On the optimal selection of generalized Nash equilibria in linearly coupled aggregative games

Benenati, Emilio; Ananduta, Wicak; Grammatico, Sergio

DOI
10.1109/CDC51059.2022.9993415

Publication date
2022

Document Version
Final published version

Published in
Proceedings of the IEEE 61st Conference on Decision and Control, CDC 2022

2022 IEEE 61st Conference on Decision and Control (CDC)
December 6-9, 2022. Cancún, Mexico

On the optimal selection of generalized Nash equilibria in linearly coupled aggregative games

Emilio Benenati, Wicak Ananduta, and Sergio Grammatico

2022 IEEE 61st Conference on Decision and Control (CDC) | 978-1-6654-6761-2/22/$31.00 ©2022 IEEE | DOI: 10.1109/CDC51059.2022.9993415

***Abstract*— Monotone aggregative games may admit multiple (variational) generalized Nash equilibria, yet currently there is no algorithm able to provide an a-priori characterization of the equilibrium solution actually computed. In this paper, we formulate for the first time the problem of selecting a specific variational equilibrium that is optimal with respect to a given objective function. We then propose a semi-decentralized algorithm for optimal equilibrium selection in linearly coupled aggregative games and prove its convergence.**

I. Introduction

Multiple decision makers (agents) are engaged in an aggregative game when each agent aims at solving an optimization problem that is coupled with the strategies of the remaining agents through some aggregate effect, e.g. a congestion effect [1, pp. 90–92]. Interestingly, aggregative games effectively model various engineering problems, such as bandwidth allocation [2], power distribution [3], and vehicle traffic control [4]. In this context, a generalized Nash equilibrium (GNE) is a particularly favourable and stable set of strategies from which no agent has interest in unilaterally deviating.

GNE problems are commonly studied under the monotone operator theory framework, see [1] for a general overview. When the game in consideration is monotone, i.e., it has a jointly convex feasible set [5, Def. 3.6] and a monotone pseudogradient mapping, the GNE seeking problem can be solved via operator splitting techniques [6] in order to compute a variational GNE (v-GNE) [5, Def. 3.10]. For instance, [7] proposes multiple semi-decentralized methods, where decision makers do not communicate among each other but with a coordinator, for monotone aggregative games. Meanwhile, some works consider a particular class, namely strongly monotone aggregative games, which admit a unique v-GNE, and propose algorithms with a semi-decentralized [8] or distributed structure [9]–[11], where agents exchange information with each other. Recently, fully distributed algorithms have also been introduced for the class of monotone aggregative games [12], [13].

Merely monotone games are of particular interest, as they are more general than their strongly monotone counterpart and monotonicity is one of the weakest conditions under which globally convergent algorithmic solutions can be obtained [5, Sec. 5]. Nonetheless, they present additional hurdles to their solution. Among these complications, we focus on the non-uniqueness of v-GNE solutions. To the best of our knowledge, the existing GNE-seeking algorithms for monotone games, e.g. [7], [12]–[14], solely guarantee convergence to an arbitrary point in the v-GNE set, with no characterization of the computed equilibrium. A noteworthy exception is the Tikhonov method [15], [16] that is not based on operator splitting techniques and finds a minimum norm v-GNE. Therefore, an open challenge in the monotone game literature concerns finding an equilibrium, among infinitely many, with a desirable property, which is not necessarily the minimum norm as in [15], [16]. This problem is crucial from a practical standpoint as unpredictability of the computed equilibrium might lead to arbitrarily inefficient performance relatively to some system-level metrics, e.g. social welfare.

In order to address this deficiency for the class of linearly coupled aggregative games, we pose the GNE selection problem, that is, the problem of computing a specific v-GNE such that it is optimal with respect to a convex selection function. The selection function encodes a preference criterion for the GNEs and it can be defined on the basis of a system-level performance metrics. For instance, in power distribution systems, a preference criterion can be the deviation from an operating set point desired by the system operator [17]. We cast the GNE selection problem as a Variational Inequality (VI) defined by the gradient of the selection function and the v-GNE set of the game.

Next, we propose an algorithmic solution to solve the GNE selection problem for the class of linearly coupled aggregative games. Our algorithm has a semi-decentralized structure in order to exploit the aggregative feature of the games. It is based on combining the preconditioned proximal point (PPP) method [7], [12], which provides fast convergence under the (non-strict) monotonicity assumption, with the hybrid steepest descent method (HSDM) [18], which can solve fixed-point selection problems. We guarantee convergence to an optimal equilibrium by showing the equivalence of the GNE selection problem to that of fixed-point selection of the PPP operator. We then prove that the PPP operator satisfies the conditions under which the HSDM converges to a solution of the corresponding fixed-point selection problem, which, in turn, is also a solution of the GNE selection problem. Finally, we show the advantages of the proposed algorithm by comparing the performance of an optimal v-GNE computed by our proposed method with that of a non-characterized v-GNE obtained using the standard PPP algorithm in randomly generated numerical examples.

Notation: The set of (non-negative) real numbers is denoted by $\mathbb{R}$ ($\mathbb{R}_{\geq 0}$). The Euclidean inner product and norm are denoted respectively by $\langle x, y\rangle$ and $\|\cdot\|$. Nonlinear set-valued operators are denoted in calligraphic letters, e.g. $\mathcal{T} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$. For $Q \succ 0$, we denote the Q-weighted norm by $\|\cdot\|_Q = \langle \cdot, Q\cdot\rangle$. For a matrix A, $\|A\|$ denotes

its largest singular value. We denote the vector of all 1 (0) with dimension n by $\mathbf{1}_n$ ($\mathbf{0}_n$). The operator $\mathrm{col}(\cdot)$ stacks the arguments column-wise. The column concatenation of the vectors x_i indexed in $i \in \mathcal{I}$ is denoted in bold, i.e., $\boldsymbol{x} := \mathrm{col}(\{x_i\}_{i\in\mathcal{I}})$. The operator $\mathrm{diag}(\cdot)$ denotes a (block) diagonal matrix with the arguments on the diagonal. The operator $\mathrm{avg}(\cdot)$ returns the element-wise average of the arguments, e.g., when $x_i \in \mathbb{R}^{\bar{n}}$ for all $i \in \mathcal{I}$, $\mathrm{avg}(\{x_i\}_{i\in\mathcal{I}}) = \frac{1}{|\mathcal{I}|}\sum_{i\in\mathcal{I}} x_i$.

