Connecting Automatic Generation Control and Economic Dispatch From an Optimization View

Na Li, *Member, IEEE*, Changhong Zhao, *Member, IEEE*, and Lijun Chen, *Member, IEEE*

***Abstract*—Automatic generation control (AGC) regulates mechanical power generation in response to load changes through local measurements. Its main objective is to maintain system frequency and keep energy balanced within each control area in order to maintain the scheduled net interchanges between control areas. The scheduled interchanges as well as some other factors of AGC are determined at a slower time scale by considering a centralized economic dispatch (ED) problem among different generators. However, how to make AGC more economically efficient is less studied. In this paper, we study the connections between AGC and ED by reverse engineering AGC from an optimization view, and then we propose a distributed approach to slightly modify the conventional AGC to improve its economic efficiency by incorporating ED into the AGC automatically and dynamically.**

***Index Terms*—Automatic generation control, distributed algorithm, economic dispatch, network optimization and control, power networks.**

I. Introduction

An interconnected electricity system can be described as a collection of subsystems, each of which is called a control area. Within each control area, the mechanical power input to the synchronous generators is automatically regulated by automatic generation control (AGC). AGC uses the local control signals, deviations in frequency, and net power interchanges between the neighboring areas to invoke appropriate valve actions of generators in response to load changes. The main objectives of the conventional AGC are to: 1) maintain system nominal frequency and 2) let each area absorb its own load changes in order to maintain the scheduled net interchanges between control areas [2], [3]. The scheduled interchanges between control areas, as well as the participation factors of each generator unit within each control area, are determined at a much slower time scale than the AGC by generating companies considering a centralized economic dispatch (ED) problem among different generators.

Since the traditional loads (which are mainly passive) change slowly and are predictable with high accuracy, the conventional AGC does not incur much efficiency loss by following the schedule made by the slower time scale ED after the load changes. However, due to the proliferation of renewable energy resources as well as demand response in the future power grid, the aggregate net loads, for example, traditional passive loads plus electric vehicle loads minus renewable generations, can fluctuate fast and by a large amount. Therefore, the conventional AGC can become much less economically efficient. We thus propose a novel modification of the conventional AGC to automatically: 1) maintain nominal frequency and 2) reach optimal power dispatch between generator units among all control areas while balancing supply and demand within the whole interconnected electricity system to achieve economic efficiency. We call this modified AGC the economic AGC. We further develop a hybrid of the conventional AGC and the economic AGC where the power interchanges among certain control areas are maintained at the nominal value but the power is dispatched optimally to different generator units within each control area. The purpose of this hybrid AGC is to prevent disturbance propagating between control areas which might lead to a system-wide blackout. Note that in the hybrid AGC, we allow the flexibility to choose where the power interchanges should be maintained.

In order to keep the modification minimal and to keep the decentralized structure of AGC, we take a reverse and forward engineering approach to develop the economic AGC.[1] We first reverse-engineer the conventional AGC by showing that the power system dynamics with the conventional AGC can be interpreted as a partial primal-dual gradient algorithm to solve a certain optimization problem. We then engineer the optimization problem to include general generation costs and general power-flow balance (which will guarantee supply-demand balance within the whole interconnected electricity system), and propose a distributed generation control scheme that is integrated into the AGC. The engineered optimization problem shares the same optima as the ED problem and, thus, the resulting distributed control scheme incorporates ED into AGC automatically. Combined with [4] on distributed load control, this work lends the promise to develop a modeling framework and solution approach for systematic design of distributed,

Manuscript received September 28, 2014; revised March 2, 2015; accepted June 21, 2015. Date of publication July 21, 2015; date of current version September 15, 2016. Some preliminary results of this work were originally presented in Americal Control Conference 2014 [1]. Recommended by Associate Editor M. Chertkov.

N. Li is with the School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138 USA (e-mail: nali@seas.harvard.edu).

C. Zhao is with the Division of Engineering and Applied Sciences, California Institute of Technology, Pasadena, CA 91125 USA (e-mail: czhao@caltech.edu).

L. Chen is with the College of Engineering and Applied Sciences, University of Colorado, Boulder, CO 80309 USA (e-mail: lijun.chen@colorado.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCNS.2015.2459451

[1]A similar approach has been used to design decentralized optimal-load control in our previous work [4].

low-complexity generation and load control to achieve system-wide efficiency and robustness.

A. Literature Review

There has been a large amount of work on AGC in the last few decades, including, for example, stability and optimum parameter setting [5]; optimal or adaptive controller design [6]–[8]; decentralized control [9], [10]; and multilevel or multitime-scale control [11], [12]; see also [3] and the references therein for a thorough and up-to-date review on AGC. Most of these works focus on improving the control performance of AGC, such as stability and transient dynamics, but not on improving the economic efficiency. References [13] and [14] introduce approaches for AGC that also support an ED feature which operates at a slower time scale and interacts with the AGC frequency stabilization function. For instance, [14] brings in the notion of minimal regulation which reschedules the entire system generation and minimizes generation cost with respect to system-wide performance. Our work aims to improve the economic efficiency of AGC in response to the load changes as well; the difference is that instead of using different hierarchical control to improve AGC, we incorporate ED automatically and dynamically into AGC. Moreover, our controller is a decentralized and closed-loop one, where each control area updates its generation based only on measurements of local physical variables that are easy to measure and information of local auxiliary variables that are easy to compute and communicate with neighboring areas. This means that the controller does not need any information on the system disturbance which is the change of the net loads in this paper.

Recently, there is a increasing interest to study frequency control from the same perspective of this paper, that is, to bridge the gap between different layers of hierarchical control, especially the dynamical control and optimal dispatch. An incomplete list includes [15]–[23]. One main difference between our work and the others is that we focus on AGC and model the built-in control mechanisms explicitly, that is, the turbine-governor control and ACE-based control. Our objective is not only to design distributed algorithms to improve the economic efficiency of AGC, but also to keep the *modification* as minor as possible in order to facilitate the implementation of the new control algorithms. The reverse and forward engineering approach adopted in this paper allows us to maximally take into account the existing system dynamics and built-in control mechanisms. As a result, the economic (or hybrid) AGC only requires minor modifications that are implementable via introducing new local auxiliary variables which are easy to compute.

This paper is organized as follows. Section II introduces a dynamic power network model with AGC, the ED problem, and the objective of the economic AGC. Section III reverse-engineers the conventional AGC, Section IV provides an economic AGC scheme based on the insight obtained by the reverse engineering, and Section V provides a hybrid of the conventional AGC and economic AGC. Finally, Section VI simulates and compares the convention AGC, the economic AGC, and the hybrid AGC.

