DISTRIBUTED OPTIMIZATION ALGORITHMS FOR GAME OF POWER GENERATION IN SMART GRID

CHUONG VAN NGUYEN
School of Mechanical Engineering
Gwangju Institute of Science and Technology
Gwangju 500-712, Republic of Korea

PHUONG HUU HOANG
Center for Advanced Power Systems
Florida State University
2000 Levy Ave, Tallahassee, FL 32310, USA

HYO-SUNG AHN*
School of Mechanical Engineering
Gwangju Institute of Science and Technology
Gwangju 500-712, Republic of Korea

Abstract. In this paper, we consider a problem of finding optimal power generation levels for electricity users in Smart Grid (SG) with the purpose of maximizing each user’s benefit selfishly. As the starting point, we first develop a generalized model based on the framework of IEEE 118 bus system, then we formulate the problem as an aggregative game, where its Nash Equilibrium (NE) is considered as the collection of optimal levels of generated powers. This paper proposes three distributed optimization strategies in forms of singularly perturbed systems to tackle the problem under limited control authority concern, with rigorous analyses provided by game theory, graph theory, control theory, and convex optimization. Our analysis shows that without constraints in power generation, the first strategy provably exponentially converges to the NE from any initializations. Moreover, under the constraint consideration, we achieve locally exponential convergence result via the other proposed algorithms, one of them is more generalized. Numerical simulations in the IEEE 118 bus system are carried out to verify the correctness of the proposed algorithms.

1. Introduction. Recent decades have witnessed an explosive growth in electricity demand under limited energy sources causing a significant environmental and economic concerns for energy management system (EMS) in power grids. In order for utility companies to enhance energy management efficiency, there are a number of strategies have been proposed. These strategies include game-theoretic power consumption scheduling program and load billing [17], [5], advanced direct load program [22] and interruptible load schedule [21], to name only a few.

2010 Mathematics Subject Classification. Primary: 58F15, 58F17; Secondary: 53C35.

Key words and phrases. Convex optimization, distributed algorithm, aggregative game, limited control authority, smart grid.

*Corresponding author: Hyo-Sung Ahn. This work was supported by the National Research Foundation (NRF) of Korea under the grant NRF-2017R1A2B3007034.
We study here a problem of controlling the power generation at electricity users (consumers) side, in which each user competes others to seek its optimal generation level in order to selfishly maximize its own benefit or minimize cost function while respecting distributed constraints as follows

\[ x_i^* = \arg \min_{x_i \in A_i} f_i(x_i, \bar{x}, \alpha), \quad i \in P, \]

where \( x_i \) is the produced power constrained in the set \( A_i \) known only by user \( i \), \( \alpha \) presents global parameters or information of the grid, \( \bar{x} := \sum_{i \in P} x_i \) and \( P \) is the set of users with its cardinality \( |P| \). The problem (1) refers to an aggregative game, a special class of non-cooperative game, in which the individual cost function of each player (or player in the context of game) depends on its action and the sum of all players’ actions, and the considered set of optimal levels \( \{x_1^*, ..., x_{|P|}^*\} \) constitutes a Nash equilibrium (NE) of the game [10]. As the power network’s size increases, solving the optimization task (1) by centralized strategies might be impractical; hence there is a need for distributed algorithms. This motivates us to design distributed strategies exponentially converging to the optimal solution (1).

Our work finds relevance in the larger context of energy networked problems [17], [5], game theoretical works [10]-[31], limited control authority (LCA) [12], [13] and singularly perturbed system (SPS) [11]. The reference [17] studies a selfish optimization of each user’s power cost, and proves that it is equivalent to the minimization of whole grid power consumption. To seek collective optimal solutions forming a NE in a game in a distributed fashion, each user executes interior point method. It, however, is assumed that each user directly knows detailed scheduled power consumption of others by exchanging information via a fully connected communication network. For tackling this drawback, the work [5] formulates the optimization problem in [17] as an aggregative game with a generalized billing scheme, then solves it by designing a distributed asynchronous gossip-based algorithm employed from [4]. This work only requires the average power consumption of the whole grids to be estimated by each user through communicating with its immediate neighbors. The authors in [5] and [14] share the similar idea that only one player is supposed to be wake up at each tick of a virtual clock to randomly communicate with one neighbor for estimating the time-varying sum of all players’ decisions while forcing its action toward NE of aggregative games under constraints. The work [25] further extends [14] to more general games by excluding each agent’s decision from its estimation vector on players’ actions and taking constant step size into account. Another discrete-time algorithm, alternating direction method of multiplier (ADMM), is recently considered as an effective tool for dealing with distributed seeking NE problem [26]. Compared with a gossip-based method in [25], the designed inexact-ADMM in [26] performs faster convergence rate by defining an extra penalty term as an error between the estimates of each player’s neighbors and using synchronous updates on players’ actions and their estimate vectors. In a recent work [28], the authors propose linearized ADMM-like NE seeking approach using asynchronous updates and fixed step size on gradient descent while ensuring a fast convergence rate as synchronous updates.

None of the approaches mentioned above, however, studies the seeking NE of aggregative games under the LCA. For a network of \( n \) agents, in which each one governed by a distributed dynamic as \( \dot{x}_i = d_i \), where \( d_i \) is the driving command, for \( i \in \{1, 2, ..., n\} \), the LCA is naturally associated with physical limitations on agents’ actuations [12]. Specifically, in dynamic average consensus (DAC) problem, when
each agent has an authority to access time-varying input signal \( u_i(t) : [0, \infty) \rightarrow \mathbb{R} \), and attempts to track the average of the network’s inputs: \( x_i \rightarrow \frac{1}{n} \sum_{j=1}^{n} u_j(t) \), the LCA requires different slow desired convergence rates of agents towards the average if \( f_i \) is implemented in real time based on a consensus algorithm. This requirement yields, for instance, \( |x_i - \frac{1}{n} \sum_{j=1}^{n} u_j(t)| \leq |\kappa_i(0)|e^{-\gamma_i t} \), where \( \kappa_i(0) \) is an initial value, meaning that the tracking rate needs to be less than or equal \( \gamma_i > 0 \). The work \cite{18} addresses the LCA when designing algorithm to deal with the problem (1). It, however, assumes the identity of convergence rate, i.e., \( \gamma_1 = \gamma_2 = \ldots = \gamma_n \), which might be impractical in networked systems. In contrast with the previous work \cite{18}, this work considers different convergence rates toward dynamic agreement under the LCA, hereby seeking the NE at each player’s own pace. Our work herein is also related to SPS comprised of slow and fast motions, which often occur when we mathematically model some natural processes, for example, biochemistry, fluid dynamics, nonlinear kinetics, and so on \cite{27}. We design algorithms in the form of SPS to seek NE of the game (1). Moreover, different from the aforementioned discrete-time algorithms \cite{17}, \cite{5}, \cite{14}-\cite{28}, we develop distributed continuous-time algorithm to seek NE of an aggregative game under constraints. One of proposed methods in our work herein is similar to networked communication protocol of \cite{28} and \cite{31} where all players only agree to exchange message with their immediate neighbors rather than directly observe the remaining players’ states as be assumed in \cite{17}. This leads us to define an aggregate vector associated with each player to keep the time-varying estimation on other players’ information.

Consequently, as the first contribution of this paper, we reformulate the selfish optimization problem of power generation in SG as a game. The physical layer is based on the infrastructure of IEEE 118 bus system, whereas distributed communication in the cyber layer is assumed possible due to smart meters at all units in the network, which helps update global information by distributed fashion. Remarkably, the model in this paper is a generalized one over of energy network model as we already mentioned in \cite{18} and \cite{19}. As our second contribution, we study the optimal energy generation from the view of game theory. It is proved in this paper that different desired rates toward dynamic consensus value among users do not effect to the NE. In comparison with other works in the literature \cite{17},\cite{5},\cite{14}-\cite{31}, our work strongly touches on the LCA issue when designing distributed algorithms. To the best of our knowledge, the seeking NE issue under the concern of LCA has not been studied before. Our third contribution refers to the generalized algorithm in Section 6. Here, since each player can indirectly estimate other players’ information by only communicating with its neighbors, this method is able to be applied for seeking NE of more general games.

In Section 2, some preliminary results in the SPS, limiting equation theorem and convex optimization are provided. We formulate the problem as a game in Section 3. Section 4, 5 and 6 are dedicated to present our distributed strategies. Numerical simulation are illustrated in Section 7. We conclude this paper in Section 8, where we discuss future research directions. All proofs are given in Appendix, together with a basic graph theory.

We denote \( 1_n \in \mathbb{R}_n \) as a column vector with all entries be 1, and \( 1_{n \times n} \in \mathbb{R}_{n \times n} \) be a matrix with all elements be 1. Matrix and vectors are in bold faces. The cardinality of set \( D \) is \( |D| \). Let \((D, d)\) be a metric space. A close \( \delta \)-neighborhood of a set \( B \), which is the set of all nonempty compact subsets of \( D \), is defined by \( N^\delta([B]) = \{ x \in \mathbb{R}^n : \text{dist}(x, B) \leq \delta \} \), where \( \text{dist}(x, B) = \min_{b \in B} d(x, b) \) is the
distance from a point \( x \in \mathbb{R}^n \) to a nonempty compact set \( B \). The relative interior of the set \( D \), denoted \( \text{relint} D \), is defined as \( \text{relint} D = \{ x \in D | N^D(x) \cap \text{aff} D \subseteq D \} \), where \( \text{aff}(D) \) is the affine hull of a finite point set \( D \) in Euclidean space \( \mathbb{R}^n \). The subscript \( r \) and \( q \) refer to reduced states and quasi-steady states. The terms G.A.S., G.E.S. and E.S. refer to globally asymptotically stable, globally exponentially stable and locally exponential stable, resp.

2. Preliminary.

2.1. Singularity perturbed systems. The SPS of dynamical states \( (\theta^*, \vartheta^*) \) around one equilibrium \( (\theta^*, \vartheta^*) \) is derived as

\[
\dot{\theta} = f(t, \vartheta, \dot{\vartheta}, \epsilon), \quad c\dot{\vartheta} = p(t, \theta, \dot{\vartheta}, \epsilon),
\]

with \( \dot{\theta} = \theta - \theta^* \), \( \dot{\vartheta} = \vartheta - \vartheta^* \). The solution of the equation \( \theta = \theta(t, \vartheta, \dot{\vartheta}, \epsilon) \) is defined as the quasi-steady state \( \dot{\vartheta}^q = g(t, \theta) \) such that \( g(t, 0) = 0 \). Let \( \xi = \vartheta - \dot{\vartheta}^q \) and \( B_\epsilon = \{ \theta \in \mathbb{R}^m \mid |D_{\vartheta}g(t, \theta)| < \epsilon \} \). For all \( (t, \vartheta, \epsilon) \in [0, \infty) \times B_\epsilon \times [0, \epsilon_0] \), let the following assumptions (B1-B5) be satisfied:

(B1) \( f(t, 0, 0, \epsilon) = 0 \) and \( p(t, 0, 0, \epsilon) = 0 \),

(B2) \( \dot{\vartheta}^q = g(t, \theta) \) is an isolated root,

(B3) the functions \( f, p, g \) and their first and second derivatives are bounded for \( \dot{\vartheta} - \dot{\vartheta}^q \in B_\epsilon \),

(B4) the origin of the reduced model \( \dot{\theta}^r = f(t, \dot{\theta}^r, \dot{\vartheta}^q, 0) \) is exponential stable (here, we denote \( \dot{\theta}^r \) instead of \( \theta \) thus \( \dot{\vartheta}^q = g(t, \dot{\theta}^r) \) when considering the reduced model),

(B5) the origin of the boundary-layer model \( \frac{d\vartheta}{dt} = p(t, \theta, \xi + g(t, \theta, 0), \epsilon) \), where \( \tau = t/\epsilon \), is exponentially stable, uniformly in \((t, \theta)\), then there exists \( \epsilon^* > 0 \) such that for all \( 0 < \epsilon < \epsilon^* \), the origin of (2) is E.S.. Moreover, if these assumptions satisfy on any compact set of \((\dot{\theta}, \dot{\vartheta})\), a global result is derived \([\text{[11]} - \text{Theorem 11.4}].\)

2.2. Limiting equation theorem. We consider the nonlinear non-autonomous system \( \dot{x} = f(x, t) \), where \( f(x, t) = g(x) + d(x, t) \) with \( d(x, t) \) is a perturbation term to be guaranteed as \( \lim_{t \to +\infty} d(x, t) = 0 \), and this system satisfies one of the following assumptions:

(S1) \( |f(x(t), t)| \) is bounded for any bounded \( x \).

