

# Linear convergence in time-varying generalized Nash equilibrium problems

Mattia Bianchi, Emilio Benenati, Sergio Grammatico

**Abstract**—We study generalized games with full row rank equality coupling constraints and we provide a strikingly simple proof of strong monotonicity of the associated KKT operator. This allows us to show linear convergence to a variational equilibrium of the resulting primal-dual pseudo-gradient dynamics. Then, we propose a fully-distributed algorithm with linear convergence guarantee for aggregative games under partial-decision information. Based on these results, we establish stability properties for online GNE seeking in games with time-varying cost functions and constraints. Finally, we illustrate our findings numerically on an economic dispatch problem for peer-to-peer energy markets.

## I. INTRODUCTION

Generalized Nash equilibrium (GNE) problems arise in many multi-agent applications, where the agents are coupled not only because of their conflicting objectives, but also via shared constraints – operational limits of the system, that the agents should respect. Among others, GNE seeking is used in energy markets [1], radio communication [2] and formation control [3] problems.

The networked structure of these applications naturally calls for distributed solution methods. In fact, part of the recent literature focuses on semi-decentralized GNE seeking algorithms [4], [5], [6], where the agents update their decision locally, with the help of a coordinator that gathers and broadcasts information over the systems (a setup also named full-information scenario). Other works [7], [8], [9], [10] deal with applications where the agents can only rely on fully-distributed peer-to-peer communication and local data. In this so-called partial-decision information scenario, the agents compensate for the lack of global knowledge by estimating the unknown quantities and by embedding consensus dynamics in their local decision processes.

In both scenarios, to cope with the presence of coupling constraints and to distribute the computation among the agents, one should resort to Lagrangian reformulations. In fact, all the references above leverage primal-dual pseudo-gradient algorithms, aimed at solving the Karush–Kuhn–Tucker (KKT) optimality conditions of the GNE problem.

In general, primal-dual algorithms fail to achieve linear convergence, even for the class of strongly-monotone generalized games [4]. Importantly, together with linear convergence, some crucial input-to-state stability (ISS) properties of pseudo-gradient iterations are also not guaranteed. This lack

of robustness is a critical issue for methods in the partial-decision information scenario, where convergence should be ensured despite the estimation error. To overcome this complication, vanishing step sizes can be used to drive the error to zero [11], at the price of slow convergence. Alternatively, several fixed-step algorithms for GNE seeking were derived based on operator-theoretic methods and on the use of preconditioning [8], [9], [12], [7]. Unfortunately, this approach comes with important limitations, such as extending the analysis to time-varying setups.

For instance, there is no available fixed-step fully-distributed method to solve GNE problems when the agents can only exchange information over switching communication networks (while methods are available for games without coupling constraints [13], [14]). Furthermore, in many decision processes with real-world applications, the cost functions of the agents and the system constraints can vary over time [15], for instance in cognitive radio networks and demand response in smart grids [16]. In such domains, linearly convergent algorithms become particularly desirable, as the solver needs to quickly update its solution in response to changes in the environment. Despite its practical relevance, there are very few works that study the online GNE problem. The paper [17] proposes a regularized algorithm, which only achieves inexact convergence, and which is not fully-distributed. Instead, the authors of [18] develop an algorithm for the partial-decision information scenario, but that achieves sublinear regret only when the solution is asymptotically constant and for diminishing step sizes.

*Contribution:* In this paper we study generalized games with full row rank coupling equality constraints –as those arising in resource allocation and transportation problems [19], where demand-matching [10] and flow [1] constraints are ubiquitous. For the first time, we show that, in this setup, linear convergence to a GNE can be achieved via primal-dual dynamics, both for the full- and partial-decision information scenario. Thanks to this result, we can also adapt the dynamics to online equilibrium seeking in time-varying games. Here we focus on the prominent class of aggregative games [4], for its desirable scalability properties, but the analysis carries over to generally-coupled costs. We summarize the novelties of our work as follows:

- 1) We provide a simple, constructive proof of the strong monotonicity of the KKT operator in games with full-row rank equality coupling constraints. As a consequence, we show linear convergence to a GNE of the pseudo-gradient ascent-descent method (Section III);
- 2) We design a linearly convergent algorithm for GNE seeking in partial-decision information, via a tracking

E. Benenati and S. Grammatico are with the the Delft Center for Systems and Control, TU Delft, The Netherlands. M. Bianchi is with the Automatic Control Laboratory, ETH Zürich, Switzerland. E-mail: {e.benenati, s.grammatico}@tudelft.nl, mbianch@ethz.ch. This work was supported by the ERC under research project COSMOS (802348) and by ETH Zürich funds.

technique [11] that avoids the need for slack variables. Our proof is based on a change of coordinates and a small gain argument: due to its generality, the argument also applies to the case of (Q-connected) time-varying communication graphs (Section IV);

- 3) We exploit our linear convergence results to study the tracking properties of the proposed methods with respect to the solution of a game with time-varying costs and constraints. In particular, for the fully-distributed algorithm, we show that the extra error induced in the dynamic tracking procedure does not jeopardize stability (Section V)

*Notation:*  $\mathbf{0}_n$  ( $\mathbf{1}_n$ ) denotes the vector of dimension  $n$  with all elements equal to 0 (1);  $I_n$  the identity matrix of dimension  $n$ ; the subscripts are omitted when there is no ambiguity. If  $A$  is symmetric,  $\lambda_{\min}(A) = \lambda_1(A) \leq \dots \leq \lambda_n(A) =: \lambda_{\max}(A)$  denote its eigenvalues.  $\otimes$  denotes the Kronecker product.  $\text{diag}(A_1, \dots, A_N)$  denotes the block diagonal matrix with  $A_1, \dots, A_N$  on its diagonal;  $\text{col}(x_1, \dots, x_N) = [x_1^\top \dots x_N^\top]^\top$ . For a positive definite symmetric matrix  $P \succ 0$ ,  $\langle x, y \rangle_P = x^\top P y$  denotes the  $P$ -weighted inner product,  $\|x\|_P$  the corresponding norm; we omit the subscript if  $P = I$ . An operator  $\mathcal{F} : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is ( $\mu$ -strongly) monotone in  $\mathcal{H}_P$  if, for any  $x, y \in \mathbb{R}^n$ ,  $\langle \mathcal{F}(x) - \mathcal{F}(y), x - y \rangle_P \geq 0$  ( $\geq \mu \|x - y\|_P^2$ ); is contractive in  $\mathcal{H}_P$  if it is Lipschitz with constant smaller than 1, i.e., for some  $\ell < 1$  and for any  $x, y \in \mathbb{R}^n$ ,  $\|\mathcal{F}(x) - \mathcal{F}(y)\|_P \leq \ell \|x - y\|_P$ ; we omit the indication “in  $\mathcal{H}_P$ ” if  $P = I$ . If  $\mathcal{F}$  is differentiable,  $D\mathcal{F}$  denotes its Jacobian.