Operator theory: We denote by Id the identity operator. For a closed convex set $C \subset \mathbb{R}^n$, the normal cone operator is denoted by $\mathrm{N}_C(\cdot) : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ [6, Def. 6.38] and the indicator function by $\iota_C : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$. We then denote the projection onto C by $\mathrm{proj}_C(x) = \mathrm{argmin}_{z\in C}\|x - z\|$. For a convex function f with subdifferential ∂f, we define the operator $\mathrm{prox}_{\partial f}(x) := \mathrm{argmin}_z\, f(z) + \frac{1}{2}\|z - x\|^2$ [6, Def. 12.23]. In particular, as $\partial\iota_C = \mathrm{N}_C$ [6, Example 16.13], $\mathrm{prox}_{N_C}(x) = \mathrm{proj}_C(x)$. For an operator $\mathcal{T}\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, we denote its zero set by $\mathrm{zer}(\mathcal{T}) := \{x \in \mathrm{dom}(\mathcal{T}) \mid 0 \in \mathcal{T}(x)\}$ and its fixed point set by $\mathrm{fix}(\mathcal{T}) := \{x \in \mathrm{dom}(\mathcal{T}) \mid x \in \mathcal{T}(x)\}$. Finally, an operator $\mathcal{T} : C \rightrightarrows \mathbb{R}^n$ is:
Monotone: if, for any two pairs $(x, y), (x', y') \in \mathrm{gph}(\mathcal{T}) := \{(x, y) | x \in C, y \in \mathcal{T}(x)\}$, it holds that $\langle y - y', x - x'\rangle \geq 0$;
Lipschitz continuous: if there exists $L > 0$, such that, for all $x, x' \in C$, $\|\mathcal{T}(x) - \mathcal{T}(x')\| \leq L\|x - x'\|$;
Nonexpansive: if Lipschitz continuous with $L = 1$;
Attracting nonexpansive: if nonexpansive with $\mathrm{fix}(\mathcal{T}) \neq \varnothing$ and $\|\mathcal{T}(x) - z\| < \|x - z\|$, for all $z \in \mathrm{fix}(\mathcal{T})$, $x \notin \mathrm{fix}(\mathcal{T})$;
α-averaged nonexpansive: if for $\alpha \in (0, 1)$ there exists $\mathcal{R}\colon C \to \mathbb{R}^n$ nonexpansive such that $\mathcal{T} = (1 - \alpha)\mathrm{Id} + \alpha\mathcal{R}$;
Firmly nonexpansive: if for all $x, x' \in \mathbb{R}^n$, $\|Tx - Ty\|^2 + \|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2 \leq \|x - y\|^2$.

II. Optimal generalized Nash equilibrium selection problem in linearly coupled games

We consider N agents, indexed by the set $\mathcal{I} := \{1, 2, \ldots, N\}$, which are engaged in a generalized game, i.e., each agent aims at solving an optimization problem which is coupled, both in the cost and in the constraints, with the decision variables of the remaining agents. Let us denote by $x_i \in \mathbb{R}^{n_i}$ the decision variable of agent i and by $\boldsymbol{x}_{-i} = \mathrm{col}(\{x_j\}_{j\in\mathcal{I}\setminus\{i\}})$ the concatenated decision variables of all agents except agent i. Let us further denote the local feasible set of each agent i by $\mathcal{X}_i \subseteq \mathbb{R}^{n_i}$ and the cost function by $J_i\colon \mathbb{R}^n \to \mathbb{R}$. We focus on the class of linearly coupled games, where the cost function of each agent $i \in \mathcal{I}$ takes the form [7, Eq. (1)]:

$$J_i(x_i, \boldsymbol{x}_{-i}) := \ell_i(x_i) + \langle \textstyle\sum_{j\in\mathcal{I}\setminus\{i\}} C_{ij}x_j, x_i\rangle, \qquad (1)$$

where ℓ_i denotes the local cost function of agent i. Furthermore, the matrix $C_{ij} \in \mathbb{R}^{n_i\times n_j}$ encodes the weight of the influence of the decision variable of agent j with respect to the cost function of agent i and is assumed to be local information known only by agent i. Additionally, we let the matrices $A_i \in \mathbb{R}^{m\times n_i}$ and the vectors $b_i \in \mathbb{R}^m$, for all $i \in \mathcal{I}$, encode a linear coupling constraint of the form $\sum_{i\in\mathcal{I}}(A_ix_i - b_i) \leq 0$. Therefore, the interdependent optimization problems are given by:

$$\forall i \in \mathcal{I}\colon \begin{cases} \min\limits_{x_i\in\mathcal{X}_i} & J_i(x_i, \boldsymbol{x}_{-i}) & \text{(2a)}\\ \text{s. t.} & \textstyle\sum_{j\in\mathcal{I}}(A_jx_j - b_j) \leq 0. & \text{(2b)}\end{cases}$$

The collective feasible set of the game in (2) is defined as

$$\boldsymbol{\Omega} := \boldsymbol{\mathcal{X}} \cap \Big\{\boldsymbol{x} \mid \textstyle\sum_{j\in\mathcal{I}}(A_jx_j - b_j) \leq 0\Big\}, \qquad (3)$$

where $\boldsymbol{\mathcal{X}} := \prod_{i\in\mathcal{I}} \mathcal{X}_i$. Let us consider the following:

Assumption 1. *For each $i \in \mathcal{I}$, the function $\ell_i(x_i)$ in* (1) *is convex and lower semicontinuous. For each $i \in \mathcal{I}$, the set $\mathcal{X}_i$ in* (2a) *is nonempty, compact, and convex. The set $\boldsymbol{\Omega}$ in* (3) *satisfies Slater's constraint qualification condition.*

Assumption 2. *[12, Assm. 2] The matrices $C_{ij} \in \mathbb{R}^{n_i\times n_j}$ in* (1)*, for all $i, j \in \mathcal{I}$, satisfy $C_{ij} = C_{ji}^\top$.*