(S2) \( \int_a^b f(x(\theta), t) d\theta \) is bounded for any finite time interval \( \Delta t = b - a \).

If the trajectories of the unforced system \( \dot{x} = g(x) \) asymptotically approach the limiting set \( S \) and \( S \) is bounded, then this set is also asymptotically approached by all bounded solutions of the system \( \dot{x} = f(x, t) \) \([\text{[2]}] \).

2.3. Convex optimization with distributed constraints. We consider the problem of solving convex optimization associated with a network of agents in term of distributed inequality constraints.

\[
\text{minimize } \mathcal{H}(x), \quad \text{s.t. } h_i(x) \leq 0, i = 1, 2, ..., m,
\]

where \( \mathcal{H} : \mathbb{R}^n \to \mathbb{R} \), \( h_i : \mathbb{R} \to \mathbb{R} \) are continuous differential and convex functions, \( m \) is a number of constraints. Let \( D \) be the intersection of the domains of all function \( h_i \). Regarding to Slater’s theorem, if Slater’s condition holds, i.e., there exists a strictly feasible point \( x \in \text{relint} D \) such that \( h_i(x) < 0 \), \( i \in \{1, 2, ..., m\} \),
then strong duality holds. The Slater condition can be refined if some $s$ of the functions $h_i(x)$ are affine and there exists $x \in \text{relint}D$ with $h_i(x) \leq 0$, $i = 1, 2, \ldots, s$ and $h_i(x) < 0$, $i = s+1, 2, \ldots, m$. The strong duality implies that $x^*$ is the solution of (3) if and only if there exists $\lambda^* = [\lambda_1^*, \lambda_2^*, \ldots, \lambda_m^*]^T \in \mathbb{R}^m$, together with $x^*$, satisfies the Karush-Kuhn-Tucker (KKT) condition [3]:

$$h_i(x^*) \leq 0, \lambda_i^* \geq 0, \lambda_i^* h_i(x^*) = 0, \quad i = \{1, 2, \ldots, m\}, \quad (4a)$$

$$\nabla H(x^*) + \sum_{i=1}^m \lambda_i^* \nabla h_i(x^*) = 0, \quad (4b)$$

where $\lambda_i^* h_i(x^*) = 0$ refers to complementary slackness condition. Specifically, we express strict complementary slackness condition if the following assumption is satisfied: $\lambda_i^* = 0 \Rightarrow h_i(x^*) < 0$ and $h_i(x^*) = 0 \Rightarrow \lambda_i^* > 0$.

3. Game formulation. We model a network as a combination of $n$ buses, where each bus belongs to one of four sets: a power load unit set $L = \{L_1, L_2, \ldots, L_{|L|}\}$ consisting of only electricity loads, a power generation unit set $G = \{G_1, G_2, \ldots, G_{|G|}\}$ comprising of only generators, a power load-generation unit set $P = \{P_1, P_2, \ldots, P_{|P|}\}$ presented for a set of electric users equipped with generators and considered as a set of players in the game, and finally a distribution transformers unit denoted as $T = \{T_1, T_2, \ldots, T_{|T|}\}$. Let $V = P \cup G \cup L \cup T$, then $n = |V|$. Each unit is equipped with a smart meter integrated with a CPU and peripherals, thus each meter directly accesses to real-time data of its consumption or generation energy, and exchanges information with its neighboring meters via data links. The proposed distributed algorithms implemented in CPUs enable each unit to compute a consensus information in real-time, then based on that, each player calculates an optimal set point to operate its generator in order to get benefit selfishly in the game. The network of the smart meters, e.g. Local Area Network with Ethernet technology, can be modeled as a connected and undirected graph; hence, graph theory can be applied to analyze the interactions between them (see Appendix A). Remarkably, the model developed by graph theory concept can treat the transformers’ losses and does not require any additional centralized communication channel to broadcast the information of the total consumed power of the grid to all users as in our previous works [18] and [19].

Without loss of generality, we suppose that the sampling time is slotted into hour-long intervals. At each slot, we respectively denote $c^L_i$, $c^G_i$, $c^P_i$, $c^T_i$ as the consumed power of load $i \in L$, the produced power of generator $i \in G$, the consumed power of player $i \in P$, and the power loss at transformer $i \in T$. These values are positive constants and considered as uncontrollable variables, and also assumed to be exactly measured by smart meters at each slot of time. For the sake of analysis, we consider $c_i$ as one of $c^L_i, -c^G_i, c^P_i$ or $c^T_i$ if $i \in L, G, P, T$ correspondingly. When the player $i$ consumes a power, it needs to pay the charge. Thus, in order to curtail that monetary power expense, instead of only purchasing power, each player $i$ needs to produce a power denoted as $x_i$, selling it to the grid and getting some benefits. However, it comes at a cost that each player has to pay for power generation (e.g., fuel cost), yielding trade-offs between electricity bill payment and operational cost. Thus, each player is at most profit if it can seek its optimal power generation and operate at that level defined as $x_i^*$. From a game theory perspective, these optimal values of all players would form a NE, which is the point that none of the players (or decision makers) gets profit by unilaterally moving their actions given the actions
of the others in the game. It means that if we denote $x = [x_i, x_{-i}]^T \in \mathbb{R}^{|P|}$ as a strategy vector in which player $i$ plays an action $x_i \in \mathbb{R}$ and the remaining players play an action profile $x_{-i} \in \mathbb{R}^{|P|-1}$, then the NE is defined as $x^* = [x^*_i, x^*_{-i}]^T \in \mathbb{R}^{|P|}$ mathematically satisfying $f_i(x_i^*, x_{-i}^*, \alpha) \leq f_i(x_i, x_{-i}^*, \alpha)$ for $\forall i \in P$, where $f_i(x)$ is an associated cost function of player $i$ and $\alpha$ is a parameter. Specifically, in our game formulation, the NE is the solution of the following optimization problem

$$\begin{align*}
\text{minimize } & f_i(x_i, \bar{x}, \alpha), \\
\text{s.t. } & x_i \in A_i, \ i \in P,
\end{align*}$$

(5)

where $f_i(x_i, \bar{x}, \alpha) = \mathcal{U}_i(x_i) + \mathcal{J}_i(\bar{x}, \alpha)(c_i - x_i)$ and $A_i$ is the convex action set. The term $\mathcal{U}_i : A_i \rightarrow \mathbb{R}$, capturing the operational cost associated with player $i$, is $\mathcal{U}_i = \phi_i x_i^2 + \psi_i x_i + m_i$. The other term $\mathcal{J}_i(\bar{x}, \alpha)$ refers to power price unit corresponding to billing scheme as in [17], [18] and [16], and is of form $\mathcal{J}_i(\bar{x}, \alpha) = \rho(\alpha - \bar{x}) + \sigma_i$, with $\rho > 0$, $\alpha = \sum_{i \in V} c_i$ and $\bar{x} = \sum_{i \in P} x_i$ be sum of all players’ action in the game. Note that the relation between parameters in $\mathcal{U}_i(x_i)$ and $\mathcal{J}(\bar{x}, \alpha)$ is described as $\phi_i > \frac{(|P|-3)\rho}{2}$ ([18], [19]). To solve (5), the global information $\sum_{i \in V} c_i - \bar{x}$ needs to be precisely estimated at each player. We then trivially obtain

$$\frac{\partial f_i(x_i, \bar{x}, \alpha)}{\partial x_i} = \rho(\bar{x} - \alpha) + (2\phi_i + \rho)x_i + \psi_i - \sigma_i - \rho c_i. \quad (6)$$

Since $\frac{\partial^2 f_i(x_i, \bar{x}, \alpha)}{\partial x_i^2} = 2(\rho + \phi_i) > 0$ for $\forall i \in P$, $f_i(x_i, \bar{x}^*, \alpha)$ is convex in $A_i$. Then under the assumption on convexity of the set $A_i$, the game possess at least one NE [24]. Furthermore, since the Hessian matrix $\mathbf{B} = \left[ \frac{\partial^2 f_i(x_i, \bar{x}, \alpha)}{\partial x_i^2} \right] \in \mathbb{R}^{|P|\times|P|}$ is symmetric matrix and $\frac{\partial^2 f_i(x_i, \bar{x}, \alpha)}{\partial x_i^2} > \sum_{j \neq i, j \in P} \left| \frac{\partial^2 f_i(x_i, x_{-i}, \alpha)}{\partial x_i \partial x_j} \right|$ under $\phi_i > \frac{(|P|-3)\rho}{2}$, one can derive that $\mathbf{B} + \mathbf{B}^T > 0$, leading to the unique existence of NE of the game [24]. In the following sections, distributed algorithms are presented to seek the unique NE, in which the sum $\frac{1}{n}(\bar{x} - \sum_{i \in V} c_i)$ is estimated in distributed fast dynamics, and the NE is simultaneously calculated in slow dynamics. We address the problem that each player is adopted to the LCA, and thus it cannot track the dynamic consensus $\frac{1}{n}(\bar{x} - \sum_{i \in V} c_i)$ at high rate but at its own slow feasible rate.

4. A distributed algorithm for the game without constraints in power generation. In this section, we design distributed algorithms on the basis of the SPS in Section 2.1 and DAC protocol from [13]. In contrast with static average consensus (SAC), in which each agent reaches the mean of constant inputs, the DAC is applied when each agent tracks the mean of time-variant inputs. For the game without constraints in the power generation, we propose the Algorithm 1.