## II. MATHEMATICAL SETUP

We consider a set of agents,  $\mathcal{I} := \{1, \dots, N\}$ , where each agent  $i \in \mathcal{I}$  shall choose its decision variable (i.e., strategy)  $x_i \in \mathbb{R}^{n_i}$ . Let  $x := \text{col}((x_i)_{i \in \mathcal{I}}) \in \mathbb{R}^n$ , with  $n := \sum_{i=1}^N n_i$ . The goal of each agent  $i \in \mathcal{I}$  is to minimize its objective function  $J_i(x_i, x_{-i})$ , which depends on both the local variable  $x_i$  and on the decision variables of the other agents  $x_{-i} := \text{col}((x_j)_{j \in \mathcal{I} \setminus \{i\}})$ . Furthermore, the feasible decisions of each agent  $i$  depends on the action of the other agents via affine equality coupling constraints. Specifically, the feasible set is  $\mathcal{X} := \{x \in \mathbb{R}^n \mid Ax = b\}$ , where  $A := [A_1, \dots, A_N]$  and  $b := \sum_{i=1}^N b_i$ ,  $A_i \in \mathbb{R}^{m \times n_i}$  and  $b_i \in \mathbb{R}^m$  being locally available information. The game is then represented by the inter-dependent optimization problems:

$$(\forall i \in \mathcal{I}) \quad \min_{y_i \in \mathbb{R}^{n_i}} J_i(y_i, x_{-i}) \quad \text{s.t. } (y_i, x_{-i}) \in \mathcal{X}. \quad (1)$$

The technical problem we consider here is the computation of a GNE, namely a set of decisions that simultaneously solve all the optimization problems in (1).

*Definition 1:* A collective strategy  $x^* = \text{col}((x_i^*)_{i \in \mathcal{I}})$  is a generalized Nash equilibrium if, for all  $i \in \mathcal{I}$ ,  $J_i(x_i^*, x_{-i}^*) \leq \inf\{J_i(y_i, x_{-i}^*) \mid (y_i, x_{-i}^*) \in \mathcal{X}\}$ .  $\square$

Next, we postulate some standard regularity and convexity assumptions for the constraint sets and cost functions.

*Assumption 1 (Convexity):* In (1),  $\mathcal{X}$  is non-empty. For each  $i \in \mathcal{I}$ ,  $J_i$  is continuous and  $J_i(\cdot, x_{-i})$  is convex and continuously differentiable for every  $x_{-i}$ .  $\square$

As common in the literature [20], [4], among all the GNEs, we focus on the subclass of variational GNEs (v-GNEs) [2, Def. 3.11], which are more economically justifiable, as well as computationally tractable [21]. Under Assumption 1 and defining the *pseudo-gradient* mapping of the game

$$F(x) := \text{col}((\nabla_{x_i} J_i(x_i, x_{-i}))_{i \in \mathcal{I}}), \quad (2)$$

$x^*$  is a v-GNE of the game in (1) if and only if there exists a dual variable  $\lambda^* \in \mathbb{R}^m$  such that the following KKT conditions are satisfied [2, Th. 4.8]:

$$\mathbf{0}_n \in F(x^*) + A^\top \lambda^*, \quad \mathbf{0}_m \in -(Ax^* - b). \quad (3)$$

Let us restrict our attention to strongly monotone games.

*Assumption 2 (Strong monotonicity):* The game mapping  $F$  in (2) is  $\mu_F$ -strongly monotone and  $\ell_F$ -Lipschitz continuous, for some  $\mu_F, \ell_F > 0$ .  $\square$

The strong monotonicity of  $F$  is sufficient to ensure existence and uniqueness of a v-GNE [22, Th. 2.3.3]; it implies strong convexity of each function  $J_i(\cdot, x_{-i})$  for any fixed  $x_{-i}$ , but not joint convexity of  $J_i(\cdot)$ . Strong monotonicity is a standard condition for algorithms with linear convergence. In addition, we make the following assumption.

*Assumption 3 (Full rank constraints):*  $A$  is full row rank.  $AA^\top \geq \mu_A I$ ,  $\|A\| \leq \ell_A$  for some scalars  $\mu_A, \ell_A > 0$ .  $\square$  Assumption 3 postulates that there are no redundant constraints (or equivalently that redundant constraints are removed). This condition is well known in duality theory and optimization, as it ensures the uniqueness of dual solutions (as it can be inferred by (3)).

### A. Aggregative games

For ease of presentation, we will specialize our results to the prominent class of aggregative games, which arises in a variety of engineering applications, e.g., network congestion control and demand-side management [23]. In particular, we assume that the cost function  $J_i$  of each agent  $i$  depends only on the local decision  $x_i$  and on an aggregation value

$$\sigma(x) = \frac{1}{N} \sum \phi_i(x_i), \quad (4)$$

where  $\phi_i : \mathbb{R}^{n_i} \rightarrow \mathbb{R}^q$  is a local function of agent  $i$ . In short, overloading the function  $J_i$  with some abuse of notation, we also write

$$J_i(x_i, x_{-i}) = J_i(x_i, \sigma(x)). \quad (5)$$

## III. LINEAR CONVERGENCE IN GENERALIZED GAMES

We start by showing that the KKT operator is strongly monotone, in a suitable norm, under Assumptions 2 and 3.

*Lemma 1:* Let  $\omega := \text{col}(x, \lambda)$ . The operator

$$\omega \mapsto \mathcal{A}(\omega) = \begin{bmatrix} F(x) + A^\top \lambda \\ -Ax + b \end{bmatrix} \quad (6)$$

is  $\mu_A$ -strongly monotone in  $\mathcal{H}_P$ , for some  $P \succ 0$  and  $\mu_A > 0$ .  $\square$

*Proof:* For some  $0 < \nu < \ell_A$ , let

$$P = \begin{bmatrix} I & \nu A^\top \\ \nu A & I \end{bmatrix} \succ 0. \quad (7)$$

---

**Algorithm 1** Semi-decentralized GNE seeking

---

Iterate to convergence: for all  $k \in \mathbb{N}$ ,

- Each  $i \in \mathcal{I}$ : receive  $\lambda^k, \sigma^k$  from coordinator; update

$$x_i^{k+1} = x_i^k - \alpha (\nabla_{x_i} J_i(x_i^k, \sigma^k) + A_i^\top \lambda^k)$$

- Coordinator: receive  $\{\phi_i(x_i^{k+1}), A_i x_i^k - b_i\}_{i \in \mathcal{I}}$ ; update

$$\lambda^{k+1} = \lambda^k + \alpha \sum_{i \in \mathcal{I}} (A_i x_i^k - b_i),$$

$$\sigma^{k+1} = \frac{1}{N} \sum_{i \in \mathcal{I}} \phi_i(x_i^{k+1})$$

---

For any  $x, x^* \in \mathbb{R}^n$  and  $\lambda, \lambda^* \in \mathbb{R}^m$ , we have

$$\begin{aligned} & \langle \mathcal{A}(\omega) - \mathcal{A}(\omega^*), \omega - \omega^* \rangle_P \\ &= \langle F(x) - F(x^*), x - x^* \rangle \\ & \quad + \nu \langle F(x) - F(x^*), A^\top (\lambda - \lambda^*) \rangle \\ & \quad + \langle A^\top (\lambda - \lambda^*), x - x^* \rangle \\ & \quad + \nu \langle A^\top (\lambda - \lambda^*), A^\top (\lambda - \lambda^*) \rangle \\ & \quad - \nu \langle A(x - x^*), A(x - x^*) \rangle - \langle A(x - x^*), \lambda - \lambda^* \rangle \\ & \geq (\mu_F - \nu \ell_A^2) \|x - x^*\|^2 \\ & \quad - \nu \ell_F \ell_A \|x - x^*\| \|\lambda - \lambda^*\| + \nu \mu_A \|\lambda - \lambda^*\|^2 \end{aligned}$$

which is positive definite for  $0 < \nu < \frac{4\mu_F\mu_A}{\ell_F^2\ell_A^2 + 4\mu_A\ell_A^2}$ . The conclusion follows by equivalence of norms. ■

Based on Lemma 1 we can prove linear convergence of classic primal-dual iterations.