Assumption 1 is standard (see e.g., [7], [12], [14]), while Assumption 2 is a technical assumption that implies that the game is potential [12, Lem. 1]. This assumption is necessary since our method relies on preconditioning [19]. This technique requires convergence to be proven on the norm induced by the preconditioning matrix, whose structure [12, Sec. IVB] makes Assumption 2 a necessary condition for such norm to be well-defined. The class of games described in (2) along with Assumptions 1–2 includes linearly-coupled aggregative games [7, Sec. IVB], obtained with $C_{ij} = \frac{1}{N}C$, for all $i, j \in \mathcal{I}$, and $\ell_i(x_i) = \bar{\ell}_i(x_i) + \frac{1}{N}x_i^\top Cx_i$, implying that (1) can be written as

$$J_i(x_i, \boldsymbol{x}_{-i}) = \bar{\ell}_i(x_i) + \langle C\,\mathrm{avg}(\{x_j\}_{j\in\mathcal{I}}), x_i\rangle, \quad \forall i \in \mathcal{I}. \qquad (4)$$

The set of solutions to the game in (2) that we consider is that of generalized Nash equilibria (GNE), i.e., a set of decisions from which no agent finds an advantage in unilaterally deviating, as formally defined next.

Definition 1. *A set of strategies $\boldsymbol{x}^* := \mathrm{col}((x_i^*)_{i\in\mathcal{I}})$ is a generalized Nash equilibrium (GNE) of the game in* (2) *if $\boldsymbol{x}^* \in \boldsymbol{\Omega}$ and, for each $i \in \mathcal{I}$,*

$$J_i(\boldsymbol{x}^*) \leq J_i(x_i, \boldsymbol{x}^*_{-i}), \qquad (5)$$

for any $x_i \in \mathcal{X}_i \cap \{y \mid A_iy - b_i \leq -\sum_{j\in\mathcal{I}\setminus\{i\}}(A_j(x_j^*) - b_j)\}$.

As in [7], [11]–[14] we focus on the computation of a subset of the GNEs of the problem in (2), namely, variational GNEs (v-GNEs), where, roughly speaking, each agent is penalized equally in meeting the coupling constraints. Under Assumption 1, a sufficient condition for the existence of a v-GNE is the monotonicity of the pseudodifferential/game mapping [20, Prop. 12.11], which is a standard assumption in the GNE seeking literature (see [7, Assm. 2], [21, Assm. 2], [12, Assm. 3] among others), and we postulate next:

Assumption 3. *The game mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$*

$$F(\boldsymbol{x}) := \begin{bmatrix} \partial_{x_1} J_1(x_1, \boldsymbol{x}_{-1}) \\ \vdots \\ \partial_{x_N} J_N(x_N, \boldsymbol{x}_{-N}) \end{bmatrix} = \mathrm{diag}(\{\partial_{\boldsymbol{x}_i}\ell_i\}_{i\in\mathcal{I}}) + \boldsymbol{C}, \qquad (6)$$

where $\boldsymbol{C} \in \mathbb{R}^{n\times n}$ is a block matrix with $\mathbf{0}_{n_i\times n_i}$ on the i-th block element of the diagonal, and C_{ij} as defined in (1) *on the (i,j)-th block, is a maximally monotone operator.*

A v-GNE of the game in (2) can be characterized as the solution to a generalized variational inequality [20, Prop. 12.4] which, in turn, is solved by the pair $(\boldsymbol{x},\lambda) =: \boldsymbol{\omega} \in \mathbb{R}^n \times \mathbb{R}^m$, where λ denotes the dual variable associated with the coupling constraint (2b), that satisfies the Karush-Kuhn-Tucker (KKT) optimality conditions [22, Thm. 3.1]:

$$\boldsymbol{\omega} \in \operatorname{zer}(\mathcal{T}_{\mathrm{e}}), \tag{7}$$

$$\mathcal{T}_{\mathrm{e}}(\boldsymbol{\omega}) := \begin{bmatrix} \prod_{i\in\mathcal{I}} \mathrm{N}_{\mathcal{X}_i}(x_i) + F(\boldsymbol{x}) + \operatorname{col}(\{A_i^\top \lambda\}_{i\in\mathcal{I}}) \\ \mathrm{N}_{\mathbb{R}^m_{\geq 0}}(\lambda) - \sum_{i\in\mathcal{I}}(A_i x_i - b_i) \end{bmatrix}.$$

Assumptions 1 and 3 are not enough to guarantee the uniqueness of the solution to the inclusion in (7). Furthermore, we may want to select, among the potentially infinite equilibria of the game in (2), a v-GNE with some desirable features. Here, we propose to seek a v-GNE that optimizes a common objective function, denoted by ϕ. As pointed out in [23], [24], this objective function can be the distance from some desired strategies, or a norm as in the Tikhonov method [15]. In engineering applications, such functions can represent system-level objectives that the agents are willing to achieve, provided that these objectives do not strongly interfere with those of all agents. Thus, we formalize the GNE selection problem as follows:

$$\begin{cases} \underset{\boldsymbol{\omega}}{\operatorname{argmin}} & \phi(\boldsymbol{\omega}) \\ \text{s.t.} & \boldsymbol{\omega} \in \operatorname{zer}(\mathcal{T}_{\mathrm{e}}), \end{cases} \tag{8}$$

with $\mathcal{T}_e$ as in (7) and the following assumption on the objective function:

Assumption 4. *The function ϕ in* (8) *is convex and differentiable while its gradient is L_ϕ-Lipschitz continuous.*

We note that, by [7, Lemma 2], the operator $\mathcal{T}_{\mathrm{e}}$ in (8) is maximally monotone, which implies that the set $\operatorname{zer}(\mathcal{T}_{\mathrm{e}})$ is closed and convex [6, Prop. 23.39]. We conclude that under Assumption 4 the optimization problem in (8) is convex.

III. Optimal equilibrium selection algorithm

In this section, we present a semi-decentralized algorithm for selecting an optimal v-GNE of the game in (2). The algorithm is semi-decentralized in the sense that the agents locally compute the updates to their strategies and communicate them to an aggregator, which updates the dual and the aggregative variables and broadcasts them to the agents.