In the Algorithm 1, $l_{ij}$ is an entry of Laplacian matrix $\mathbf{L}$ (see Appendix A) and $k > 0$ is a constant. The action $x_i$ and its derivative $\dot{x}_i$ are considered as slowly time-variant inputs for the fast dynamics. When $\epsilon_i$ is 0, the fast dynamics behave similarly as SAC in terms of the convergence of $\dot{z}_i$ and $w_i$ toward agreement states $\dot{z}_i^q = -(1/n)(\bar{x} - \sum_{i \in V} c_i)$ and $\dot{w}_i^q = -(1/n)\sum_{i \in P} \dot{x}_i$, which are referred to quasi-steady states in the SPS. In practice, the exchange of information in network under the LCA requires a certain amount of time, leading that each player can only track the agreement at its own positive pre-determined rate. Since the designed slow dynamics rely on the agreement value and gradient-based method to seek the NE, the rate of convergence towards NE is bounded above by the pre-determined rate. Associated with each player $i$, we define $y_i$ as its tracking value of the agreement
Algorithm 1 Without power generation limits

**Fast Dynamic:** 
\(i \in P\) in (7a)-(7b) and \(i \in L \cup G \cup T\) in (7c)-(7d)
\[
\begin{align*}
\epsilon_i \dot{z}_i &= -(z_i + x_i - c_i) - \sum_{j \in V} l_{ij}(z_j + \mu_j), \\
\epsilon_i \dot{w}_i &= -(w_i + \hat{x}_i) - \sum_{j \in V} l_{ij}(w_j + \nu_j), \\
\epsilon_i \dot{\hat{x}}_i &= -(z_i - c_i) - \sum_{j \in V} l_{ij}(z_j + \mu_j), \\
\epsilon_i \dot{\hat{\nu}}_i &= -(w_i - w_i^*) - \sum_{j \in V} l_{ij}(w_j + \nu_j), \\
\end{align*}
\] (7a)

**Slow Dynamic:** 
\(i \in P\)
\[
\begin{align*}
\dot{y}_i &= -\gamma_i y_i - \gamma_i z_i - w_i, \\
\dot{x}_i &= -k \frac{\partial f_i(x_i, ny_i)}{\partial x_i} \\
\end{align*}
\] (8)

at a distributed rate \(\gamma_i\) under the LCA. One can achieve a reduced equation for \(y_i\) by substituting the quasi-steady states \(z_i^q\) and \(w_i^q\) into equation (8) as \(\dot{y}_i^q = -\gamma_i y_i - \gamma_i z_i^q - w_i^q\). It yields a solution \(|y_i^q - \frac{1}{n}(\bar{x} - \sum_{i \in V} c_i)| = |y_i^q - \frac{1}{n}(\bar{x}(0) - \sum_{i \in V} c_i)|e^{-\gamma_i t}\). Then, according to Theorem 11.2 [11], we have
\[
|y_i - \frac{1}{n}(\bar{x} - \sum_{i \in V} c_i)| \leq O(\epsilon_i) + |\kappa_i(0)|e^{-\gamma_i t},
\] (9)

where \(\kappa_i(0) = y_i(0) - \frac{1}{n}(\bar{x}(0) - \sum_{i \in V} c_i)\). This motivates us to define \(y_i\) as a value tracking \(\frac{1}{n}(\bar{x} - \sum_{i \in V} c_i)\) and to utilize the function \(f_i(x_i, ny_i)\) in (8) as \(f_i(x_i, ny_i) = U_i(x_i) + \tilde{J}_i(ny_i)(c_i - x_i)\) with its derivative being of a form
\[
\frac{\partial f_i(x_i, ny_i)}{\partial x_i} = \rho ny_i + (2\phi_i + \rho)x_i + \psi_i - \sigma_i - \rho c_i.
\] (10)

We shift the system (7)-(8) to equilibrium points \((z^*, \mu^*, w^*, \nu^*, y^*)\) corresponding to the NE \(x^*\) by considering variable changes \(\hat{z}_i = z_i - z_i^*, \hat{\mu}_i = \mu_i - \mu_i^*, \hat{w}_i = w_i - w_i^*, \hat{\nu}_i = \nu_i - \nu_i^*\) for \(i \in V\) and \(\hat{x}_i = x_i - x_i^*\) \(\hat{y}_i = y_i - y_i^*\) for \(\forall i \in P\) to obtain the following transformed system

**Fast Dynamic:** 
\(i \in P\) in (11a)-(11b), \(i \in L \cup G \cup T\) in (11c)-(11d)
\[
\begin{align*}
\epsilon_i \dot{\hat{z}}_i &= -(\hat{z}_i + \hat{x}_i - \hat{c}_i) - \sum_{j \in V} l_{ij}(\hat{z}_j + \hat{\mu}_j), \\
\epsilon_i \dot{\hat{w}}_i &= -(\hat{w}_i + \hat{x}_i) - \sum_{j \in V} l_{ij}(\hat{w}_j + \hat{\nu}_j), \\
\epsilon_i \dot{\hat{x}}_i &= -(\hat{z}_i - \hat{c}_i) - \sum_{j \in V} l_{ij}(\hat{z}_j + \hat{\mu}_j), \\
\epsilon_i \dot{\hat{\nu}}_i &= -(\hat{w}_i - \hat{w}_i^*) - \sum_{j \in V} l_{ij}(\hat{w}_j + \hat{\nu}_j), \\
\end{align*}
\] (11a)

**Slow Dynamic:** 
\(i \in P\)
\[
\dot{\hat{y}}_i = -\gamma_i \hat{y}_i - \gamma_i \hat{z}_i - \hat{w}_i, \\
\dot{\hat{x}}_i = -k (\rho ny_i + (2\phi_i + \rho)\hat{x}_i),
\] (12)

which can be verified to follow the condition (B1), and \(\hat{c}_i = 0\) as \(c_i\) is constant at each slot of time. To derive boundary-layer model, we rely on the fact that the fast dynamics (11) globally exponentially converge to \(\hat{z}_i^q = -\frac{1}{n}1_{[p]} \hat{x}\) and globally converge to \(\hat{\mu}_i^q\) [8]. These states are isolated roots of (11a) when \(\epsilon_i = 0\), satisfying
the condition (B2). From (11a), by defining $\tilde{\xi}_i^t = \tilde{z}_i^t - \tilde{z}_i^q$, $\tilde{\xi}_i^u = \tilde{\mu}_i - \tilde{\mu}_i^q$, $\tau_i = t/\epsilon_i$, we arrive at a boundary-layer model as follows

$$d\tilde{\xi}_i^t/d\tau_i = -\tilde{\xi}_i^t - \sum_{j \in V} l_{ij}(\tilde{\xi}_i^t + \tilde{\xi}_j^u), \quad d\tilde{\xi}_i^u/d\tau_i = \sum_{j \in V} l_{ij}\tilde{\xi}_j^t,$$

(13)

and also obtain an analogous boundary-layer model for (11b)-(11d). A system constructed from (13) for $\forall i \in P, j \in V$ is G.E.S and uniformly in $(t, \bar{x})$, which is originally analyzed in [8] and thus omitted. Thus, we can obtain the following result

**Lemma 4.1.** (Boundary-layer model) The origin of the boundary-layer model constructed from (11) is G.E.S., uniformly in $(t, \bar{x})$. We then study the convergence and stability property of reduced model for the transformed system (11)-(12) around the NE.

**Lemma 4.2.** The origin of a reduced model for the transformed system (11)-(12) is G.E.S.

**Proof.** See Appendix B.

Now, we introduce a result about the convergence of the Algorithm 1.

**Theorem 4.3.** There exists $\epsilon^* > 0$ such that for all $0 < \epsilon_i < \epsilon^*$, $i \in V$, the origin of the transformed system (11)-(12) is G.E.S. Therefore, if each player unit $i \in P$ adopts the proposed distributed algorithm (7a)-(7b) and (8), whereas others follow (7c)-(7d), then its action eventually reaches the optimal power generation level in a distributed fashion from any initialization.

**Proof.** See Appendix B.

5. A distributed algorithm for the game under constraints in power generation.

**Lemma 5.1.** The non-cooperative problem (5) is equivalent to the convex optimization problem with distributed inequality constraints.

$$\text{minimize } \mathcal{H}(x, \alpha),$$

$$\text{s.t. } h_i(x) \leq 0, \quad h_{ui}(x) \leq 0, \quad i \in P.$$

(14)

Here, $\mathcal{H} : \mathbb{R}^{|P|} \rightarrow \mathbb{R}, \mathcal{H}(x, \alpha) = \mathcal{H}_1(x, \alpha) + \mathcal{H}_2(x, \alpha)$, with $\mathcal{H}_1(x, \alpha) = \sum_{i \in P} f_i(x, \alpha)$ and $\mathcal{H}_2(x, \alpha) = -\frac{\mu}{2} \sum_{i \in P} \left( (c_i - x_i) \sum_{j \neq i, j \in P} (c_j - x_j) \right)$. In addition, $h_i(x) = x_i - c_i$, and $h_{ui}(x) = x_i - x_i^u$. Under the assumption $\phi_i > \frac{2|P|-3\mu}{\mu}$, then $\mathcal{H}(x, \alpha)$ is strictly convex and its optimizer $x^*$ is unique. Moreover, under the Slater’s condition in Section 2.3, there exists $\lambda^*$ so that $(x^*, \lambda^*)$ satisfies the KKT condition (4), and the existence is unique according to Lemma 1 [18].

**Proof.** See Appendix B.

It is worth distinguishing the optimization problem (14) from distributed energy resources (DERs) [1] or economic dispatch problem (EDP) [6], in which all the agents in an interconnected networked system work cooperatively to minimize the global cost $\sum_{i \in P} f_i(x)$ under constraints. Our previous work in [18] was developed under the assumption that the identity of the feasible convergence rates of all players towards the agreement estimation are same, which yields a major restriction in
practical issues. Therefore, in this work, we generalize the previous work by consider
different convergence rates among them. For the game under constraints in power
generation, we propose the Algorithm 2.

**Algorithm 2** With respecting power generation limits

**Fast Dynamic:** (7)

**Slow Dynamic:** for \( i \in P \)

\[
\begin{align*}
y_i &= -\gamma_i y_i - \gamma_i z_i - w_i, \quad (15a) \\
\dot{\lambda}_i &= \lambda_i (x_i^d - x_i), \dot{\lambda}_u = \lambda_u (x_i - x_i^u), \dot{x}_i = -k \left( \frac{\partial f_i(x_i, y_i)}{\partial x_i} - \lambda_i + \lambda_u \right) \quad (15b)
\end{align*}
\]

We will show that the trajectories of the dynamical system in the Algorithm 2
converge to the \((y^\ast, x^\ast, x^\ast)\), where \(x^\ast\) is the unique NE of the game. For the sake of
analysis, we denote the set \(D^A = D^A_1 \cup D^A_u\), where \(D^A_1 = \{ i \in P : \lambda_i^r = 0 \}\) is composed of
\(\zeta\) elements and \(D^A_u = \{ i \in P : \lambda_u^r \neq 0 \}\) composed of \(\zeta\) elements, then \(\zeta + \zeta \leq |P|\),
and \(D^A \cap D^A_u = \emptyset\) because generator cannot operate at the minimum and maximum
value at the same time. We also define \(D^f = D^f_1 \cap D^f_u\) where \(D^f_1 = \{ i \in P : \lambda_i^f = 0 \}\),
\(D^f_u = \{ i \in P : \lambda_u^f = 0 \}\), and assume strictly complementary slackness condition is satisfied,
\(i.e., x_i^r > x_i^f\) for \(i \in D^f_1\) and \(x_i^r < x_i^f\) for \(i \in D^f_u\), which leads to \(P = D^A \cup D^f\)
and \(D^A \cap D^f = \emptyset\). By following an analogous analysis as presented in Section 4,
we obtain a transformed system \((15)\) around \((y^\ast, x^\ast, x^\ast)\) as follows

\[
\begin{align*}
\dot{y}_i &= -\gamma_i y_i - \gamma_i z_i - w_i, \quad (16a) \\
\dot{\lambda}_i &= f_i(\hat{\lambda}_i, \hat{x}_i, \hat{\lambda}_u), \dot{\lambda}_u = f_u(\hat{\lambda}_u, \hat{x}_i), \dot{x}_i = -k \left( \rho n y_i + (2 \phi_i + \rho) \hat{x}_i - \hat{\lambda}_i + \hat{\lambda}_u \right). \quad (16b)
\end{align*}
\]

where \(f_i(\hat{\lambda}_i, \hat{x}_i, \hat{\lambda}_u) = \hat{\lambda}_i (x_i^d - x_i) - \hat{\lambda}_i \hat{x}_i - \hat{\lambda}_i \hat{x}_i, f_u(\hat{\lambda}_u, \hat{x}_i) = \hat{\lambda}_u (x_i^r - x_i^u) + \hat{\lambda}_u \hat{x}_i + \hat{\lambda}_u \hat{x}_i \). Moreover, the above transformed system in slow dynamic can be verified to satisfy condition (B1) and a reduced model is obtained as