II. System Model

A. Dynamic Network Model With AGC

Consider a power transmission network, denoted by a graph $(\mathcal{N}, \mathcal{E})$, with a set $\mathcal{N} = \{1, \dots, n\}$ of buses and a set $\mathcal{E} \subset \mathcal{N} \times \mathcal{N}$ of transmission lines connecting the buses. Here, each bus may denote an aggregated bus or a control area. We make the following assumptions:

- The lines $(i, j) \in \mathcal{E}$ are lossless and characterized by their reactance x_{ij}.
- The voltage magnitudes $|V_i|$ of buses $i \in \mathcal{N}$ are constants.
- Reactive power injections at the buses and reactive power flows on the lines are ignored.

We assume that $(\mathcal{N}, \mathcal{E})$ is connected and directed, with an arbitrary orientation such that if $(i, j) \in \mathcal{E}$, then $(j, i) \notin \mathcal{E}$. We use $i : i \to j$ and $k : j \to k$, respectively, to denote the set of buses i such that $(i, j) \in \mathcal{E}$ and the set of buses j such that $(j, k) \in \mathcal{E}$. We study generation control when there is a step change in net loads from their nominal (operating) points, which may result from a change in demand or in nondispatchable renewable generation. To simplify the notation, all of the variables in this paper represent deviations from their nominal (operating) values. Note that in practice those nominal values are usually determined by the last ED problem, which will be introduced later.

Frequency Dynamics: For each bus j, let ω_j denote the frequency, P_j^M be the mechanical power input, and P_j^L as the total load. For a link (i, j), let P_{ij} denote the transmitted power form bus i to bus j. The frequency dynamics at bus j is given by the swing equation

$$\dot{\omega}_j = -\frac{1}{M_j}\left(D_j\omega_j - P_j^M + P_j^L + \sum_{k:j\to k} P_{jk} - \sum_{i:i\to j} P_{ij}\right) \quad (1)$$

where M_j is the generator inertia and D_j is the damping constant at bus j.

Branch Flow Dynamics: Assume that the frequency deviation ω_j is small for each bus $j \in \mathcal{N}$. Then, the deviations P_{ij} from the nominal branch flows follow the dynamics:

$$\dot{P}_{ij} = B_{ij}(\omega_i - \omega_j) \quad (2)$$

where

$$B_{ij} := \frac{|V_i||V_j|}{x_{ij}} \cos\left(\theta_i^0 - \theta_j^0\right)$$

is a constant determined by the nominal bus voltages and the line reactance. Here, θ_i^0 is the nominal voltage phase angle of bus $i \in \mathcal{N}$. The detailed derivation is given in [4].

Turbine-Governor Control: For each generator, we consider a governor-turbine control model, where a speed governor senses a speed deviation and/or a power change command and converts it into an appropriate valve action, and then a turbine converts the change in the valve position into the change in mechanical power output. The governor-turbine control is usually modeled as a two-state dynamic system. One state corresponds to the speed governor and the other state corresponds to the turbine. Since the time constant of the governor is much smaller than the turbine for most systems, we simplify the governor-turbine control model from two states to a single state P_j^M

$$\dot{P}_j^M = -\frac{1}{T_j}\left(P_j^M - P_j^C + \frac{1}{R_j}\omega_j\right) \tag{3}$$

where P_j^C is the power change command and T_j and R_j are costant parameters. See [2] for a detailed introduction of governor-turbine control.

ACE-Based Control: In the conventional AGC, the power change command P_j^C is adjusted automatically by the tie-line bias control which drives the area control errors (ACEs) to zero. For a bus j, the ACE is defined as

$$\text{ACE}_j = B_j\omega_j + \sum_{k:j\to k} P_{jk} - \sum_{i:i\to j} P_{ij}.$$

The adjustment of power change command is given as follows:

$$\dot{P}_j^C = -K_j\left(B_j\omega_j + \sum_{k:j\to k} P_{jk} - \sum_{i:i\to j} P_{ij}\right) \tag{4}$$

where B_j and K_j are positive constant parameters. In this paper, we also call this AGC the ACE-based AGC.

In summary, the dynamic model with power control over a transmission network is given by (1)–(4). If the system is stable given certain load changes, then by simple analysis, we can show that the ACE-based AGC drives the system to a new steady state where the load change in each control area is absorbed within each area, that is, $P_j^M = P_j^L$ for all $j \in \mathcal{N}$, and the frequency is returned to the nominal value, that is, $\omega_j = 0$ for all $j \in \mathcal{N}$; as shown in Proposition 1 in Section III. Notice that the ACE-based AGC has a decentralized structure, namely, that it only uses local control signals, that is, deviations in frequency and the net power interchanges with neighboring buses.

B. Economic Dispatch (ED)

Due to the proliferation of renewable energy resources, such as solar and wind in the future power grid, the aggregate net loads will fluctuate much faster and by large amounts. The ACE-based AGC that requires each control area to absorb its own load changes may be economically inefficient. Therefore, we proposed to modify the ACE-based AGC to: 1) maintain the nominal frequency and 2) drive the mechanical power output $P_j^M, j \in \mathcal{N}$ to the optimum of the following ED problem:

$$\min \quad \sum_{j\in\mathcal{N}} C_j\left(P_j^M\right) \tag{5a}$$

$$\text{s.t.} \quad \sum_{j\in\mathcal{N}} P_j^M = \sum_{j\in\mathcal{N}} P_j^L \tag{5b}$$

$$\text{over} \quad P_j^M,\ j\in\mathcal{N}$$

where each generator at j incurs certain cost $C_j(P_j^M)$ when its power generation is P_j^M.[2] Equation (5b) imposes power balanced within the global system. The cost function $C_j(\cdot)$ is assumed to be continuous, differentiable, and convex. In the rest of this paper, we call this modified AGC the economic AGC and we will show how to reverse- and forward-engineer the ACE-based AGC to design an economic AGC scheme.

Remark: In the conventional ACE-based AGC, if a control area j has multiple generator units, the generation change P_j^C in (4) is allocated by a central regulator (e.g., the ISO) to individual generator units via participation factors. The participation factors are inversely proportional to the units' incremental cost of production which are determined by the last ED performed. See [24] for a detailed description. Thus, if the net loads fluctuate fast and dramatically due to the large penetration of renewable energy, this centralized allocation plan by using constant participation factors also becomes economically inefficient. The results developed in this paper can also be applied to improve the economic efficiency of the generation control for each unit within one area and the allocation is performed in a distributed way as shown in Section V of the hybrid AGC. In fact, a system operator can apply our results to the generation control at different levels of the power system (e.g., different control areas, different generators within one area, etc, according to the practical requirements of the system). For the simplicity of illustration, at the beginning, we do not specify the level of the generation control that we study. We will focus on the abstract model in (1)–(4) and treat each bus j as a generator bus.