A. Semi-decentralized optimal equilibrium selection

We propose Algorithm 1, whose derivation is deferred to Section III-B, for solving (8). The iterated steps of the algorithm are summarized as follows:

Communication 1: Each player $i \in \mathcal{I}$ receives the updated dual variable $\lambda^{(k)}$ and the decision variables of the remaining agents $\boldsymbol{x}_{-i}^{(k)}$ to compute the current estimate of the local Lagrangian function

$$\Lambda(x_i, \boldsymbol{x}_{-i}^{(k)}, \lambda^{(k)}) = J_i(x_i, \boldsymbol{x}_{-i}^{(k)}) + \langle \lambda^{(k)}, A_i x_i \rangle.$$

Algorithm 1 Optimal v-GNE selection for linearly coupled games

Initialization. Set $x_i^{(0)} \in \mathcal{X}_i$ and $\lambda_i^{(0)} \in \mathbb{R}^m_{\geq 0}$, for all $i \in \mathcal{I}$.
Iteration at stage $k \in \mathbb{N}$:

1) Each agent $i \in \mathcal{I}$ receives $x_j^{(k)}$, for all $j \in \mathcal{I}$, and $\lambda^{(k)}$ from the aggregator.
2) Each agent $i \in \mathcal{I}$ updates in parallel:
$$\mathring{x}_i^{(k)} = \underset{y\in\mathcal{X}_i}{\operatorname{argmin}}\ J_i(y, \boldsymbol{x}_{-i}^{(k)}) + \langle \lambda^{(k)}, A_i y\rangle + \tfrac{\rho_i}{2}\|y - x_i^{(k)}\|^2. \tag{9}$$
3) Each agent $i \in \mathcal{I}$ sends $\mathring{x}_i^{(k)}$ to the aggregator.
4) The aggregator updates:
$$\mathring{\lambda}^{(k)} = \operatorname{proj}_{\mathbb{R}^m_{\geq 0}}\Big(\lambda^{(k)} + \tau \textstyle\sum_{i\in\mathcal{I}} \Big(A_i \mathring{x}_i^{(k)} - b_i\Big)\Big), \tag{10}$$
$$\lambda^{(k+1)} = \mathring{\lambda}^{(k)} - \beta^{(k)} \nabla_\lambda \phi(\mathring{\boldsymbol{x}}^{(k)}, \mathring{\lambda}^{(k)}). \tag{11}$$
5) Each agent receives $\nabla_{\boldsymbol{x}_i}\phi(\mathring{\boldsymbol{x}}^{(k)}, \mathring{\lambda}^{(k)})$ from the aggregator.
6) Each agent updates:
$$x_i^{(k+1)} = \mathring{x}_i^{(k)} - \beta^{(k)} \nabla_{\boldsymbol{x}_i}\phi(\mathring{\boldsymbol{x}}^{(k)}, \mathring{\lambda}^{(k)}), \tag{12}$$
and sends $x_i^{(k+1)}$ to the aggregator.

Regularized optimal response: Each agent computes a strategy $\mathring{x}_i^{(k)}$ by finding the optimizer of $\Lambda(x_i, \boldsymbol{x}_{-i}^{(k)}, \lambda^{(k)})$, with a quadratic penalization term on the deviation from the current strategy weighted by the parameter ρ_i, as in (9).
Communication 2 and dual ascent: The proposed strategies $\mathring{x}_i^{(k)}$ are gathered by an aggregator, which computes the gradient $\nabla_{\boldsymbol{x}_i}\phi(\mathring{\boldsymbol{x}}^{(k)}, \mathring{\lambda}^{(k)})$ and updates the dual variable via a dual ascent step as in (10) with step size τ. The results are communicated back to the agents.
Step towards the optimal selection: The agents and the aggregator update the primal and the dual variables by performing a gradient descent step of ϕ with vanishing step size $\beta^{(k)}$ as in (11) and (12), respectively.

Remark 1. *If the matrices C_{ij} are all equal, then the local cost functions read as in* (4)*, thus they depend on the remaining agents only through the aggregate value* $\operatorname{avg}(\{x_j\}_{j\in\mathcal{I}})$*. In such case, the agents only suffice to receive* $\operatorname{avg}(\{x_j\}_{j\in\mathcal{I}})$ *and the gradient directions $\nabla_{\boldsymbol{x}_i}\phi(\mathring{\boldsymbol{x}}^{(k)}, \mathring{\lambda}^{(k)})$.*

Remark 2. *If the selection function is separable, that is, $\phi(\boldsymbol{x}) = \sum_{i\in\mathcal{I}} \phi_i(x_i, \lambda_i)$, then the second round of communications is not needed, as the step in* (12) *can be immediately computed using local information only.*

Under a choice of step sizes that satisfies Assumptions 5 and 6, let us now present the main result of this paper in Theorem 1, which states that the sequence $(\boldsymbol{\omega}^{(k)})_{k\in\mathbb{N}}$ converges to the set of the optimal solutions to Problem (8), implying that $(\boldsymbol{x}^{(k)})_{k\in\mathbb{N}}$ converges to the optimal v-GNE set.

Assumption 5. *The step sizes ρ_i and τ satisfy:*

(i) $\rho_i > \sum_{j\in\mathcal{I}\setminus\{i\}} \|C_{ij}\| + \|A_i^\top\|$*, for all* $i \in \mathcal{I}$*,*
(ii) $\tau < (\sum_{i\in\mathcal{I}} \|A_i\|)^{-1}$.

Assumption 6. *The sequence $(\beta^{(k)})_{k\in\mathbb{N}} \in \mathbb{R}_{\geq 0}$ satisfies*

$\lim_{k\to\infty}\beta^{(k)}=0$, $\sum_{k\geq 1}\beta^{(k)}=\infty$, *and* $\sum_{k\geq 1}(\beta^{(k)})^2<\infty$.

