theta \in \Theta_i^\nu} d_H \left(\mathcal{X}_\theta, \tfrac{1}{\mu_i^\nu} \mathcal{X}_i^\nu\right) , \tag{12}$$

and $\operatorname{span} \mathcal{X}_i^\nu = \operatorname{span} \mathcal{X}_\theta$ for all $\theta \in \Theta_i^\nu$.

ii) $\overline{\lambda}^\nu \triangleq \max_{i \in \mathcal{I}^\nu} \lambda_i^\nu \longrightarrow 0$, where λ_i^ν measures the difference between the subdifferential of a finite player's cost function when she takes or not into account the impact of her own action on the aggregate profile:

$$\lambda_i^\nu \triangleq \sup_{(\boldsymbol{x}, \boldsymbol{Y}) \in \mathcal{M}^2} \sup_{\mathbf{g} \in \partial_1 \hat{f}_i^\nu(\mu_i^\nu \boldsymbol{x}, \boldsymbol{Y} - \mu_i^\nu \boldsymbol{x})} d\left(\mathbf{g}, \partial_1 f_i^\nu(\mu_i^\nu \boldsymbol{x}, \boldsymbol{Y})\right) \tag{13}$$

(Recall that $\hat{f}_i^\nu(\boldsymbol{x}_i, \boldsymbol{X}_{-i}) \triangleq f_i^\nu(\boldsymbol{x}_i, \boldsymbol{X})$. The subgradients of $\hat{f}_i^\nu$ with respect to the first variable count the impact of $\boldsymbol{x}_i$ on $\boldsymbol{X}$ while the subgradients of f_i^ν with respect to the first variable do not.)

iii) $\overline{d}^\nu \triangleq \max_{i \in \mathcal{I}^\nu} d_i^\nu \longrightarrow 0$, where d_i^ν measures the Hausdorff distance between the subdifferential of nonatomic players' cost functions and that of the finite players' cost functions:

$$d_i^\nu \triangleq \sup_{\theta \in \Theta_i} \sup_{(\boldsymbol{x}, \boldsymbol{Y}) \in \mathcal{M}^2} d_H \left(\partial_1 f_i^\nu (\mu_i^\nu \boldsymbol{x}, \boldsymbol{Y}), \partial_1 f_\theta(\boldsymbol{x}, \boldsymbol{Y}) \right). \tag{14}$$

iv) $D^\nu \longrightarrow 0$, where $D^\nu \triangleq d_H(A^\nu, A)$ 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, $\operatorname{span} A = \operatorname{span} A^\nu$ for all $\nu \in \mathbb{N}^*$.

Remark 4.1. Roughly speaking, along an AAS, the impact of a finite player's action on the aggregate profile gradually disappears. Besides, the pure-action set of a finite player converges to the pure-action set of a nonatomic player in the subset of Θ that the finite player represents, whereas the subgradients of her cost with respect to her own action also tend to those of a nonatomic player that she represents. Finally, the aggregate-profile constraint sets in the AAS converge to the one in the nonatomic game.

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

Remark 4.2. Without loss of generality, we assume $\operatorname{ri}(A \cap \overline{\mathcal{X}}) \neq \emptyset$ 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 4.3, we will construct an AAS for two fairly general cases of nonatomic games.

In order to compare a pure-action profile in a finite-player game and one in a nonatomic game, we introduce the following linear mappings which define an equivalent nonatomic action profile for a finite-player action profile and vice versa.

First, define $\psi^\nu : \mathcal{M}^{I^\nu} \to L^2([0,1], \mathcal{M})$ for each $\nu \in \mathbb{N}^*$ by

$$\forall \boldsymbol{x}^\nu \in \mathcal{M}^{I^\nu}, \ \psi^\nu(\boldsymbol{x}^\nu) = (\psi_\theta^\nu(\boldsymbol{x}^\nu))_{\theta \in \Theta}, \ \text{where} \ \psi_\theta^\nu(\boldsymbol{x}^\nu) \triangleq \frac{\boldsymbol{x}_i^\nu}{\mu_i^\nu}, \ \forall i \in \mathcal{I}^\nu, \ \forall \theta \in \Theta_i^\nu. \tag{15}$$

The interpretation of ψ^ν is the following. If one considers finite player $i \in \mathcal{I}^\nu$ as a coordinator of the coalition formed by a group of nonatomic players represented by set Θ_i^ν equipped with Lebesgue measure, then $\psi_\theta^\nu(\boldsymbol{x}^\nu)$ dictates the behavior of each of them.

Then, define mapping $\overline{\psi}^\nu : L^2([0,1], \mathcal{M}) \to \mathcal{M}^{\mathcal{I}^\nu}$ for each $\nu \in \mathbb{N}^*$ by

$$\forall \boldsymbol{x} \in L^2([0,1], \mathcal{M}^T), \ \overline{\psi}^\nu(\boldsymbol{x}) = \left(\overline{\psi}_i^\nu(\boldsymbol{x}) \right)_{i \in \mathcal{I}^\nu}, \ \text{where} \ \overline{\psi}_i^\nu(\boldsymbol{x}) = \int_{\Theta_i^\nu} \boldsymbol{x}_\theta \, \mathrm{d}\theta. \tag{16}$$

The interpretation of $\overline{\psi}^\nu$ is that the nonatomic players in Θ_i^ν form a coalition which behaves like a finite player, so that $\overline{\psi}_i^\nu(\boldsymbol{x})$ is just her pure-action.

Finally, let us make the following assumption for this section.

Assumption 7. *There is a strictly positive constant η and a pure-action profile $\bar{\boldsymbol{x}} \in \widetilde{\mathcal{X}}$ such that, for almost all $\theta \in \Theta$, $d(\bar{\boldsymbol{x}}_\theta, \operatorname{rbd} \mathcal{X}_\theta) > \eta$.*

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

4.2 Convergence of equilibrium profiles and aggregate equilibrium profiles.

The following Theorem 4.1 gives the main result of this paper. It shows that a VWE in a strongly monotone nonatomic aggregative game can be approximated by VNE of an AAS, both in the case with and without aggregative constraints.

Recall that, according to Theorem 3.4, 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 4.1 (Convergence of VNE to VWE). *Under Assumptions 1 to 7, let $(\mathcal{G}^\nu(A^\nu))_\nu$ be an AAS of nonatomic aggregative game $G(A)$ with an aggregative constraint. Let $\boldsymbol{x}^*$ be the VWE of $G(A)$, $\hat{\boldsymbol{x}}^\nu \in \widetilde{\mathcal{X}}^\nu(A^\nu)$ a VNE of $\mathcal{G}^\nu(A^\nu)$ for each $\nu \in \mathbb{N}^*$, and $\boldsymbol{X}^*$, $\hat{\boldsymbol{X}}^\nu$ their respective aggregate-action profiles. Then there exists constants $\rho > 0$ and $\bar{\rho} > 0$ such that the following results hold with $K_A \triangleq \frac{M+1}{\min\{\rho, \bar{\rho}\}}$:*

(1) 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 $\max(\overline{\delta}^\nu, D^\nu) < \min\{\overline{\rho}, \rho\}$,

$$\|\hat{\boldsymbol{X}}^\nu - \boldsymbol{X}^*\|^2 \leq \frac{1}{\beta}\left((3B_{\mathbf{f}} + 1)K_A \max(D^\nu, \overline{\delta}^\nu) + (2M + 1)(\overline{d}^\nu + \overline{\lambda}^\nu)\right). \tag{17}$$

(2) If G is strongly monotone with modulus α, then $(\psi^\nu(\hat{\boldsymbol{x}}^\nu))_\nu$ (cf. Eq. (15)), converges to $\boldsymbol{x}^*$ in L^2-norm: for all $\nu \in \mathbb{N}^*$ such that $\max(\overline{\delta}^\nu, D^\nu) < \min\{\overline{\rho}, \rho\}$,

$$\|\psi^\nu(\hat{\boldsymbol{x}}^\nu) - \boldsymbol{x}^*\|_2^2 \leq \frac{1}{\alpha}\left((3B_{\mathbf{f}} + 1)K_A \max(D^\nu, \overline{\delta}^\nu) + (2M + 1)(\overline{d}^\nu + \overline{\lambda}^\nu)\right). \tag{18}$$

If there are no aggregate constraints, replace K_A and D^ν all by 0 in (17) and (18).

Some notions and a series of lemmas are needed for the proof of Theorem 4.1.

Firstly, for $\nu \in \mathbb{N}^*$, define a new norm $\|\cdot\|_\nu$ on $\mathbb{R}^{T\mathcal{I}^\nu}$ by $\|\boldsymbol{x}^\nu\|_\nu \triangleq \left(\sum_{i \in \mathcal{I}^\nu} \frac{\|\boldsymbol{x}_i^\nu\|^2}{\mu_i^\nu}\right)^{\frac{1}{2}} = \langle \boldsymbol{x}^\nu, D_{\mu^\nu}\boldsymbol{x}^\nu\rangle^{\frac{1}{2}}$, where D_{μ^ν} is the symmetric positive definite matrix $\text{diag}(\frac{1}{\mu_1^\nu}I_T, \ldots, \frac{1}{\mu_{N^\nu}^\nu}I_T)$, with I_T being the T-dimensional identity matrix. Denote by d_ν the associated metric, and by diam_ν the diameter of a set with norm $\|\cdot\|_\nu$.