*Theorem 1 (GNE seeking in full-decision information):*

Let Assumptions 1-3 hold and  $\ell_A := (\ell_F + \ell_A) \sqrt{\frac{\lambda_{\max}(P)}{\lambda_{\min}(P)}}$ . For any  $0 < \alpha < \frac{2\mu_A}{\ell_A^2}$  and any initial condition  $(x^0, \lambda^0)$ , the iteration

$$x^{k+1} = x^k - \alpha (F(x^k) + A^\top \lambda^k) \quad (8a)$$

$$\lambda^{k+1} = \lambda^k + \alpha (Ax^k - b) \quad (8b)$$

converges linearly to the unique solution  $\omega^* = (x^*, \lambda^*)$  of the KKT conditions in (3): for all  $k \in \mathbb{N}$

$$\|\omega^{k+1} - \omega^*\|_P^2 \leq \rho \|\omega^k - \omega^*\|_P^2, \quad (9)$$

where  $\rho = 1 - 2\alpha\mu_A + \alpha^2\ell_A^2 < 1$ . □

*Proof:* By Lemma 1,  $\|\omega - \alpha\mathcal{A}(\omega) - (\omega^* - \alpha\mathcal{A}(\omega^*))\|_P^2 = \|\omega - \omega^*\|_P^2 - 2\alpha\langle \mathcal{A}(\omega) - \mathcal{A}(\omega^*), \omega - \omega^* \rangle_P + \alpha^2\|\mathcal{A}(\omega) - \mathcal{A}(\omega^*)\|_P^2 \leq \rho\|\omega - \omega^*\|_P^2$ , and the conclusion follows because (8) can be rewritten as  $\omega^{k+1} = \omega^k - \alpha\mathcal{A}(\omega^k)$ . ■

For the case of aggregative games, (8a) can be implemented in a semi-decentralized way, as in Algorithm 1.

#### IV. PARTIAL-DECISION INFORMATION

In this section, we consider aggregative games in the so-called partial-decision information scenario, where there is no central coordinator, and the agents can only exchange information via peer-to-peer communication over a communication graph  $\mathcal{G} = (\mathcal{I}, \mathcal{E})$ , with weight matrix  $W \in \mathbb{R}^{N \times N}$ , and  $w_{i,j} := [W]_{i,j} > 0$  if and only if  $(j, i)$  belongs to the set of edges  $\mathcal{E} \subseteq \mathcal{I} \times \mathcal{I}$ .

*Assumption 4 (Communication):* The graph  $\mathcal{G}$  is strongly connected. The weight matrix  $W$  satisfies:

---

**Algorithm 2** Fully-distributed GNE seeking

---

Initialization: choose  $\alpha > 0$  as in Theorem 2; for all  $i \in \mathcal{I}$ , set  $x_i^0 \in \mathbb{R}^{n_i}$ ,  $\sigma_i^0 = \phi_i(x_i^0)$ ,  $z_i^0 \in \mathbb{R}^m$ ,  $\lambda_i^0 = z_i^0$ ,  $r_i^0 = A_i x_i^0 - b_i$ .

Iterate to convergence: for all  $k \in \mathbb{N}$ , for all  $i \in \mathcal{I}$ ,

- Local variables update:

$$x_i^{k+1} = x_i^k - \alpha \left( F_i(x_i, \sigma_i^k) + A_i^\top \lambda_i^k \right)$$

$$z_i^{k+1} = z_i^k + \alpha N r_i^k$$

- Tracking: Agent  $i$  exchanges the variables  $(\sigma_i, \lambda_i, r_i)$  with its neighbors, and does

$$\sigma_i^{k+1} = \sum_{j \in \mathcal{N}_i} w_{i,j} \sigma_j^k + \phi_i(x_i^{k+1}) - \phi_i(x_i^k)$$

$$r_i^{k+1} = \sum_{j \in \mathcal{N}_i} w_{i,j} r_j^k + A_i x_i^{k+1} - A_i x_i^k$$

$$\lambda_i^{k+1} = \sum_{j \in \mathcal{N}_i} w_{i,j} \lambda_j^k + z_i^{k+1} - z_i^k$$

---

- Double stochasticity:  $\mathbf{1}^\top W = \mathbf{1}^\top, W\mathbf{1} = W\mathbf{1}$ ;

- Self-loops:  $w_{i,i} > 0$  for all  $i \in \mathcal{I}$ .

We denote  $\theta := \|W - \frac{1}{N}\mathbf{1}\mathbf{1}^\top\| < 1$ . □

To remedy the lack of global knowledge, we let each agent  $i \in \mathcal{I}$  keep:

- $\sigma_i \in \mathbb{R}^q$ : estimate of the aggregation  $\sigma(x)$ ;
- $\lambda_i \in \mathbb{R}^m$ : estimate of the dual variable  $\lambda$ ;
- $r_i \in \mathbb{R}^m$ : estimate of the residual  $Ax - b$ ;
- $z_i \in \mathbb{R}^m$ : additional dual variable.

We make the following standard assumption.

*Assumption 5:* For each  $i \in \mathcal{I}$ , the function  $\phi_i$  in (4) is continuously differentiable and  $\ell_\sigma$ -Lipschitz continuous; furthermore, the mapping

$$F_i(x_i, \sigma_i) := \nabla_{x_i} J_i(x_i, \sigma_i) + \frac{1}{N} [D\phi_i(x_i)]^\top \nabla_{\sigma_i} J_i(x_i, \sigma_i) \quad (10)$$

is  $\tilde{\ell}_\sigma$ -Lipschitz continuous with respect to  $\sigma_i$ . □

Our proposed dynamics are illustrated in Algorithm 2.

Let us define  $\sigma = \text{col}((\sigma_i)_{i \in \mathcal{I}})$ ,  $r = \text{col}((r_i)_{i \in \mathcal{I}})$ ,  $\lambda = \text{col}((\lambda_i)_{i \in \mathcal{I}})$  and the extended game mapping

$$F(x, \sigma) := \text{col}((F_i(x_i, \sigma_i))_{i \in \mathcal{I}}). \quad (11)$$

Note that  $F(x, \mathbf{1} \otimes \sigma(x)) = F(x)$ . Furthermore, let

$$\bar{\lambda} = \frac{1}{N} \sum_{i \in \mathcal{I}} \lambda_i, \quad \tilde{\lambda} = \lambda - \mathbf{1} \otimes \bar{\lambda} \quad (12a)$$

$$\bar{\sigma} = \frac{1}{N} \sum_{i \in \mathcal{I}} \sigma_i, \quad \tilde{\sigma} = \sigma - \mathbf{1} \otimes \bar{\sigma} \quad (12b)$$

$$\bar{r} = \frac{1}{N} \sum_{i \in \mathcal{I}} r_i, \quad \tilde{r} = r - \mathbf{1} \otimes \bar{r}. \quad (12c)$$

The following lemma shows an invariance property typical of tracking dynamics, as the one in Algorithm 2.

*Lemma 2:* For all  $k \in \mathbb{N}$ , it holds that  $\bar{\lambda}^k = \frac{1}{N} \sum_{i \in \mathcal{I}} z_i^k$ ,  $\bar{r}^k = \frac{1}{N} \sum_{i \in \mathcal{I}} (A_i x_i^k - b_i)$ ,  $\bar{\sigma}^k = \sigma(x^k)$ . □

*Proof:* Via induction, by the initialization and double stochasticity of  $W$ . ■

To study the convergence of Algorithm 2, we first need the following crucial reformulation.

