Game-Theoretic Learning for Power System Dynamic Ancillary Service Provisions

Xie, Haiwei; Cremer, Jochen L.

DOI
10.1109/LCSYS.2024.3409608

Publication date
2024

Document Version
Final published version

Published in
IEEE Control Systems Letters

Citation (APA)
Xie, H., & Cremer, J. L. (2024). Game-Theoretic Learning for Power System Dynamic Ancillary Service Provisions. *IEEE Control Systems Letters*, *8*, 1307-1312. https://doi.org/10.1109/LCSYS.2024.3409608

Game-Theoretic Learning for Power System Dynamic Ancillary Service Provisions

Haiwei Xie and Jochen L. Cremer, *Member, IEEE*

***Abstract*—This letter studies the problem of coordinating aggregators in the power system to provide fast frequency response as dynamic ancillary services. We approach the problem from the perspective of suboptimal $\mathcal{H}_\infty$ control, and propose an efficient and tractable formulation. We further develop a distributed solution method for the investigated problem, which enables aggregator agents to learn their optimal provisions in an adaptive way. More precisely, we reformulate the original problem into a state-based potential game, where the agents interact with each other towards our designed Nash equilibrium. The proposed game-theoretic learning approach decouples the coupling Linear Matrix Inequality constraint, guarantees the convergence to the equilibrium which is close enough to the original optimum. The learning process is also robust to the changes in communication graphs. We demonstrate the efficacy of our proposed approach with a case study on a 3-aggregator system.**

***Index Terms*—Distributed optimization, game theory, power system, fast frequency response, $\mathcal{H}_\infty$ control.**

I. Introduction

The INTEGRATION of renewable energy resources deteriorates power system inertia, and consequently poses severe challenges to the system's frequency stability under large contingencies. A rapid delivery of active power support is thus expected immediately after the contingency to arrest the frequency deviation. Aggregators appear as promising providers for such Fast Frequency Response (FFR) dynamic ancillary services [1], [2]: one aggregator accommodates a large number of small distributed resources (usually inverter-based) and controls them to provide aggregated rapid post-fault active power. The stringent activation time constraints require the automatic FFR to be programmed in advance before contingencies occur, demonstrating specific post-fault active power dynamics under measured frequency deviations.

The demanded FFR ancillary services should be defined to ensure the system frequency stability under the possibly most severe contingency: the demand varies with the regional grid codes, contingency sets, and the system status (e.g., the existing inertia/ damping). The post-fault system frequency response needs to be derived in order to check the critical index of the response (e.g., Nadir, ROCOF) is acceptable according to the grid codes [3]. It still remains a problem coordinating different FFR providers (i.e., aggregators) thereafter to achieve this demand. Authors in [4], [5] advocated considering the coordination together with the computation of system FFR demands, while these approaches are limited to accommodate multiple aggregators having diverse provisions, as the derivation of the constraints becomes handy and highly customized. The time-domain simulation results with regard to the providers' deliverables are generally non-convex, which makes the optimization problem intractable without specific approximation or relaxation [6], [7]. Authors in [2], [8] investigated the dynamic flexibility underlying one aggregator, while a possible collaboration among multiple aggregators to pursue the overall economic efficiency have not been considered. To address the challenges, we approach the problem of dynamic ancillary service provisions from a suboptimal $\mathcal{H}_\infty$ control perspective: $\mathcal{H}_\infty$ norm is employed as the metric to assess the consistency between the aggregators' total FFR provisions and the system FFR requirements. To this end, a convex problem is formulated, where the overall economic efficiency is optimized together with a control on the aggregators' performance.

When coordinating the aggregators for ancillary service provisions, it is crucial to safeguard their local data, such as operation costs and device limits, to address end-user privacy concerns [9]. In addition, due to the uncertainty encompassed by the participants, one aggregator has to go through frequent changes in their local information. The system may suffer from unexpected quit or re-connection of aggregators for their service provisions. Researchers resorted to distributed optimization methodologies for aggregators to deviate from the need of sharing all local information to a central agent: primal-dual algorithms, Alternating Direction Method of Multipliers are well-investigated to decompose the coupling constraints, and have successful applications in energy management problem [9], [10]. However, these approaches usually still require a central *collect and broadcast* step to update the dual variables which prevents a fully distributed graph-based communication. Therefore they lack robustness to information loss and communication uncertainties [11], which poses a problem in the case of aggregators. In order to address these drawbacks, previous research efforts have shed insights into utilizing game-theoretic control techniques [11], [12], [13]. However, these existing design paradigms are not ready for use

Manuscript received 8 March 2024; revised 7 May 2024; accepted 25 May 2024. Date of publication 4 June 2024; date of current version 24 June 2024. Recommended by Senior Editor C. Prieur.
(Corresponding author: Haiwei Xie.)

The authors are with the Department of Electrical Sustainable Energy, Delft University of Technology, 2628 CD Delft, The Netherlands (e-mail: H.Xie-2@tudelft.nl; J.L.Cremer@tudelft.nl).
Digital Object Identifier 10.1109/LCSYS.2024.3409608

in our problem due to the specialty of the coupling constraint which depicts a semidefinite cone. In this vein, we extend the game-theoretic approach [12] to a semidefinite program setting.