Remark 3. *The sequence* $\beta^{(k)}=\beta_0/k^\gamma$, *for any* $\beta_0>0$ *and* $\gamma\in(1/2,1]$, *satisfies Assumption 6.*

Theorem 1. *Let Assumptions 1–6 hold. Let* $\Omega^\star$ *be the set of solutions to Problem* (8). *Furthermore, let* $(\boldsymbol{\omega}^{(k)})_{k\in\mathbb{N}}$, *where* $\boldsymbol{\omega}^{(k)}=(\boldsymbol{x}^{(k)},\lambda^{(k)})$ *be the sequence generated by Algorithm 1. Then, we have*

$$\lim_{k\to\infty}\operatorname{dist}(\boldsymbol{\omega}^{(k)},\Omega^\star)=0.$$

B. Algorithm derivation

As the first step towards obtaining Algorithm 1, let us consider the preconditioned proximal point (PPP) method, which was proposed in [7], [12] to solve the GNE seeking problem for the class of linearly coupled aggregative games. The PPP method can be compactly written as the iteration

$$\boldsymbol{\omega}^{(k+1)}=\mathcal{T}_{\mathrm{PPP}}(\boldsymbol{\omega}^{(k)}),\tag{13}$$

where the operator $\mathcal{T}_{\mathrm{PPP}}$ is defined as

$$\mathcal{T}_{\mathrm{PPP}}(\boldsymbol{\omega}):=(\mathrm{Id}+\Gamma^{-1}\mathcal{T}_{\mathrm{e}})^{-1}(\boldsymbol{\omega}),\tag{14}$$

with preconditioning matrix

$$\Gamma:=\begin{bmatrix}\rho-\boldsymbol{C} & -\boldsymbol{A}^\top\\ -\boldsymbol{A} & \tau^{-1}I\end{bmatrix},$$

$\rho:=\operatorname{diag}(\{\rho_i I_{n_i}\}_{i\in\mathcal{I}})$ and $\boldsymbol{A}^\top:=\operatorname{col}(\{A_i^\top\}_{i\in\mathcal{I}})$. The operator $\mathcal{T}_{\mathrm{PPP}}$ enjoys the following property, which we leverage for proving Theorem 1:

Lemma 1. *If Assumptions 1, 2, 3, and 5 hold, then* $\mathcal{T}_{\mathrm{PPP}}$ *is attracting nonexpansive in the* Γ*-induced norm* $\|\cdot\|_\Gamma$.

We note that $\Gamma=\Gamma^\top$ is a necessary condition for the Γ-induced norm to be well-defined, which explains the need for Assumption 2. Next, we transform Problem (8) into a fixed-point selection problem:

Lemma 2. *Let Assumptions 1–5 hold. Problem* (8) *is equivalent to*

$$\textit{find } \boldsymbol{\omega}^\star \textit{ s.t. } \inf_{\boldsymbol{\omega}\in\operatorname{fix}(\mathcal{T}_{\mathrm{PPP}})}\langle\boldsymbol{\omega}-\boldsymbol{\omega}^\star,\nabla\phi(\boldsymbol{\omega}^\star)\rangle\geq 0.\tag{15}$$

By finding a connection between Problem (8) and a class of fixed-point selection problems that has been studied in the literature, e.g., [25]–[27], we can then resort to the algorithmic solutions available for the latter. Specifically, we consider the hybrid steepest descent method (HSDM) [18]. As formally stated next, indeed Algorithm 1 is a particular instance of the HSDM.

Lemma 3. *Let Assumptions 1–6 hold. Then, Algorithm 1 and the HSDM iterations, for all* $k\in\mathbb{N}$,

$$\boldsymbol{\omega}^{(k+1)}=\mathcal{T}_{\mathrm{PPP}}(\boldsymbol{\omega}^{(k)})-\beta^{(k)}\nabla\phi(\mathcal{T}_{\mathrm{PPP}}(\boldsymbol{\omega}^{(k)})),\tag{16}$$

with $\mathcal{T}_{\mathrm{PPP}}$ *as defined in* (14)*, are equivalent.*

IV. Illustrative example

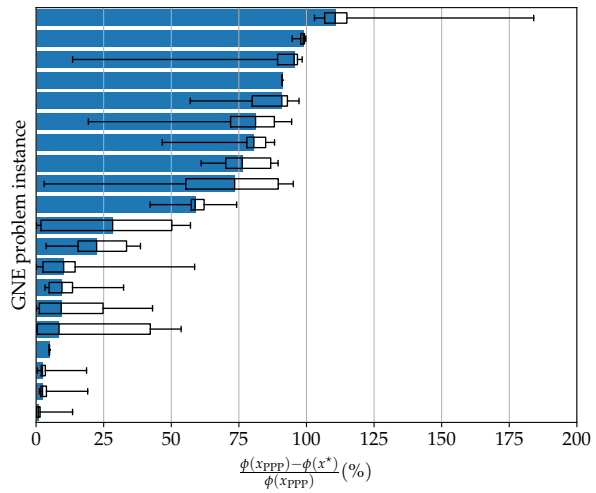
We illustrate the advantages of the proposed methodology on a numerical example. Let $N=6$ agents compete over the usage of 3 utilities, where the cost of each utility grows linearly with its aggregate usage. This setting is modelled by the game in (2) with local cost function given by (4) and C diagonal, whose non-zero elements are randomly sampled from the uniform distribution with support set $[0,1]$. Let $\ell_i(x_i)=q^\top x_i$ represent the cost of agent i incurred by employing the contended utilities, where q is randomly sampled from the uniform distribution with support set $[-10,0]$. Let $\mathcal{X}_i=\Pi_{j=1}^3[a_{i,j},100]$, where $a_{i,j}$ is drawn from the uniform distribution with support set $[-1,1]$. The shared constraints in (2b) are given by $A_i=I_3$ and $b_i=5\mathbf{1}_3$ for all $i\in\mathcal{I}$, that is, the sum of each utility is less than 5. Finally, the selection function ϕ is given by

$$\phi(\boldsymbol{x})=\textstyle\sum_{i\in\mathcal{R}}\|x_i\|_{Q_i}^2+\tilde{q}_i^\top x_i,\tag{17}$$