*Lemma 3:* The iteration in Algorithm 2 is equivalent to

$$\begin{aligned} \begin{bmatrix} x^{k+1} \\ \bar{\lambda}^{k+1} \end{bmatrix} &= \underbrace{\xi^k - \alpha \mathcal{A}(\xi^k)}_{:=\mathcal{B}_1(\xi^k)} + \underbrace{\left[ \alpha \left( F(x^k, \mathbf{1} \otimes \bar{\sigma}^k) - F(x^k, \sigma^k) + A^\top \bar{\lambda}^k \right) \right]}_{:=\mathcal{B}_2(\omega^k)} \\ \begin{bmatrix} \bar{\sigma}^{k+1} \\ \tilde{r}^{k+1} \\ \tilde{\lambda}^{k+1} \end{bmatrix} &= \underbrace{\begin{bmatrix} W \bar{\sigma}^k \\ W \tilde{r}^k \\ W \tilde{\lambda}^k \end{bmatrix}}_{:=\mathcal{B}_3(\chi^k)} + \underbrace{\begin{bmatrix} \tilde{\Pi} \text{col}((\phi_i(x_i^{k+1}) - \phi_i(x_i^k))_{i \in \mathcal{I}}) \\ \tilde{\Pi} \text{col}((A_i(x_i^{k+1} - x_i^k))_{i \in \mathcal{I}}) \\ \alpha N \tilde{r}^k \end{bmatrix}}_{:=\mathcal{B}_4(\omega^k)} \end{aligned} \quad (13)$$

where  $\omega = (\xi, \chi)$ ,  $\xi = (x, \bar{\lambda})$ ,  $\chi = (\bar{\sigma}, \tilde{r}, \tilde{\lambda})$ , and  $\tilde{\Pi} := I - (\frac{1}{N} \mathbf{1} \mathbf{1}^\top) \otimes I_N$ ,  $W = W \otimes I$ ,  $A = \text{diag}((A_i)_{i \in \mathcal{I}})$ : the sequence  $(x^k, \sigma^k, r^k, \lambda^k)_{k \in \mathbb{N}}$  generated by Algorithm 2 and the sequence  $(x^k, \bar{\sigma}^k + \mathbf{1} \otimes \sigma(x^k), \tilde{r}^k + \mathbf{1} \otimes (Ax^k - b), \bar{\lambda}^k + \tilde{\lambda}^k)_{k \in \mathbb{N}}$  generated by (13) coincide.  $\square$

*Proof:* The update of  $\chi$  follows by noting that  $\bar{\sigma} = \tilde{\Pi} \sigma$ ,  $\tilde{r} = \tilde{\Pi} r$ ,  $\tilde{\lambda} = \tilde{\Pi} \lambda$  and  $\tilde{\Pi} W \tilde{\Pi} = W \tilde{\Pi}$ . Then, the proof follows by definition of  $\mathcal{A}$ , by using Lemma 2, and finally by noting that  $F(x^k, \mathbf{1} \otimes \bar{\sigma}^k) = F(x^k)$  by Lemma 2. Note that  $\sigma = \bar{\sigma} + \mathbf{1} \otimes \sigma(x)$  and  $r = \tilde{r} + \frac{1}{N} \otimes (Ax - b)$ , which allows us to eliminate the variables  $\bar{\sigma}$  and  $\tilde{r}$  in the iteration, and similarly for  $z_i$ , since  $\bar{\lambda}^{k+1} = \bar{\lambda}^k + \frac{1}{N} \sum_{i \in \mathcal{I}} (z_i^{k+1} - z_i^k) = \bar{\lambda}^k + \frac{1}{N} \sum_{i \in \mathcal{I}} (\alpha N r_i^k) = \bar{\lambda}^k + \alpha (Ax^k - b)$  by Lemma 2.  $\blacksquare$

*Theorem 2 (GNE seeking in partial-decision information):*

Let Assumptions 1-4 hold. Then, there is  $\alpha_{\max} > 0$  such that, for all  $0 < \alpha < \alpha_{\max}$ , the iteration in (13) converges linearly to  $\omega^* = (\xi^*, \mathbf{0})$ , with  $\xi^* = (x^*, \lambda^*)$ : for all  $k \in \mathbb{N}$

$$V(\omega^{k+1}) \leq \eta V(\omega^k),$$

where  $V(\omega) := \|\xi - \xi^*\|_P^2 + \|\chi\|^2$ , for some  $\eta < 1$ .  $\square$

*Proof:* Note that the operator  $\mathcal{B}_3$  in (13) is a contraction by Assumption 4, definition of  $\chi$  and (12); instead  $\mathcal{B}_1$  is a contraction for  $\alpha$  small enough as in Theorem 1. Moreover, the mappings  $\mathcal{B}_2$  and  $\mathcal{B}_4$  are Lipschitz continuous with constant proportional to the step size  $\alpha$ , by the assumptions. Therefore, with Lemma 1 and the reformulation in (13) in place, the proof can be carried out via standard small-gain arguments, and is hence only sketched here. By the Cauchy–Schwarz inequality, we can bound

$$\begin{aligned} & \|\mathcal{B}_1(\xi^k) + \mathcal{B}_2(\omega^k) - \xi^*\|_P^2 \\ & \leq \|\mathcal{B}_1(\xi^k) - \mathcal{B}_1(\xi^*)\|^2 \\ & \quad + \|\mathcal{B}_1(\xi^k) - \mathcal{B}_1(\xi^*)\|_P \|\mathcal{B}_2(\omega^k)\|_P + \|\mathcal{B}_2(\omega^k)\|_P^2 \\ & \leq (1 - 2\alpha\mu_A + \alpha^2\ell_A^2) \|\xi^k - \xi^*\|_P^2 \\ & \quad + \alpha\ell_1 \|\xi^k - \xi^*\|_P \|\chi^k\| + \alpha^2\ell_2 \|\chi^k\|^2 \end{aligned}$$

and

$$\begin{aligned} & \|\mathcal{B}_3(\chi^k) + \mathcal{B}_4(\omega^k)\|^2 \\ & \leq \theta \|\chi^k\|^2 + \alpha\ell_3 \|\xi^k - \xi^*\|_P \|\chi^k\| + \alpha\ell_4 \|\chi^k\|^2 \\ & \quad + \alpha^2\ell_5 \|\xi^k - \xi^*\|_P^2 + \alpha^2\ell_6 \|\chi^k\|^2, \end{aligned}$$

where  $\ell_1, \ell_2, \ell_3, \ell_4, \ell_5, \ell_6 > 0$  are parameters independent of  $\alpha$ , and  $\|\mathcal{B}_2(\omega^k)\|$  is bounded independently of  $\|\chi^k\|$  (hence, of  $\|\xi^k\|$ ) by definition of  $\mathcal{B}_2$ . Therefore, for the norm  $V$ ,

$$V(\omega^{k+1}) = \begin{bmatrix} \|\xi^{k+1} - \xi^*\|_P \\ \|\chi^{k+1}\| \end{bmatrix}^\top \begin{bmatrix} \|\xi^{k+1} - \xi^*\|_P \\ \|\chi^{k+1}\| \end{bmatrix}$$

$$\begin{aligned} & \leq \begin{bmatrix} \|\xi^k - \xi^*\|_P \\ \|\chi^k\| \end{bmatrix}^\top M \begin{bmatrix} \|\xi^k - \xi^*\|_P \\ \|\chi^k\| \end{bmatrix} \\ & \leq \lambda_{\max}(M) V(\omega^k), \end{aligned}$$

where

$$M := \begin{bmatrix} 1 - 2\alpha\mu_A + \alpha^2(\ell_A + \ell_5) & \alpha\frac{1}{2}(\ell_1 + \ell_3) \\ \alpha\frac{1}{2}(\ell_1 + \ell_3) & \theta + \alpha\ell_4 + \alpha^2(\ell_2 + \ell_6) \end{bmatrix}$$

and  $\lambda_{\max}(M) = \|M\| < 1$  for small enough  $\alpha > 0$ .  $\blacksquare$

Note that, for the special case  $A = \mathbf{1}^\top$ , a linearly convergent continuous-time method for GNE seeking in partial-decision information was studied in [10]. Yet, to our knowledge, Theorem 2 is the first result to ensure linear convergence in the case of more general general full-rank constraints. Due to space limitations, we do not derive here an explicit expression for  $\alpha_{\max}$  and  $\eta$  (or  $\mu_A$  in Lemma 1).