The main contributions of this letter are: (i) propose an efficient and tractable formulation of the optimal dynamic ancillary service provision problem; (ii) develop a distributed solution method for the problem using game-theoretic control, enabling the aggregators to learn their optimal provisions without sharing important local information; (iii) conduct a case study to verify the robustness of our approach.

II. Problem Formulation

We consider a set of aggregators that are providing their FFR ancillary service to the system operator, denoted by $I = \{1, \dots, \mathrm{n}\}$. For any aggregator $i \in I$, we assume its FFR provision is in the form of the transfer function (in s)

$$T_i(s) = \sum_{k=1}^{\mathrm{K}} T_i^k(s) = \sum_{k=1}^{\mathrm{K}} \frac{-D_i^k - H_i^k s}{\tau_k s + 1}, \tag{1}$$

where τ_k is the latency of the FFR delivery. We assume that there are overall K levels of latency, which are known and standardized in the system. Following the definition of (1), H_i^k and D_i^k are the controllable parameters that the aggregator could adjust within its limits and deliver at τ_k. We denote each aggregator's set of decision variables as $v_i \doteq \{H_i^1, D_i^1, \dots, H_i^{\mathrm{K}}, D_i^{\mathrm{K}}\}$. The FFR ancillary service demand is assumed computed by the system operator following standard grid codes, considering system situation, fault size, and closed-loop stability requirements. The demand is given as a target $T(s) = (-D - Hs)/(\tau s + 1)$, where the latency τ, the equivalent inertial H and damping D are all constants.

Remark 1: The assumption of (1) is demonstrated to be achievable by aggregators [2], [14]. We adopt this assumption as it aligns with the advanced inverter control paradigms [15], which means the inverter-based participants in aggregators can be well accommodated. Indeed, H_i^k is usually named as (virtual) inertia, and D_i^k the damping.

With the above problem setting, we compute the residual between the system operator's FFR demands and the collective response from all aggregators as follows

$$\begin{aligned}\mathcal{R}(s) &= T(s) - \sum_{k \in K} \sum_{i \in I} T_i^k(s)\\ &= \frac{-D - Hs}{\tau s + 1} - |I| \sum_{k=1}^{\mathrm{K}} \frac{-\bar{D}^k - \bar{H}^k s}{\tau_k s + 1}\\ &= \frac{b_{\mathrm{K}+1} s^{\mathrm{K}+1} + b_{\mathrm{K}} s^{\mathrm{K}} + \dots + b_1 s + b_0}{a_{\mathrm{K}+1} s^{\mathrm{K}+1} + a_{\mathrm{K}} s^{\mathrm{K}} + \dots + a_1 s + a_0},\end{aligned}$$

where we use top bar $\bar{(\cdot)}$ to denote the average among all the aggregators, e.g., $\bar{D}^k = \sum_{i \in I} D_i^k / |I|$, same for $\bar{H}^k$.

Remark 2: The coefficient vectors $\boldsymbol{a} = [a_0, \dots, a_{\mathrm{K}+1}]$ and $\boldsymbol{b} = [b_0, \dots, b_{\mathrm{K}+1}]$ are composed of the coefficients of the corresponding polynomial with the degree $(\mathrm{K} + 1)$. It is not hard to observe that $\boldsymbol{a}$ is a constant vector defined by the given parameters τ and τ_k, while the elements b_k are linear combinations on the system decision variables in v_i. We can also derive the corresponding state-space model of the residual system $\mathcal{R}(s)$ in Controllable Canonical Form as ($\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, $\mathcal{D}$ as system matrices):

$$\begin{aligned}\dot{q} &= \mathcal{A} q + \mathcal{B} w\\ y &= \mathcal{C} q + \mathcal{D} w\end{aligned}$$

where q, w, and y denote the states, the frequency deviation, and the residual, respectively. System matrices $\mathcal{C}$ and $\mathcal{D}$ are dependent on $\boldsymbol{b}$ and thus the set of all aggregators' decision variables, $v := \{v_1, \dots, v_n\}$. Separating all the constant elements in $\mathcal{C}_0$ and $\mathcal{D}_0$, we have $\mathcal{C}(v) = \mathcal{C} - \mathcal{C}_0$ and $\mathcal{D}(v) = \mathcal{D} - \mathcal{D}_0$. Note that the elements in $\mathcal{C}(v)$ and $\mathcal{D}(v)$ are thus the inner product of a constant vector with v.