III. Reverse Engineering of ACE-Based AGC

In this section, we reverse-engineer the dynamic model with the ACE-based AGC (1)–(4). We show that the equilibrium points of (1)–(4) are the optima of a properly defined optimization problem and furthermore the dynamics (1)–(4) can be interpreted as a partial primal-dual gradient algorithm to solve this optimization problem. The reverse-engineering suggests a way to modify the ACE-based AGC to incorporate ED into the AGC scheme.

We first characterize the equilibrium points of the power system dynamics with AGC (1)–(4). Let $\omega = \{\omega_j, j \in \mathcal{N}\}$, $P^M = \{P_j^M, j\in\mathcal{N}\}$, $P^C = \{P_j^C, j\in\mathcal{N}\}$, and $P = \{P_{i,j}, (i,j)\in\mathcal{E}\}$.

Proposition 1: (ω, P^M, P^C, P) is an equilibrium point of the system (1)–(4) if and only if $\omega_j = 0$, $P_j^C = P_j^M = P_j^L$, and $\sum_{i:i\to j} P_{ij} = \sum_{k:j\to k} P_{jk}$ for all $j \in \mathcal{N}$.

Proof: At a fixed point

$$\dot{P}_{ij} = B_{ij}(\omega_i - \omega_j) = 0.$$

Therefore, $\omega_i = \omega_j$ for all $i, j \in \mathcal{N}$, given that the transmission network is connected. Moreover

$$\text{ACE}_j = B_j\omega_j + \sum_{k:j\to k} P_{jk} - \sum_{i:i\to j} P_{ij} = 0.$$

[2] Since all variables denote the deviations in this paper, it may be not straightforward to interpret this ED problem, for example, its connection with the slower timescale ED problem which is defined on the absolute value of each variable instead of the deviated value. In the appendix, we provide two interpretations of the ED problem in (5).

Thus, $\sum_{j\in\mathcal{N}} \text{ACE}_j = \sum_{j\in\mathcal{N}} B_j\omega_j = \omega_i \sum_{j\in\mathcal{N}} B_j = 0$, so $\omega_i = 0$ for all $i \in \mathcal{N}$. The rest of the proof is straightforward. We omit it due to space limitations. ■

Consider the following optimization problem:

OGC-1

$$\min \quad \sum_{j\in\mathcal{N}} C_j\left(P_j^M\right) + \sum_{j\in\mathcal{N}} \frac{D_j}{2}|\omega_j|^2 \tag{6a}$$

$$\text{s.t.} \quad P_j^M = P_j^L + D_j\omega_j + \sum_{k:j\to k} P_{jk} - \sum_{i:i\to j} P_{ij}$$

$$P_j^M = P_j^L \tag{6b}$$

$$\text{over} \quad \omega_j, P_j^M, P_{ij},\ j\in\mathcal{N},\ (i,j)\in\mathcal{E} \tag{6c}$$

where (6c) requires that each control area absorbs its own load changes. The following result is straightforward.

Lemma 2: (ω^*, P^{M^*}, P^*) is an optimum of OGC-1 if and only if $\omega_j^* = 0$, $P_j^{M^*} = P_j^L$, and $\sum_{k:j\to k} P_{jk}^* = \sum_{i:i\to j} P_{ij}^*$ for all $j \in \mathcal{N}$.

Proof: First, the constraints (6b) and (6c) imply that $D_j\omega_j + \sum_{k:j\to k} P_{j,k} - \sum_{i:i\to j} P_{i,j} = 0$ for all $j \in \mathcal{N}$. Then, we can use contradiction to prove that $\omega_i^* = \omega_j^*$ for all $(i,j) \in \mathcal{E}$. By following similar arguments in Proposition 1, we can prove the statement in the lemma. ■

Note that problem OGC-1 appears simple, as we can easily identify its optima if we know all of the information on the objective function and the constraints. However, in practice, this information is unknown. Moreover, even if we know an optimum, we cannot just set the system to the optimum. Since the power network is a physical system, we have to find a way that respects the power system dynamics to steer the system to the optimum. Though the cost function $C_j(P_j^M)$ does not play any role in determining the optimum of OGC-1, it will become clear later that the choice of the cost function does have an importantimplication to the algorithm design and the system dynamics.

We now show that the dynamic system (1)–(4) is actually a partial primal-dual gradient algorithm for solving OGC-1 with $C_j(P_j^M) = (\beta_j/2)(P_j^M)^2$ where $\beta_j > 0$:

Introducing Lagrangian multipliers λ_j and μ_j for the constraints in OGC-1, we obtain the following Lagrangian function:

$$\begin{aligned} L = &\sum_{j\in\mathcal{N}} \frac{\beta_j}{2}\left(P_j^M\right)^2 + \sum_{j\in\mathcal{N}} \frac{D_j}{2}|\omega_j|^2 \\ &+ \sum_{j\in\mathcal{N}} \lambda_j \left(P_j^M - P_j^L - D_j\omega_j - \sum_{k:j\to k} P_{jk} + \sum_{i:i\to j} P_{ij}\right) \\ &+ \sum_{j\in\mathcal{N}} \mu_j\left(P_j^M - P_j^L\right). \end{aligned} \tag{7}$$

Based on the above Lagrangian function, we can write down a partial primal-dual subgradient algorithm of OGC-1 as follows:

$$\omega_j = \lambda_j \tag{8a}$$

$$\dot{P}_{ij} = \epsilon_{P_{ij}}(\lambda_i - \lambda_j) \tag{8b}$$

$$\dot{P}_j^M = -\epsilon_{P_j}\left(\beta_j P_j^M + \lambda_j + \mu_j\right) \tag{8c}$$

$$\dot{\lambda}_j = \epsilon_{\lambda_j}\left(P_j^M - P_j^L - D_j\omega_j - \sum_{k:j\to k} P_{jk} + \sum_{i:i\to j} P_{ij}\right) \tag{8d}$$

$$\dot{\mu}_j = \epsilon_{\mu_j}\left(P_j^M - P_j^L\right) \tag{8e}$$

where $\epsilon_{P_{ij}}$, ϵ_{P_j}, ϵ_{λ_j}, and ϵ_{μ_j} are positive stepsizes. Note that (8a) solves $\max_{\omega_j}(D_j/2)\omega_j^2 - \lambda_j D_j\omega_j$ rather than follows the primal gradient algorithm with respect to w_j; hence, the algorithm (8) is called the "partial" primal-dual gradient algorithm. See the Appendix for a description of the general form of the partial primal-dual gradient algorithm and its convergence.