It is easy to see that $\|\boldsymbol{x}^\nu\| \leq \sqrt{\overline{\mu}^\nu}\|\boldsymbol{x}^\nu\|_\nu$, where $\overline{\mu}^\nu \triangleq \max_{i \in \mathcal{I}^\nu} \mu_i^\nu$.

The advantage of using ν-norm for game $\mathcal{G}^\nu$ is that it keeps the norm of pure-action profiles at the same order of those in the nonatomic game, instead of $\sqrt{\overline{\mu}^\nu}$ times smaller. Indeed, ψ^ν is an isomorphism from $\widetilde{\mathcal{X}}^\nu$ (with ν-norm) to $\psi^\nu(\widetilde{\mathcal{X}}^\nu)$ (with L^2-norm):

$$\|\psi^\nu(\boldsymbol{x}^\nu)\|_2^2 = \int_\Theta \|\psi_\theta^\nu(\boldsymbol{x}^\nu)\|^2\,\mathrm{d}\theta = \sum_{i \in \mathcal{I}^\nu} \int_{\Theta_i^\nu} \|\frac{\boldsymbol{x}_i^\nu}{\mu_i^\nu}\|^2\,\mathrm{d}\theta = \sum_{i \in \mathcal{I}^\nu} \frac{\|\boldsymbol{x}_i^\nu\|^2}{\mu_i^\nu} = \|\boldsymbol{x}^\nu\|_\nu^2.$$

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

Notations. Let $\Pi_i^\nu(\cdot)$ denote the projection function onto $\frac{1}{\mu_i^\nu}\mathcal{X}_i^\nu$ 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 projection function onto $\widetilde{\mathcal{X}}^\nu(A^\nu) \subset \mathbb{R}^{I^\nu T}$, and Π the projection function onto $\widetilde{\mathcal{X}}(A) \subset L^2([0,1], \mathbb{R}^T; \mu)$.

The following Lemma 4.1 shows that the players become infinitesimal along an ASS. When scaled by $\frac{1}{\overline{\mu}^\nu}$, they stay at the same order as a nonatomic player.

Lemma 4.1. *Under Assumption 3, for all $\nu \in \mathbb{N}^*$, $\|\boldsymbol{x}^\nu\|_\nu \leq \overline{\delta}^\nu + M$ and $\|\boldsymbol{x}^\nu\| \leq \sqrt{\overline{\mu}^\nu}(\overline{\delta}^\nu + M)$ for all $\boldsymbol{x}^\nu \in \widetilde{\mathcal{X}}^\nu$.*

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

The following lemma shows that the convergence of the pure-action set of a player in finite-player game G^ν to that of her corresponding nonatomic player in G, assumed by Eq. (12), can be described by the convergence of spaces of pure-action profiles. Although $\widetilde{\mathcal{X}}^\nu$ is of finite dimension while $\widetilde{\mathcal{X}}$ is of infinite dimension, the isomorphism ψ^ν is used to here so that the analysis is done on $L^2([0,1], \mathbb{R}^T)$.

Lemma 4.2 (Convergence of $\widetilde{\mathcal{X}}^\nu$ to $\widetilde{\mathcal{X}}$). *Under Assumption 3, for all $\nu \in \mathbb{N}^*$,*
(1) for each $\boldsymbol{x}^\nu \in \widetilde{\mathcal{X}}^\nu$, $d_2(\psi^\nu(\boldsymbol{x}^\nu), \widetilde{\mathcal{X}}) \leq \overline{\delta}^\nu$;
(2) for each $\boldsymbol{x} \in \widetilde{\mathcal{X}}$, $d_\nu(\overline{\psi}^\nu(\boldsymbol{x}), \widetilde{\mathcal{X}}^\nu) \leq \overline{\delta}^\nu$, $d(\overline{\psi}^\nu(\boldsymbol{x}), \widetilde{\mathcal{X}}^\nu) \leq \sqrt{\overline{\mu}^\nu}\overline{\delta}^\nu$;
(3) for each $i \in \mathcal{I}^\nu$ and each $\boldsymbol{x}_i^\nu \in \mathcal{X}_i^\nu$, if $d(\frac{\boldsymbol{x}_i^\nu}{\mu_i^\nu}, rbd\frac{\mathcal{X}_i^\nu}{\mu_i^\nu}) > \delta_i^\nu$, then $\frac{\boldsymbol{x}_i^\nu}{\mu_i^\nu} \in \mathcal{X}_\theta$ for all $\theta \in \Theta_i^\nu$;
(4) for each $i \in \mathcal{I}^\nu$, each $\theta \in \Theta_i^\nu$, and each $\boldsymbol{x}_\theta \in \mathcal{X}_\theta$, if $d(\boldsymbol{x}_\theta, rbd\,\mathcal{X}_\theta) > \delta_i^\nu$, then $\mu_i^\nu\boldsymbol{x}_\theta \in \mathcal{X}_i^\nu$.

Proof. (1) Let $\boldsymbol{x}^\nu \in \widetilde{\mathcal{X}}^\nu$. For each $i \in \mathcal{I}^\nu$ and each $\theta \in \Theta_i^\nu$, define $\boldsymbol{y}_\theta = \Pi_\theta(\psi_\theta^\nu(\boldsymbol{x}^\nu)) \in \mathcal{X}_\theta$, so that $\|\boldsymbol{y}_\theta - \psi_\theta^\nu(\boldsymbol{x}^\nu)\| \leq \delta_i^\nu$. Let us show that $\boldsymbol{y}$ is measurable on each Θ_i^ν, hence measurable on Θ so that $\boldsymbol{y} \in \mathcal{X}$. 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, $\|\psi^\nu(\boldsymbol{x}^\nu) - \boldsymbol{y}\|_2 \leq \overline{\delta}^\nu$, which shows that $d_2(\psi^\nu(\boldsymbol{x}^\nu), \widetilde{\mathcal{X}}) \leq \overline{\delta}^\nu$.

(2) Let $\boldsymbol{x} \in \widetilde{\mathcal{X}}$. 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 $\frac{1}{\mu_i^\nu}\mathcal{X}_i^\nu$ is a convex subset in $\mathbb{R}^T$, $\frac{1}{\mu_i^\nu}\int_{\Theta_i^\nu} \Pi_i^\nu(\boldsymbol{x}_\theta)\,\mathrm{d}\theta \in \frac{1}{\mu_i^\nu}\mathcal{X}_i^\nu$. Define $\boldsymbol{y} \in \widetilde{\mathcal{X}}^\nu$ by $\boldsymbol{y}_i \triangleq \int_{\Theta_i^\nu} \Pi_i^\nu(\boldsymbol{x}_\theta)\,\mathrm{d}\theta \in \mathcal{X}_i^\nu$ for $i \in \mathcal{I}^\nu$. Then,

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

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

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

(3) 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 $d(\frac{\boldsymbol{x}_i^\nu}{\mu_i^\nu}, \mathrm{rbd}\,\frac{\mathcal{X}_i^\nu}{\mu_i^\nu}) > \eta$ for some $\eta > \delta_i^\nu$.

Assume that $\frac{\boldsymbol{x}_i^\nu}{\mu_i^\nu} \notin \mathcal{X}_\theta$ i.e. $\|\frac{\boldsymbol{x}_i^\nu}{\mu_i^\nu} - \Pi_{\mathcal{X}_\theta}(\frac{\boldsymbol{x}_i^\nu}{\mu_i^\nu})\| > 0$. Let $\boldsymbol{y}_i^\nu = \frac{\boldsymbol{x}_i^\nu}{\mu_i^\nu} + \eta\frac{\frac{\boldsymbol{x}_i^\nu}{\mu_i^\nu} - \Pi_{\mathcal{X}_\theta}(\frac{\boldsymbol{x}_i^\nu}{\mu_i^\nu})}{\|\frac{\boldsymbol{x}_i^\nu}{\mu_i^\nu} - \Pi_{\mathcal{X}_\theta}(\frac{\boldsymbol{x}_i^\nu}{\mu_i^\nu})\|} \in \frac{\mathcal{X}_i^\nu}{\mu_i^\nu}$. Since $\mathcal{X}_i^\nu$ is convex,

$d(\boldsymbol{y}_i^\nu, \mathcal{X}_\theta) = \|\frac{\boldsymbol{x}_i^\nu}{\mu_i^\nu} - \Pi_{\mathcal{X}_\theta}(\frac{\boldsymbol{x}_i^\nu}{\mu_i^\nu})\| + \|\boldsymbol{y}_i^\nu - \frac{\boldsymbol{x}_i^\nu}{\mu_i^\nu}\| > \eta > \overline{\delta}^\nu$, which contradicts the fact that $d(\mathcal{X}_\theta, \frac{\boldsymbol{x}_i^\nu}{\mu_i^\nu}) \leq \delta_i^\nu$. Hence $\frac{\boldsymbol{x}_i^\nu}{\mu_i^\nu} \in \mathcal{X}_\theta$.