Ideally, we would like the residual to be *small* enough, meaning the aggregators' provisions are well achieving the system-level requirements. We therefore employ $\mathcal{H}_\infty$ norm to quantify this *smallness*, as $\mathcal{H}_\infty$ norm of a system provides the maximum energy gain of the output signal y (the residual) for a given input signal w (the frequency deviation). This quantification further endows us with a measure of the system-level performance. To see this, we introduce a variable γ here to constrain the $\mathcal{H}_\infty$ norm of the residual system $\mathcal{R}(s)$, i.e., $|\mathcal{R}(s)|_\infty \leq \gamma$, and γ should be minimized as a term in the objective function. As γ goes to 0, the residual system's output will also reduce to 0, which means the responses from the aggregators are perfectly matching the target. In order to achieve economic optimality, we also introduce $\sum_{i \in I} g_i(v_i)$ in the objective function, with g_i the continuous differentiable convex function representing the cost function of aggregator i regarding its decision variables. To sum up, with a parameter w controlling the trade-off between two terms, we have:

$$\min_{v_i, \gamma, \mathcal{Q}} \quad \sum_{i \in I} g_i(v_i) + w\gamma \tag{2a}$$

$$\text{s.t.} \quad v_i \in V_i, \quad \forall i \in I \tag{2b}$$

$$|\mathcal{R}(s)|_\infty \leq \gamma. \tag{2c}$$

Eq. (2b) encodes the local constraints limiting the decision variables of each aggregator, where V_i is a polytope representing the domain of v_i. Based on the Bounded Real Lemma (BRL) [16], we transform the constraint (2c) into a Linear Matrix Inequality (LMI) constraint (with $\mathcal{Q} \succeq 0$):

$$\underbrace{\begin{bmatrix} \mathcal{Q}\mathcal{A}^T + \mathcal{Q}\mathcal{A} & \mathcal{Q}\mathcal{B} & \mathbf{0}\\ \mathcal{B}^T \mathcal{Q} & -\gamma I & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & -\gamma I \end{bmatrix}}_{:=L(\mathcal{Q},\gamma)} + \underbrace{\begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathcal{C}(v)^T\\ \mathbf{0} & \mathbf{0} & \mathcal{D}(v)^T\\ \mathcal{C}(v) & \mathcal{D}(v) & \mathbf{0} \end{bmatrix}}_{:=\sum_{i \in I} L_i(v_i)}$$

$$\preceq \underbrace{\begin{bmatrix} \mathbf{0} & \mathbf{0} & -\mathcal{C}_0^T\\ \mathbf{0} & \mathbf{0} & -\mathcal{D}_0^T\\ -\mathcal{C}_0 & -\mathcal{D}_0 & \mathbf{0} \end{bmatrix}}_{:=F_0}.$$

The LMI constraint defines a semidefinite cone, which preserves the nice property of convexity. Each element of matrix $L_i(v_i)$ is a linear combination of only v_i.[1] All the constant elements are included in the constant matrix F_0.

[1] We add the regulator as agent 0, and consider the variables γ and $\mathcal{Q}$ as its decision variables, i.e., $v_0 = (\gamma, \mathcal{Q})$ tuple, and $L_0(v_0) := L(\mathcal{Q}, \gamma)$. Moreover, the cost function of the regulator is defined as $g_0(v_0) := w\gamma$.

III. Game-Theoretic Control and State-Based Potential Games

The existence of coupling constraint (2c) implies that the decision made by one agent is constrained by the other agents' decisions. This is inherently consistent with the feature of a game, where the utility of one agent is affected not only by its own decisions, but also by the other agent's decisions. Stemming from this similarity, game-theoretic control provides a distributed alternative to handle the coupled optimization problems concerning multiple agents [17]. The procedures of game-theoretic control design is to reformulate a control problem into a game, via i) endowing each agent a utility function (which should well encode the original costs), and ii) endowing each agent a learning rule (which will be programmed as control laws). It is expected that such a successful reformulation could bring desirable collective behavior for the system designer. In our context, the desirable collective behavior is that all the aggregators achieve their optimal provision as in (2). Before we proceed, let us revisit some preliminaries on the state-based potential game, the properties of which attract us to formulate problem (2) into such a game.

A. Preliminaries: State-Based Potential Games

Definition 1 (State-Based Game [18]): A deterministic game is regarded as a *state-based game* when there exists i) agent set I; ii) state space X; iii) state-dependent action sets of the form $U_i(x), x \in X$ for each agent $i \in I$; iv) state-dependent cost functions $J_i(x, u) \in \mathbb{R}, u \in U(x) = \prod U_i(x)$ for each agent $i \in I$; and v) deterministic state transition function $\tilde{x} = f(x, u) \in X$ for $\forall x \in X$ and $u \in U(x)$. For any $x \in X$, it holds true that the state won't change if all the agents take the null action $0 \in U(x)$, i.e., $x = f(x, 0)$. We denote such a state based game by the tuple $G = \{I, X, \{U_i(x)\}, \{J_i(x, u)\}, f\}$.

Definition 2 (State-Based Potential Game [18]): A state-based game is a *state-based potential game* when there exists a potential function $\Phi : X \times U \to \mathbb{R}$. The potential function satisfies the following two conditions for every state $x \in X$:

C-1 For every agent $i \in I$, $u \in U(x)$, and $u_i' \in U_i(x)$, it holds that $J_i(x, u) - J_i(x, u_i', u_{-i}) = \Phi(x, u) - \Phi(x, u_i', u_{-i})$, where $-i$ denotes the set of all the agents other than agent i, and u_{-i} includes the actions of agents in $-i$.

C-2 for every action $u \in U(x)$ and ensuing state following the dynamic $\tilde{x} = f(x, u)$, we have $\Phi(x, u) = \Phi(\tilde{x}, 0)$.