Let $\epsilon_{\lambda_j} = 1/M_j$ for all $j \in \mathcal{N}$. By applying linear transformation from (λ_j, μ_j) to (ω_j, P_j^C)

$$\omega_j = \lambda_j$$

$$P_j^C = K_j M_j\left(\lambda_j - \frac{1}{\epsilon_{\mu_j} M_j}\mu_j\right)$$

the partial primal-dual gradient algorithm (8) becomes

$$\dot{\omega}_j = -\frac{1}{M_j}\left(D_j\omega_j - P_j^M + P_j^L + \sum_{k:j\to k} P_{jk} - \sum_{i:i\to j} P_{ij}\right) \tag{9a}$$

$$\dot{P}_{ij} = \epsilon_{P_{ij}}(\omega_i - \omega_j) \tag{9b}$$

$$\dot{P}_j^M = -\epsilon_{P_j}\beta_j\left(P_j^M - \frac{\epsilon_{\mu_j}}{K_j\beta_j}P_j^C + \frac{1+\epsilon_{\mu_j}M_j}{\beta_j}\omega_j\right) \tag{9c}$$

$$\dot{P}_j^C = -K_j\left(D_j\omega_j + \sum_{k:j\to k} P_{jk} - \sum_{i:i\to j} P_{ij}\right). \tag{9d}$$

If we set $\epsilon_{P_{ij}} = B_{ij}$, $\epsilon_{\mu_j} = R_jK_j/(1 - R_jK_jM_j)$, $\beta_j = R_j/(1 - R_jK_jM_j)$, and $\epsilon_{P_j} = 1/\beta_jT_j$, then the partial primal-dual algorithm (9) is exactly the power system dynamics with AGC (1)–(4) if $B_j = D_j$, $j \in \mathcal{N}$. Note that the assumption of $B_j = D_j$ looks restrictive. But since B_j is a design parameter, we can set it to D_j. However, in reality, D_j is uncertain and/or hard to measure because it does not only account for damping of the generator but also contains a component due to the frequency-dependent loads. In Section VI, the simulation results demonstrate that even if $B_j \neq D_j$, the algorithm still converges to the same equilibrium point. It remains as one of our future works to characterize the range of B_j, which guarantees the convergence of the algorithm. Nonetheless, the algorithm in (9) provides a tractable and easy way to choose parameters for the ACE-based AGC in order to guarantee its convergence.

Theorem 3: If $1 > R_jK_jM_j$ for all $j \in \mathcal{N}$, with the aforementioned chosen ϵ_{λ_j}, ϵ_{μ_j}, $\epsilon_{P_{ij}}$, and ϵ_{Pj}, the partial primal-dual gradient algorithm (9) (i.e., the system dynamics (1)–(4)) converges to a fixed point $(\omega^*, P^*, P^{M^*}, P^{C^*})$ where (ω^*, P^*, P^{M^*}) is an optimum of problem OGC-1 and $P^{C^*} = P^{M^*}$.

Proof: The proof is deferred in the Appendix. ■

Remark: We have made an equivalence transformation in the above: from algorithm (8) to algorithm (9). The reason for doing this transformation is to derive an algorithm that admits physical interpretation and can thus be implemented as the system dynamics. In particular, P_j^L is unknown and, hence, μ_j cannot be directly observed or estimated, while P_j^C can be estimated/calculated based on the observable variables ω_j and P_{ij}. Since the control should be based on observable or estimable variables, the power system implements algorithm (9) instead of (8) for the ACE-based AGC.

The aforementioned reverse-engineering, that is, the power system dynamics with AGC as the partial primal-dual gradient algorithm solving an optimization problem, provides a modeling framework and systematic approach to design new AGC mechanisms that achieve different (and potentially improved) objectives by engineering the associated optimization problem. The new AGC mechanisms would also have different dynamic properties (such as responsiveness) and incur different implementation complexity by choosing different optimizing algorithms to solve the optimization problem. In the next section, we will engineer problem OGC-1 to design an AGC scheme that achieves economic efficiency.

IV. Economic AGC by Forward Engineering

We have seen that the power system dynamics with the ACE-based AGC (1)–(4) is a partial primal-dual gradient algorithm solving a cost minimization problem OGC-1 with a "restrictive" constraint $P_j^M = P_j^L$ that requires supply-demand balance within each control area. As mentioned before, this constraint may render the system economically inefficient. Based on the insight obtained from the reverse-engineering of the conventional AGC, we relax this constraint and propose an AGC scheme that: 1) keeps the frequency deviation to 0 (i.e., $\omega_j = 0$ for all $j \in \mathcal{N}$) and 2) achieves economic efficiency, that is, the mechanical power generation solves the ED problem (5).

Consider the following optimization problem:

OGC-2

$$\min \quad \sum_{j\in\mathcal{N}} C_j\left(P_j^M\right) + \sum_{j\in\mathcal{N}} \frac{D_j}{2}|\omega_j|^2 \tag{10a}$$

$$\text{s.t.} \quad P_j^M = P_j^L + D_j\omega_j + \sum_{k:j\to k} P_{jk} - \sum_{i:i\to j} P_{ij} \tag{10b}$$

$$P_j^M = P_j^L + \sum_{k:j\to k}\gamma_{jk} - \sum_{i:i\to j}\gamma_{ij}$$

$$\text{over} \quad \omega_j, P_j^M, P_{ij}, \gamma_{ij},\ j\in\mathcal{N},\ (i,j)\in\mathcal{E} \tag{10c}$$

where γ_{ij} are auxiliary variables introduced to relax the restrictive constraint $P_j^M = P_j^L$, which allows different control areas to change their power-flow interchanges in order to promote the global system economic efficiency. Though **OGC-2** looks much more complicated than the simple ED problem (5), as shown in Lemma 4, the optimal solution P_j^{M*} of **OGC-2** is equal to the optimal solution of ED (5) for which we call the AGC derived in this section as economic AGC. As shown later, the reason to focus on **OGC-2** is to keep $\omega_j = 0$ for all $j \in \mathcal{N}$ and to derive an implementable control algorithm which requires minor modifications on the ACE-based AGC (3), (4).

Lemma 4: Let $(\omega^*, P^{M*}, P^*, \gamma^*)$ be an optimum of OGC-2, then $\omega_j^* = 0$ for all $j \in \mathcal{N}$, and P^{M*} is the optimal solution of the ED problem (5).