(4) The proof is similar to that of (3). $\qquad\square$

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

Lemma 4.3. *Under Assumption 3, for $\nu \in \mathbb{N}^*$,*
(1) $d_H(\overline{\mathcal{X}}^\nu, \overline{\mathcal{X}}) \leq \overline{\delta}^\nu$;
(2) 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 $\boldsymbol{X} \in \overline{\mathcal{X}}$;
(3) 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 $\boldsymbol{X}^\nu \in A$;
(4) 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. (1) Define mapping S^ν from $\mathcal{M}^{I^\nu}$ to $\mathbb{R}^T$ by $S^\nu(\boldsymbol{x}^\nu) = \sum_{i \in \mathcal{I}^\nu} \boldsymbol{x}_i^\nu$. Then S^ν is nonexpansive hence it is continuous. To see this, let $\boldsymbol{z}^\nu$ and $\boldsymbol{w}^\nu$ be in $\mathcal{M}^{I^\nu}$, then by the Cauchy-Schwarz inequality,

$$\|\sum_{i \in \mathcal{I}^\nu} \boldsymbol{z}_i^\nu - \sum_{i \in \mathcal{I}^\nu} \boldsymbol{w}_i^\nu\|^2 = \|\sum_{i \in \mathcal{I}^\nu} \frac{\boldsymbol{z}_i^\nu - \boldsymbol{w}_i^\nu}{\sqrt{\mu_i^\nu}}\sqrt{\mu_i^\nu}\|^2 \leq \sum_{i \in \mathcal{I}^\nu} \frac{\|\boldsymbol{z}_i^\nu - \boldsymbol{w}_i^\nu\|^2}{\mu_i^\nu} \cdot \sum_{i \in \mathcal{I}^\nu} \mu_i^\nu = \|\boldsymbol{z}^\nu - \boldsymbol{w}^\nu\|_\nu^2.$$

Now fix $\boldsymbol{x} \in \widetilde{\mathcal{X}}$. Consider $\boldsymbol{y} \in \widetilde{\mathcal{X}}^\nu$ such that $\|\overline{\psi}^\nu(\boldsymbol{x}) - \boldsymbol{y}\|_\nu \leq \overline{\delta}^\nu$ (cf. Lemma 4.2). Then $\|\sum_{i \in \mathcal{I}^\nu} \overline{\psi}_i^\nu(\boldsymbol{x}) - \sum_{i \in \mathcal{I}^\nu} \boldsymbol{y}_i\|^2 \leq \|\overline{\psi}^\nu(\boldsymbol{x}) - \boldsymbol{y}\|_\nu^2 \leq (\overline{\delta}^\nu)^2$. Hence $d(\int \boldsymbol{x}, \overline{\mathcal{X}}^\nu) \leq \|\sum_{i \in \mathcal{I}^\nu} \overline{\psi}_i^\nu(\boldsymbol{x}) - \sum_{i \in \mathcal{I}^\nu} \boldsymbol{y}_i\| \leq \overline{\delta}^\nu$.

On the other hand, let $\boldsymbol{x}^\nu \in \widetilde{\mathcal{X}}^\nu$, thus $\boldsymbol{X}^\nu \triangleq \sum_{i \in \mathcal{I}^\nu} \boldsymbol{x}_i^\nu \in \overline{\mathcal{X}}^\nu$. For each $i \in \mathcal{I}^\nu$ and each $\theta \in \Theta_i^\nu$, define $\boldsymbol{y}_\theta = \Pi_\theta\left(\frac{1}{\mu_i^\nu}\boldsymbol{x}_i^\nu\right) \in \mathcal{X}_\theta$, so that $\|\frac{1}{\mu_i^\nu}\boldsymbol{x}_i^\nu - \boldsymbol{y}_\theta\| \leq \delta_i^\nu$. Then, $\|\sum_{i \in \mathcal{I}^\nu} \boldsymbol{x}_i^\nu - \int \boldsymbol{y}\| \leq \sum_{i \in \mathcal{I}^\nu} \|\int_{\Theta_i^\nu} \frac{1}{\mu_i^\nu}\boldsymbol{x}_i^\nu - \boldsymbol{y}_\theta\,\mathrm{d}\theta\| \leq \sum_{i \in \mathcal{I}^\nu} \int_{\Theta_i^\nu} \|\frac{1}{\mu_i^\nu}\boldsymbol{x}_i^\nu - \boldsymbol{y}_\theta\|\,\mathrm{d}\theta \leq \sum_{i \in \mathcal{I}^\nu} \mu_i^\nu \delta_i^\nu \leq \overline{\delta}^\nu$, which shows that $d\left(\boldsymbol{X}^\nu, \overline{\mathcal{X}}\right) \leq \overline{\delta}^\nu$ for all $\boldsymbol{X}^\nu \in \overline{\mathcal{X}}^\nu$.

(2-3) The proof is similar to that for Lemma 4.2.(3).

(4) These are corollaries of (2) and (3). $\qquad\square$

Lemma 4.4. *Under Assumptions 3 and 7, there is a strictly positive constant ρ and a nonatomic pure-action profile $\boldsymbol{z} \in \widetilde{\mathcal{X}}$ 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 pure-action profile in Assumption 7 and an arbitrary $\boldsymbol{y} \in \widetilde{\mathcal{X}}(A)$ 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 $\boldsymbol{z} \in \widetilde{\mathcal{X}}$ by $\boldsymbol{z} = \boldsymbol{y} - t(\boldsymbol{y} - \bar{\boldsymbol{x}})$.

Firstly, $\|\int \boldsymbol{y} - \int \boldsymbol{z}\| = t\|\int \boldsymbol{y} - \int \bar{\boldsymbol{x}}\| \le t2M \le \frac{2}{3}d(\int \boldsymbol{y}, \mathrm{rbd}\,(\overline{\mathcal{X}} \cap A))$, 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 4.4. Define $\bar{\rho} = \frac{1}{3}d(\boldsymbol{Z}, \mathrm{rbd}\,\overline{\mathcal{X}} \cap A) > 0$. Then, for $\nu \in \mathbb{N}^*$ such that $\max(\bar{\delta}^\nu, D^\nu) < 3\bar{\rho}$, $d(\boldsymbol{Z}, \mathrm{rbd}\,(\overline{\mathcal{X}}^\nu \cap A^\nu)) \ge 3\bar{\rho} - \max(\bar{\delta}^\nu, D^\nu)$.
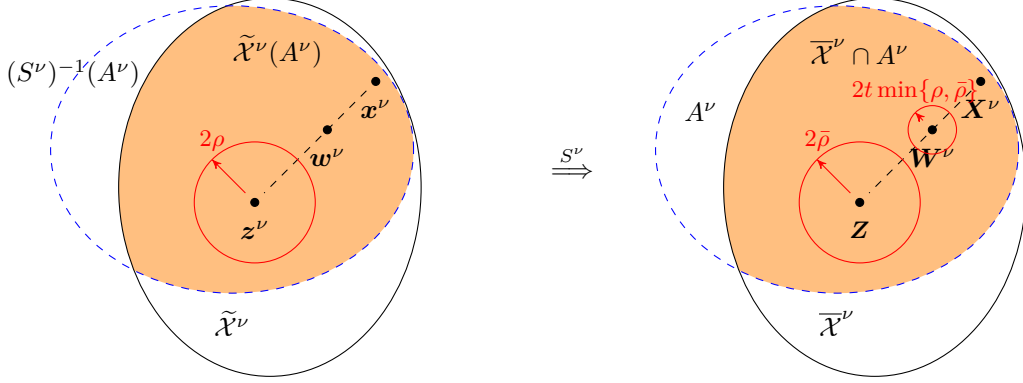

Figure 1: Illustration of the mapping S^ν between $\widetilde{\mathcal{X}}^\nu(A^\nu)$ and $\overline{\mathcal{X}} \cap A$ used in Lemma 4.5

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

Lemma 4.5 (Convergence of $\widetilde{\mathcal{X}}^\nu(A^\nu)$ to $\widetilde{\mathcal{X}}(A)$). *Under Assumptions 3 and 7, let $K_A = \frac{M+1}{\min\{\rho, \bar{\rho}\}}$. Then, for all $\nu \in \mathbb{N}^*$ such that $\max(\bar{\delta}^\nu, D^\nu) < \min\{\bar{\rho}, \rho\}$,*
(1) for each $\boldsymbol{x}^\nu \in \widetilde{\mathcal{X}}^\nu(A^\nu)$, $d_2(\psi^\nu(\boldsymbol{x}^\nu), \widetilde{\mathcal{X}}(A)) \le 2K_A \max(D^\nu, \bar{\delta}^\nu)$;
(2) for each $\boldsymbol{x} \in \widetilde{\mathcal{X}}(A)$, $d_\nu(\overline{\psi}^\nu(\boldsymbol{x}), \widetilde{\mathcal{X}}^\nu(A^\nu)) \le K_A \max(D^\nu, \bar{\delta}^\nu)$ and $d(\overline{\psi}^\nu(\boldsymbol{x}), \widetilde{\mathcal{X}}^\nu(A^\nu)) \le \sqrt{\overline{\mu}^\nu} K_A \max(D^\nu, \bar{\delta}^\nu)$.

Proof. (1) Consider $\boldsymbol{x}^\nu \in \widetilde{\mathcal{X}}^\nu(A^\nu)$ and $\boldsymbol{X}^\nu = \sum_i \boldsymbol{x}_i^\nu$, i.e. $S^\nu(\boldsymbol{x}^\nu)$. Consider $\boldsymbol{z}$ in Lemma 4.4 and $\boldsymbol{Z} = \int \boldsymbol{z}$. Let $\boldsymbol{z}^\nu \triangleq \overline{\psi}^\nu(\boldsymbol{z})$. Since for each θ, $d(\boldsymbol{z}_\theta, \mathrm{rbd}\,\mathcal{X}_\theta) > 3\rho > \bar{\delta}^\nu$, $\boldsymbol{z}^\nu \in \widetilde{\mathcal{X}}^\nu$ according to Lemma 4.2.(4).