Definition 3 (Stationary State Nash Equilibrium [19]): A state-action pair $[x^*, u^*]$ is a *stationary state Nash equilibrium* for a state-based game if it satisfies the following two conditions:

C-1 We have $u_i^* \in \text{argmin}\{J_i(x^*, u_i, u_{-i}^*)|u_i \in U_i(x^*)\}$ for $\forall i \in I$,

C-2 x^* is a fixed point of the state transition function with u^*, i.e., $x^* = f(x^*, u^*)$.

B. Connection to the Game-Theoretic Control

State-based potential games possess two properties that are attractive in game-theoretic control:

- *Existence of Nash equilibrium.* The learning algorithms of the game ideally end up in the Nash equilibrium where no agent can improve their utility by updating their decisions anymore. We therefore wish to find a way to engineer the Nash equilibrium of the game, such that it is exactly located at the optimum we desire. The premise here is to have the guarantee that the Nash equilibrium does exist. Since the stationary state Nash equilibrium in a state-based potential game is guaranteed to exist [18], it is theoretically appealing to design a state-based potential game for the original problem, and accurately locate the equilibrium to be equivalent or close to the optimum.
- *Existence of distributed Nash equilibrium-seeking algorithms.* It is established that for a state-based potential game, there are feasible choices of distributed learning algorithms that are theoretically guaranteed to converge to equilibrium [13], [19]. This distributed fashion is preferred in the sense that the agents can make use of only local information to make the decision. The local information includes i) the agent's own information and ii) the information communicated from its neighbours. Therefore, each agent is exchanging the information with only its neighbours.

IV. A Distributed Solution Method

In this section, we show how we fit the problem (2) into a state-based potential game, and provide a distributed learning algorithm such that the agents achieve the Nash equilibrium (i.e., optimum) following the learning rules.

A. Game Design

1) The Agent Set and the Communication Graph: The agent set (including regulator agent 0) is $I = \{0, 1, \ldots, n\}$. Each agent has a neighbor set N_i which includes all the neighboring agents it can directly communicate with. The agent itself is regarded to be in N_i, i.e., $i \in N_i$. We assume that the communication graph of the system denoted by $\{N_i\}$ is undirected and connected.

2) States and State Space: We denote matrix e_i as the agent i's estimation for the coupling constraint (average share), i.e., $e_i \sim \frac{1}{|I|}(\sum_{j \in I} L_j(v_j) - F_0)$. The system states tuple $x = (v, e)$ is then composed of two parts, i) all the decision variables $v = (v_0, v_1, \ldots, v_n)$; and ii) the estimation from all the agents $e = (e_0, e_1, \ldots, e_n)$. The feasible space of the states x is

$$X = \{(v, e)|v_i \in V_i, e_i \prec 0\}. \quad (3)$$

By defining the estimation state e_i, we are taking the first step to decouple the coupling constraint, as all the agents now have an independent estimation state for this constraint.

3) Action Tuples and the State-Dependent Action Space: The action tuple $u_i = (\hat{v}_i, \hat{e}_{i\to out})$ is defined for each agent. It is composed of two parts similar to the states, i) the changes to be made on the local decision variables $\hat{v}_i$, and ii) the set of respective estimation matrices to be delivered to the corresponding neighbors $\hat{e}_{i\to out} = \{\hat{e}_{i\to j}|j \in N_i.\}$. An intuitive interpretation of the *delivered estimation* $\hat{e}_{i\to j}$ is, the *amount of estimation* that agent i would spare from its own estimation to assist its neighbor agent j. Imagine now the estimation matrix e_i from agent i is largely negative definite, which means all the eigenvalues of e_i are relatively far below 0. In this case, it would be helpful if agent i can assist some other agent j by supporting them with a negative definite estimation matrix $\hat{e}_{i\to j}$. With this support, agent j can be *more bold* when it adjusts v_j to get higher utility, while still maintaining its estimation e_j being negative definite.

The action's feasible space is state-dependent, as it needs to ensure the ensuing states lie in the feasible space (3). Given a system state x, the action's feasible set/region is

$$U_i(x) = \left\{ \left(\hat{v}_i, \hat{e}_{i\to out}\right) \middle| \begin{array}{c} v_i + \hat{v}_i \in V_i, \\ \hat{e}_{i\to out} \preceq 0, \\ e_i + L_i\left(\hat{v}_i\right) - \hat{e}_{i\to out} \prec 0 \end{array} \right\}. \quad (4)$$

4) State Transition Function: We denote the local dynamic for agent i as $\tilde{x} = f_i(x, u_i)$. It is as follows with $\tilde{v}_i$, $\tilde{e}_i$ and $\tilde{e}_{out}$ being the ensuing states

$$\begin{cases} \tilde{v}_i = v_i + \hat{v}_i \\ \tilde{e}_i = e_i + L_i\left(\hat{v}_i\right) - \sum_{out\in N_i} \hat{e}_{i\to out} + \sum_{in\in N_i} \hat{e}_{in\to i} \\ \tilde{e}_{out} = e_{out} + \hat{e}_{i\to out}, \ \forall out \in N_i. \end{cases} \quad (5)$$

The transition function on v_i follows that $\hat{v}_i$ is the changes in the decision variables. The transition function on e contains contributions from three parts: i) $L_i(\hat{v}_i)$ be the estimation change caused by the change in v_i; ii) $-\sum \hat{e}_{i\to out}$ is the sum of estimation delivered to all neighbors; iii) $\sum \hat{e}_{in\to i}$ is the sum of estimation received from all neighbors.