Proof: First, note that at the optimum, $\omega_i^* = \omega_j^*$ for all $(i,j) \in \mathcal{N}$. Second, combining (10b) and (10c) gives

$$D_j\omega_j + \sum_{k:j\to k}(P_{jk}-\gamma_{jk}) - \sum_{i:i\to j}(P_{ij}-\gamma_{ij}) = 0$$

for all $j \in \mathcal{N}$. Following similar arguments as in Proposition 1, we have $\omega_i^* = 0$ for all $i \in \mathcal{N}$. Therefore, the constraint (10c) is redundant and can be removed. So, problem OGC-2 reduces to the ED problem (5). ■

Following the same procedure as in Section III, we derive the following partial prime-dual algorithm to solve OGC-2:

$$\omega_j = \lambda_j \tag{11a}$$

$$\dot{P}_{i,j} = \epsilon_{P_{ij}}(\lambda_i - \lambda_j) \tag{11b}$$

$$\dot{P}_j^M = -\epsilon_{P_j}\left(C_j'\left(P_j^M\right) + \lambda_j + \mu_j\right) \tag{11c}$$

$$\dot{\gamma}_{ij} = \epsilon_{\gamma_{ij}}(\mu_i - \mu_j) \tag{11d}$$

$$\dot{\lambda}_j = \epsilon_{\lambda_j}\left(P_j^M - P_j^L - D_j\omega_j - \sum_{k:j\to k}P_{jk} + \sum_{i:i\to j}P_{ij}\right) \tag{11e}$$

$$\dot{\mu}_j = \epsilon_{\mu_j}\left(P_j^M - P_j^L - \sum_{k:j\to k}\gamma_{jk} + \sum_{i:i\to j}\gamma_{ij}\right). \tag{11f}$$

Let $\epsilon_{\lambda_j} = 1/M_j$, $\epsilon_{P_{ij}} = B_{ij}$, $\epsilon_{\mu_j} = R_jK_j/(1 - R_jK_jM_j)$, and $\epsilon_{P_j} = (1 - R_jK_jM_j)/T_jR_j$ as in Section III. By using linear transformation $\omega_j = \lambda_j$ and $P_j^C = K_jM_j(\lambda_j - (1/\epsilon_{\mu_j}M_j)\mu_j)$, the partial primal-dual gradient algorithm (11) becomes

$$\dot{\omega}_j = -\frac{1}{M_j}\left(D_j\omega_j - P_j^M + P_j^L + \sum_{k:j\to k}P_{jk} - \sum_{i:i\to j}P_{ij}\right) \tag{12a}$$

$$\dot{P}_{ij} = B_{ij}(\omega_i - \omega_j) \tag{12b}$$

$$\dot{P}_j^M = -\frac{1}{T_j}\left(\frac{1 - R_jK_jM_j}{R_j}C_j'\left(P_j^M\right) - P_j^C + \frac{1}{R_j}\omega_j\right) \tag{12c}$$

$$\dot{P}_j^C = -K_j\left(D_j\omega_j + \sum_{k:j\to k}(P_{jk}-\gamma_{jk}) - \sum_{i:i\to j}(P_{ij}-\gamma_{ij})\right) \tag{12d}$$

$$\dot{\gamma}_{ij} = \epsilon_{\gamma_{ij}}\left(\left(M_i\omega_i - \frac{P_i^C}{K_i}\right)\epsilon_{\mu_i} - \left(M_j\omega_j - \frac{P_j^C}{K_j}\right)\epsilon_{\mu_j}\right). \tag{12e}$$

Compared with algorithm (9) (i.e., the power system dynamics with the ACE-based AGC), the difference in algorithm (12) is the new variables γ_{ij} and the marginal cost $C_j'(\cdot)$ in the generation control (12c). Note that γ_{ij} can be calculated based on the observable/measurable variables. So the above algorithm is implementable. However, it might not be practical to add an additional variable γ_{ij} for each branch $(i,j) \in \mathcal{E}$. To further facilitate the implementation, we can remove $\gamma_{i,j}$ by introducing γ_j for each bus j and replace (12d) and (12e) by the following dynamics:

$$\dot{P}_j^C = -K_j\left(D_j\omega_j + \sum_{k:j\to k}(P_{jk} - \gamma_j + \gamma_k) - \sum_{i:i\to j}(P_{ij} - \gamma_i + \gamma_j)\right) \tag{13a}$$

$$\dot{\gamma}_i = \epsilon_\gamma\left(\left(M_i\omega_i - \frac{P_i^C}{K_i}\right)\epsilon_{\mu_i}\right) \tag{13b}$$

which tells us that the power change command P_j^C can be controlled using local measurements ω_j, P_{jk}, γ_j, and local communications on γ_i, γ_k with the neighbors i, k where $(i,j), (j,k) \in \mathcal{E}$. Here, γ_j is a local auxiliary variable which is updated using local information at each bus $j \in \mathcal{N}$.

Similarly, we have the following result.

Theorem 5: The algorithm [(12a)–(12c), (13a), (13b)] converges to a fixed point $(\omega^*, P^*, P^{M^*}, P^{C^*}, \gamma^*)$, where $(\omega^*, P^*, P^{M^*}, \gamma^*)$ is an optimum of problem OGC-2, which is also optimal to the ED problem in (5), and $P_j^{C^*} = ((1 - R_jK_jM_j)/R_j)C_j'(P_j^{M^*})$.

Proof: Please see the Appendix for the convergence of the partial primal-dual gradient algorithm. ■

With Lemma 4 and Theorem 5, we can implement algorithm [(12a)–(12c), (13a), (13b)] as an economic AGC for the power system. By comparing with the ACE-based AGC in (1)–(4) and the economic AGC in [(12a)–(12c), (13a), (13b)], we note that economic AGC has only a slight modification to the ACE-based AGC and keeps the decentralized structure of AGC. In other words, adding local communication about the new local auxiliary variable γ_j based on (13a) and (13b) can improve the economic efficiency of AGC.

Remark: We can actually derive a simpler and yet implementable algorithm without introducing variable $\gamma_{ij}, (i,j) \in \mathcal{E}$ (or $\gamma_i, i \in \mathcal{N}$). However, we choose to derive the algorithm (12) and (13) in order to have minimal modification to the existing conventional AGC and keep the resulting control decentralized, where each control area updates its generation based on measurements of local physical variables that are easy to measure and information on local auxiliary variables that are easy to compute and communicate with neighboring areas.

V. HYBRID OF ACE-BASED AGC AND ECONOMIC AGC

In the ACE-based AGC, each control area absorbs its own energy fluctuation in order to prevent the disturbance propagation which might lead to a system-wide blackout. In the economic AGC, all of the control areas share the energy fluctuations in order to improve the economic efficiency. The side effect is that this sharing could propagate the disturbance and might lead to a blackout due to the potential outage of some transmission lines. Therefore, in this section, we propose a hybrid of ACE-based AGC and economic AGC which takesthe safety and efficiency into account. We call this AGC a hybrid AGC.