Also, $d(\frac{\boldsymbol{z}_i^\nu}{\mu_i^\nu}, \mathrm{rbd}\,\frac{\mathcal{X}_i^\nu}{\mu_i^\nu}) \ge 2\rho$. Indeed, if $\frac{\boldsymbol{y}_i}{\mu_i^\nu} \in N_{2\rho}(\frac{\boldsymbol{z}_i^\nu}{\mu_i^\nu})$, then let $\boldsymbol{y}_\theta \triangleq \boldsymbol{z}_\theta + (\frac{\boldsymbol{y}_i}{\mu_i^\nu} - \frac{\boldsymbol{z}_i^\nu}{\mu_i^\nu})$ for $\theta \in \Theta_i^\nu$. Then, $\boldsymbol{y}_\theta \in N_{2\rho}(\boldsymbol{z}_\theta) \subset \mathcal{X}_\theta$ as $\mathrm{span}\,\mathcal{X}_\theta = \mathrm{span}\,\mathcal{X}_i^\nu$, and $d(\boldsymbol{y}_\theta, \mathrm{rbd}\,\mathcal{X}_\theta) \ge d(\boldsymbol{z}_\theta, \mathrm{rbd}\,\mathcal{X}_\theta) - \|\boldsymbol{z}_\theta - \boldsymbol{y}_\theta\| > 3\rho - 2\rho = \rho$. Hence, from 4.2.(4), one has $\frac{\boldsymbol{y}_i}{\mu_i^\nu} \in \frac{\mathcal{X}_i^\nu}{\mu_i^\nu}$.

Besides, since $\max(\bar{\delta}^\nu, D^\nu) < \bar{\rho}$, one has $d(\boldsymbol{Z}, \mathrm{rbd}\,(\overline{\mathcal{X}}^\nu \cap A^\nu)) \ge 2\bar{\rho}$.

Define $\boldsymbol{w}^\nu \triangleq \boldsymbol{x}^\nu + t(\boldsymbol{z}^\nu - \boldsymbol{x}^\nu)$ with $t \triangleq \frac{\max(D^\nu, \bar{\delta}^\nu)}{\min\{\rho, \bar{\rho}\}} < 1$. Let $\boldsymbol{W}^\nu = S^\nu(\boldsymbol{w}^\nu)$. Then, $\|\boldsymbol{w}^\nu - \boldsymbol{x}^\nu\|_\nu = \max(D^\nu, \bar{\delta}^\nu) \frac{\|\boldsymbol{z}^\nu - \boldsymbol{x}^\nu\|_\nu}{\min\{\rho, \bar{\rho}\}} \le \max(D^\nu, \bar{\delta}^\nu) \frac{2(M+1)}{\min\{\rho, \bar{\rho}\}}$.

The linear mapping S^ν maps the segment linking $\boldsymbol{x}^\nu$ and $\boldsymbol{z}^\nu$ in $\widetilde{\mathcal{X}}^\nu(A^\nu)$ to a segment linking $\boldsymbol{X}^\nu$ and $\boldsymbol{Z}$ in $\overline{\mathcal{X}}^\nu \cap A^\nu$. Hence, by the definition of $\boldsymbol{w}^\nu$, $N_{2t\min\{\rho, \bar{\rho}\}}(\boldsymbol{W}^\nu) \cap \mathrm{span}\,(\overline{\mathcal{X}}^\nu \cap A^\nu) \subset \overline{\mathcal{X}}^\nu \cap A^\nu$, because each point in $N_{2t\min\{\rho, \bar{\rho}\}}(\boldsymbol{W}^\nu) \cap \mathrm{span}\,(\overline{\mathcal{X}}^\nu \cap A^\nu)$ is on the segment linking $\boldsymbol{X}^\nu$ and some point in $N_{2\min\{\rho, \bar{\rho}\}}(\boldsymbol{Z}) \cap \mathrm{span}\,(\overline{\mathcal{X}}^\nu \cap A^\nu) \subset \overline{\mathcal{X}}^\nu \cap A^\nu$. In particular, $2t\min\{\rho, \bar{\rho}\} = 2\max(D^\nu, \bar{\delta}^\nu)$ means that $d(\boldsymbol{W}^\nu, \mathrm{rbd}\,(\overline{\mathcal{X}}^\nu \cap A^\nu)) \ge 2\max(D^\nu, \bar{\delta}^\nu) > \max(D^\nu, \bar{\delta}^\nu)$. Consequently, $\boldsymbol{W}^\nu \in \overline{\mathcal{X}} \cap A$ according to Lemma 4.3(4).

Then, in the same manner, one has $d(\frac{\boldsymbol{w}_i^\nu}{\mu_i^\nu}, \mathrm{rbd}\,\frac{\mathcal{X}_i^\nu}{\mu_i^\nu}) > 2t\rho > \bar{\delta}_i^\nu$. Thus, Lemma 4.3(3) implies that $\frac{\boldsymbol{w}^\nu}{\mu_i^\nu} \in \mathcal{X}_\theta$ for each $\theta \in \Theta_i^\nu$, hence $\psi^\nu(\boldsymbol{w}^\nu) \in \widetilde{\mathcal{X}}^\nu$ and $\psi^\nu(\boldsymbol{w}^\nu) \in \widetilde{\mathcal{X}}^\nu(A)$.

Finally, $\|\boldsymbol{x}^\nu - \boldsymbol{w}^\nu\|_\nu \le \max(D^\nu, \bar{\delta}^\nu) \frac{2(M+1)}{\min\{\rho, \bar{\rho}\}}$. Hence, $d_2(\psi^\nu(\boldsymbol{x}^\nu), \widetilde{\mathcal{X}}(A)) \le 2K_A \max(D^\nu, \bar{\delta}^\nu)$.

The proof for (2) is similar and omitted.

$\quad\square$

Remark 4.3 (Difference between unilateral projections of actions and collective projection of the action profile). Lemma 4.5 shows that $d_2(\psi^\nu(\boldsymbol{x}^\nu), \Pi(\psi^\nu(\boldsymbol{x}^\nu))) \le 2K_A \max(D^\nu, \bar{\delta}^\nu)$ and $d(\overline{\psi}^\nu(\boldsymbol{x}), \Pi^\nu(\overline{\psi}^\nu(\boldsymbol{x}))) \le \sqrt{\overline{\mu}^\nu} K_A \max(D^\nu, \bar{\delta}^\nu)$. Recall that Π and Π^ν stand for the projection functions onto $\widetilde{\mathcal{X}}(A)$ and $\widetilde{\mathcal{X}}^\nu(A^\nu)$.
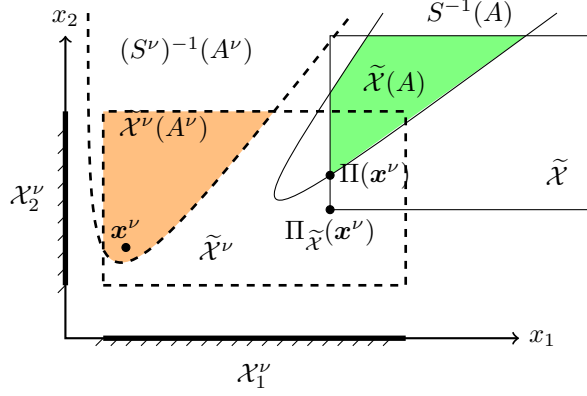
Figure 2: Difference between projections on $\widetilde{\mathcal{X}}$ and on $\widetilde{\mathcal{X}}(A)$. (Since it is impossible to draw the graph of a L^2 pure-action-profile space with a continuum of players, we illustrate the idea with two players.)

Lemma 4.5 is of first importance for our proof of Theorem 4.1. Without this lemma, we only have the convergence of individual, i.e. unilateral pure-action spaces in the AAS to the unilateral pure-action spaces in the nonatomic game, as shown in Lemma 4.2. Without coupling constraints, this should be sufficient in the proof of the convergence of NE. However, in the presence of coupling aggregative constraints, this convergence of unilateral pure-action spaces is not enough. Given a profile in $\widetilde{\mathcal{X}}^\nu(A^\nu)$, unilateral projection of each atomic player's pure-action onto the corresponding nonatomic players' pure-action spaces, i.e. from $\mathcal{X}_i^\nu$ to $\mathcal{X}_\theta$, cannot guarantee that the resulting profile of pure-actions is in $\widetilde{\mathcal{X}}(A)$, and vice versa. Lemma 4.5 shows that, for each pure-action profile $\boldsymbol{x}^\nu \in \widetilde{\mathcal{X}}^\nu(A^\nu)$, its projection on the space of finite-player pure-action profiles in $\widetilde{\mathcal{X}}(A)$ very close to $\boldsymbol{x}^\nu$, and vice versa.

We are finally ready to prove Theorem 4.1.