The design of the state transition function provides a further step towards the decoupling of the coupling constraint. As we can see, during the dynamic transition, we always have $\sum e_i(t) = \sum e_i(0) + \sum L_i(v_i(t) - v_i(0))$. Therefore, if we nicely select the initial estimations as $\sum e_i(0) = \sum L_i(v_i(0)) - F_0$, we will have $\sum e_i(t) = \sum L_i(v_i(t)) - F_0$ for all t in the later iterations. This means that we can now evaluate the coupling constraint with the sum of the agents' estimation states. The feasible space of the states (3) ensures that all the estimations from agents are negative definite, and thus the coupling constraint is strictly satisfied.

5) State-Dependent Cost Function: The agent's cost functions depend on the ensuing system state $\tilde{x}$, which is deterministically determined by $x \in X$ and $u \in \prod_{i\in I} U_i(x)$ as a result of (5),

$$J_i(\tilde{x}|x, u_i, u_{-i}) = g_i(\tilde{v}_i) - \mu \sum_{j\in N_i} \log\det\left(-\tilde{e}_j\right). \quad (6)$$

The additional terms added to the original decoupled local cost function are logarithmic barrier functions for positive definite matrix. These additional terms implicitly enforce the constraints that $\tilde{e}_j \prec 0$ for every $j \in N_i$, which is satisfied naturally following (3) and (4).

Theorem 1 (Optimality of the Nash Equilibrium): Consider a game $\mathcal{G}$ defined with (3)-(6). we have that, as μ goes to 0, the stationary Nash equilibrium of $\mathcal{G}$ converges to the solution of the original optimization problem (2).

Proof: The proof is built on three preparatory Lemmas.

Lemma 1: The defined state-based game with the agents' utility function (6) is a potential game with the potential function

$$\Phi(x, u) = \sum_{i\in I} g_i(\tilde{v}_i) - \mu \sum_{i\in I} \log\det(-\tilde{e}_i).$$

Lemma 2: The state-action pair $[x^*, u^*] = [(v^*, e^*), (\hat{v}^*, \hat{e}^*)]$ is a stationary state Nash equilibrium of the game $\mathcal{G}$, if and only if:

C-1 The decision variables v^* in states is the optimum of the following optimization problem with only uncoupled local constraints:

$$\begin{aligned} \min_{v} \quad & \sum_{i\in I} g_i(v_i) - |I|\mu \log\det\left(F_0 - \sum_{i\in I} L_i(v_i)\right) \\ \text{s.t.} \quad & v_i \in V_i \qquad \forall i \in I. \end{aligned} \quad (7)$$

C-2 The estimation variables e^* in states satisfies that for all $i \in I$,

$$e_i^* = \frac{1}{|I|}\left(\sum_{i\in I} L_i\left(v_i^*\right) - F_0\right).$$

C-3 For all $i \in I$, $\hat{v}_i{}^* = 0$ and $\sum \hat{e}^*_{i\leftarrow in} - \sum \hat{e}^*_{i\to out} = 0$.

Lemma 3: The optimum of the optimization problem (7) converges to the solution of the original optimization problem (2) as μ goes to 0.

Lemma 1 identifies the potential function of the designed game and therefore proves the existence of the stationary Nash Equilibrium. Lemma 2 identifies the equivalence between the equilibrium and v^* the optimal solution of (7). Lemma 3 shows the convergence of v^* to the solution of (2) as μ goes to 0, and therefore completes the proof of Theorem 1.

Lemma 1 is easy to verify. The proofs of Lemma 2 and Lemma 3 are provided in Appendices. ■

B. A Distributed Learning Algorithm

Proposition 1 (Better response dynamics [19]): In a state-based potential game, if the myopic agents all make attempts to improve their own utility function based on their local information, and take turns to update their action pairs, it is guaranteed that the Nash equilibrium defined in Lemma 2 is attained when no more agents can improve their utility.

Based on Proposition 1, we here provide a distributed learning algorithm based on the Curnot adjustment. Each agent follows (8) and takes a turn to update its localized state-action pair

$$\begin{aligned} u_i(k) &= \underset{u_i\in U_i(x)}{\operatorname{argmin}}\ J_i(\tilde{x}|x(k), u_i, 0), \\ x(k+1) &= f_i(x(k), u_i(k)). \end{aligned} \quad (8)$$

The agent assumes others are taking action 0 ($u_{-i} = 0$) when making the decisions, which means the information needed in (8) is all locally available by agent i. The learning process is continued until every agent confirms that it does not have the will to change its actions. Note that the initial states are chosen in the feasible set (3) satisfying $\sum e_i(0) = \sum L_i(v_i(0)) - F_0$.