\[
\begin{align*}
\dot{y}_i^r &= \frac{1}{n} 1^\top P X^r - \gamma_i \hat{\delta}_i^r(0) e^{-\gamma_i t}, \quad (17a) \\
\dot{\lambda}_i^r &= f_i(\hat{\lambda}_i^r, \hat{x}_i^r, \hat{\lambda}_u^r) = f_u(\hat{\lambda}_u^r, \hat{x}_i^r), \dot{x}_i^r = -k \left( f_2(\hat{\lambda}_i^r, \hat{\lambda}_u^r) + v_i(t) \right) \quad (17b)
\end{align*}
\]

where \(f_2(\hat{\lambda}_i^r, \hat{\lambda}_u^r) = \rho P \hat{\delta}_i^r(0) X^r + (2 \phi_i + \rho) \hat{x}_i^r - \hat{\lambda}_i^r + \hat{\lambda}_u^r, v_i(t) = v_i(t) e^{-\gamma_i t}\) with \(v_i(0) = \rho \hat{\delta}_i^r(0) P\). To proceed, we denote \(\hat{\sigma}_i^r = [\hat{\lambda}_i^r, \hat{\lambda}_u^r, \hat{x}_i^r]^\top \in \mathbb{R}^3\) such that the set of equations (17b) is written as \(\dot{\hat{\sigma}}_i^r = \hat{f}_i(\hat{\sigma}_i^r, t)\). We then define \(\hat{\sigma}_r^r = [\hat{\sigma}_1^r, \ldots, \hat{\sigma}_{|P|}^r]^\top \in \mathbb{R}^{3|P|}\) and \(\hat{f}_r(\hat{\sigma}_r, t) = [\hat{f}_1(\hat{\sigma}_1, t), \ldots, \hat{f}_{|P|}(\hat{\sigma}_{|P|}, t)]^\top \in \mathbb{R}^{3|P|}\). By following the analysis in Lemma 4.2, one can obtain a reduced model in the Algorithm 2 as

\[
\begin{align*}
\dot{y}_i^r &= \hat{f}_r(y_i^r, \sigma_r^r, t), \quad (18a) \\
\dot{\sigma}_r^r &= \hat{\tau}_r(\sigma_r^r, t). \quad (18b)
\end{align*}
\]

where \(\hat{f}_r(y_r^r, \sigma_r^r, t) = -k B y_r^r + \frac{1}{n} 1^\top P (\hat{\lambda}_i^r - \hat{\lambda}_u^r) + k D \hat{\delta}_i^r(0) r(t), \hat{\lambda}_i^r = [\hat{\lambda}_{i1}^r, \ldots, \hat{\lambda}_{i|P|}^r]^\top \in \mathbb{R}^{3|P|}, \hat{\lambda}_u^r = [\hat{\lambda}_{u1}^r, \ldots, \hat{\lambda}_{u|P|}^r]^\top \in \mathbb{R}^{3|P|}\). For the sake of analysis, we separate the right-hand side of two subsystems (18) as \(\hat{f}_y(y_r^r, \sigma_r^r, t) = f_y(y_r^r, \sigma_r^r) + f_u(y_r^r, \sigma_r^r)\).
Next, we revisit some notations in [18] as follows. We define \( \hat{\phi} \) vector to store temporary estimation on others' information which is computed actions. Related to this topic, in [25], [28] and [31], each player utilizes an aggregate here is that each player can solve more general games if it knows other players’ A generalized distributed algorithm for the game via leader-following

Theorem 5.4. Assume Slater condition, strict complementary slackness condition and \( \phi_i > \frac{(|P|−3)/2}{2} \) are satisfied. There exists \( \epsilon^* > 0 \) such that \( 0 < \epsilon_i < \epsilon^* \), \( \forall i \in V \), if every unit \( i \in V \) adopts the proposed Algorithm 2 and the initialization of \( \lambda \) is positive, then each player’s action eventually reaches its optimal power generation level locally in a distributed fashion.

Proof. See Appendix B.

6. A generalized distributed algorithm for the game via leader-following consensus approach. The algorithm in the Section 5 is specifically designed for the aggregative game since only dynamic average of information is required to be estimated by all players. This restriction motivates our more general design, which can be applied for more general games, and the aforementioned game, therefore, is only one special case. In general problems, \( f_i(x) \) can be a quadratic function [29], an aggregate of logarithmic functions in congestion game [30], or specific potential functions in potential game [15], etc. To tackle these problems, the idea we consider here is that each player can solve more general games if it knows other players’ actions. Related to this topic, in [25], [28] and [31], each player utilizes an aggregate vector to store temporary estimation on others’ information which is computed
through distributed methods. But, different from these works, we consider the LCA that enables each player to track the information of others at its own pace.

We apply a leader-follower consensus approach in [23] to our design. Each \(i \in V\) is considered as a virtual leader broadcasting a reference information denoted as \(\text{ref}_i(t)\), with \(\text{ref}_i(x) = c_i\) if \(i \in P\) and \(\text{ref}_i(t) = -c_i\) otherwise. That reference cannot be observed directly by nonneighboring units but estimated via message exchanged among neighbors in the communication graph. To this end, we define a local vector associated with player \(i\) as \(y_i = [y_{i1}, y_{i2}, \ldots, y_{in}] \in \mathbb{R}^n\), where \(y_{ij}\) is its estimation on the unit \(j\)'s reference, \(j \in V\). Under the LCA, we aim to design an algorithm satisfying

\[
|y_{ij} - \text{ref}_j(t)| \leq O(\epsilon_i) + |\kappa_{ij}(0)|e^{-\gamma_i t},
\]

where \(\kappa_{ij}(0)\) is an initial value. We propose the Algorithm 3 as follows

**Algorithm 3** With respecting power generation limits for a larger game

**Fast Dynamic:** \(i, j \in V\)

\[
\epsilon_i \dot{z}_{ij} = -z_{ij} + a_{ij} - \sum_{k \in V} a_{ik}(\chi_{i,k}^z + \chi_{i,k}^\mu), \quad \epsilon_i \dot{\mu}_{ij} = \sum_{k \in V} a_{ik} \chi_{i,k}^z, \quad \epsilon_i \dot{w}_{ij} = -w_{ij} + g_{ij} - \sum_{k \in V} a_{ik}(\chi_{i,k}^w + \chi_{i,k}^\nu), \quad \epsilon_i \dot{v}_{ij} = \sum_{k \in V} a_{ik} \chi_{i,k}^w, \quad \epsilon_i \dot{d}_{ij} = -d_{ij} + r_{ij} - \sum_{k \in V} a_{ik}(\chi_{i,k}^d + \chi_{i,k}^\eta), \quad \epsilon_i \dot{\eta}_{ij} = \sum_{k \in V} a_{ik} \chi_{i,k}^d.
\]

**Slow Dynamic:** \(i \in P, j \in V\)

\[
\dot{y}_{ij} = -\gamma_i w_{ij} - \gamma_i w_{ij}/z_{ij} - d_{ij}/z_{ij},
\]

\[
\dot{\lambda}_i = \lambda_i(x_i^a - x_i), \quad \dot{\lambda}_u = \lambda_u(x_i - x_u), \quad \dot{x}_i = -k \left( \frac{\partial f_i(x_i, y_i)}{\partial x_i} - \lambda_i + \lambda_u \right).
\]

In the Algorithm 3, \(a_{ij}\) is an entry of adjacent matrix \(A_a\) (see Appendix A) and \(k > 0\) is a constant. The difference between these estimate vectors \(i\) and \(k\) on the information of a unit \(j\) is defined as \(\chi_{i,j}^z = z_{ij} - z_{kj}, \chi_{i,j}^\mu = \mu_{ij} - \mu_{kj}, \chi_{i,j}^w = w_{ij} - w_{kj}, \chi_{i,j}^\nu = \nu_{ij} - \nu_{kj}, \chi_{i,j}^d = d_{ij} - d_{kj}\) and \(\chi_{i,j}^\eta = \eta_{ij} - \eta_{kj}\). Additionally, \(g_{ij} = a_{ij}(x_j + c_j)\) and \(r_{ij} = -a_{ij} x_j\) for \(j \in P\), and \(q_{ij} = a_{ij} c_j\) and \(r_{ij} = 0\) for \(j \in L \cup G\). We denote \(z_i = [z_{i1}, \ldots, z_{in}]^{T} \in \mathbb{R}^n\) as auxiliary estimation vector associated with player \(i\) and similarly define for \(\mu_i, w_i, v_i, d_i, \eta_i\). Moreover, the fast dynamic (20) exponentially converge to \(z_{ij} = \frac{1}{n} \sum_{k \in V} a_{kj}(x_k - \chi_{i,k}^z + \chi_{i,k}^\mu), w_{ij} = \frac{1}{n} \sum_{k \in V} a_{kj}(x_k - \chi_{i,k}^w + \chi_{i,k}^\nu), \) and \(d_{ij} = \frac{1}{n} \sum_{k \in V} a_{kj}(x_k - \chi_{i,k}^d + \chi_{i,k}^\eta)\). Thus, \(w_{ij}/z_{ij}\) converges to \(w_{ij}^{\text{ref}}/z_{ij}^{\text{ref}} = -(x_j - c_j)\) for \(j \in P\) and to \(c_j\) for \(j \in L \cup G\). Therefore, \(d_{ij}/z_{ij}\) converges to \(d_{ij}^{\text{ref}}/z_{ij}^{\text{ref}} = -x_j\) for \(j \in P\) and to \(0\) otherwise, under initialization \(z_{ij}(0)\) satisfying Lemma 6.2. By substituting these quasi-steady states into (21a), we have \(y_{ij}^{\text{ref}} = (x_j - c_j) + \kappa_i p(0) e^{-\gamma_i t}\) for \(j \in P\) and \(y_{ij} = -c_j + \kappa_i p(0) e^{-\gamma_i t}\) for \(j \in L \cup G\).
Then,
\[
|y_{ij} - (x_j - c_j)| \leq O(\varepsilon_i) + |\varepsilon_{ij}| e^{-\gamma t}, \quad j \in P \tag{22a}
\]
\[
|y_{ij} + c_j| \leq O(\varepsilon_i) + |\varepsilon_{ij}| e^{-\gamma t}, \quad j \in L \cup G \cup T \tag{22b}
\]
satisfying the LCA requirement.