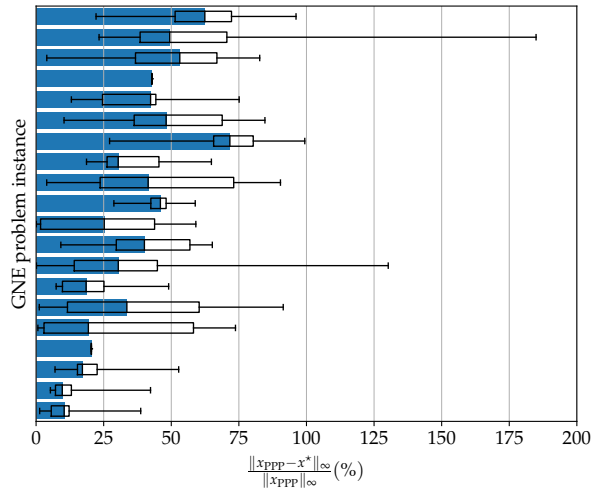
where $\mathcal{R}\subset\mathcal{I}$ is randomly generated with $|\mathcal{R}|=2$, Q_i is a diagonal matrix whose nonzero elements are randomly drawn from the uniform distribution with support set $[1,2]$ and the elements of $\tilde{q}_i$ are randomly drawn from the uniform distribution with support set $[-1,1]$. This choice for ϕ includes the case where the equilibrium point is selected in order to minimize a weighted distance from a reference point for the decision variables of 2 randomly selected agents. We sample 20 GNE selection problems. For each of the problem, we compare the result obtained by Algorithm 1 with the outcome of the standard PPP method [7, Alg. 6], which is obtained from Algorithm 1 by fixing $\beta^{(k)}=0$ in (11) and (12). On the other hand, for Algorithm 1, we set $\beta^{(k)}=\beta_0/k^\gamma$, with $\beta_0=0.1$ and γ selected from the set $\{0.6,0.8,1\}$ (see Remark 3). For each problem and each value of γ, both algorithms are run 20 times from a randomly generated initial condition. The results in Figure 1a show that the GNEs computed by the standard PPP are most of the time suboptimal with respect to ϕ in (17) as the values of the selection function are higher from those of the GNEs computed by Algorithm 1. As can also be seen from Figure 1a, the advantage gained by Algorithm 1 depends on the primitives of the problem and the performance of the standard PPP algorithm with respect to the selection function are strongly dependent on the initial condition. Figure 1b shows that the distance between the equilibrium points is not correlated with the reduction in the value of ϕ. Figure 2 compares the convergence rates to the set of GNEs by means of the residual r of the KKT conditions in (7) defined as

$$r(\boldsymbol{\omega}^{(k)})=\left\|\boldsymbol{\omega}^{(k)}-\begin{bmatrix}\operatorname{proj}_{\boldsymbol{\mathcal{X}}}\left(\boldsymbol{x}^{(k)}-F(\boldsymbol{x}^{(k)})-\operatorname{col}(\{A_j^\top\lambda\}_{j\in\mathcal{I}})\right)\\ \operatorname{proj}_{\mathbb{R}^m_{\geq 0}}\left(\lambda^{(k)}+\sum_{i\in\mathcal{I}}(A_i x_i^{(k)}-b_i)\right)\end{bmatrix}\right\|_\infty.$$

Figure 2 shows that Algorithm 1 presents slower convergence to the set of GNEs compared to PPP. This is expected since, although the updates in (9) and (10) lead the decision variables to the set of GNEs, the gradient step in (11) and (12) may lead the decision variables away from it until the step size $\beta^{(k)}$ is small enough; thus slowing down the convergence. Such an observation hints, as possible future research directions, the exploration of higher-order or accelerated methods inspired by the HSDM to achieve



(a)



(b)

Fig. 1: (a) Reduction in the selection function value for the GNEs computed by Algorithm 1 ($\boldsymbol{x}^\star$) using $\gamma = 0.6$ with respect to that of the solution computed by the standard PPP algorithm ($\boldsymbol{x}_{\mathrm{PPP}}$). (b) Distance between the computed solutions.

a faster convergence to the optimal GNE. Moreover, we observe from Figure 2 that a smaller value of γ results in a slower convergence to the GNE set (the residual) (see the top plot of Figure 2). This is due to the fact that a smaller value of γ implies a slower convergence of the diminishing step size $\beta^{(k)}$ to 0. This increases the weight of the gradient descent steps in (11)–(12) during the transient, which further slows the convergence down. On the other hand, it is observed that a high value of γ, despite having a fast residual convergence, might result in a slow convergence of the objective function $\phi(\boldsymbol{x}^{(k)})$ to the optimal value (see the bottom plot of Figure 2). This trade-off suggests that a careful choice of the step size $\beta^{(k)}$ is crucial for the performance of the algorithm.

V. Conclusion

For linearly coupled aggregative games with multiple equilibria, it is possible to select a particular solution that optimizes a convex preference function. The equilibrium selection can be achieved via a semi-decentralized computation

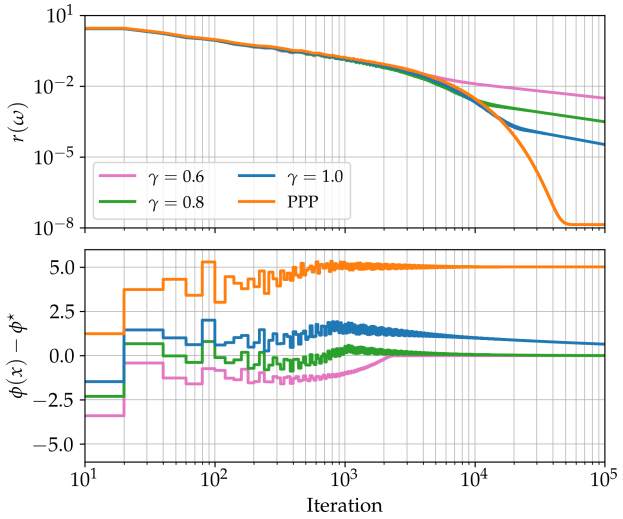


Fig. 2: Comparison of Algorithm 1 for various values of γ with the standard PPP algorithm in terms of the convergence of the residual r (top plot) and the objective function ϕ (bottom plot). Here, $\phi^\star$ is the optimal ϕ computed by Algorithm 1 for $\gamma = 0.6$ (which obtains the minimum cost). Each line represents the average of simulation results from all the 20 GNE selection problems with 20 randomly sampled initial conditions.

by combining an instance of the preconditioned proximal point algorithm with the hybrid steepest descent method. This framework can be exploited to enforce a system-level objective among the set of equilibrium strategies. We identify as future work the extension of the results to more general monotone games, the characterization of the convergence rate, and the dependency of the latter on the design parameters.