Remark 3: We can see that in Proposition 1, there are no strict requirements for the order of updating during the learning process. Therefore, the learning process is naturally robust to changes of agent update orders, or even occasional missing out. Furthermore, the learning rules are based on current local information, with no specifications on the communication graph as long as it is connected. Therefore, the whole learning process is robust to communication loss or changes.

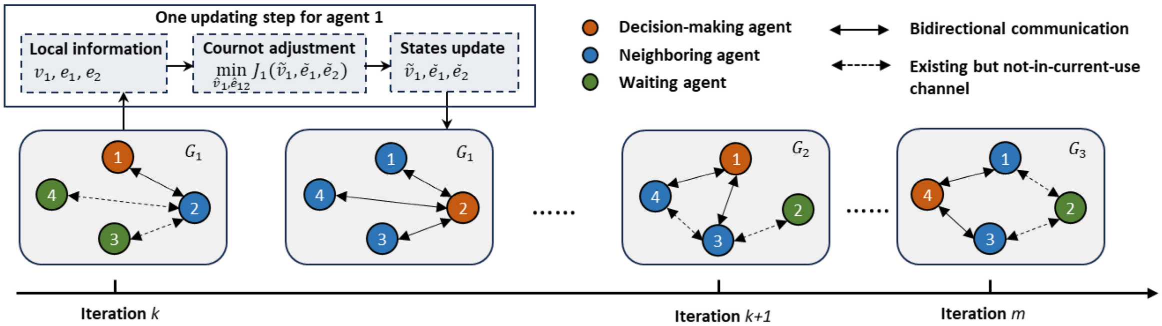

Fig. 1. The schematic of the proposed game-theoretic learning approach.

V. Case Study

We consider a system with 3 aggregators, offering their FFR services at two standardized latency levels, $K = 2$: $\tau_1 = 0.5s$ and $\tau_2 = 5s$. The cost function related to each aggregator's service provision, and their corresponding local limits are:

1) $g_1(v_1) = 0.2H_1^1 + 2D_1^1; \quad 0 \leq H_1^1, D_1^1 \leq 5$
2) $g_2(v_2) = 0.3H_2^2 + 3D_2^1 + D_2^2; \quad 0 \leq H_2^2, D_2^1, D_2^2 \leq 5$
3) $g_3(v_3) = 1.5D_3^2; \quad 0 \leq D_3^2 \leq 10.$

The stability-economic weight w is set as 5. The setting for the system FFR target is $H = 15$ p.u., $D = 20$ p.u. and a latency $\tau = 3.2$ s. Fig. 2 presents our results obtained under $\mu = 0.02$, and under a varying communication graph. The communication graph G_1 encountered sudden changes twice, at iteration 1000 and iteration 2000. Fig. 2 (a)-(d) demonstrate that the approach is indeed robust to the changes happening in the communication graph during the learning. The decision variables converged to the solution of (7), which is close to but not exactly equal to the optimum solution of (2) (as shown in dashed lines). The errors between are due to the non-zero hyperparameter μ that steers the solution slightly away from the exact optimum, as a price for decoupling the constraint. μ could be adjusted lower to get results closer to the optimum, while it will deteriorate the convergence rate as the agents are *less driven* by the instructions coming from coupling constraints. Fig. 2 (e) illustrates the decrease of the potential during the learning process. Fig. 2 (f) shows the coupling constraint is strictly satisfied during the whole learning process. This implies that the approach is safe to be applied online, as when the agents are adjusting their decisions, all the constraints are guaranteed to be satisfied. The convergence of the algorithm can take a large number of iterations, showing the algorithm's limitation on the computational time. In the FFR coordination problem, though there may not be a very strict requirement on the computational time: the aggregators implement the output of the algorithm by programming their devices ahead of real-time operation; the FFR response is thereafter automatically provided in real-time when the system contingency occurs. When the number of agents is large, to address the time limitations, a solution set for multiple common FFR targets could be prepared in advance. Subsequently, when one FFR target is identified in operation, the solution pre-computed for a "close" target can be initialized for aggregators to accelerate the convergence to the optimum.

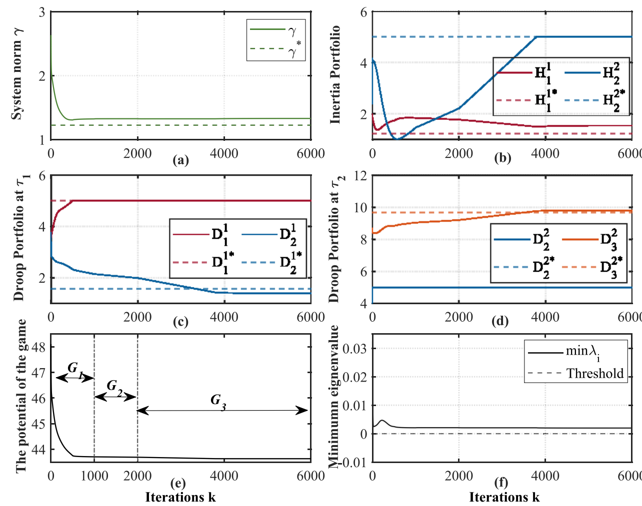

Fig. 2. Results during the game-theoretic learning process. In (a)-(d), the color represents the decision variables coming from different agents (green-0, red-1, blue-2, orange-3).