Given an interconnected power network $\mathcal{N}$, which is divided (/partitioned) into different areas, denoted by $\mathcal{A} = \{A_1, A_2, \ldots, A_m\}$.[3] Here, each $A_l \subseteq \mathcal{N}$ is a control area. The objective of the hybrid AGC is to: 1) maintain the nominal frequency and 2) each area A_i absorbs its own energy disturbance in an economically efficient way such that $\{P_j^M\}_{j\in A_l}$ is the optimum to the following optimization problem: for each $A_l \in \mathcal{A}$

$$\min_{P_j^M, j\in A_l} \quad \sum_{j\in A_l} C_j\left(P_j^M\right) \tag{14a}$$

$$\text{s.t.} \quad \sum_{j\in A_l} P_j^M = \sum_{j\in A_l} P_j^L. \tag{14b}$$

Let $\mathcal{E}_{\text{in}}$ be the subset of links that connect buses within a same area. Now consider the following optimization problem:

$$\min \quad \sum_{j\in\mathcal{N}} C_j\left(P_j^M\right) + \sum_{j\in\mathcal{N}} \frac{D_j}{2}|\omega_j|^2 \tag{15a}$$

$$\text{s.t.} \quad P_j^M = P_j^L + D_j\omega_j + \sum_{(j,k)\in\mathcal{E}} P_{jk} - \sum_{(i,j)\in\mathcal{E}} P_{ij} \tag{15b}$$

$$P_j^M = P_j^L + \sum_{(j,k)\in\mathcal{E}_{\text{in}}} \gamma_{jk} - \sum_{(i,j)\in\mathcal{E}_{\text{in}}} \gamma_{ij}$$

$$\text{over} \quad \omega_j, P_j^M, P_{ij}, \gamma_{ij}, \ j\in\mathcal{N}, \ (i,j)\in\mathcal{E}. \tag{15c}$$

We have the following Lemma regarding the optimal solution of OGC-3.

Lemma 6: Let $(\omega^*, P^{M^*}, P^*, \gamma^*)$ be an optimum of OGC-3, then 1) $\omega_j^* = 0$ for all $j \in \mathcal{N}$ and 2) for each area A_l, $\{P_j^M\}_{j\in A_l}$ is the optimal solution to problem (14).

Proof: Note that $\mathcal{A} = \{A_1, A_2, \ldots, A_m\}$ forms a partition $\mathcal{N}$ and the network $(\mathcal{N}, \mathcal{E})$ is connected. By using a similar argument in the proof of Lemma 4, we can obtain the statement in this Lemma. We omit the details here. ■

Following the same procedure as in Section IV, we can derive the following partial prime-dual algorithm to solve OGC-2, which is the hybrid AGC we need to obtain:

$$\dot{\omega}_j = -\frac{1}{M_j}\left(D_j\omega_j - P_j^M + P_j^L + \sum_{(j,k)\in\mathcal{E}} P_{jk} - \sum_{(i,j)\in\mathcal{E}} P_{ij}\right) \tag{16a}$$

$$\dot{P}_{ij} = B_{ij}(\omega_i - \omega_j) \tag{16b}$$

$$\dot{P}_j^M = -\frac{1}{T_j}\left(\frac{1 - R_jK_jM_j}{R_j}C_j'\left(P_j^M\right) - P_j^C + \frac{1}{R_j}\omega_j\right) \tag{16c}$$

[3]We assume that $\mathcal{A}$ forms a partition of $\mathcal{N}$, that is, $A_{l_1} \cap A_{l_2} = \emptyset$ for any $l_1 \neq l_2$ and $\cup_{l=1,\ldots,m} A_l = \mathcal{N}$.

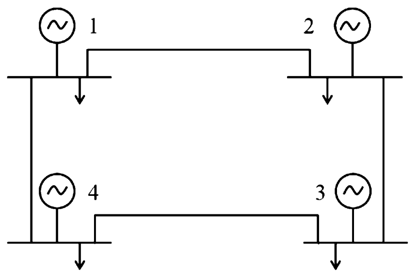

Fig. 1. Four-area interconnected system.

TABLE I
GENERATOR PARAMETERS

Area, j	M_j	D_j	$\lvert V_j \rvert$	T_j	R_j	K_j	B_j
1	3	1	1.045	4	0.05	2	2
2	2.5	1.5	0.98	4	0.05	2	3
3	4	1.2	1.033	4	0.05	2	2
4	3.5	1.4	0.997	4	0.05	2	3

$$\dot{P}_j^C = -K_j \Bigg(D_j\omega_j + \sum_{(j,k)\in\mathcal{E}} P_{jk} - \sum_{(i,j)\in\mathcal{E}} P_{ij} - \sum_{(j,k)\in\mathcal{E}_{\rm in}} (\gamma_j - \gamma_k) + \sum_{(i,j)\in\mathcal{E}_{\rm in}} (\gamma_i - \gamma_j) \Bigg) \tag{16d}$$

$$\dot{\gamma}_i = \epsilon_\gamma \left(\left(M_i\omega_i - \frac{P_i^C}{K_i} \right) \epsilon_{\mu_i} \right). \tag{16e}$$

Similarly, we can guarantee the stability of the hybrid AGC.

Theorem 7: The algorithm (16) converges to a fixed point $(\omega^*, P^*, P^{M^*}, P^{C^*}, \gamma^*)$, where $(\omega^*, P^*, P^{M^*}, \gamma^*)$ is an optimum of problem OGC-3, which is given in Lemma 6.

Proof: Please see the Appendix for the convergence of the partial primal-dual gradient algorithm. ■

Compared with the economic AGC in [(12a)–(12c), (13a), (13b)], the only difference of the hybrid AGC is that if bus i and j are connected but do not belong to the same area, then they do not communicate the auxiliary variable γ_i and γ_j with each other. As a result, the hybrid AGC possesses all of the nice properties of the economic AGC. It requires only local measurement, local computation, and local communication. Moreover, the modification to the conventional ACE-based AGC is moderate.

VI. CASE STUDY

For illustrative purpose, we consider a simple interconnected system with four buses, as shown in Fig. 1. The values of the generator and transmission-line parameters are shown in Tables I and II. Notice though that our theoretical results require that $B_j = D_j$ for each j, here we choose B_j differently from D_j since D_j is usually uncertain in reality. To make it easy to compare the simulation results, we choose a same cost function for each area, where $C_i(P_{M_i}) = aP_{M_i}^2$. Therefore, we know that the optimal dispatch is to equally share the load.