We associate these equilibrium points \((z^*, \mu^*, w^*, v^*, d^*, \eta^*, y^*, \lambda^*)\) with the NE \(x^*\), and analyze the system (20)-(21) around these equilibriums. Let us obtain the following transformed system:

**Fast Dynamic:** for \(i, j \in V\)
\[
\varepsilon_i \ddot{z}_{ij} = -\ddot{z}_{ij} - \sum_{k \in V} a_{ik} (\ddot{\lambda}_{i,k} + \ddot{\lambda}_{i,k}^\mu), \varepsilon_i \ddot{\mu}_{ij} = \sum_{k \in V} a_{ik} \ddot{\lambda}_{i,k}^\mu, \tag{23a}
\]
\[
\varepsilon_i \ddot{w}_{ij} = -\ddot{w}_{ij} - \ddot{\mu}_{ij} - \sum_{k \in V} a_{ik} (\ddot{w}_{i,k} + \ddot{w}_{i,k}^\nu), \varepsilon_i \ddot{\nu}_{ij} = \sum_{k \in V} a_{ik} \ddot{w}_{i,k}^\nu, \tag{23b}
\]
\[
\varepsilon_i \ddot{x}_{ij} = -\ddot{x}_{ij} - \ddot{\nu}_{ij} - \sum_{k \in V} a_{ik} (\ddot{x}_{i,k}^d + \ddot{x}_{i,k}^\eta), \varepsilon_i \ddot{\eta}_{ij} = \sum_{k \in V} a_{ik} \ddot{x}_{i,k}^d, \tag{23c}
\]

**Slow Dynamic:** for \(i \in P, j \in V\)
\[
\dot{y}_{ij} = -\gamma_i y_{ij} - \gamma_i \left( \frac{w_{ij}}{z_{ij}} - \frac{v_{ij}}{z_{ij}} \right) - \left( \frac{d_{ij}}{z_{ij}} - \frac{d_{ij}}{z_{ij}} \right), \tag{24a}
\]
\[
\dot{\lambda}_{ij} = f_{\lambda_i}(\dot{\lambda}_{i,j}, \dot{x}_i), \dot{\lambda}_{ui} = f_{\lambda_u}(\dot{\lambda}_{u,i}, \dot{x}_i), \tag{24b}
\]
\[
\dot{x}_i = -k \left( \rho \sum_{j \in V, j \neq i} \ddot{y}_{ij} + 2(\phi + \rho) \ddot{x}_i - \ddot{\lambda}_{i,j} + \ddot{\lambda}_{ui} \right). \tag{24c}
\]

We first derive a reduced equation for \(\dot{y}_{ij}\) as \(\ddot{y}_{ij} = -\gamma_i y_{ij} + \gamma_i \ddot{x}_{ij} + \ddot{x}_j\) if \(j \in P\) and \(\ddot{y}_{ij} = \ddot{y}_{ij}\) if \(j \in L \cup G \cup T\). Thus, \(\ddot{y}_{ij} = \dot{x}_j + \theta_{ij}(0)e^{-\gamma t}\) for \(j \in P\) and \(\ddot{y}_{ij} = \theta_{ij}(0)e^{-\gamma t}\) for \(j \in L \cup G \cup T\), where \(\theta_{ij}(0) = \ddot{y}_{ij}(0) - \ddot{x}_j(0)\). Similar to the analysis in Section 5, a reduced model is finally obtained as
\[
\ddot{y}_{ij} = \begin{cases} 
\dot{x}_j - \gamma_i \theta_{ij}(0)e^{-\gamma t} & \text{if } j \in P \\
-\gamma_i \theta_{ij}(0)e^{-\gamma t} & \text{if } j \in L \cup G \cup T
\end{cases} \tag{25a}
\]
\[
\ddot{\lambda}_{ij} = f_{\lambda_i}(\dot{\lambda}_{i,j}, \dot{x}_i), \ddot{\lambda}_{ui} = f_{\lambda_u}(\dot{\lambda}_{ui}, \dot{x}_i), \dot{x}_i = -k \left( \rho \sum_{j \in V, j \neq i} \ddot{y}_{ij} + 2(\phi + \rho) \ddot{x}_i - \ddot{\lambda}_{i,j} + \ddot{\lambda}_{ui} \right) \tag{25b}
\]

with \(h_i(t) = \rho \sum_{j \in V, j \neq i} \theta_{ij}(0)e^{-\gamma t}\).

We then construct a boundary-layer model from (23a) by applying the changes of variable \(\dot{s}_{ij} = \ddot{z}_{ij} - \ddot{\mu}_{ij}, \dot{s}_{ij} = \ddot{x}_{ij} - \ddot{z}_{ij}\)
\[
\varepsilon_i \dot{s}_{ij} = \sum_{k \in V} a_{ik} (\dot{s}_{ik} - \dot{s}_{ek}^\mu), \varepsilon_i \dot{s}_{ij} = \sum_{k \in V} a_{ik} (\dot{s}_{ik} - \dot{s}_{ek}^\mu + \dot{s}_{ij} - \dot{s}_{ij}^\mu). \tag{26}
\]

Let \(\dot{s}_1 = [\dot{s}_{1j}, \dot{s}_{12}, \ldots, \dot{s}_{1n}]^\top, \dot{s}_2 = [\dot{s}_{2j}, \dot{s}_{22}, \ldots, \dot{s}_{2n}]^\top, \dot{s}_3 = [\dot{s}_{3j}, \dot{s}_{32}, \ldots, \dot{s}_{3n}]^\top, \dot{s}_4 = [\dot{s}_{4j}, \dot{s}_{42}, \ldots, \dot{s}_{4n}]^\top \in \mathbb{R}^n, \dot{s}_5 = [\dot{s}_1^\top, (\dot{s}_2^\top)^\top, \ldots, (\dot{s}_4^\top)^\top] \in \mathbb{R}^{4n}\) and consider the time exchange expression \(\tau = tf/\epsilon\) to write the system (26) in a matrix form as
\[
\frac{d\dot{s}_1}{dt} = (L \otimes I)\dot{s}_2, \quad \frac{d\dot{s}_2}{dt} = -\dot{s}_1 - (L \otimes I)\dot{s}_2 - (L \otimes I)\dot{s}_3. \tag{27}
\]

In an analogous manner, one also can obtain a similar boundary-layer model for (23b)-(23c).

**Lemma 6.1.** (Boundary-layer model) The origin of the boundary-layer model for (23) is G.E.S., uniformly in \((t, \dot{x})\).
Proof. See Appendix B 

In order to guarantee \( z_{ij} \neq 0 \) in (21a), we establish a condition as follows

**Lemma 6.2.** If \( z_{ij} \) is initialized at the neighborhood of their quasi-steady-state \( z_{ij}^q \) respectively, i.e., there exists a small constant \( \rho > 0 \) such that \( \max|z_{ij}(0) - z_{ij}^q| \leq \rho \), then \( z_{ij} > 0 \) for \( \forall i, j \in V \).

**Proof.** See Appendix B 

Based on the analysis of boundary-layer model and reduced model, the final seeking result is stated as

**Theorem 6.3.** Assume Slater condition, strict complementary slackness condition and \( \phi_i > (|P_i| - 3)p \) are satisfied. There exists \( \epsilon^* > 0 \) such that \( 0 < \epsilon_i < \epsilon^* \) for \( \forall i \in V \), if every unit in set \( V \) adopts the proposed Algorithm 3 and their initializations satisfy Lemma 6.2 with \( \lambda_i(0), \lambda_u(0) > 0 \), then each player's action eventually reaches the optimal power generation level locally in a distributed fashion.

**Proof.** See Appendix B 

**Remark 1.** (Comparison between the Algorithm 2 and the Algorithm 3) Although Algorithm 3 can be applied to solve more general games, it desires more computational resources than Algorithm 2, because each unit in Algorithm 3 maintains \( n \)-dimension real-time estimation vectors to exchange them with its neighbors, while in Algorithm 2 each unit needs only one variable, an estimation of average sum, to be exchanged.

7. Simulation in IEEE 118 bus system. IEEE 118 bus system provides a simple approximation of the American electric power system, comprising of 118 nodes, 91 loads, 54 generators, 9 transformers and 186 branches [9]. Regarding to this system, the size of each set is characterized by \( |L| = 54 \), \( |T| = 9 \), \( |G| = 17 \), \( |P| = 37 \). The local cost function for operating generator \( i \) is given as \( \phi_i x_i^2 + \psi_i x_i + m_i \), where its coefficients are belong to the set \( \phi_i = [0.0024, 0.0697] \), \( \psi_i = [8.3391, 37.6968] \), and \( m_i = [6.78, 74.33] \). The consumed power of each unit \( i \in L \cup P \) belongs to the range \( c_i = [2, 277] \) (MW), and the power loss of unit \( i \in T \) happens randomly from uniform distribution \( c_i^p \sim \mathcal{U}(0.01, 0.05) \) (MW). The convergence rate under the LCA is chosen from uniform distribution \( \gamma_i \sim \mathcal{U}(0.1, 10) \). Simulation results for Algorithm 1 can be seen in Fig. 1A, 1B and 1C, whereas Fig. 2A, 2B and 2C are devoted for Algorithm 2. For Algorithm 3, the simulation results are illustrated in Fig. 3A, 3B and 3C. For the sake of visualization, a number of player units in a set \( \{x_4, x_{12}, x_{49}, x_{70}, x_{74}, x_{110}\} \) are also taken into consideration, and their convergence trajectories to optimal power generation levels are in Fig. 1C, 2C and 3C. The NE achieved by Algorithm 3 is at the same value and as fast as Algorithm 2.

With Algorithm 1, estimation \( y_i \) on the average of aggregate sum \( \frac{1}{2} (\bar{x} - \sum_{i \in V} c_i) \) is calculated in Fig. 1A, converging to the exact consensus value at \( -31.3568 \) (MW). In Fig. 1C, the NE is determined as \( \{x_4^*, x_{12}^*, x_{49}^*, x_{70}^*, x_{74}^*, x_{110}^*\} = [0.1908, 1.8883, 0.4739, 0.1560, 0.0274, 0.1312] \) (MW). For Algorithm 2, the Fig. 2A presents the estimation of \( \frac{1}{n} (\bar{x} - \sum_{i \in V} c_i) \) as \( -29.8339 \) (MW), which is achieved by all players in the game. Player units adopting this algorithm can seek their optimal power generation levels while respecting the given constraints as illustrated in Fig. 2B. For instance, in Fig. 2C, we consider limit capacity constraints for some players as \( x_4 \in [3.1, 30] \) (MW), \( x_{12}^* \in [4.5, 40] \) (MW), \( x_{49}^* \in [8.5, 80] \) (MW),
$x^*_70 \in [6.5, 60]$ (MW), $x^*_74 \in [6.5, 60]$ (MW), $x^*_110 \in [3.8, 40]$ (MW), and NE is found as $[x^*_4, x^*_12, x^*_49, x^*_70, x^*_74, x^*_110] = [3.1, 4.5, 8.5, 6.5, 6.3, 8]$ (MW), satisfying the corresponding constraints. This verifies the correctness of Algorithm 2. For Algorithm 3, each player unit is able to estimate information of other players by exchanging message. For example, Fig. 3A shows that all units’ estimations on information of a player 4, which is $x_4 - c_4$, exactly converge to the value value $x^*_4 - c_4 = -28.75$ (MW). Moreover, Fig. 3B illustrates how the player 4 estimates information of other players in the game.

(Figure 1. Simulation results for Algorithm 1)

8. Conclusion. In this paper, we developed a model for the SG based on the framework of IEEE 118 bus system. We showed that the problem of seeking the optimal set point for operating generators to achieve benefits selfishly turns out to be considered as an aggregative game in the interconnected network. To tackle the problem, we propose three distributed algorithms to reach that optimal set point, which is considered as a Nash Equilibrium of the game. The first one is designed when constraints are ignored, while the others take constraints into account. The simulation in the IEEE 118 bus system demonstrates the effectiveness of the proposed algorithms. There are some directions for future work. Firstly, other technical constraints associated to power flows and thermal phenomena may be considered. Secondly, we will extend our model for the network with uncertainties and switching topologies.

Appendix A. (Communication graph) The connected and undirected graph consists of $n$ vertices, in which each vertex is presented for individual unit’s smart meter,
and there always exists a path of data links between each pair of smart meters. The graph is denoted as \( G = (V, E) \), where \( V = \{ L_1, ..., L_{|L|}, G_1, ..., G_{|G|}, P_1, ..., P_{|P|}, T_1, ..., T_{|T|} \} \) is the vertex set and \( E \subseteq V \times V \) is the edge set. If \((i, j) \in E\) then \((j, i) \in E\), and \(i, j\) are neighboring meters. We define \( \mathcal{N}_i = \{ j \in V | (i, j) \in E \} \) as a set of neighbors of the smart meter \( i \). Associated with the graph, the adjacency matrix \( A_n = \{ a_{ij} \} \in \mathbb{R}^{n \times n} \) is constructed according to interaction between meters, where \( a_{ij} = 1 \) if \((i, j) \in E\), otherwise, \( a_{ij} = 0 \). Moreover, the associated Laplacian matrix \( L = [l_{ij}]_{n \times n} \) is computed from \( L = A_d - A_n \), where \( A_d \in \mathbb{R}^{n \times n} \) is a diagonal matrix with \( i \)-th element on the diagonal being \( \sum_{j \in \mathcal{N}_i} a_{ij} \). Here, \( L1_n = 0_n \) and \( 1_n^T L = 0_n^T \) [20].