Appendix

A. Proof of Lemma 1

The matrix Γ is positive definite under Assumptions 2 and 5, following the generalized Gerschgorin disc theorem [28, Thm. 2] as in [7, Lemma 8]. It can be proven that this, together with the maximal monotonicity of $\mathcal{T}_{\mathrm{e}}$, implies that $\Gamma^{-1}\mathcal{T}_{\mathrm{e}}$ is maximally monotone in the Γ-induced norm $\|\cdot\|_\Gamma$, and the proof follows verbatim the one of [21, Lemma 7(*ii*)]. By noting that $\mathcal{T}_{\mathrm{PPP}}$ is the resolvent of the operator $\Gamma^{-1}\mathcal{T}_{\mathrm{e}}$ as defined in [6, Def. 23.1], it follows from [6, Prop. 23.8] that $\mathcal{T}_{\mathrm{PPP}}$ is firmly nonexpansive. From the existence of a v-GNE [20, Prop. 12.11], $\mathrm{fix}(\mathcal{T}_{\mathrm{PPP}}) \neq \varnothing$ and, therefore, the thesis follows immediately by applying [6, Remark 4.36]. ■

B. Proof of Lemma 2

By Assumption 5, Γ is positive definite and $\Gamma^{-1}\mathcal{T}_{\mathrm{e}}$ is maximally monotone (see Appendix A). Therefore, it holds that [6, Prop. 23.38]:

$$\boldsymbol{\omega} \in \mathrm{zer}(\mathcal{T}_{\mathrm{e}}) \Leftrightarrow \boldsymbol{\omega} \in \mathrm{fix}(\mathcal{T}_{\mathrm{PPP}}). \tag{18}$$

Since ϕ is differentiable by Assumption 4, the claim follows immediately as (15) is the stationary point problem associated to (8) [29, Sec. 1.3.1] . ■

C. Proof of Lemma 3

Let $\boldsymbol{y} := [\boldsymbol{p}, d]^\top = \mathcal{T}_{\mathrm{PPP}}(\boldsymbol{\omega}^{(k)})$. From (14), we obtain

$$\mathcal{T}_{\mathrm{e}}(\boldsymbol{y}) \ni \Gamma(\boldsymbol{\omega}^{(k)} - \boldsymbol{y}).$$

Substituting the definition of $\mathcal{T}_{\mathrm{e}}$, we obtain, for $\boldsymbol{p}$,

$$\left(\prod_{i\in\mathcal{I}} \mathrm{N}_{\mathcal{X}_i} + F\right)(\boldsymbol{p}) \ni (\rho - \boldsymbol{C})(\boldsymbol{x}^{(k)} - \boldsymbol{p}) - \boldsymbol{A}^\top \lambda^{(k)}. \quad (19)$$

By the definition of J_i in (1) and rearranging (19), we obtain

$$\forall\, i \in \mathcal{I}:\ \ p_i + \left(\mathrm{N}_{\mathcal{X}_i} + \rho_i^{-1}\partial \ell_i\right)(p_i) \ni x_i^{(k)} - \rho_i^{-1}\left(\textstyle\sum_{j\in\mathcal{I}\setminus\{i\}} C_{ij}x_j^{(k)} + A_i^\top \lambda^{(k)}\right).$$

Following $\mathrm{N}_{\mathcal{X}_i} = \partial \iota_{\mathcal{X}_i}$ and [6, Prop. 16.44], then the update in (13), for each agent $i \in \mathcal{I}$, reads as

$$p_i = \mathrm{prox}_{\iota_{\mathcal{X}_i} + \rho_i^{-1}\ell_i}\left(x_i^{(k)} - \rho_i^{-1}(A_i^\top \lambda^{(k)} + \textstyle\sum_{j\in\mathcal{I}\setminus\{i\}} C_{ij}x_j^{(k)})\right),$$

which is equivalent to the update in (9) by the definition of proximal operator; thus, $\boldsymbol{p} = \mathring{\boldsymbol{x}}^{(k)}$. We similarly prove that the expression for $d = \mathring{\lambda}^{(k)}$ is equivalent to the update in (10). We can then finally observe that (11) and (12), for all $i \in \mathcal{I}$, are an expanded form of (16). ■

D. Proof of Theorem 1

We note that $\boldsymbol{\Omega}$ is bounded by Assumption 1. Thus, the set of v-GNEs is bounded as it is a subset of $\boldsymbol{\Omega}$ and it is nonempty under Assumption 3 [20, Prop. 12.11]. Under Assumption 1, the set of dual variables that solve (7) is bounded [22, Prop. 3.3], thus $\mathrm{zer}(\mathcal{T}_{\mathrm{e}})$ (as well as $\mathrm{fix}(\mathcal{T}_{\mathrm{PPP}})$, from (18)) is nonempty and bounded. From Lemma 1, $\mathcal{T}_{\mathrm{PPP}}$ is attracting nonexpansive. Therefore, the iteration in (16), which is equivalent to Algorithm 1 by Lemma 3, satisfies all the assumptions of [30, Thm. 3] and the thesis follows. ■

REFERENCES

[1] G. Belgioioso, P. Yi, S. Grammatico, and L. Pavel, "Distributed generalized Nash equilibrium seeking: An operator-theoretic perspective," *IEEE Control Systems Magazine*, vol. 42, no. 4, pp. 87–102, 2022.

[2] P. Zhou, W. Wei, K. Bian, D. O. Wu, Y. Hu, and Q. Wang, "Private and truthful aggregative game for large-scale spectrum sharing," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 2, pp. 463–477, 2017.

[3] G. Belgioioso, W. Ananduta, S. Grammatico, and C. Ocampo-Martinez, "Operationally-safe peer-to-peer energy trading in distribution grids: A game-theoretic market-clearing mechanism," *IEEE Transactions on Smart Grid*, vol. 13, no. 4, pp. 2897–2907, 2022.