TABLE II
LINE PARAMETERS

line	1-2	2-3	3-4	4-1
r	0.004	0.005	0.006	0.0028
x	0.0386	0.0294	0.0596	0.0474

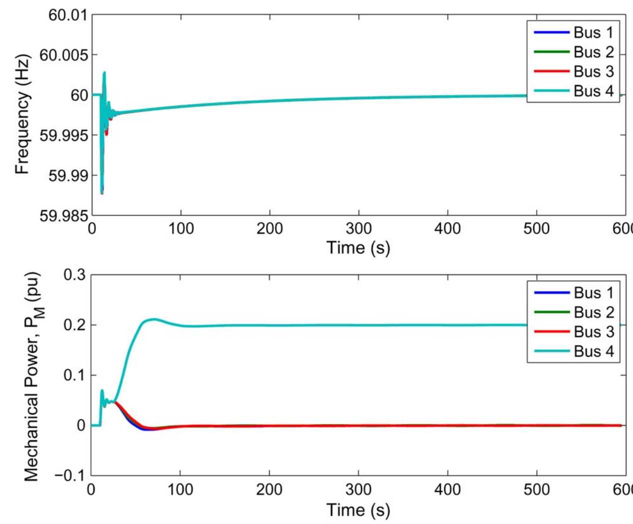

Fig. 2. ACE-based AGC.

In the model used for simulation, we relax some of the assumptions made in the previous analysis. For each transmission line, we consider nonzero line resistance and do not assume small differences between phase-angle deviations, which means that the power-flow model is in the form of

$$P_{ij} = \frac{|V_i||V_j|}{x_{ij}^2 + r_{ij}^2} \left(x_{ij}\left(\sin\theta_{ij} - \sin\theta_{ij}^0\right) - r_{ij}\left(\cos\theta_{ij} - \cos\theta_{ij}^0\right) \right).$$

Simulation results show that our proposed AGC scheme works well even in these nonideal, practical systems.

At time $t = 10$ s, a step change of load occurs at bus 4 where $P_4^L = 1$ p.u. In the simulation, to be consistent with the real practice in the conventional AGC, the signal for the AGC is only reset at every 15 s. Fig. 2 shows the dynamics of the frequencies and mechanical power outputs for the four buses using ACE-based AGC (1)–(4), which indicates that bus 4 absorbs all of the disturbance eventually. Fig. 3 shows the dynamics of the frequencies and mechanical power outputs for the four buses using the economic AGC [(12a)–(12c), (13a), (13b)], which tells that all of the buses share the disturbance equally and, thus, optimally. Fig. 4 shows the dynamics of the frequencies and mechanical power outputs for the four buses using the hybrid AGC where buses 1 and 4 form one control area and buses 2 and 3 form another control area. Gradually buses 1 and 4 share the disturbance equally and buses 2 and 3 are not affected by the disturbance. Fig. 5 compares the total generation costs using the ACE-based AGC, the economic AGC, and the hybrid AGC with the minimal generation cost of the ED problem (5). We see that the economic AGC tracks the optimal value of the ED problem and the hybrid AGC dispatches the power generation optimally within each area. An interesting observation is that the frequency dynamics are very similar. One possible explanation is the fast frequency synchronization. Since the

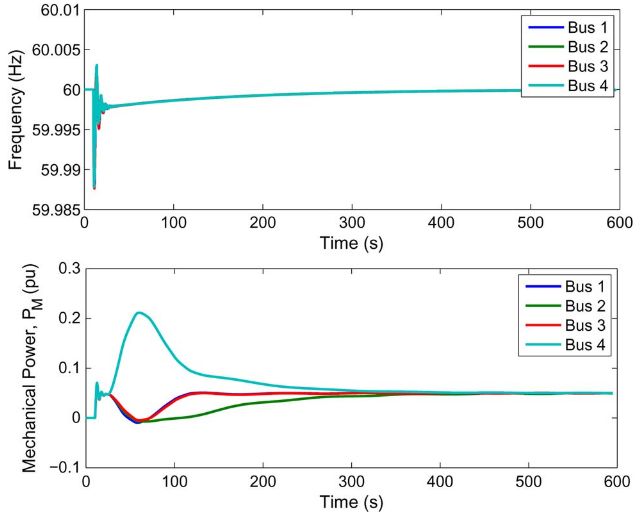

Fig. 3. Economic AGC.

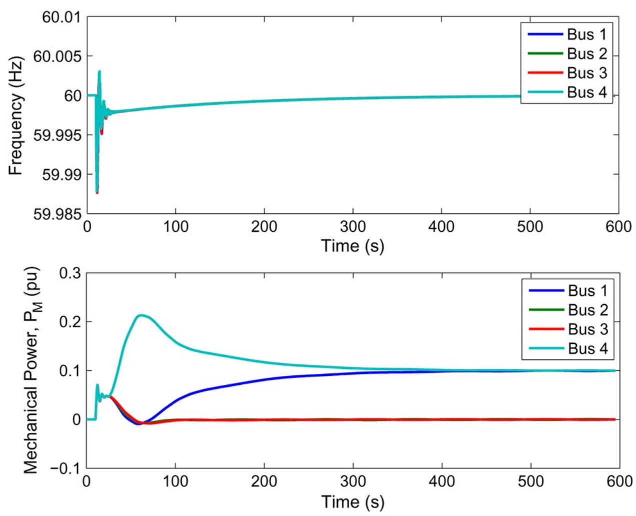

Fig. 4. Hybrid AGC.

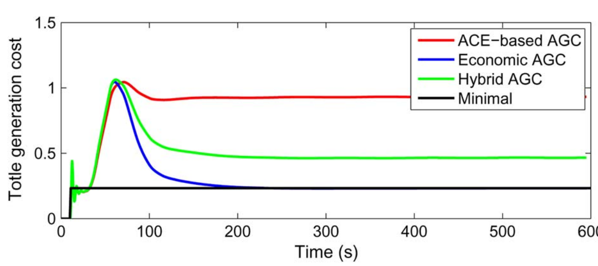

Fig. 5. Generation cost.

AGC control signal is reset every 15 s, before the AGC takes action, the frequency has been synchronized within the first 15 s, when the frequency has the most dramatic transient dynamics.

VII. Conclusion

We reverse-engineer the conventional AGC, and based on the insight obtained from the reverse engineering, we design a decentralized generation control scheme that integrates the ED into the AGC and achieves economic efficiency. Combined with the previous work [4] on distributed load control, this work lends the promise of developing a modeling framework and solution approach for systematic design of distributed, low-complexity generation and load control to achieve system-wide efficiency and robustness.