*Remark 1:* The proof of Theorem 1 directly applies to the case of a time-varying graph with weight matrix  $W^k$ , provided that Assumption 4 holds for each  $k \in \mathbb{N}$ . With some modification, the argument can be extended also to the case of doubly stochastic graphs that are not strongly connected at each step, but such that  $\|W^{kQ} W^{(k+1)Q} \dots W^{(k+1)Q} - \frac{1}{N} \mathbf{1} \mathbf{1}^\top\| \leq \theta < 1$ , for a  $Q > 0$  and all  $k \in \mathbb{N}$ .  $\square$

## V. EQUILIBRIUM TRACKING IN TIME-VARYING GAMES

We now consider the case where the game in (1) varies over time at a rate such that we can not assume a time-scale separation between the game evolution and the GNE seeking iterations. For each time index  $t \in \mathbb{N}$ , the agents acquire a new instance of the game:

$$(\forall i \in \mathcal{I}) \quad \min_{y_i \in \mathbb{R}^{n_i}} J_i^t(y_i, x_{-i}) \quad \text{s.t.} \quad (y_i, x_{-i}) \in \mathcal{X}^t. \quad (14)$$

We consider the case when the constraints of the game vary only in their affine part, that is,  $\mathcal{X}^t := \{x \in \mathbb{R}^n | Ax = b^t\}$ , and Assumptions 1–3 hold for each  $t$ . The games in (14) define a primal-dual GNE pair sequence  $(\omega_t^*)_{t \in \mathbb{N}}$ , corresponding to the zero set of the KKT operators  $(\mathcal{A}_t)_{t \in \mathbb{N}}$  defined for all  $t$  as in (3), with  $F, b$  replaced respectively by  $F^t := \text{col}((\nabla_{x_i} J_i^t)_{i \in \mathcal{I}})$  and  $b^t$ . The GNE sequence is unique for each  $t$  following the strong monotonicity of  $\mathcal{A}_t$  and [24, Ex. 22.12]. As the rate at which the problem varies is comparable to the agents' computation time, the agents can only compute an approximation of the GNE at time  $t$  before they are presented with a new instance of the problem. The goal of the agents is then to find a sequence  $(\omega_t^*)_{t \in \mathbb{N}}$  which asymptotically tracks relatively well the GNE sequence. We formulate the following assumption, which is standard in the literature of online optimization [25, Assm. 1], [15, Eq. 9] and is verified, for example, for games affected by a bounded process noise in the linear constraints [17, Lemma 5].

*Assumption 6:* For some  $\delta \geq 0$ , it holds that the solution  $\omega_t^*$  of the game in (14) satisfies

$$\sup_{t \in \mathbb{N}} \|\omega_{t+1}^* - \omega_t^*\| \leq \delta. \quad \square$$

Assumption 6 implies that the solution at time  $t$  is an approximate solution for the problem at time  $t+1$ . Given an

estimate of the solution at time  $\omega^{t-1}$  for some time step  $t$ , we then propose to compute  $\omega^t$  by performing  $K$  iterations of the iteration in (8), warm-started at  $\omega^t$ , that is:

$$y^{t,0} = \omega^{t-1} \quad (15a)$$

$$y^{t,k+1} = y^{t,k} - \alpha A_t(y^{t,k}) \quad \text{for } k = 0, \dots, K-1 \quad (15b)$$

$$\omega^t = y^{t,K}, \quad (15c)$$

where  $(y^{t,k})_{k \in \{1, \dots, K\}}$  are auxiliary variables. The following lemma shows that, for an appropriately chosen step size, the proposed algorithm tracks the GNE trajectory up to an asymptotic error which depends on  $\delta$  and  $K$ .

*Theorem 3:* For any  $0 < \alpha < \frac{2\mu_A}{\ell_A^2}$ ,  $\omega^0$ ,  $K \in \mathbb{N}_{>0}$ , the sequence  $(\omega^t)_{t \in \mathbb{N}}$  generated by the iteration in (15) satisfies

$$\limsup_{t \rightarrow \infty} \|\omega^t - \omega_t^*\|_P \leq \frac{\rho^{K/2}}{1 - \rho^{K/2}} \delta \sqrt{\lambda_{\max}(P)}. \quad (16)$$

where  $\rho$  is as in Theorem 1.  $\square$

*Proof:* Following Theorem 1, for  $\alpha \leq \frac{2\mu_A}{\ell_A^2}$ ,

$$\|\omega^t - \omega_t^*\|_P \leq \rho^{K/2} \|\omega^{t-1} - \omega_{t-1}^*\|_P.$$

From the latter, the triangle inequality and the fact  $\|z\|_P^2 \leq \lambda_{\max}(P) \|z\|^2$  for all  $z$ :

$$\begin{aligned} \|\omega^t - \omega_t^*\|_P &\leq \rho^{K/2} \|\omega^{t-1} - \omega_{t-1}^*\|_P + \rho^{K/2} \|\omega_t^* - \omega_{t-1}^*\|_P \\ &\leq \rho^{K/2} \|\omega^{t-1} - \omega_{t-1}^*\|_P + \rho^{K/2} \delta \sqrt{\lambda_{\max}(P)}. \end{aligned}$$

Iterating the latter  $t$  times, we obtain

$$\|\omega^t - \omega_t^*\|_P \leq \rho^{Kt/2} \|\omega^0 - \omega_0^*\|_P + \sum_{\tau=1}^t \rho^{K\tau/2} \delta \sqrt{\lambda_{\max}(P)}.$$

Since  $\rho^{K/2} < 1$ , the thesis follows by the convergence of the geometric sequence.  $\blacksquare$

Let us now turn our attention to the time-varying counterpart of the partial-decision information setup described in Section IV. Again, we consider aggregative games in the form

$$J_i^t(x_i, x_{-i}) = J_i^t(x_i, \sigma^t(x))$$

where  $\sigma^t(x) := \frac{1}{N} \sum_{i \in \mathcal{I}} \phi_i^t(x_i)$  and we postulate that  $\phi_i^t$  satisfies Assumption 5 for all  $t$ . As in Section IV, we augment the state of each agent with an estimate of  $\sigma^t(x)$ , of the dual variable and of the residual  $Ax - b^t$ . For every  $t$ , denote  $\omega_t^* = (\omega_t^*, \mathbf{0})$ , where  $\omega_t^* = (x_t^*, \lambda_t^*)$  is a primal-dual solution of the game at time  $t$  and the vector of zeros represents the target estimation error. We then define the reference trajectory as  $(\omega_t^*)_{t \in \mathbb{N}}$ . At each time-step, we propose to appropriately re-initialize the dynamic tracking of the estimated variables and, in the spirit of the iteration in (15), to apply a finite number of iterations of Algorithm 2. The resulting method is illustrated in Algorithm 3. We obtain the following counterpart of Lemma 2 for the re-initialized dynamic tracking.