Proof of Theorem 4.1. Fix $\nu \in \mathbb{N}^*$, define $\hat{\boldsymbol{y}}^\nu = \psi(\hat{\boldsymbol{x}}^\nu) \in L^2([0,1],\mathbb{R}^T)$ and $\hat{\boldsymbol{z}}^\nu \triangleq \Pi(\hat{\boldsymbol{y}}^\nu)$. Then $\hat{\boldsymbol{z}}^\nu \in \widetilde{\mathcal{X}}(A)$ is a pure-action profile in nonatomic game $G(A)$. By the definition of VWE (cf. Eq. (11)), there is $\boldsymbol{g} \in H(\boldsymbol{x}^*)$ such that $\int_0^1 \langle \boldsymbol{g}_\theta(\boldsymbol{x}_\theta^*, \boldsymbol{X}^*), \ \boldsymbol{x}_\theta^* - \hat{\boldsymbol{z}}_\theta^\nu \rangle \, \mathrm{d}\theta \leq 0$.

Secondly, by the definition of VNE (cf. Eq. (3)), there is $\mathbf{h}_i \in \partial_1 \hat{f}_i^\nu(\hat{\boldsymbol{x}}_i^\nu, \hat{\boldsymbol{X}}_{-i}^\nu)$ for each $i \in \mathcal{I}^\nu$ such that $\sum_{i \in \mathcal{I}^\nu} \langle \mathbf{h}_i(\hat{\boldsymbol{x}}_i^\nu, \hat{\boldsymbol{X}}_{-i}^\nu), \hat{\boldsymbol{x}}_i^\nu - \boldsymbol{z}_i^\nu \rangle \leq 0$ for all $\boldsymbol{z}_i^\nu \in \mathcal{X}_i^\nu$. For all $i \in \mathcal{I}^\nu$ and $\theta \in \Theta_i^\nu$, by the definition of λ_i^ν and d_i^ν (cf. Eqs. (13) and (14)), there exists $\boldsymbol{r} \in H(\hat{\boldsymbol{y}}^\nu)$ with $\boldsymbol{r}_\theta \in \partial_1 f_\theta(\hat{\boldsymbol{y}}_\theta^\nu, \hat{\boldsymbol{Y}}^\nu) = \partial_1 f_\theta(\frac{\hat{\boldsymbol{x}}_\theta^\nu}{\mu_i^\nu}, \hat{\boldsymbol{X}}^\nu)$ such that $\|\boldsymbol{r}_\theta(\hat{\boldsymbol{y}}_\theta^\nu, \hat{\boldsymbol{Y}}^\nu) - \mathbf{h}_i(\hat{\boldsymbol{x}}_\theta^\nu, \hat{\boldsymbol{X}}_{-i}^\nu)\| \leq \lambda_i^\nu + d_i^\nu$.

Thirdly, $\|\hat{\boldsymbol{y}}^\nu - \hat{\boldsymbol{z}}^\nu\|_2 \leq 2K_A \max(D^\nu, \overline{\delta}^\nu)$ by Lemma 4.5.

With these two results, while noticing that $\hat{\boldsymbol{y}}_\theta^\nu \leq M + \overline{\delta}^\nu$ for all θ by Lemma 4.1 and that $\hat{\boldsymbol{X}}^\nu = \hat{\boldsymbol{Y}}^\nu$, one has:

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

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

$$+ \sum_{i \in \mathcal{I}^\nu} \int_{\Theta_i^\nu} \langle \boldsymbol{r}_\theta(\hat{\boldsymbol{y}}_\theta^\nu, \hat{\boldsymbol{Y}}^\nu) - \mathbf{h}_i(\hat{\boldsymbol{x}}_i^\nu, \hat{\boldsymbol{X}}_{-i}^\nu), \ \hat{\boldsymbol{y}}_\theta^\nu - \boldsymbol{x}_\theta^* \rangle \, \mathrm{d}\theta + \sum_{i \in \mathcal{I}^\nu} \int_{\Theta_i^\nu} \langle \mathbf{h}_i(\hat{\boldsymbol{x}}_i^\nu, \hat{\boldsymbol{X}}_{-i}^\nu), \ \hat{\boldsymbol{y}}_\theta^\nu - \boldsymbol{x}_\theta^* \rangle \, \mathrm{d}\theta$$

$$\leq 0 + \int_\Theta \|\mathbf{g}_\theta(\boldsymbol{x}_\theta^*, \boldsymbol{X}^*)\| \, \|\hat{\boldsymbol{z}}_\theta^\nu - \hat{\boldsymbol{y}}_\theta^\nu\| \, \mathrm{d}\theta + \sum_{i \in \mathcal{I}^\nu} \int_{\Theta_i^\nu} \|\boldsymbol{r}_\theta(\hat{\boldsymbol{y}}_\theta^\nu, \hat{\boldsymbol{Y}}^\nu) - \mathbf{h}_i(\hat{\boldsymbol{x}}_\theta^\nu, \hat{\boldsymbol{X}}_{-i}^\nu)\| \, \|\hat{\boldsymbol{y}}_\theta^\nu - \boldsymbol{x}_\theta^*\| \, \mathrm{d}\theta + J^\nu$$

$$\leq 2B_{\mathbf{f}} K_A \max(D^\nu, \overline{\delta}^\nu) + (2M + \overline{\delta}^\nu)(\overline{d}^\nu + \overline{\lambda}^\nu) + J^\nu \tag{19}$$

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

Next, for the VWE $\boldsymbol{x}^* \in \widetilde{\mathcal{X}}(A)$, let $\boldsymbol{y}^{*\nu} = \overline{\psi}(\boldsymbol{x}^*) \in \mathcal{M}^{\mathcal{I}^\nu}$ and $\boldsymbol{z}^{*\nu} \triangleq \Pi^\nu(\boldsymbol{y}^{*\nu}) \in \widetilde{\mathcal{X}}^\nu(A^\nu)$:

$$
\begin{aligned}
J^\nu &= \sum_{i \in \mathcal{I}^\nu} \left\langle \mathbf{h}_i(\hat{\boldsymbol{x}}_i^\nu, \hat{\boldsymbol{X}}_{-i}^\nu), \ \hat{\boldsymbol{x}}_i^\nu - \boldsymbol{y}^*{}_i^\nu \right\rangle \\
&= \sum_{i \in \mathcal{I}^\nu} \left\langle \mathbf{h}_i(\hat{\boldsymbol{x}}_i^\nu, \hat{\boldsymbol{X}}_{-i}^\nu), \ \hat{\boldsymbol{x}}_i^\nu - \boldsymbol{z}^*{}_i^\nu \right\rangle + \sum_{i \in \mathcal{I}^\nu} \left\langle \mathbf{h}_i(\hat{\boldsymbol{x}}_i^\nu, \hat{\boldsymbol{X}}_{-i}^\nu), \ \boldsymbol{z}^*{}_i^\nu - \boldsymbol{y}^*{}_i^\nu \right\rangle \\
&\leq 0 + (B_{\mathbf{f}} + \overline{\lambda}^\nu) \|\boldsymbol{z}^{*\nu} - \boldsymbol{y}^{*\nu}\| \leq (B_{\mathbf{f}} + \overline{\lambda}^\nu) \sqrt{\overline{\mu}^\nu} K_A \max(D^\nu, \overline{\delta}^\nu),
\end{aligned}
\tag{20}
$$

because of the definition of VNE $\hat{\boldsymbol{x}}^\nu$, the definition of $\overline{\lambda}^\nu$ and Lemma 4.5.(i).

Let us summarize by combining (19) and (20):

$$
\begin{aligned}
\int_\Theta \left\langle \mathbf{g}_\theta(\boldsymbol{x}_\theta^*, \boldsymbol{X}^*) - \boldsymbol{r}_\theta(\hat{\boldsymbol{y}}_\theta^\nu, \hat{\boldsymbol{Y}}^\nu), \boldsymbol{x}_\theta^* - \hat{\boldsymbol{y}}_\theta^\nu \right\rangle \mathrm{d}\theta \\
\leq (3B_{\mathbf{f}} + 1) K_A \max(D^\nu, \overline{\delta}^\nu) + (2M + 1)(\overline{d}^\nu + \overline{\lambda}^\nu).
\end{aligned}
\tag{21}
$$

Hence, if G is strongly monotone with modulus α, then $\alpha \|\hat{\boldsymbol{y}}^\nu - \boldsymbol{x}^*\|_2^2 \leq (2B_{\mathbf{f}} + 1) K_A (D^\nu + \overline{\delta}^\nu) + (2M + 1)(\overline{d}^\nu + \overline{\lambda}^\nu)$. If G is aggregatively strongly monotone with modulus β, then $\beta \|\hat{\boldsymbol{X}}^\nu - \boldsymbol{X}^*\|^2 \leq (3B_{\mathbf{f}} + 1) K_A \max(D^\nu, \overline{\delta}^\nu) + (2M + 1)(\overline{d}^\nu + \overline{\lambda}^\nu)$. $\qquad \square$

Remark 4.4. The strong monotonicity of the nonatomic game G, either with respect to pure-action profile or with respect to aggregate-action profile, is essential in this result. Strict monotonicity is not enough, in contrast to finite-player games (cf. [50]). Indeed, since $L^2([0,1], \mathcal{M}^T)$ is only weakly compact, one cannot ensure that $\int_\Theta \langle \mathbf{g}_\theta(\boldsymbol{x}_\theta^*, \boldsymbol{X}^*) - \boldsymbol{r}_\theta(\hat{\boldsymbol{y}}_\theta^\nu, \hat{\boldsymbol{Y}}^\nu), \boldsymbol{x}_\theta^* - \hat{\boldsymbol{y}}_\theta^\nu \rangle \mathrm{d}\theta$ tends to $\int_\Theta \langle \mathbf{g}_\theta(\boldsymbol{x}_\theta^*, \boldsymbol{X}^*) - \boldsymbol{r}_\theta(\hat{\boldsymbol{z}}_\theta, \hat{\boldsymbol{Z}}), \boldsymbol{x}_\theta^* - \hat{\boldsymbol{z}}_\theta \rangle \mathrm{d}\theta$ in (21), where $\hat{\boldsymbol{z}}$ is an accumulation point of $(\hat{\boldsymbol{y}}^\nu)_\nu$ in the weak topology.