**Appendix B.**

**Proof of Lemma 4.2:** We firstly substitute quasi-steady states \( \hat{\xi}^q_i = -\frac{1}{n} 1_i^T \hat{P} \hat{x}, \hat{\omega}^q_i = -\frac{1}{n} 1_i^T \hat{P} \hat{\dot{x}} \) into slow dynamics (12), which yields a reduced solution \( \hat{y}^q_i = \frac{1}{n} 1_i^T \hat{P} \hat{x}^r + \hat{\delta}^r_0(0) e^{-\gamma_1 t} \) and its derivative as \( \hat{\dot{y}}^q_i = \frac{1}{n} 1_i^T \hat{P} \hat{\dot{x}}^r - \gamma_i \hat{\delta}^r_i(0) e^{-\gamma_i t} \), where \( \hat{\delta}^r_0(0) = y^r_i(0) - \frac{1}{n} 1_i^T \hat{P} \hat{x}^r(0) \) for \( \forall i \in P \). For convenience, we combine all exponentially decreasing functions to a collective vector \( \mathbf{r}(t) = [e^{-\gamma_1 t}, e^{-\gamma_2 t}, ..., e^{-\gamma_{|P|} t}]^T \in \mathbb{R}^{|P|} \), then write \( \hat{y}^r_i \) and \( \hat{\dot{y}}^r_i \) for all \( i \in P \) in a system as

\[
\hat{y}^r = \frac{1}{n} 1_{|P|\|P|} \hat{P} \hat{x}^r + \text{diag} \{ \hat{\delta}^r_i(0) \} \mathbf{r}(t), \quad \hat{\dot{y}}^r = \frac{1}{n} 1_{|P|\|P|} \hat{P} \hat{\dot{x}}^r - \text{diag} \{ \gamma_i \hat{\delta}^r_i(0) \} \mathbf{r}(t). \tag{28}
\]
We then derive a dynamic equation for $\hat{x}_i$ by substituting $\hat{y}_r$ into $\dot{\hat{x}}_i$ in (12) as
\[
\dot{\hat{x}}_r = -k \left( \frac{\partial^2 f_i(\hat{x}^-_r)}{\partial (\hat{x}_r)^2} \hat{x}_r + \frac{\partial^2 f_i(\hat{x}^-_r)}{\partial \hat{x}_r \partial \hat{x}_j} \sum_{j \neq i, j \in P} \hat{x}_j \right) - k \rho n \hat{\delta}_i(0) e^{-\gamma_i t}.
\]
Finally, we collect all $\dot{\hat{x}}_r$ for $\forall i \in P$ to form $\dot{\hat{x}}_r$ and consider (28) to obtain the following reduced model.
\[
\dot{\hat{y}}_r = -kB \hat{y}_r + kD \text{diag}\{\delta_i(0)\}r(t), \quad \dot{\hat{x}}_r = -kB \hat{x}_r - k \rho n \text{diag}\{\delta_i(0)\}r(t).
\]
(29)

Proof of Theorem 4.3: Here, the proof technique is based on singular perturbed algorithm in Section 2.1. We already mentioned that the conditions (B1) and (B2) are satisfied for the transformed system (11)-(12). Next, we are going to verify the condition (B3). Indeed, the boundary-layer model (13) is G.E.S, then for any $\theta > 0$ such that $\hat{z}_i - \frac{1}{\theta} \frac{1}{n} \mathbf{1}_{[P]} \hat{x}, \hat{\mu}_i - \hat{\mu}_i^0, \hat{\nu}_i - \frac{1}{\theta} \frac{1}{n} \mathbf{1}_{[P]} \hat{x}, \hat{\nu}_i - \hat{\nu}_i^0 \in B_{\theta}$, then $\hat{z}_i, \hat{x}_i, \hat{\mu}_i, \hat{\mu}_i^0, \hat{\nu}_i$ are bounded, and $\hat{y}_i$ is bounded according to (12) as a result. It leads to that the right-hand side of system (11)-(12) and quasi-steady states $\hat{z}_i^q, \hat{\mu}_i^q, \hat{\nu}_i^q, \hat{\nu}_i^0$ are bounded. To prove the first partial derivatives of the right-hand side are bounded, we notice that the derivatives $\hat{z}_i, \hat{\mu}_i, \hat{\nu}_i, \hat{\nu}_i^0$ are bounded, thus partial derivatives of the right-hand side according to variable $t$ is bounded, so are partial derivatives according to other variables due to a chain rule, i.e., $\frac{\partial \hat{z}_i}{\partial x_i} = \frac{\partial \hat{z}_i}{\partial t} \frac{\partial t}{\partial x_i}$, which is bounded.
Also, the first partial derivatives of quasi-steady states are bounded. Similarly, one can conclude that second partial derivatives of the right-hand side and quasi-steady states are bounded, which satisfy condition (B3). To conclude the proof, we utilize the result in the globally exponential stable property of boundary-layer model and reduced model as proved in Lemma 4.1 and 4.2 correspondingly. Since all conditions (B1)-(B5) are guaranteed for all subsets \( B_i, B_2 \subset \mathbb{R} \), the convergence is global. The proof is completed.

**Proof of Lemma 5.1:** Taking derivative at each component of \( H(x, \alpha) \) yields \( \frac{\partial H_i(x, \alpha)}{\partial x_i} \) \( = \frac{\partial f_i(x, \alpha)}{\partial x_i} + \sum_{j \neq i} \frac{\partial f_i(x, \alpha)}{\partial x_j} \frac{\partial H_j(x, \alpha)}{\partial x_j} - \rho \sum_{j \neq i} (c_j - x_j) \frac{\partial H_j(x, \alpha)}{\partial x_j} - \frac{\partial \rho \sum_{j \neq i} (c_j - x_j)}{\partial x_i} \), \( \frac{\partial H_i(x, \alpha)}{\partial x_i} = -\rho \frac{\partial}{\partial x_i} \left( (c_i - x_i) \sum_{j \neq i} (c_j - x_j) \right) \). Proof is completed.

**Proof of Lemma 5.2:** Here, we begin with the unforced system:

\[
\dot{y}^r = f_y(y^r, \sigma^r), \quad \dot{\sigma}^r = f_\sigma(\sigma^r),
\]

where \( \sigma^r \) is bounded and \( \sigma^r \) globally asymptotically converges to the origin ([18]-Lemma 1). In addition, the system \( \dot{y}^r = f_y(y^r, \sigma^r) = -k \dot{y}^r + \frac{1}{n} \sum_{p \in P} (\dot{\lambda}^r_p - \dot{\lambda}_r^p) \),

with \( \dot{\lambda}^r_p - \dot{\lambda}_r^p \) considered as an input and \(-k \dot{y} \) is Hurwitz matrix, is input-to-state stable since there exists a class \( KL \) function \( \beta \) and a class \( K \) function \( \upsilon \) so that its solution satisfies: \( \| \dot{y}^r(t) \| \leq \beta(\| y^r(t) \|, t - t_0) + \upsilon(\sup_{t_0 \leq s \leq t} | \dot{\lambda}^r(s) - \dot{\lambda}_r(s) |) \) for any initial state \( y(t_0) \) and any bounded input \( \dot{\lambda}^r - \dot{\lambda}_r \). Thus, since \( \lim_{t \to \infty} (\dot{\lambda}^r - \dot{\lambda}_r) = 0 \), the system \( \dot{y}^r = f_y(y^r, \sigma^r) \) is G.A.S [11], so is (30), meaning that the trajectories of the unforced system (30) is proved to globally converge to the origin.

Because it can be directly verified from (18) that \( f_y(y^r, \sigma^r, t) \) and \( f_\sigma(\sigma^r, t) \) are bounded for bounded \( y^r \) and \( \sigma^r \) as \( \tau(t) \) and \( \Xi(t) \) are exponentially decaying terms, in order to apply Limiting equation theorem in Section 2.2 to prove the reduced model (18) be G.A.S., it remains to show the trajectories \( y^r \) and \( \sigma^r \) of the reduced model (18) are bounded.

Firstly, we prove that the trajectory of \( \dot{\sigma}^r = f_\sigma(\sigma^r, t) \) is bounded by using Lyapunov function. The Lyapunov function is of the form \( V(\sigma^r, t) = V_1 + V_2 + V_2 \), where \( V_1 = \frac{1}{2} (\dot{\lambda}^r)^T \dot{\lambda}^r \), \( V_2 = \sum_{i \in D_f^t} \dot{\lambda}^r_i \) \( + \sum_{i \in D_f^t} \left( \dot{\lambda}^r_i \frac{\lambda^r_i - \lambda^r_i}{\alpha^r_i} \ln(1 + \lambda^r_i/\alpha^r_i) \right) \) \( + \sum_{i \in D_\partial} \dot{\lambda}^r_i \) \( + \sum_{i \in D_\partial} \left( \dot{\lambda}^r_i - \lambda^r_i \ln(1 + \lambda^r_i/\alpha^r_i) \right) \). Taking derivatives of \( V(\sigma^r, t) \) along the trajectories of system \( \dot{\sigma}^r = f_\sigma(\sigma^r, t) \), yields \( \dot{V}_1 = \frac{1}{2} (\dot{\lambda}^r)^T \dot{\lambda}^r - k \sum_{i \in P} \dot{\lambda}^r_i \dot{\lambda}^r_i + (2 \phi_i + \rho) \dot{\lambda}^r_i - \dot{\lambda}^r_i + \dot{\lambda}^r_i + \rho \dot{r}_i ) \) \( = -k \rho (\dot{\lambda}^r_i)^2 - (\dot{\lambda}^r_i)^2 \Gamma \dot{\lambda}^r - \sum_{i \in P} \dot{\lambda}^r_i \dot{r}_i \) \( - \sum_{i \in P} \dot{\lambda}^r_i \dot{r}_i \), \( \dot{V}_2 = \sum_{i \in P} \dot{\lambda}^r_i (\dot{\lambda}^r_i - \lambda^r_i (x^r_i - x^r_i)) \) \( - \sum_{i \in P} \dot{\lambda}^r_i (\dot{\lambda}^r_i - \lambda^r_i (x^r_i - x^r_i)) \), where \( \Gamma = k \text{diag}(\rho + 2 \phi_1, \ldots, \rho + 2 \phi_\alpha) \) is positive definite. By substituting \( \dot{\lambda}^r_i \) and
\( \dot{x}_u \) from (17b) into \( V_2 \) and \( V_2 \), we have \( \dot{V}(\sigma^r, t) = V_1 + V_2 + V_3 = -\rho(1 | x^r \|^2 - (x^r)^T T x + \sum_{i \in P} \lambda_i^r (x_i^r - x_i^r) + \sum_{i \in P} \chi_i^r (x_i^r - x_i^r) + \sum_{i \in P} \chi_i^r r_i(t) \). Derived from slow dynamics, the equation for reduced states on (15b) at equilibrium yields \( \lambda_i^r (x_i^r - x_i^r) = 0, \lambda_i^r (x_i^r - x_i^r) = 0 \). Thus, \( \lambda_i^r (x_i^r - x_i^r) = \lambda_i^r (x_i^r - x_i^r) \leq 0, \lambda_i^r (x_i^r - x_i^r) = \lambda_i^r (x_i^r - x_i^r) \leq 0 \) for \( \lambda_i^r \geq 0, \lambda_i^r \geq 0 \). Therefore, \( \dot{V}(\sigma^r, t) \leq -(x^r)^T T x - \sum_{i \in P} \chi_i^r r_i(t) \leq -(x^r)^T T x + \max |v_i(0)| e^{-\gamma_{\text{min}} t} \sum_{i \in P} |x_i^r|, \) where \( \gamma_{\text{min}} = \min \{\gamma_1, \gamma_2, ..., \gamma_{|P|} \} \). The existence of a perturbation term causes \( V(\sigma^r, t) \) not always be negative semi-definite, resulting in a possible divergence of trajectories \( (\hat{x}_i^r, \hat{x}_u^r, \hat{x}_i^r) \). The point here is that \( \hat{x}_i^r \) is always bounded since the term \( -(x^r)^T T x^r \) will dominate the perturbation term and forces \( V(\sigma^r, t) \) negative if \( \hat{x}_i^r \) becomes large. Moreover, when \( (\hat{x}_i^r, \hat{x}_u^r, \hat{x}_i^r) \) diverge, we always have

\[
\dot{V}(\sigma^r, t) \leq \max |v_i(0)| e^{-\gamma_{\text{min}} t} \sum_{i \in P} |x_i^r| \\
\leq \max |v_i(0)| e^{-\gamma_{\text{min}} t} (\frac{1}{2} (x^r)^T T x^r + V_2 + V_3) \\
\leq \max |v_i(0)| e^{-\gamma_{\text{min}} t} V(\sigma^r, t)
\]