*Lemma 4:* For all  $t \in \mathbb{N}$ , it holds that  $\bar{r}^t = \sum_{i \in \mathcal{I}} (A_i x_i^t - b_i^t)$ ,  $\bar{\sigma}^t = \sigma^t(x^t)$ ,  $\bar{\lambda}^t = \sum_{i \in \mathcal{I}} (z_i^t)$ .  $\square$

*Proof:* Let for some  $t$ :

$$\bar{r}^t = \frac{1}{N} \sum_{i \in \mathcal{I}} A_i x_i^t - b_i^t. \quad (17)$$

### Algorithm 3 Time-varying fully-distributed GNE seeking

*Initialization:* choose  $\alpha > 0$  as in Theorem 2; for all  $i \in \mathcal{I}$ , set  $x_i^0 \in \mathbb{R}^{n_i}$ ,  $\sigma_i^0 = \mathbf{0}$ ,  $z_i^0 \in \mathbb{R}^m$ ,  $\lambda_i^0 = z_i^0$ ,  $r_i^0 = A_i x_i^0$ ,  $b_i^0 = \mathbf{0}$ ,  $\phi_i^0(\cdot) = \mathbf{0}$ .

*Iteration:* at time  $t \in \mathbb{N}_{>0}$ , for each agent  $i \in \mathcal{I}$ ,

- 1) Acquire  $J_i^t, \phi_i^t, b_i^t$
- 2) Re-initialize  $\hat{x}_i^{t,0} = x_i^{t-1}$ ,  $\hat{z}_i^{t,0} = z_i^{t-1}$ ,  $\hat{\sigma}_i^{t,0} = \sigma_i^{t-1} - \phi_i^{t-1}(x_i^{t-1}) + \phi_i^t(x_i^{t-1})$ ,  $\hat{r}_i^{t,0} = r_i^{t-1} + b_i^t - b_i^{t-1}$ ,  $\hat{\lambda}_i^{t,0} = \lambda_i^{t-1}$
- 3) For all  $k \in \{0, \dots, K-1\}$ , for all  $i \in \mathcal{I}$ :
  - Local variables update:

$$\begin{aligned} \hat{x}_i^{t,k+1} &= \hat{x}_i^{t,k} - \alpha \left( F_i^t(\hat{x}_i^{t,k}, \hat{\sigma}_i^{t,k}) + A_i^\top \hat{\lambda}_i^{t,k} \right) \\ \hat{z}_i^{t,k+1} &= \hat{z}_i^{t,k} + \alpha N \hat{r}_i^{t,k} \end{aligned}$$

- Estimation update: Agent  $i$  exchanges the variables  $(\hat{\sigma}_i^{t,k}, \hat{\lambda}_i^{t,k}, \hat{r}_i^{t,k})$  with its neighbors, and updates

$$\begin{aligned} \hat{\sigma}_i^{t,k+1} &= \sum_{j \in \mathcal{N}_i} w_{i,j} \hat{\sigma}_j^{t,k} + \phi_i^t(\hat{x}_i^{t,k+1}) - \phi_i^t(\hat{x}_i^{t,k}) \\ \hat{r}_i^{t,k+1} &= \sum_{j \in \mathcal{N}_i} w_{i,j} \hat{r}_j^{t,k} + A_i \hat{x}_i^{t,k+1} - A_i \hat{x}_i^{t,k} \\ \hat{\lambda}_i^{t,k+1} &= \sum_{j \in \mathcal{N}_i} w_{i,j} \hat{\lambda}_j^{t,k} + \hat{z}_i^{t,k+1} - \hat{z}_i^{t,k} \end{aligned}$$

- 4) Set  $x_i^t = \hat{x}_i^{t,K}$ ,  $z_i^t = \hat{z}_i^{t,K}$ ,  $\sigma_i^t = \hat{\sigma}_i^{t,K}$ ,  $r_i^t = \hat{r}_i^{t,K}$ ,  $\lambda_i^t = \hat{\lambda}_i^{t,K}$

From the update step 4) and the re-initialization step 2), we obtain  $x_i^t = \hat{x}_i^{t,K} = \hat{x}_i^{t+1,0}$  and

$$\begin{aligned} \frac{1}{N} \sum_{i \in \mathcal{I}} \hat{r}_i^{t+1,0} &= \bar{r}^t - \frac{1}{N} \sum_{i \in \mathcal{I}} (b_i^{t+1} - b_i^t) \\ &\stackrel{(17)}{=} \frac{1}{N} \sum_{i \in \mathcal{I}} (A_i \hat{x}_i^{t+1,0} - b_i^{t+1}). \end{aligned}$$

Then, by Lemma 2,  $\frac{1}{N} \sum_{i \in \mathcal{I}} \hat{r}_i^{t+1,K} = \frac{1}{N} \sum_{i \in \mathcal{I}} A_i \hat{x}_i^{t+1,K} - b_i^{t+1}$ . From the latter and the update step 4),

$$\bar{r}^{t+1} = \frac{1}{N} \sum_{i \in \mathcal{I}} \hat{r}_i^{t+1,K} = \frac{1}{N} \sum_{i \in \mathcal{I}} A_i x_i^{t+1} - b_i^{t+1}.$$

The result follows by induction and similarly for  $\bar{\sigma}^t, \bar{\lambda}^t$ .  $\blacksquare$  The re-initialization of the dynamic tracking introduces an additional error term in the solution tracking, which requires the following technical assumption on the time variation of the functions  $(\phi_i^t)_{t \in \mathbb{N}}$ :

*Assumption 7:* For some  $\delta_\phi > 0$ , it holds that

$$\sup_{x,t} \|\text{col}(\phi_i^t(x_i) - \phi_i^{t+1}(x_i))_{i \in \mathcal{I}}\| \leq \delta_\phi. \quad \square$$

*Theorem 4:* Let  $\omega^t = \text{col}(x^t, \bar{\lambda}^t, \bar{\sigma}^t, \bar{r}^t, \bar{\lambda}^t)$ . There exists  $\alpha_{\max}$  such that, for every  $0 < \alpha < \alpha_{\max}$ , the sequence  $(\omega^t)_{t \in \mathbb{N}}$  generated by Algorithm 3 satisfies

$$\limsup_{t \rightarrow \infty} \|\omega^t - \omega_t^*\|_Q \leq \frac{\eta^{K/2}}{1 - \eta^{K/2}} \sqrt{\lambda_{\max}(Q)} ((1 + N\ell_A)\delta + \delta_\phi)$$

where  $Q = \text{diag}(P/2, I)$ , for some  $\eta \in (0, 1)$ .  $\square$

*Proof:* Following the same steps as in Lemma 3, the inner iteration (Step 3) of Algorithm 3 is equivalent to the iteration in (13), where  $F$  and  $\phi_i$  are substituted by their time-varying counterpart. Now denote

$$\begin{aligned} \psi^t &= \text{col}(\mathbf{0}, \mathbf{0}, \text{col}((\phi_i^t(x_i^t) - \phi_i^{t-1}(x_i^t))_{i \in \mathcal{I}}), \\ &\quad \text{col}(b_i^t - b_i^{t-1})_{i \in \mathcal{I}}, \mathbf{0}). \end{aligned}$$

From Assumptions 6, 7 and from  $Ax_t^* = b^t$ ,

$$\begin{aligned} \|\psi^t\| &\leq \delta_\phi + \sum_{i \in \mathcal{I}} \|b_i^t - b_i^{t-1}\| \\ &\leq \delta_\phi + N \|b^t - b^{t-1}\| \leq \delta_\phi + N \ell_A \delta \end{aligned} \quad (18)$$

From Theorem 2 and accounting for the re-initialization step, we find for every  $t$ :

$$\|\omega^t - \omega_t^*\|_Q \leq \eta^{K/2} \|\omega^{t-1} + \psi^t - \omega_t^*\|_Q \quad (19)$$

for some  $\eta \in (0, 1)$ . By the triangle inequality, Assumption 6 and (18), and from the fact  $\|z\|_Q^2 \leq \lambda_{\max}(Q) \|z\|^2$ , we have

$$\begin{aligned} \|\omega^{t-1} + \psi^t - \omega_t^*\|_Q &\leq \|\omega^{t-1} - \omega_t^*\|_Q + \|\psi^t\|_Q \\ &\leq \|\omega^{t-1} - \omega_{t-1}^*\|_Q + \|\omega_t^* - \omega_{t-1}^*\|_Q + \|\psi^t\|_Q \\ &\leq \|\omega^{t-1} - \omega_{t-1}^*\|_Q + \sqrt{\lambda_{\max}(Q)} ((1 + N \ell_A) \delta + \delta_\phi). \end{aligned}$$