VI. Conclusion and Future Work

In this letter, we provide an efficient and tractable approach for the aggregators to learn their optimal dynamic ancillary service provisions, based on $\mathcal{H}_\infty$ system norm and a proposed game-theoretic learning algorithm. Although our work is brought up and introduced in a specific setting, it could be directly applied to other distributed use cases with coupling semidefinite constraints. The designed approach works as a distributed protocol. Future research can investigate this protocol for power system FFR market mechanism design, such that the agents are well incentivized and voluntarily participate in the game. We also recommend future research to develop approaches that further boost the convergence time.

Appendix

A. Proof of Lemma 2

($\Rightarrow$) We first prove that if $[x^*, u^*]$ is a stationary state Nash equilibrium, then it satisfies [C-1]- [C-3]. Based on the second condition in Definition 3, we have $x^* = f(x^*, u^*)$. It is only possible if [C-3] is satisfied. Then, based on the first condition in Definition 3, we have $J_i(x^*, u_i^*, u_{-i}^*) = \min_{u_i \in U_i(x^*)} J_i(x^*, u_i, u_{-i}^*)$ for $\forall i \in I$. Since $J_i(x^*, u_i, u_{-i}^*)$ is a convex function on the tuple $u_i = (\hat{v}_i, \hat{e}_{i\to out}) \in U_i(x^*)$, we have the following equalities based on the sufficient and necessary optimality condition of convex optimization:

$$\operatorname{tr}\left(\frac{\partial J_i\left(x^*, u_i, u_{-i}^*\right)}{\partial \hat{e}_{i \to j}}\Bigg|_{u_i = u_i^*}^T \left(\hat{e}_{i \to j} - \hat{e}_{i \to j}^*\right)\right) \geq 0,$$
$$\forall i \in I,\ \forall j \in N_i,\ \forall \hat{e}_{i \to j} \preceq 0 \tag{9}$$
$$\left[\frac{\partial g_i}{\partial \hat{v}_{i,d}}\bigg|_{\tilde{v}^*} + \mu \operatorname{tr}\left(\left(\tilde{e}_i^*\right)^{-1} \frac{\partial L_i(v_i)}{\partial \hat{v}_{i,d}}\right)\right]\left(\hat{v}_{i,d} - \hat{v}_{i,d}^*\right) \geq 0,$$
$$\forall i \in I,\ \forall \hat{v}_{i,d} \in \hat{v}_i,\ \forall \hat{v}_i \in U_i\left(x^*\right). \tag{10}$$

Eq. (9) implies that $\tilde{e}_j^* - \tilde{e}_i^* = 0, \forall i \in I, j \in N_i$. Because the communication graph is connected, we then have $\tilde{e}_j^* - \tilde{e}_i^* = 0,\ \forall i, j \in I$. Since $\sum_i \tilde{e}_i^* = \sum_{i \in I} L_i(\tilde{v}_i^*) - F_0$, we thus have

$$\tilde{e}_i^* = \frac{1}{|I|}\left(\sum_{i \in I} L_i\left(\tilde{v}_i^*\right) - F_0\right), \forall i \in I \qquad \text{([C-2] proved)}. \tag{11}$$

Substituting (11) into (10), we have

$$\left[\frac{\partial g_i}{\partial \hat{v}_{i,d}}\bigg|_{\tilde{v}^*} + |I|\mu \operatorname{tr}\left(\left(\sum L_i\left(\tilde{v}_i^*\right) - F_0\right)^{-1} \frac{\partial L_i(v_i)}{\partial \hat{v}_{i,d}}\right)\right]$$
$$\cdot\left(\hat{v}_{i,d} - \hat{v}_{i,d}^*\right) \geq 0, \forall i \in I,\ \forall \hat{v}_{i,d} \in \hat{v}_i,\ \forall \hat{v}_i \in U_i\left(x^*\right), \tag{12}$$

which is exactly the sufficient optimality condition of [C-1].

($\Leftarrow$) First, if u^* satisfies [C-3], then $x^* = f(x^*, u^*)$, x^* is a fixed point of the state transition function with u^*. In the second step, we could easily verify (9) from (11) [C-2]. Then we substitute (11) [C-2] into (12) [C-3], we get (10). This completes the optimality condition of

$$u_i^* \in \operatorname{argmin}_{u_i \in U_i(x^*)} J_i\left(x^*, u_i, u_{-i}^*\right), \quad \forall i \in I.$$

B. Proof of Lemma 3

The logarithmic barrier function for the positive definite cone is defined on $X \in S_{++}^n$.

$$B(X) := -\log\det(X) = -\ln \prod_{i=1}^{n} \lambda_i(X) = -\sum_{i=1}^{n} \ln \lambda_i(X).$$

$B(X)$ exerts a repelling force against the boundary of the positive definite cone ∂S_+^n

$$\partial S_+^n = \{X \in S^n \mid \lambda_i(X) \geq 0, i = 1, \ldots, n,$$
$$\text{and } \lambda_i(X) = 0 \text{ for some } i \in \{1, \ldots, n\}\}.$$

As X gets closer to boundary ∂S_+^n, $B(X)$ approaches ∞.