[4] Z. Liu, Q. Wu, S. Huang, L. Wang, M. Shahidehpour, and Y. Xue, "Optimal day-ahead charging scheduling of electric vehicles through an aggregative game model," *IEEE Transactions on Smart Grid*, vol. 9, pp. 5173–5184, 2018.

[5] F. Facchinei and C. Kanzow, "Generalized Nash equilibrium problems," *Annals of Operations Research*, vol. 175, no. 1, pp. 177–211, 2010.

[6] H. H. Bauschke and P. L. Combettes, *Convex analysis and monotone operator theory in Hilbert spaces*. Springer, 2011.

[7] G. Belgioioso and S. Grammatico, "Semi-decentralized generalized Nash equilibrium seeking in monotone aggregative games," *IEEE Transactions on Automatic Control*, 2021. available at https://doi.org/10.1109/TAC.2021.3135360.

[8] G. Belgioioso and S. Grammatico, "Projected-gradient algorithms for generalized equilibrium seeking in aggregative games are preconditioned forward-backward methods," pp. 2188–2193, 2018.

[9] J. Koshal, A. Nedić, and U. V. Shanbhag, "Distributed algorithms for aggregative games on graphs," *Operations Research*, vol. 64, no. 3, pp. 680–704, 2016.

[10] S. Liang, P. Yi, and Y. Hong, "Distributed Nash equilibrium seeking for aggregative games with coupled constraints," *Automatica*, vol. 85, pp. 179–185, 2017.

[11] D. Gadjov and L. Pavel, "Single-timescale distributed GNE seeking for aggregative games over networks via forward–backward operator splitting," *IEEE Transactions on Automatic Control*, vol. 66, no. 7, pp. 3259–3266, 2020.

[12] G. Belgioioso and S. Grammatico, "A distributed proximal-point algorithm for Nash equilibrium seeking in generalized potential games with linearly coupled cost functions," in *2019 18th European Control Conference (ECC)*, pp. 1–6, 2019.

[13] G. Belgioioso, A. Nedić, and S. Grammatico, "Distributed generalized Nash equilibrium seeking in aggregative games on time-varying networks," *IEEE Transactions on Automatic Control*, vol. 66, no. 5, pp. 2061–2075, 2020.

[14] B. Franci, M. Staudigl, and S. Grammatico, "Distributed forward-backward (half) forward algorithms for generalized Nash equilibrium seeking," in *Proceedings of 2020 European Control Conference (ECC)*, pp. 1274–1279, IEEE, 2020.

[15] H. Yin, U. V. Shanbhag, and P. G. Mehta, "Nash equilibrium problems with scaled congestion costs and shared constraints," *IEEE Transactions on Automatic Control*, vol. 56, no. 7, pp. 1702–1708, 2011.

[16] J. Lei, U. V. Shanbhag, and J. Chen, "Distributed computation of Nash equilibria for monotone aggregative games via iterative regularization," *59th IEEE Conference on Decision and Control (CDC)*, pp. 2285–2290, 2020.

[17] X. Chen, E. Dall'Anese, C. Zhao, and N. Li, "Aggregate power flexibility in unbalanced distribution systems," *IEEE Transactions on Smart Grid*, vol. 11, no. 1, pp. 258–269, 2020.

[18] I. Yamada and N. Ogura, "Hybrid steepest descent method for variational inequality problem over the fixed point set of certain quasi-nonexpansive mappings," *Numerical Functional Analysis and Optimization*, vol. 25, no. 7-8, pp. 619–655, 2005.

[19] P. Yi and L. Pavel, "Distributed generalized Nash equilibria computation of monotone games via double-layer preconditioned proximal-point algorithms," *IEEE Transactions on Control of Network Systems*, vol. 6, no. 1, pp. 299–311, 2018.

[20] D. P. Palomar and Y. C. Eldar, *Convex optimization in signal processing and communications*. Cambridge university press, 2010.

[21] P. Yi and L. Pavel, "An operator splitting approach for distributed generalized Nash equilibria computation," *Automatica*, vol. 102, pp. 111–121, 2019.

[22] A. Auslender and M. Teboulle, "Lagrangian duality and related multiplier methods for variational inequality problems," *SIAM Journal on Optimization*, vol. 10, no. 4, pp. 1097–1115, 2000.

[23] A. Dreves, "How to select a solution in generalized Nash equilibrium problems," *Journal of Optimization Theory and Applications*, vol. 178, no. 3, pp. 973–997, 2018.

[24] A. Dreves, "An algorithm for equilibrium selection in generalized Nash equilibrium problems," *Computational Optimization and Applications*, vol. 73, no. 3, pp. 821–837, 2019.

[25] I. Yamada, "The hybrid steepest descent method for the variational inequality problem over the intersection of fixed point sets of nonexpansive mappings," in *Inherently Parallel Algorithms in Feasibility and Optimization and their Applications* (D. Butnariu, Y. Censor, and S. Reich, eds.), vol. 8 of *Studies in Computational Mathematics*, pp. 473–504, Elsevier, 2001.

[26] A. Cegielski, A. Gibali, S. Reich, and R. Zalas, "An algorithm for solving the variational inequality problem over the fixed point set of a quasi-nonexpansive operator in euclidean space," *Numerical Functional Analysis and Optimization*, vol. 34, no. 10, pp. 1067–1096, 2013.

[27] S. Hirstoaga, "Iterative selection method for common fixed point problems," *Journal of Mathematical Analysis and Applications*, vol. 324, no. 2, pp. 1020–1035, 2006.

[28] D. G. Feingold and R. S. Varga, "Block diagonally dominant matrices and generalizations of the Gerschgorin circle theorem," *Pacific Journal of Mathematics*, vol. 12, no. 4, pp. 1241–1250, 1962.

[29] F. Facchinei and J.-S. Pang, *Finite-dimensional variational inequalities and complementarity problems*. Springer, 2007.

[30] N. Ogura and I. Yamada, "Nonstrictly convex minimization over the bounded fixed point set of a nonexpansive mapping," *Numerical Functional Analysis and Optimization*, vol. 24, no. 1-2, pp. 129–135, 2003.