According to the Grönwall-Bellman Lemma [2],

\[
V(\sigma^r, t) \leq V(t_0) e^{\max |v_i(0)| t} e^{-\gamma_{\text{min}} n} dt.
\]

Since \( V(\sigma^r, t) \) is bounded, so is \( \sigma^r \).

Next, the boundedness of trajectory for \( \dot{y}^r = \tilde{f}_y(\hat{x}^r, \sigma^r, t) \) is proved by noticing that its solution is originated from (28), in which \( \dot{y}^r \) is bounded as it is a linear function of bounded \( \tilde{x}^r \) and \( \tilde{X}(t) \). Therefore, applying limiting equation theorem shows that all trajectories of (18) converge to the origin from any initialization. This completes our proof.

**Proof of Lemma 5.3:** Linearizing the reduced model (18) yields

\[
\dot{y}^r = -k B y^r + \frac{k}{n} 1_{nxn} (\hat{x}_i^r - \hat{x}_u^r) + u(t),
\]

\[
\begin{bmatrix}
\dot{\lambda}_i^r \\
\dot{\lambda}_u^r
\end{bmatrix} = \begin{bmatrix}
\Phi^{r,I} & 0 \\
0 & \Phi^{r,I}
\end{bmatrix}
\begin{bmatrix}
\lambda_i^r \\
\lambda_u^r
\end{bmatrix},
\]

\[
\begin{bmatrix}
\dot{\lambda}_i^r \\
\dot{\lambda}_u^r
\end{bmatrix} = \begin{bmatrix}
\chi^{r,A} & 0 \\
0 & \chi^{r,A}
\end{bmatrix}
\begin{bmatrix}
\lambda_i^r \\
\lambda_u^r
\end{bmatrix} + \begin{bmatrix}
0 \\
0
\end{bmatrix} \begin{bmatrix}
v_i(t) \\
v_i(t)
\end{bmatrix},
\]

where \( \Phi = \begin{bmatrix}
0 & 0 & \Phi^A \\
0 & 0 & \Phi^A \\
- k \Delta^A_i & - k \Delta^A_i & - k B
\end{bmatrix}, \Psi = \begin{bmatrix}
0 & 0 \\
0 & 0 \\
- k \Delta^A_i & - k \Delta^A_i
\end{bmatrix}, u(t) = k D \]

\( \text{diag} \{ |y_i(t) | \} \) and \( \{ v_i(t) \} \) \( \in P \). We define matrix \( Q_1 = \text{diag} \{ \sqrt{ \lambda_i^r | k, ... \} \} \) for \( i \in D^r_i \) and \( Q_2 = \text{diag} \{ \sqrt{ \lambda_u^r | k, ... \} \} \) for \( i \in D^r_u \), which yields \( Q_1 \Phi^A Q_1^T = k \Delta^A_i Q_1 \) \( \Phi^A Q_2^T = k \Delta^A_i Q_2^T \). By considering a similarity transformation as
system (31c) is also E.S. [11] and thus \( \hat{\lambda} \) and \( \hat{\lambda} \) for \( i \)
Consequently, the subsystem (31a) is also E.S.. Therefore, (31) is E.S.

(B3) is satisfied. The Slater condition and \( \phi \) derivatives are bounded as well by applying the chain rule. Hence, the condition (11) and (16), which is shown to be satisfied condition (B1)-(B2). In a similar way the theorem. The proposed algorithm has its transformed system formulated as

We also utilize the singularly perturbed algorithm to prove

Proof of Theorem 5.4: We also utilize the singularly perturbed algorithm to prove the theorem. The proposed algorithm has its transformed system formulated as (11) and (16), which is shown to be satisfied condition (B1)-(B2). In a similar way as in Theorem 4.1, one can have the right-hand side of (16) is bounded and its derivatives are bounded as well by applying the chain rule. Hence, the condition (B3) is satisfied. The Slater condition and \( \phi_i > \frac{\|\rho_i(\Delta t)\|}{2} \) guarantee the uniqueness of \((x^*, \lambda^*)\) as shown in Lemma 5.1 and the \textit{strict complementary slackness} condition assures the local exponential stability of the reduced model (18) as proved in Lemma 5.3. The positive initialization of \( \lambda \) makes \( \lambda_i(t) \) never be negative ([18],[7]); hence the KKT condition \( \lambda_i^* \geq 0 \) is satisfied. Because the fast dynamic of Algorithm 1 and Algorithm 2 are identical, the boundary-layer model for the latter algorithm is similar to that of the former, thus it is E.S. Then, since the reduced model is E.S. according to Lemma 5.3, the transformed system for the Algorithm 2 is also E.S.. Consequently, each player action converges to its optimal value, hence the proof.

Proof of Lemma 6.1: We construct \( \mathbf{P} = [\mathbf{P}_1 \mathbf{P}_2 \ldots \mathbf{P}_n] \in \mathbb{R}^{n \times n^2} \) as an orthogonal matrix, where \( \mathbf{P}_1 = \theta \mathbf{I} \quad \mathbf{I} \in \mathbb{R}^{n \times n}, \theta > 0 \), and rewrite \( \mathbf{P} = [\mathbf{P}_1 \mathbf{P}_\Delta] \). We also define a new vector \( \mathbf{s}_1^b \in \mathbb{R}^{n^2} \) as \( \mathbf{s}_1^b = \mathbf{P}_1 \mathbf{s}_b^1 \), then the system (27), which is recalled as the boundary-layer model for (23a), becomes

\[
\frac{d\mathbf{s}_1^b}{d\tau} = [\mathbf{P}_1 \mathbf{P}_\Delta]^\top (\mathbf{L} \otimes \mathbf{I})\mathbf{s}_2^2, \quad \frac{d\mathbf{s}_2^2}{d\tau} = -\mathbf{s}_2^2 - (\mathbf{L} \otimes \mathbf{I})\mathbf{s}_2^2 - (\mathbf{L} \otimes \mathbf{I})[\mathbf{P}_1 \mathbf{P}_\Delta]\mathbf{s}_1^b. \tag{32}
\]

For the sake of analysis, we present a new component vector \( \mathbf{s}_4^1 = \mathbf{s}_4^1(n+1:n^2) \) and \( \mathbf{s}_3^1 = \mathbf{s}_3^1(1:n) \). By using these new vectors and considering properties of the Laplacian matrix in graph theory [20], the system (32) is simplified as

\[
\begin{bmatrix}
\frac{d\mathbf{s}_1^1}{d\tau} \\
\frac{d\mathbf{s}_2^1}{d\tau} \\
\frac{d\mathbf{s}_3^1}{d\tau} \\
\frac{d\mathbf{s}_4^1}{d\tau}
\end{bmatrix} = \begin{bmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & \mathbf{P}_\Delta^\top(\mathbf{L} \otimes \mathbf{I}) & 0 \\
0 & -\mathbf{L} \otimes \mathbf{I} & -\mathbf{I} - (\mathbf{L} \otimes \mathbf{I}) & \mathbf{s}_4^1 \\
0 & 0 & 0 & 0
\end{bmatrix} \begin{bmatrix}
\mathbf{s}_1^1 \\
\mathbf{s}_2^1 \\
\mathbf{s}_3^1 \\
\mathbf{s}_4^1
\end{bmatrix}. \tag{33}
\]

It can be seen that \( \mathbf{s}_1^1 \) is a constant vector as \( \frac{d\mathbf{s}_1^1}{d\tau} = 0 \). Hence, it is equivalent to consider the reduced boundary-layer model constructed from (23a) with a non-constant vector \( \mathbf{s} = [(\mathbf{s}_1^1)^\top (\mathbf{s}_4^1)^\top]^\top \)

\[
\frac{d\mathbf{s}}{d\tau} = \mathbf{C}\mathbf{s}, \tag{34}
\]
where $C = \begin{bmatrix} 0 & P_\Delta^T (L \otimes I) \\ -(L \otimes I)P_\Delta & -I \end{bmatrix}$. One can verify that $(L \otimes I)P_\Delta$ be full rank and $-I - (L \otimes I) < 0$, thus $C$ is Hurwitz. In addition, the origin of (34) is uniform in $(t, \bar{x})$. Thus, the origin of (34) is G.E.S. and uniform in $(t, \bar{x})$.

Likewise, one can conclude that the origin of the boundary-layer model for (23b)-(23c), which is derived from (20b)-(20c) around the equilibrium $(x^*, \nu^*, \nu^*, d^*, \eta^*)$, is also G.E.S. and uniform in $(t, \bar{x})$. This completes our proof.

Proof of Lemma 6.2: The fast dynamic (20a) exponentially converge to $z_{ij}^q = \frac{1}{n} \sum_{k \in V} a_{kj}(x)$ in the large and globally converge to $\mu_{ij}^q$. By considering the change of variables $\delta_{ij} = z_{ij} - z_{ij}^q$ and $\xi_{ij} = \mu_{ij} - \mu_{ij}^q$, one can obtain the following fast dynamics

$$
\begin{align*}
\epsilon \delta_{ij} &= \sum_{k \in V} a_{ik}(\delta_{ij} - \delta_{kj}), \\
\epsilon \dot{\delta}_{ij} &= -\delta_{ij} - \sum_{k \in V} a_{ik}(\delta_{ij} - \delta_{kj} + \xi_{ij} - \xi_{kj}).
\end{align*}
$$

We define $\xi_i = [\xi_i, \xi_i, ..., \xi_i]^T \in \mathbb{R}^n$, $\delta_i = [\delta_{i1}, \delta_{i2}, ..., \delta_{in}]^T \in \mathbb{R}^n$, $\xi = [\xi_1, \xi_2, ..., \xi_n]^T \in \mathbb{R}^n$, and $\delta = [\delta_1, \delta_2, ..., \delta_n]^T \in \mathbb{R}^n$. We also define a new vector $\bar{\xi} \in \mathbb{R}^{n_2}$ as $\bar{\xi} = P\xi$ and denote $\bar{\delta} = \xi(n+1:n^2)$. Moreover, in order to evaluate $z_i$ and a radius be $\min_{k \in V} \{ \frac{1}{n} \sum_{k \in V} a_{kj} \} - \Delta > 0$. It can be verified that if

$$
|\delta(0)| \leq \min\{r_{ij}^q\}/\chi,
$$

then $|\delta(t)| \leq \min\{r_{ij}^q\}$ for all $i, j \in V$. Thus, $|z_{ij} - z_{ij}^q| \leq \min\{r_{ij}^q\}$. Hence, trajectory $z_{ij}$ never escape a circle with a center located at $z_{ij}^q$ and a radius be $\min\{r_{ij}^q\}$. In other words, $z_{ij}$ never reaches 0 for all $i, j \in V$. Moreover, in order to evaluate $\varphi$, one can begin with $|\bar{\delta}(0)| \leq n^2 \max\{z_{ij}(0) - z_{ij}^q\}$. Hence, if $\varphi = (\min\{r_{ij}^q\}/n\chi)^2$ and $\max\{z_{ij}(0) - z_{ij}^q\} \leq \varphi$, then (37) follows.