By substituting the latter in (19) and by iterating the resulting inequality, we obtain:  $\|\omega^t - \omega_t^*\|_Q \leq \eta^{Kt/2} \|\omega^0 - \omega_0^*\|_Q + \sum_{\tau=1}^t \eta^{K\tau/2} \sqrt{\lambda_{\max}(Q)} ((1 + N \ell_A) \delta + \delta_\phi)$ . As  $\eta^{K/2} < 1$ , the result follows by convergence of the geometric series. ■

*Remark 2:* Considering a time-varying matrix constraint  $A_t$  (instead of  $A$ ) would generate some complications, as the matrix  $P$  in (7) would be time-varying as well. This case can be dealt with under the extra assumption that  $\bar{A}A_t^\top \geq \mu_A$  for a matrix  $\bar{A}$  and all  $k \in \mathbb{N}$ ; but it is not discussed here. □

## VI. NUMERICAL EXAMPLE<sup>1</sup>

We demonstrate the proposed algorithms on a market clearing problem for peer-to-peer energy market inspired by [26]. We consider  $N = 6$  prosumers that aim at determining their energy portfolio. At each time-step  $t$ , the agents can either purchase power from a main energy operator, produce it from a dispatchable energy source or trade it with their respective neighbors over a randomly generated undirected graph  $\mathcal{G}^E$ . Furthermore, the agents can exchange information over an undirected connected communication graph  $\mathcal{G}$ . We denote for each agent  $i$  and each time-step  $t$  the power purchased from the main operator as  $x_{i,t}^{\text{mg}}$ , the produced power as  $x_{i,t}^{\text{dg}}$  and the power that agent  $i$  purchases from agent  $j$  as  $x_{i,j,t}^{\text{tr}}$ ,  $j \in \mathcal{N}_i^E$ , with  $\mathcal{N}_i^E$  the set of neighbors of agent  $i$  over  $\mathcal{G}^E$ . The energy price posed by the main operator increases linearly with the aggregate power requested. Thus, by defining the aggregative value  $\sigma^{\text{mg}}(x^{\text{mg}}) = \sum_{i \in \mathcal{I}} x_i^{\text{mg}}$ , the cost incurred by each agent for purchasing energy from the operator is  $J^{\text{mg}}(x_{i,t}^{\text{mg}}, \sigma^{\text{mg}}(x_t^{\text{mg}})) = c^{\text{mg}} \sigma^{\text{mg}}(x_t^{\text{mg}}) x_{i,t}^{\text{mg}}$ ,  $c^{\text{mg}} > 0$ . We consider a quadratic cost on the power generation incurred by the agents [26, Eq. 2], with the form  $J_{i,t}^{\text{dg}}(x_{i,t}^{\text{dg}}) = c^{\text{dg}} (x_{i,t}^{\text{dg}} - x_{i,t}^{\text{dg,ref}})^2$  where  $x_{i,t}^{\text{dg,ref}}$  is the *time-varying* scheduled setpoint of the dispatchable generators and  $c^{\text{dg}} > 0$ . The price (or revenue) of trading energy between peers is linear [26, Eq. 10], and the agents incur a quadratic cost on the transactions for utilizing the market; thus the total objective function related to the peer-to-peer trading is given by  $J^{\text{tr}}((x_{i,j,t}^{\text{tr}})_{j \in \mathcal{N}_i^E}) = \sum_{j \in \mathcal{N}_i^E} (c^{\text{tr}} x_{i,j,t}^{\text{tr}} + \kappa^{\text{tr}} (x_{i,j,t}^{\text{tr}})^2)$ . The agents cannot sell power to the main operator and, due to physical limitations, the power generated by the dispatchable

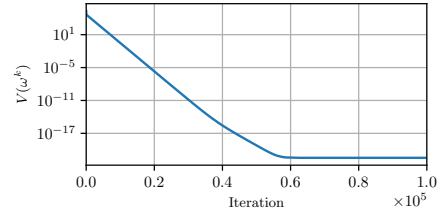


Fig. 1: Convergence of Algorithm 2 for the day-ahead market clearing problem.

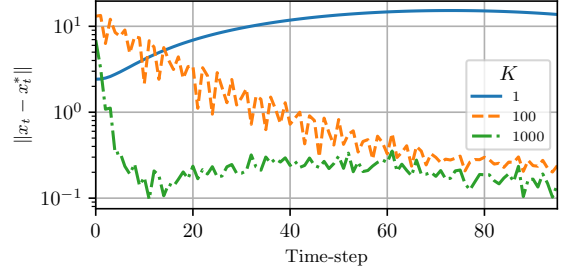


Fig. 2: Tracking error of Algorithm 3 for the real-time market clearing problem with respect to the day-ahead solution computed by  $10^5$  iterations of Algorithm 2.

units must be non-negative. As the formulation in (1) does not consider inequality constraints, this is enforced by a Lipschitz continuous approximation of the logarithmic barrier function  $J^{\text{bar}}(x_{i,t}^{\text{mg}}, x_{i,t}^{\text{dg}}) = \Gamma_\gamma(x_{i,t}^{\text{mg}}) + \Gamma_\gamma(x_{i,t}^{\text{dg}})$ , where  $\Gamma_\gamma(y) := \min(-\log(y), -\gamma y + 1 - \log(1/\gamma))$  for  $\gamma \gg 0$ . The total cost incurred by each agent is thus

$$\begin{aligned} J_i^t(x_{i,t}, x_{-i,t}) &= J^{\text{mg}}(x_{i,t}^{\text{mg}}, \sigma^{\text{mg}}(x_t^{\text{mg}})) + J_{i,t}^{\text{dg}}(x_{i,t}^{\text{dg}}) \\ &\quad + J^{\text{tr}}((x_{i,j,t}^{\text{tr}})_{j \in \mathcal{N}_i^E}) + J^{\text{bar}}(x_{i,t}^{\text{mg}}, x_{i,t}^{\text{dg}}). \end{aligned}$$

Moreover, given a power demand  $p_{i,t}^d$ , the agents need to satisfy the power balance equation [26, Eq. 1]

$$\sum_{j \in \mathcal{N}_i^E} (x_{i,j,t}^{\text{tr}}) + x_{i,t}^{\text{mg}} + x_{i,t}^{\text{dg}} = p_{i,t}^d. \quad (20)$$

As the power balance constraints are local, we do not apply the dynamic tracking method to the associated dual variables (i.e., dual variables are managed locally, see [27, Rem. 2]). Instead, coupling constraints between the agents decisions arises via trading reciprocity constraints [26, Eq. 8]:

$$x_{i,j,t}^{\text{tr}} + x_{j,i,t}^{\text{tr}} = 0. \quad (21)$$

We first consider a time-invariant scenario and compute the day-ahead market clearing solution over an entire day, with time-steps of 15 minutes: namely, the cost of agent  $i$  is given by  $\sum_{t=1}^T J_i^t$ , and the constraints in (20)-(21) are imposed for all  $t = 1, 2, \dots, T$ , with  $T = 96$ . Figure 1 shows that, as expected, Algorithm 2 exhibits a linear convergence rate with respect to the Lyapunov function in Theorem 2.