4.3 Construction of an atomic approximating sequence.

As seen in our previous results, a nonatomic player θ is characterized by two elements: her pure-action set $\mathcal{X}_\theta$, and her subdifferential correspondence $\partial_1 f_\theta$ defined from $\mathcal{M}^2$ to $\mathbb{R}^T$: $(\boldsymbol{x}, \boldsymbol{Y}) \mapsto \partial_1 f_\theta(\boldsymbol{x}, \boldsymbol{Y})$.

Note that it is the subdifferential of the cost function $\partial_1 f_\theta$, 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 an AAS for a given nonatomic game G in two particular cases: (1) the player characteristic profile $\theta \mapsto (\mathcal{X}_\theta, \partial_1 f_\theta)$ is piecewise continuous (cf. Definition 4.2) 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.

Case 1: Piecewise Continuous Characteristics – Uniform Splitting

Definition 4.2 (Continuity of nonatomic player characteristic profile). The player characteristic profile $\theta \mapsto (\mathcal{X}_\theta, \partial_1 f_\theta)$ in nonatomic game G is *continuous* at $\theta \in \Theta$ if, for all $\varepsilon > 0$, there exists $\eta > 0$ such that: for each $\theta' \in \Theta$

$$
|\theta - \theta'| \leq \eta \ \Rightarrow \begin{cases} d_H(\mathcal{X}_\theta, \mathcal{X}_{\theta'}) \leq \varepsilon \\ \sup_{(\boldsymbol{x}, \boldsymbol{Y}) \in \mathcal{M} \times \mathcal{M}} d_H\left(\partial_1 f_\theta(\boldsymbol{x}, \boldsymbol{Y}), \partial_1 f_{\theta'}(\boldsymbol{x}, \boldsymbol{Y})\right) \leq \varepsilon. \end{cases}
\tag{22}
$$

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 $\theta \mapsto (\mathcal{X}_\theta, \partial_1 f_\theta)$ of nonatomic 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 $\bar{v}_i^\nu = \frac{v_{i-1}^\nu + v_i^\nu}{2}$.

Assumption 8. *For each $\theta \in \Theta$ and each $\boldsymbol{x} \in \mathcal{M}$, $\mathcal{M}' \ni \boldsymbol{Y} \mapsto f_\theta(\boldsymbol{x}, \boldsymbol{Y})$ is convex. Denote the subdifferential of $f_\theta(\boldsymbol{x}, \cdot)$ at $\boldsymbol{Y}$ by $\partial_2 f_\theta(\boldsymbol{x}, \boldsymbol{Y})$. There is $B_{\mathbf{g}} > 0$ such that $\|\mathbf{g}\| \leq B_{\mathbf{g}}$ for all $\mathbf{g} \in \partial_2 f_\theta(\cdot, \cdot)$ on $\mathcal{M}^2$ for all θ.*

Proposition 4.1. *Under Assumptions 3 to 6 and 8, for $\nu \in \mathbb{N}^*$, consider the finite-player game $\mathcal{G}^\nu(A^\nu)$ with an aggregative constraint defined with $A^\nu \triangleq A$, player set $\mathcal{I}^\nu \triangleq \{1 \dots I^\nu\}$, where for each finite player $i \in \mathcal{I}^\nu$:*

$$\mathcal{X}_i^\nu \triangleq \mu_i^\nu \mathcal{X}_{\bar{v}_i^\nu} \quad and \quad f_i^\nu(\boldsymbol{x}, \boldsymbol{Y}) \triangleq \mu_i^\nu f_{\bar{v}_i^\nu}\left(\tfrac{1}{\mu_i^\nu}\boldsymbol{x}, \boldsymbol{Y}\right), \ \forall (\boldsymbol{x}, \boldsymbol{Y}) \in \mu_i^\nu \mathcal{M} \times \mathcal{M}.$$

Then $\left(\mathcal{G}^\nu(A)\right)_\nu = \left(\mathcal{I}^\nu, \widetilde{\mathcal{X}}^\nu, A, (f_i^\nu)_{i \in \mathcal{I}^\nu}\right)_\nu$ is an AAS of nonatomic game $G(A)$.

Proof. First notice that $\bar{\mu}^\nu \leq \frac{1}{\nu}$ hence it tends to 0.

Let us show the four points required by Definition 4.1 as follows.

i) Given an arbitrary $\varepsilon > 0$, there is a common modulus of uniform continuity η such that Eq. (22) is true 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 $d_H\left(\mathcal{X}_\theta, \frac{1}{\mu_i^\nu}\mathcal{X}_i^\nu\right) = d_H(\mathcal{X}_\theta, \mathcal{X}_{\bar{v}_i^\nu}) < \varepsilon$.

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

$$d_H \left(\partial_1 f_i^\nu\left(\mu_i^\nu \boldsymbol{x}, \boldsymbol{Y}\right) \ , \ \partial_1 f_\theta(\boldsymbol{x}, \boldsymbol{Y})\right) = d_H \left(\frac{\mu_i^\nu}{\mu_i^\nu}\partial_1 f_{\bar{v}_i^\nu}\left(\frac{1}{\mu_i^\nu}\mu_i^\nu \boldsymbol{x}, \boldsymbol{Y}\right) \ , \ \partial_1 f_\theta(\boldsymbol{x}, \boldsymbol{Y})\right) < \varepsilon.$$

To be rigorous, we need $\operatorname{span} \mathcal{X}_\theta$ to be the same for all $\theta \in \Theta_i^\nu$. For this, we can further divide Θ_i^ν into a finite number of, say n, groups so that the nonatomic players in each group have the same span $\mathcal{X}_\theta$. This is possible because $\mathcal{X}_\theta$ are all in $\mathbb{R}^T$, a finite dimensional space.

iii) Consider $\partial_1 \hat{f}_i^\nu(\boldsymbol{x}, \boldsymbol{Y} - \boldsymbol{x})$, the subdifferential of $\hat{f}_i^\nu(\cdot, \boldsymbol{Y}')$ at $\boldsymbol{x}$, given $\boldsymbol{Y}' = \boldsymbol{Y} - \boldsymbol{x}$. By definition, $\hat{f}_i^\nu(\cdot, \boldsymbol{Y}') \triangleq f_i^\nu(\cdot, \boldsymbol{Y}' + \cdot)$. Consider $\hat{f}_i^\nu(\cdot, \boldsymbol{Y}')$ as a composition function on $\mathcal{M}'$ defined by $f_i^\nu \circ L$, where L is the following affine mapping from $\mu_i^\nu \mathcal{M}'$ to $\mathbb{R}^{2T}$:

$$L(\boldsymbol{x}) = \begin{pmatrix} I_T \\ I_T \end{pmatrix} \boldsymbol{x} + \begin{pmatrix} \mathbf{0}_T \\ \boldsymbol{Y}' \end{pmatrix} \ ,$$

where I_T is the T-dimensional identity matrix and $\mathbf{0}_T$ the T-dimensional 0 vector.

On the one hand, according to [8, Proposition 16.6], $\partial_1 \hat{f}_i^\nu(\boldsymbol{x}, \boldsymbol{Y} - \boldsymbol{x}) \subset \{(I_T, I_T)\mathbf{g} : \mathbf{g} \in \partial f_i^\nu(\boldsymbol{x}, \boldsymbol{Y})\}$, where $\partial f_i^\nu(\boldsymbol{x}, \boldsymbol{Y})$ is the subdifferential of f_i^ν at $(\boldsymbol{x}, \boldsymbol{Y})$. On the other hand, according to [8, Proposition 16.7],

$$\partial f_i^\nu(\boldsymbol{x}, \boldsymbol{Y}) \subset \left\{ \begin{pmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \end{pmatrix} : \mathbf{g}_1 \in \partial_1 f_i^\nu(\boldsymbol{x}, \boldsymbol{Y}), \mathbf{g}_2 \in \partial_2 f_i^\nu(\boldsymbol{x}, \boldsymbol{Y}) \right\} \ .$$

Therefore, $\partial_1 \hat{f}_i^\nu(\boldsymbol{x}, \boldsymbol{Y} - \boldsymbol{x}) \subset \{\mathbf{g}_1 + \mathbf{g}_2 : \mathbf{g}_1 \in \partial_1 f_i^\nu(\boldsymbol{x}, \boldsymbol{Y}), \mathbf{g}_2 \in \partial_2 f_i^\nu(\boldsymbol{x}, \boldsymbol{Y})\}$.