Proof of Theorem 6.3: To begin, we define $\hat{y}_{iP} = [\hat{y}_{i1}, ..., \hat{y}_{iP}]^T \in \mathbb{R}^{P}$, $\hat{y}_{iP} = [\hat{y}_{i1}, ..., \hat{y}_{i1}, \hat{y}_{i1}, ..., \hat{y}_{i1}, \hat{y}_{i1}, ..., \hat{y}_{i1}]^T \in \mathbb{R}^{n_1 - P}$. From (25), we obtain

$$
\hat{y}_{ij} = \begin{cases} 
-k(f_x(\bar{x}, \lambda_i^r, \lambda_u), T_i(t)) - \gamma_1\theta_{ij}(0)e^{-\gamma t} & \text{if } j \in P \\
-\gamma_2\hat{y}_{ij} & \text{if } j \in L \cup G \cup T
\end{cases}
$$

We then combine (38) for all $i \in P$ and $j \in V$ to obtain

$$
\begin{align*}
\hat{y}_{iP}^r &= -kB\hat{x} + k(\lambda_i^r - \lambda_u) - \text{diag}(\gamma_1\theta_{ij}(0)) \rho(t) - k[h_i(t), ..., h_i(t)]^T, \\
\hat{y}_{iP}^r &= -\gamma_1\hat{y}_{iP}^r,
\end{align*}
$$

and $\hat{y}_{iP} = \hat{x} + \text{diag}(\theta_{ij}(0)) \rho(t)$. Thus, we can rewrite (39) as

$$
\begin{align*}
\hat{y}_{iP}^r &= -kB\hat{y}_{iP} + k(\lambda_i^r - \lambda_u) + T_i(t) \rho(t) - H(t), \\
\hat{y}_{iP}^r &= -\gamma_1\hat{y}_{iP},
\end{align*}
$$

where $C = \begin{bmatrix} 0 & P_\Delta^T (L \otimes I) \\ -(L \otimes I)P_\Delta & -I \end{bmatrix}$. One can verify that $(L \otimes I)P_\Delta$ be full rank and $-I - (L \otimes I) < 0$, thus $C$ is Hurwitz. In addition, the origin of (34) is uniform in $(t, \bar{x})$. Thus, the origin of (34) is G.E.S. and uniform in $(t, \bar{x})$.
where $T_i = k(B - \text{diag}((\frac{1}{r^T})\text{diag}(\theta_{ij}(0)))) \in \mathbb{R}^{[P] \times [P]}$, $H_i(t) = [h_i(t), ..., h_i(t)]^T \in \mathbb{R}^{[P]}$, with notice that $\lim_{t \to -\infty} r(t) = 0$ and $\lim_{t \to \infty} H_i(t) = 0$. Finally, one can obtain the concatenated form of the reduced model (25) as

\[
\begin{align*}
\dot{\hat{y}}_p &= \hat{f}_y(p)(\hat{y}_p, \hat{\sigma}_r, t), \\
\dot{\sigma}_r &= \hat{f}_\sigma(\sigma_r, t).
\end{align*}
\]

(41a) The subsystem (41a) containing $\hat{f}_y(p)(\hat{y}_p) = -\text{diag} \{\gamma_1, ..., \gamma_{[P]}\} \dot{y}_p$ is G.E.S.. In (41b), $\hat{f}_y(p)(\hat{y}_p, \hat{\sigma}_r, t) = -kB\dot{y}_p + k\hat{I}(\hat{\lambda}_l - \hat{\lambda}_u) + k\hat{I}(T_ir(t) - H_i(t))$, with

\[
\hat{B} = \begin{bmatrix}
B & 0 & \cdots & 0 \\
0 & B & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & B
\end{bmatrix}
\]

and $\hat{I} = [I \ I \ \ldots \ I]^T$. Moreover, the subsystem (41c) finds similar with (18b) and is of the form $\hat{f}_\sigma(\hat{\sigma}_r, t) = \hat{f}_\sigma(\hat{\sigma}_r) + \Pi(t)$, where $\Pi(t) = -K[0, 0, h_1(t), 0, 0, h_2(t), ..., 0, 0, h_{[P]}(t)] \in \mathbb{R}^{3|P|}$. By following analysis in Lemma 5.2 and Lemma 5.3, one can obtain that the system of (41b)-(41c) is G.A.S. and locally E.S.. Hence, the reduced model (41) is E.S.. It is also verified that the boundary-layer model is E.S. according to Lemma 6.1, and assumption (B1),(B2) and (B3) are satisfied. Additionally, the uniqueness of $(x^*, \lambda^*)$ is guaranteed by The Slater condition and $\phi_i > \frac{(|P|-3)p}{2}$ according to Lemma 5.1, and positive the initialization of $\lambda_i$ and $\lambda_u$ guarantees the KKT condition that $\lambda^*_l, \lambda^*_u \geq 0$. Therefore, the transformed system is E.S., with the initialization corresponding to Lemma 6.2. We complete the proof. □

REFERENCES

[1] H. S. Ahn, B. Y. Kim, Y. H. Lim, B. H. Lee and K. K. Oh, Distributed coordination for optimal energy generation and distribution in cyber-physical-energy networks, IEEE Trans. on Cybernetics, **48** (2018), 941–954.

[2] I. Barkana, Defending on the beauty of the invariance principle, Int. Journal of Control, **87** (2014), 186–206.

[3] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge Univ. Press, Cambridge, U.K., 2004.

[4] S. Boyd, A. Ghosh, B. Prabhakar and D. Shah, Randomized gossip algorithms, *IEEE Trans. Inf. Theory*, **52** (2006), 2508–2530.

[5] H. Chen, Y. Li, R. Louie and B. Vucetic, Autonomous demand side management based on energy consumption scheduling and instantaneous load billing: an aggregative game approach, *IEEE Trans. Smart Grid*, **5** (2014), 1744–1754.

[6] A. Cherukuri and J. Cortés, Distributed generator coordination for initialization and anytime optimization in economic dispatch, *IEEE Trans. on Control of Network Systems*, **2** (2015), 226–237.

[7] H. B. Dürr, E. Saka and C. Ebenbauer, A smooth vector field for quadratic programming, in *51st IEEE Conf. on Decision and Control*, (2012), 2515–2520.

[8] A. Freeman, P. Yang and K. M. Lynch, Stability and convergence properties of dynamic average consensus estimators, *IEEE Int. Conf. on Decision and Control*, (2006), 398–403.

[9] *IEEE 118 bus*. Available from: [http://motor.ece.iit.edu/data/](http://motor.ece.iit.edu/data/).

[10] M. K. Jensen, Aggregative games and best-reply potentials, *Econom. Theory*, **43** (2010), 45–66.

[11] H. K. Khalil, *Nonlinear Systems*, 3rd edition, Prentice Hall, New Jersey, 2002.

[12] S. S. Kia, J. Cortés and S. Martínez, Dynamic average consensus under limited control authority and privacy requirement, *International Journal of Robust and Nonlinear Control*, **25** (2015), 1941–1966.
[13] S. S. Kia, J. Cortés and S. Martínez, Singularly perturbed algorithms for dynamic average consensus, in European Control Conference, (2013), 1758–1763.
[14] J. Koshal, A. Nedić and Uday V. Shanbhag, A gossip algorithm for aggregative games on graphs, in 51st IEEE Conf. on Decision and Control, (2012), 4840–4845.
[15] N. Li and J. Marden, Designing games for distributed optimization, IEEE Journal of Selected Topics in Signal Processing, 7 (2013), 230–242.
[16] K. Ma, G. Hu and C. J. Spanos, Distributed energy consumption control via real-time pricing feedback in smart grid, IEEE Trans. on Control System, 22 (2014), 1907–1914.
[17] A. H. Mohsenian-Rad, V. W. S. Wong, J. Jatskevich, R. Schober and A. Leon-Garcia, Autonomous demand-side management based on game-theoretic energy consumption scheduling for the future smart grid, IEEE Trans. Smart Grid, 1 (2016), 320–331.
[18] C. V. Nguyen, P. H. Hoang, M. H. Trinh, B-H Lee and H-S. Ahn, Distributed Nash equilibrium seeking of an aggregative game by a singular perturbed algorithm, in Australian and New Zealand Control Conference, (2017), 145–150.
[19] C. V. Nguyen, P. H. Hoang, H. K. Kim and H-S. Ahn, Distributed learning in a multi-agent potential game, in 2017 Int. Conf. on Control, Automation and Systems, (2017), 260–271.
[20] R. Olfati-Saber, J. A. Fax and R. M. Murray, Consensus and cooperation in networked multi-agent systems, Proc. of the IEEE, 95 (2007), 215–233.
[21] M. A. A. Pedrasa, T. D. Spooner and I. F. MacGill, Scheduling of demand-side resources using binary particle swarm optimization, IEEE Trans. Power Syst., 24 (2009), 1173–1181.
[22] B. Ramanathan and V. Vittal, A framework for evaluation of advanced direct load control with minimum disruption, IEEE Trans. Power Syst., 23 (2008), 1681–1688.
[23] W. Ren and R. W. Beard, Distributed Consensus in Multi-Vehicle Cooperative Control: Theory and Applications, Springer-Verlag, London, 2008.
[24] J. B. Rosen, Existence and uniqueness of equilibrium points for concave n-person games, Econometrica, 33 (1965), 520–534.
[25] F. Salehisadaghiani and L. Pavel, Distributed Nash equilibrium seeking: a gossip-based algorithm, Automatica J. IFAC, 72 (2016), 209–216.
[26] F. Salehisadaghiani and L. Pavel, Distributed Nash equilibrium seeking via the alternating direction method of multipliers, in IFAC-Papers Online, 50 (2017), 6166–6171.
[27] E. Shchepakina, V. Sobolev and M. P. Mortell, Singular Perturbations: Introduction to System Order Reduction Methods with Applications, Springer, Switzerland, 2014.
[28] W. Shi and L. Pavel, LANA: an ADMM-like Nash equilibrium seeking algorithm in decentralized environment, in American Control Conference, (2017), 285–290.
[29] M. Stanković, K. Johansson and D. Stipanović, Distributed seeking of Nash equilibria with applications to mobile sensor networks, IEEE Trans. on Automat. Control, 57 (2012), 904–919.
[30] X. Wang, N. Xiao, T. Wongpiromsarn, L. Xie, E. Frazzoli and D. Rus, Distributed consensus in noncooperative congestion games: An application to road pricing, in 10th IEEE Conf. on Control and Automation, (2013), 1668–1673.
[31] M. Ye and G. Hu, Distributed Nash equilibrium seeking by a consensus based approach, IEEE Trans. on Automat. Control, 62 (2017), 4811–4818.

Received April 2018; 1st revision October 2018; final revision April 2019.

E-mail address: chuongnguyen@gist.ac.kr
E-mail address: phh18@my.fsu.edu
E-mail address: hyosung@gist.ac.kr