Then, we consider a real-time scenario. In particular, the agents only have access to a prediction on their load demand and generation setpoint over the coming quarter of an hour; hence the cost of agent  $i$  at each time  $t = 1, 2, \dots, T$  is given by  $J_i^t$ . Note that the agents are in fact faced with a

<sup>1</sup>[https://github.com/bemilio/Simple\\_peer\\_to\\_peer](https://github.com/bemilio/Simple_peer_to_peer)

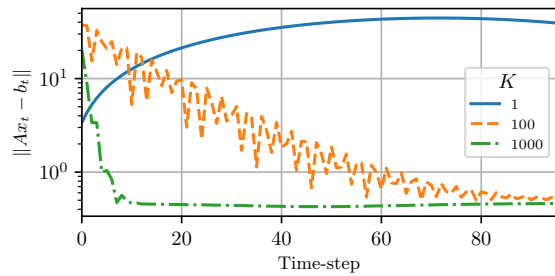


Fig. 3: Constraint violation incurred by Algorithm 3 for real-time market clearing. The inequality constraints enforced via the barrier function  $J^{\text{bar}}$  are always satisfied.

time-varying generalized game as discussed in Section V, which we address via Algorithm 3. The results are shown in Figure 2. Because of the slow convergence (i.e.,  $\eta$  is close to 1),  $K = 1$  results in a significant tracking error; however, good performance is observed already for  $K = 100$ . Finally, in Figure 3, we show the constraint violation obtained by the proposed method over the simulation horizon. As constraints are only satisfied asymptotically, performing only a finite number of iterations per time-step leads to a constraint violation, which as expected decreases with  $K$ .

## VII. CONCLUSION

Strongly monotone GNE problems with full row rank equality coupling constraints can be solved with linear convergence rate, both in semi-decentralized and fully-distributed settings, via primal-dual algorithms. The contractivity properties of the iterates also allow the tracking of the solution sequence in time-varying games; in this online setting, the asymptotic tracking accuracy can be increased by increasing the update frequency.

As our results exploit the strong monotonicity of the KKT operator in a (non-diagonally) weighted space, it is not clear how to embed projections in the proposed methods, which is the main drawback of our approach. Future work should hence focus on linear convergence in generalized games with local constraints (and with inequality coupling constraints).

## REFERENCES

- [1] 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.
- [2] F. Facchinei and C. Kanzow, "Generalized Nash equilibrium problems," *Annals of Operations Research*, vol. 175, pp. 177–211, 2010.
- [3] W. Lin, C. Li, Z. Qu, and M. A. Simaan, "Distributed formation control with open-loop Nash strategy," *Automatica*, vol. 106, pp. 266–273, 2019.
- [4] G. Belgioioso and S. Grammatico, "Semi-decentralized generalized Nash equilibrium seeking in monotone aggregative games," *IEEE Transactions on Automatic Control*, 2021.
- [5] D. Paccagnan, B. Gentile, F. Parise, M. Kamgarpour, and J. Lygeros, "Nash and wardrop equilibria in aggregative games with coupling constraints," *IEEE Transactions on Automatic Control*, vol. 64, no. 4, pp. 1373–1388, 2019.
- [6] E. Benenati, W. Ananduta, and S. Grammatico, "On the optimal selection of generalized Nash equilibria in linearly coupled aggregative games," in *2022 IEEE 61st Conference on Decision and Control (CDC)*, 2022, pp. 6389–6394.

- [7] D. Gadjev 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, 2021.
- [8] L. Pavel, "Distributed GNE seeking under partial-decision information over networks via a doubly-augmented operator splitting approach," *IEEE Transactions on Automatic Control*, vol. 65, no. 4, pp. 1584–1597, 2020.
- [9] M. Bianchi, G. Belgioioso, and S. Grammatico, "Fast generalized Nash equilibrium seeking under partial-decision information," *Automatica*, vol. 136, p. 110080, 2022.
- [10] Z. Deng and X. Nian, "Distributed generalized Nash equilibrium seeking algorithm design for aggregative games over weight-balanced digraphs," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 3, pp. 695–706, 2019.
- [11] 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, 2021.
- [12] Y. Huang and J. Hu, "Distributed solution of GNEP over networks via the Douglas-Rachford splitting method," in *2021 60th IEEE Conference on Decision and Control (CDC)*, 2021, pp. 3110–3116.
- [13] M. Bianchi and S. Grammatico, "Fully distributed Nash equilibrium seeking over time-varying communication networks with linear convergence rate," *IEEE Control Systems Letters*, vol. 5, pp. 499–504, 2021.
- [14] D. Gadjev and L. Pavel, "Distributed Nash equilibrium seeking resilient to adversaries," in *2021 60th IEEE Conference on Decision and Control (CDC)*, 2021, pp. 191–196.
- [15] E. Dall'Anese, A. Simonetto, S. Becker, and L. Madden, "Optimization and learning with information streams: Time-varying algorithms and applications," *IEEE Signal Processing Magazine*, vol. 37, no. 3, pp. 71–83, 2020.
- [16] T. Li, G. Peng, Q. Zhu, and T. Başar, "The confluence of networks, games, and learning a game-theoretic framework for multiagent decision making over networks," *IEEE Control Systems Magazine*, vol. 42, no. 4, pp. 35–67, 2022.
- [17] Y. Su, F. Liu, Z. Wang, S. Mei, and Q. Lu, "Online distributed tracking of generalized Nash equilibrium on physical networks: Closing the loop via measurement feedback," *Autonomous Intelligent Systems*, vol. 1, no. 1, p. 6, 2021.
- [18] K. Lu, G. Li, and L. Wang, "Online distributed algorithms for seeking generalized Nash equilibria in dynamic environments," *IEEE Transactions on Automatic Control*, vol. 66, no. 5, pp. 2289–2296, 2021.
- [19] O. Stein and N. Sudermann-Merx, "The noncooperative transportation problem and linear generalized Nash games," *European Journal of Operational Research*, vol. 266, no. 2, pp. 543–553, 2018.
- [20] F. Parise, B. Gentile, and J. Lygeros, "A distributed algorithm for almost-Nash equilibria of average aggregative games with coupling constraints," *IEEE Transactions on Control of Network Systems*, vol. 7, no. 2, pp. 770–782, 2020.
- [21] A. A. Kulkarni and U. V. Shanbhag, "On the variational equilibrium as a refinement of the generalized Nash equilibrium," *Automatica*, vol. 48, no. 1, pp. 45–55, 2012.
- [22] F. Facchinei and J. Pang, *Finite-dimensional variational inequalities and complementarity problems*. Springer New York, 2007.
- [23] S. Grammatico, "Dynamic control of agents playing aggregative games with coupling constraints," *IEEE Transactions on Automatic Control*, vol. 62, no. 9, pp. 4537–4548, 2017.
- [24] H. H. Bauschke and P. L. Combettes, *Convex analysis and monotone operator theory in Hilbert spaces*. Springer New York, 2017, vol. 2011.
- [25] A. Simonetto, E. Dall'Anese, S. Paternain, G. Leus, and G. B. Giannakis, "Time-varying convex optimization: Time-structured algorithms and applications," *Proceedings of the IEEE*, vol. 108, no. 11, pp. 2032–2048, 2020.
- [26] G. Belgioioso, W. Ananduta, S. Grammatico, and C. Ocampo-Martinez, "Energy management and peer-to-peer trading in future smart grids: A distributed game-theoretic approach," in *2020 European Control Conference (ECC)*, 2020, pp. 1324–1329.
- [27] M. Bianchi and S. Grammatico, "Continuous-time fully distributed generalized Nash equilibrium seeking for multi-integrator agents," *Automatica*, vol. 129, p. 109660, 2021.