Now consider the optimization problem

$$\min_{v} \sum_{i \in I} g_i(v_i) - |I|\mu \log\det\left(F_0 - \sum_{i \in I} L_i(v_i)\right)$$
$$\text{s.t.} \quad v_i \in V_i \qquad \forall i \in I.$$

As μ gets to 0, the influence of the barrier function reduces and the solution is pushing closer to (but never can achieve) the boundary. Therefore, the problem converges to

$$\min_{v} \sum_{i \in \mathcal{I}} g_i(v_i)$$
$$\text{s.t.} \quad v_i \in V_i \qquad \forall i \in I$$
$$\sum_{i \in I} L_i(v_i) \prec F_0,$$

which is a strict version of our original problem.

References

[1] J. Björk, K. H. Johansson, and F. Dörfler, "Dynamic virtual power plant design for fast frequency reserves: Coordinating hydro and wind," *IEEE Trans. Control Netw. Syst.*, vol. 10, no. 3, pp. 1266–1278, Sep. 2023.

[2] V. Häberle, M. W. Fisher, E. Prieto-Araujo, and F. Dörfler, "Control design of dynamic virtual power plants: An adaptive divide-and-conquer approach," *IEEE Trans. Power Syst.*, vol. 37, no. 5, pp. 4040–4053, Sep. 2022.

[3] L. Badesa, F. Teng, and G. Strbac, "Simultaneous scheduling of multiple frequency services in stochastic unit commitment," *IEEE Trans. Power Syst.*, vol. 34, no. 5, pp. 3858–3868, Sep. 2019.

[4] C. Feng, Y. Wang, P.-Y. Kong, H. Gao, S. Chen, and Q. Chen, "Provision of contingency frequency services for virtual power plants with aggregated models," *IEEE Trans. Smart Grid*, vol. 14, no. 4, pp. 2798–2811, Jul. 2022.

[5] Z. Chu, U. Markovic, G. Hug, and F. Teng, "Towards optimal system scheduling with synthetic inertia provision from wind turbines," *IEEE Trans. Power Syst.*, vol. 35, no. 5, pp. 4056–4066, Sep. 2020.

[6] B. K. Poolla, S. Bolognani, and F. Dörfler, "Optimal placement of virtual inertia in power grids," *IEEE Trans. Autom. Control*, vol. 62, no. 12, pp. 6209–6220, Dec. 2017.

[7] D. Groß, S. Bolognani, B. K. Poolla, and F. Dörfler, "Increasing the resilience of low-inertia power systems by virtual inertia and damping," in *Proc. IREP'2017 Symp.*, 2017, p. 64.

[8] W. Zhong, M. A. Adib Murad, M. Liu, and F. Milano, "Impact of virtual power plants on power system short-term transient response," *Elect. Power Syst. Res.*, vol. 189, Dec. 2020, Art. no. 106609.

[9] C. Feng, K. Zheng, Y. Zhou, P. Palensky, and Q. Chen, "Update scheduling for ADMM-based energy sharing in virtual power plants considering massive prosumer access," *IEEE Trans. Smart Grid*, vol. 14, no. 5, pp. 3961–3975, Sep. 2023.

[10] Z. Khalik, G. P. Padilla, T. C. J. Romijn, and M. C. F. Donkers, "Vehicle energy management with ecodriving: A sequential quadratic programming approach with dual decomposition," in *Proc. Annu. Am. Control Conf. (ACC)*, 2018, pp. 4002–4007.

[11] N. Li and J. R. Marden, "Designing games to handle coupled constraints," in *Proc. 49th IEEE Conf. Decis. Control (CDC)*, 2010, pp. 250–255.

[12] N. Li and J. R. Marden, "Decoupling coupled constraints through utility design," *IEEE Trans. Autom. Control*, vol. 59, no. 8, pp. 2289–2294, Aug. 2014.

[13] Y. Liang, F. Liu, and S. Mei,, "Distributed real-time economic dispatch in smart grids: A state-based potential game approach," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 4194–4208, Sep. 2018.

[14] J. Chen, M. Liu, and F. Milano, "Aggregated model of virtual power plants for transient frequency and voltage stability analysis," *IEEE Trans. Power Syst.*, vol. 36, no. 5, pp. 4366–4375, Sep. 2021.

[15] Y. Shen, W. Wu, B. Wang, and S. Sun, "Optimal allocation of virtual inertia and droop control for renewable energy in stochastic look-ahead power dispatch," *IEEE Trans. Sustain. Energy*, vol. 14, no. 3, pp. 1881–1894, Jul. 2023.

[16] K. Zhou and J. C. Doyle, *Essentials of Robust Control*, vol. 104. Upper Saddle River, NJ, USA: Prentice hall, 1998.

[17] J. S. Shamma, "Game theory, learning, and control systems," *Nat. Sci. Rev.*, vol. 7, no. 7, pp. 1118–1119, 2020.

[18] J. R. Marden, "State based potential games," *Automatica*, vol. 48, no. 12, pp. 3075–3088, 2012.

[19] N. Li and J. R. Marden, "Designing games for distributed optimization," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 2, pp. 230–242, Apr. 2013.