By the definition of f_i^ν, at $(\boldsymbol{x}, \boldsymbol{Y}) \in \mu_i^\nu \mathcal{M} \times \mathcal{M}$, the subdifferential of f_i^ν with respect to the second variable is $\partial_2 f_i^\nu(\boldsymbol{x}, \boldsymbol{Y}) = \mu_i^\nu f_{\bar{v}_i^\nu}(\frac{1}{\mu_i^\nu}\boldsymbol{x}, \boldsymbol{Y})$, hence each $\mathbf{g}_2 \in \partial_2 f_i^\nu(\boldsymbol{x}, \boldsymbol{Y})$ is bounded by $\mu_i^\nu B_{\mathbf{g}}$ in l^2-norm according to Assumption 8. In consequence, for each $\mathbf{h} \in \partial_1 \hat{f}_i^\nu(\boldsymbol{x}, \boldsymbol{Y} - \boldsymbol{x})$, there is $\mathbf{g}_1 \in \partial_1 f_i^\nu(\boldsymbol{x}, \boldsymbol{Y})$ such that $\|\mathbf{h} - \mathbf{g}_1\| \leq \mu_i^\nu B_{\mathbf{g}}$. Finally, set $\lambda_i^\nu \triangleq \mu_i^\nu B_{\mathbf{g}}$, and $\bar{\lambda}^\nu \triangleq \max_i \lambda_i^\nu$ tends to 0.

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

Case 2: Finite-dimensions Parameterized Characteristics – Meshgrid Approximation

Assume that nonatomic game G satisfy two conditions:

(i) The feasible pure-action sets are K-dimensional polytopes: there exist a constant real-valued $K \times T$ matrix $\boldsymbol{A}$, and a bounded mapping $\boldsymbol{b} : \Theta \to \mathbb{R}^K$, such that for any θ, $\mathcal{X}_\theta = \{\boldsymbol{x} \in \mathbb{R}^T : \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}_\theta\}$, 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 $(\boldsymbol{x}, \boldsymbol{Y}) \in \mathcal{M}^2$, $\partial_1 f(\boldsymbol{x}, \boldsymbol{Y}; \cdot)$ 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 $\underline{b}_{k,n_k} \triangleq \underline{b}_k + \frac{n_k}{\nu}(\overline{b}_k - \underline{b}_k)$ for $k \in \{1, \dots, K\}$, and $\underline{s}_{k,n_k} \triangleq \underline{s}_k + \frac{n_k}{\nu}(\overline{s}_k - \underline{s}_k)$ 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 $\underline{b}_{k,n_k} = \overline{b}_k$ or $\underline{s}_{k,n_k} = \overline{s}_k$, 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 $\mu(\Theta_{\boldsymbol{n}}^\nu) > 0$.

Remark 4.5. If there is a set of players of strictly positive measure sharing the same $\boldsymbol{b}$ and $\boldsymbol{s}$, then condition $\max_{i \in \mathcal{I}^\nu} \mu_i \to 0$ is satisfied. In this case, adding another dimension in the partition by cutting $\Theta = [0,1]$ into ν uniform segments solves the problem.

Proposition 4.2. *For $\nu \in \mathbb{N}^*$, let finite-player game $\mathcal{G}^\nu(A^\nu)$ with an aggregative constraint be defined by $A^\nu \triangleq A$, player set $\mathcal{I}^\nu \triangleq \{\boldsymbol{n} \in \Gamma^\nu : \mu(\Theta_{\boldsymbol{n}}^\nu) > 0\}$ and, for each finite player $\boldsymbol{n} \in \mathcal{I}^\nu$,*

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

$$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} \int_{\Theta_{\boldsymbol{n}}^\nu} \boldsymbol{s}_\theta \, d\theta\), \quad \forall (\boldsymbol{x}, \boldsymbol{Y}) \in \mu_i^\nu \mathcal{M} \times \mathcal{M}.$$

Then, under Assumptions 3 to 6 and 8, $(\mathcal{G}^\nu(A))_\nu = \big(\mathcal{I}^\nu, \widetilde{\mathcal{X}}^\nu, A, (f_i^\nu)_{i \in \mathcal{I}^\nu}\big)_\nu$ is an AAS of the nonatomic game $G(A)$ with an aggregative constraint.

Proof. Let us show the four properties required by Definition 4.1 as follows.

i) For each $\boldsymbol{n} \in \mathcal{I}^\nu$, $\frac{1}{\mu_{\boldsymbol{n}}^\nu} \mathcal{X}_{\boldsymbol{n}}^\nu = \Big\{\boldsymbol{x} \in \mathbb{R}^T : \boldsymbol{A}\boldsymbol{x} \leq \frac{1}{\mu_{\boldsymbol{n}}^\nu} \int_{\Theta_{\boldsymbol{n}}^\nu} \boldsymbol{b}_\theta \, d\theta\Big\}$. Then, by a result generalized from [7, Thm. 4.1], there is a constant C_0 such that, for each $\theta' \in \Theta_{\boldsymbol{n}}^\nu$: $d_H\left(\mathcal{X}_{\theta'}, \frac{1}{\mu_{\boldsymbol{n}}^\nu} \mathcal{X}_{\boldsymbol{n}}^\nu\right) \leq C_0 \left\|\boldsymbol{b}_{\theta'} - \frac{1}{\mu_{\boldsymbol{n}}^\nu} \int_{\Theta_{\boldsymbol{n}}^\nu} \boldsymbol{b}_\theta \, d\theta\right\| \leq \frac{C_0}{\nu} \left\|\overline{\boldsymbol{b}} - \underline{\boldsymbol{b}}\right\|$. Hence, $\overline{\delta}^\nu$ tends to 0.

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:

$$d_H\Big(\partial_1 f_{\boldsymbol{n}}^\nu(\mu_{\boldsymbol{n}}^\nu \boldsymbol{x}, \boldsymbol{Y}) \,, \, \partial_1 f_{\theta'}(\boldsymbol{x}, \boldsymbol{Y})\Big) = d_H\Big(\partial_1 f\big(\boldsymbol{x}, \boldsymbol{Y}; \tfrac{1}{\mu_{\boldsymbol{n}}^\nu} \int_{\Theta_{\boldsymbol{n}}^\nu} \boldsymbol{s}_\theta \, d\theta\big) \,, \, \partial_1 f\big(\boldsymbol{x}, \boldsymbol{Y}; \boldsymbol{s}_{\theta'}\big)\Big)$$

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

by the Lipschitz continuity of $\partial_1 f(\boldsymbol{x}, \boldsymbol{Y}; \cdot)$. Hence, $\overline{d}^\nu$ tends to 0.

iii) $\overline{\lambda}^\nu$ tends to 0 is proved as in Proposition 4.1;

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

Remark 4.6. In Proposition 4.2, 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 4.7. If the finite-player games constructed by these two constructions do not satisfy the condition that $\hat{f}_i^\nu(\cdot, \boldsymbol{Y}_{-i})$ is convex in $\boldsymbol{x}_i$, then one can use the pseudo-VNE of $\mathcal{G}^\nu(A^\nu)$ (cf. Appendix). The existence of pseudo-VNE is guaranteed by our construction of f_i^ν's.

Recall that most of the aggregatively games have cost functions of form Eq. (10). As long as the conditions in Proposition 3.1 (monotone per-unit cost of public products and concave private utility) are satified, $\hat{f}_i^\nu(\cdot, \boldsymbol{Y}_{-i})$ is convex in $\boldsymbol{x}_i$.

Remark 4.8. In both constructions above, the number of finite players I^ν tends to infinite and their "size" μ_i^ν tends to zero. These conditions are not required in the definition of AAS. The two constructions are made in this way for two reasons: firstly, nonatomic players' characteristics (the pure-action set and the subdifferential of their cost function with respect to the first variable) in a very small set Θ_i^ν are ensured to be close enough by construction; secondly, the subdifferential of cost function f_i of finite player i with respect to the second variable $\boldsymbol{X}$ is vanishing so that she behaves almost like a nonatomic player. If we consider pseudo-VNE instead of NE in the approximating finite-player games, then the second reason no longer exists. If the nonatomic players are fairly homogeneous or even there are finitely-many different types of them, then there is no need to divide Θ into smaller and smaller intervals to regroup sufficiently homogeneous nonatomic players, as imposed by the first reason. We shall bear in mind these two points when constructing an AAS for explicit games.

5 Conclusion.

Theorem 4.1 provides a theoretical basis for the use of finite-dimensional VNE/NE as an approximation of the VWE/WE in a strongly monotone or aggregatively strongly monotone nonatomic aggregative game with or without aggregative constraints. There are numerous research themes related to this result and our topic in general.

Firstly, one needs to find efficient algorithms for the computation of NE and VNE in finite-player aggregative games with or without aggregative constraints via the solution of their characteristic variational inequalities. 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 [22].

Secondly, the extension of evolutionary dynamics for population games and related algorithms to nonatomic games with infinitely many classes of players can be non trivial. A recent work [23] proposes online learning methods for population games with heterogeneous convex pure-action sets. The presence of aggregate constraints adds two additional difficulties as analyzed in Remarks 2.2 and 2.3. Evolutionary dynamics in population games 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 a pure-action profile violating the coupling constraint. On the other hand, a feasible deviation in the pure-action profile cannot always be decomposed into unilateral deviations of players.

Thirdly, our results are limited to monotone games and the convergence result is limited to strongly monotone games. The study of nonatomic 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.

Fourthly, other methods to deal with aggregate constraints exist, such as rationing in the case of capacity constraint of network or power grid.

Acknowledgments.

Appendix: Nonatomic behavior of finite players.

In the definition of AAS, we suppose that the cost functions of a finite player in a finite-player game is convex in her own action, i.e. $\hat{f}_i^\nu(\cdot, \boldsymbol{Y}_{-i})$ is convex on $\mathcal{M}$ for each fixed aggregate profile of her rivals $\boldsymbol{x}_{-i}$. However, this condition can be replaced by the condition that $f_i^\nu(\cdot, \boldsymbol{Y})$ is convex on $\mathcal{M}$ for each fixed aggregate profile $\boldsymbol{Y} \in \mathcal{M}$, which is naturally satisfied with the two constructions presented in Section 4.3. In this case, instead of considering VNE in the AAS, we consider another equilibrium notion called pseudo-VNE, where finite players take themselves as nonatomic ones by ignoring the impact of their action on the aggregate-action profile.

Consider a finite-player aggregative game $\mathcal{G}(A)$ with an aggregative constraint defined in Definition 2.3. Define a correspondence $H' : \widetilde{\mathcal{X}} \rightrightarrows \mathbb{R}^{IT}$ as follows: for all $\boldsymbol{x} \in \widetilde{\mathcal{X}}$, $H'(\boldsymbol{x}) \triangleq \{(\mathbf{g}_i)_{i \in \mathcal{I}} \in \mathbb{R}^{IT} : \mathbf{g}_i \in \partial_1 f_i(\boldsymbol{x}_i, \boldsymbol{X}), \forall i \in \mathcal{I}\}$, where $\boldsymbol{X} = \sum_{i \in \mathcal{I}} \boldsymbol{x}_i$ is the aggregate-action profile induced by $\boldsymbol{x}$.

Definition 5.1. A *pseudo-VNE* of a finite-player game $\mathcal{G}(A)$ with an aggregative constraint is a solution to the following GVI:

$$\text{Find } \hat{\boldsymbol{x}} \in \widetilde{\mathcal{X}}(A) \text{ s.t. } \exists \mathbf{g} \in H'(\boldsymbol{x}) \text{ s.t. } \langle \mathbf{g}, \boldsymbol{x} - \hat{\boldsymbol{x}} \rangle \geq 0, \forall \boldsymbol{x} \in \widetilde{\mathcal{X}}(A) .$$

Remark 5.1 (Pseudo-VNE is not an equilibirum). A pseudo-NE/pseudo-VNE is *not* an equilibrium from a game theoretical point of view. Defined as a solution to a certain GVI, it is an auxiliary notion for the approximation of NE/VNE.

The authors in [40] refer to this equilibrium notion as *Wardrop equilibrium*. We avoid this name here in order to distinguish it from the WE in nonatomic games.

Assumption 9 (Convex pseudo-costs). *For each $i \in \mathcal{I}$, the function $f_i(\boldsymbol{x}_i, \boldsymbol{Y})$ is continuous in both variables, and is convex in $\boldsymbol{x}_i$ for all $\boldsymbol{Y} \in \overline{\mathcal{X}}$.*

Under Assumption 9, the existence of a pseudo-VNE can be proved as for VNE in Proposition 2.5.

Define a *pseudo-AAS* for a nonatomic aggregative game $G(A)$ with an aggregative constraint be defined almost exactly like an AAS, except that condition (iv) is no longer required, and Assumption 1 is replaced by Assumption 9. Then, in Theorem 4.1 one can replace the VNE by the pseudo-VNE.

Theorem 5.1 (Convergence of pseudo-VNE to VWE)**.** *Under Assumptions 2 to 6 and 9, let $(\mathcal{G}^\nu(A^\nu))_\nu$ be a pseudo-AAS of $G(A)$. Let $\boldsymbol{x}^*$ be the VWE of $G(A)$, $\hat{\boldsymbol{x}}^\nu \in \widetilde{\mathcal{X}}^\nu(A^\nu)$ a pseudo-VNE of $\mathcal{G}^\nu(A^\nu)$ for each $\nu \in \mathbb{N}$, and $\boldsymbol{X}^*$, $\hat{\boldsymbol{X}}^\nu$ their respective aggregate-action profiles. Then, there exist constants $\rho > 0$, $\bar{\rho}$ and $K_A \triangleq \frac{M+1}{\min\{\rho, \bar{\rho}\}}$ for the following results to hold:*

(1) If G is aggregatively strongly monotone with modulus β, then $(\hat{\boldsymbol{X}}^\nu)_\nu$ converges to $\boldsymbol{X}^*$: for all $\nu \in \mathbb{N}^*$ such that $2\max(\overline{\delta}^\nu, D^\nu) < \rho$, $\|\hat{\boldsymbol{X}}^\nu - \boldsymbol{X}^*\|^2 \leq \frac{1}{\beta}\left((3B_{\mathbf{f}} + 1)K_A \max(D^\nu, \overline{\delta}^\nu) + (2M+1)\overline{d}^\nu\right)$.

(2) If G is strongly monotone with modulus α, then $(\psi^\nu(\hat{\boldsymbol{x}}^\nu))_\nu$ converges to $\boldsymbol{x}^*$ in L^2-norm: for all $\nu \in \mathbb{N}^*$ such that $2\max(\overline{\delta}^\nu, D^\nu) < \rho$, $\|\psi^\nu(\hat{\boldsymbol{x}}^\nu) - \boldsymbol{x}^*\|_2^2 \leq \frac{1}{\alpha}\left((3B_{\mathbf{f}} + 1)K_A \max(D^\nu, \overline{\delta}^\nu) + (2M+1)\overline{d}^\nu\right)$.

If there are no aggregate constraint, replace K_A and D^ν all by 0.

The proof is almost the same as for Theorem 4.1. The only difference is that, instead of considering $\mathbf{h}_i \in \partial_1 \tilde{f}_i^\nu(\hat{\boldsymbol{x}}_i^\nu, \hat{\boldsymbol{X}}_{-i}^\nu)$ for each $i \in \mathcal{I}^\nu$ such that $\sum_{i \in \mathcal{I}^\nu} \left\langle \mathbf{h}_i(\hat{\boldsymbol{x}}_i^\nu, \hat{\boldsymbol{X}}_{-i}^\nu), \hat{\boldsymbol{x}}_i^\nu - \boldsymbol{z}_i^\nu \right\rangle \leq 0$ for all $\boldsymbol{z}_i^\nu \in \mathcal{X}_i$, one has to consider $\mathbf{h}_i \in \partial_1 f_i^\nu(\hat{\boldsymbol{x}}_i^\nu, \hat{\boldsymbol{X}}^\nu)$ for each $i \in \mathcal{I}^\nu$ such that $\sum_{i \in \mathcal{I}^\nu} \left\langle \mathbf{h}_i(\hat{\boldsymbol{x}}_i^\nu, \hat{\boldsymbol{X}}^\nu), \hat{\boldsymbol{x}}_i^\nu - \boldsymbol{z}_i^\nu \right\rangle \leq 0$ for all $\boldsymbol{z}_i \in \mathcal{X}_i$.

References

[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] BAUSCHKE, H. and COMBETTES, P. L. (2011). *Convex Analysis and Monotone Operator Theory in Hilbert Spaces.* 2nd ed. Springer International Publishing.

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

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

[11] CAVAZZUTI, E., PAPPALARDO, M. and PASSACANTANDO, M. (2002). Nash equilibria, variational inequalities, and dynamical systems. **114** 491–506.

[12] CHAN, D. and PANG, J. S. (1982). The generalized quasi-variational inequality problem. *Mathematics of Operations Research*, **7** 211–222.

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

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

[15] DEBREU, G. (1952). A social equilibrium existence theorem. *Proceedings of the National Academy of Sciences of the U.S.A.*, **38** 866–893.

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

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

[18] FAN, K. (1952). Fixed-point and minimax theorems in locally convex topological linear spaces. *Proceedings of the National Academy of Sciences of the U.S.A.*, **38** 121–126.

[19] FANG, S. C. and PETERSON, E. L. (1982). Generalized variational inequalities. *Journal of Optimization Theory and Applications*, **38** 363–383.

[20] 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.

[21] GLICKSBERG, I. L. (1952). A further generalization of the Kakutani fixed point theorem, with application to Nash equilibrium points. *Proceedings of the American Mathematical Society*, **3** 170–174.

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

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

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

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

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

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

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

[29] KHAN, M., RATH, K. P. and SUN, Y. (1997). On the existence of pure strategy equilibria in games with a continuum of players. *Journal of Economic Theory*, **76** 13–46.

[30] KULKARNI, A. A. and SHANBHAG, U. V. (2012). On the variational equilibrium as a refinement of the generalized Nash equilibrium. *Automatica*, **48** 45–55.

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

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

[33] 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.

[34] 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.

[35] Mas-Colell, A. (1984). On a theorem of Schmeidler. *Journal of Mathematical Economics*, **13** 201–206.

[36] Maynard Smith, J. (1982). *Evolution and the Theory of Games*. Cambridge University Press.

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

[38] Nash, J. F. (1950). Equilibrium points in n-person games. *Proc. of the Nat. Acad. of Sci. of the U.S.A.*, **36** 48–49.

[39] Orda, A., Rom, R. and Shimkin, N. (1993). Competitive routing in multiuser communication networks. **1** 510–521.

[40] 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.

[41] 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.

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

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

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

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

[46] Smith, M. (1984). A descent algorithm for solving monotone variational inequalities and monotone complementarity problems. **44** 485–496.

[47] 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.

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

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

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

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

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

[53] 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.

[54] Zhu, D. and Marcotte, P. (1994). An extended descent framework for variational inequalities. **80** 349–366.