Appendix

A. *Interpretation of the ED in (5)*

Here, we provide two ways of constructing(/interpreting) the cost functions in (5). Now let P_j^M denote the nominal value of the mechanical power generation and ΔP_j^M denote the deviation from the nominal value. One type of cost $C_j(\Delta P_j^M)$ is the cost on the deviation, for example, $|\Delta P_j^M|^2$. This means that as long as there is a deviation from the nominal value P_j^M, there is a cost incurred. The second one is more directly related to generation cost which is used at the slow time-scale ED problem. At the slow time-scale ED problem, P_j^M is determined by minimizing the generation cost $\sum_j c_j(P_j^M)$ such that $\sum_j P_j^M = \sum_j P_j^d$. When there is a deviation ΔP_j^M, the new generation cost is $c_j(P_j^M + \Delta P_j^M)$. This gives a natural way to construct the cost function of $C_j(\Delta P_j^M)$, which is that $C_j(\Delta P_j^M) := c_j(P_j^M + \Delta P_j^M)$.

B. *Proof of Convergence*

The primal-dual gradient flow (also called saddle point flow) method for optimization problems has been studied and applied in different literature, for example, [25]–[29]. The proof techniques used in this literature can be applied to our problem with a minor modification via using the properties of the optimization problems. Instead of only proving Theorems 3, 5, and 7, we provide a partial primal-dual gradient algorithm for a general convex optimization problem and show that the algorithm converges to the optimal primal-dual point if the optimization problem satisfies certain conditions. Then, we will prove Theorems 3, 5, and 7. Focusing on the general optimization problem first allows us to illustrate the main ideas behind the detailed algorithms used in the paper, and the results have more general applications than the AGC itself.

1) Partial Primal-Dual Gradient Algorithm: Consider the following optimization problem:

$$\begin{aligned} \min_{x,y} \quad & f(x) + g(y) \\ \text{s.t.} \quad & Ax + By = C \end{aligned} \tag{17}$$

where $f(x)$ is a strict convex and twice differentiable function of x, $g(y)$ is a convex and differentiable function of y. Notice that $g(y)$ can be a constant function.

The Lagrangian function of this optimization problem is given by

$$L(x, y, \alpha) = f(x) + g(y) + \alpha^T(Ax + By - C).$$

Assume that the constraint is feasible and an optimal solution exists, then the strong duality holds. Moreover, the primal-dual optimal solution (x^*, y^*, α^*) is a saddle point of $L(x, y, \alpha)$ and vice-versa.

The partial primal-dual gradient algorithm is given by **Algorithm-1**

$$x(t) = \arg\min_x \{L(x,y,\alpha)\} = \arg\min_x \{f(x) + \alpha^T Ax\}$$
$$\dot{y} = -\Xi_y(\frac{\partial L(x,y,\alpha)}{\partial y} = -\Xi_y\left(\frac{\partial g(y)}{\partial y} + B^T\alpha\right)$$
$$\dot{\alpha} = \Xi_\alpha(\frac{\partial L(x,y,\alpha)}{\partial \alpha} = \Xi_\alpha(Ax + By - C)$$

where $\Xi_y = \mathrm{diag}(\epsilon_{y_i})$ and $\Xi_\alpha = \mathrm{diag}(\epsilon_{\alpha_j})$. In the following we will study the convergence of this algorithm.

Define

$$q(\alpha) \triangleq \min_x \{f(x) + \alpha^T Ax\}$$
$$\hat{L}(y,\alpha) \triangleq q(\alpha) + g(y) + \alpha^T(By - C).$$

The following proposition demonstrates some properties of $q(\alpha)$ and $\hat{L}(y,\alpha)$.

Proposition 8: $q(\alpha)$ is a concave function and its gradient is given as $\partial q(\alpha)/\partial\alpha = Ax$. If $\ker(A^T) = 0$, then $q(\alpha)$ is a strictly concave function. As a consequence, $\hat{L}(y,\alpha)$ is strictly concave on α.

Proof: Since $f(x)$ is a strictly convex function of x, we can directly apply in [30, Proposition 6.1.1] to conclude that $q(\alpha)$ is a concave function of α and $\partial q/\partial\alpha = Ax$. Let $H := \nabla^2 f(x)$, which is a positive definite matrix. From [30, (6.9)], we have $\nabla^2 q(\alpha) = -AH^{-1}A^T$. Therefore, we know that if $ker(A^T) = 0$, $\nabla^2 q$ is a negative definite matrix, implying that $q(\alpha)$ is a strictly concave function. The rest of the proposition follows directly. ■

Moreover, we have the following connections between $L(x,y,\alpha)$ and $\hat{L}(y,\alpha)$.

Proposition 9: If (x^*, y^*, α^*) is a saddle point of L, then (y^*, α^*) is a saddle point of $\hat{L}$ and $x^* = \arg\min_x\{f(x) + (\alpha^*)^T Ax\}$. Moreover, if (y^*, α^*) is a saddle point of $\hat{L}$, then (x^*, y^*, α^*) is a saddle point of L where $x^* = \arg\min_x\{f(x) + (\alpha^*)^T Ax\}$.

Proof: The proof is straightforward by comparing the first-order conditions of saddles points for L and $\hat{L}$. Note that the convexity of f, g, and concavity of q imply that those first-order conditions are necessary and sufficient conditions for saddle points. ■

We also have the following properties of the saddle points of $\hat{L}$,

Proposition 10: Assume $\ker(A^T) = 0$. Given any two saddle points (y_1^*, α_1^*), (y_2^*, α_2^*) of $\hat{L}$, we have $\alpha_1^* = \alpha_2^*$, and $\hat{L}(y_1^*, \alpha_1^*) = \hat{L}(y_2^*, \alpha_2^*) = \hat{L}(y_1^*, \alpha_2^*) = \hat{L}(y_2^*, \alpha_1^*)$.

Proof: If (y_1^*, α_1^*), (y_2^*, α_2^*) are two saddle points

$$\hat{L}(y_1^*, \alpha) \le \hat{L}(y_1^*, \alpha_1^*) \le \hat{L}(y, \alpha_1^*)$$
$$\hat{L}(y_2^*, \alpha) \le \hat{L}(y_2^*, \alpha_2^*) \le \hat{L}(y, \alpha_2^*)$$

for any (y,α). Thus, we have $\hat{L}(y_1^*, \alpha_2^*) \le \hat{L}(y_1^*, \alpha_1^*) \le \hat{L}(y_2^*, \alpha_1^*) \le \hat{L}(y_2^*, \alpha_2^*) \le \hat{L}(y_1^*, \alpha_2^*)$, which implies that $\hat{L}(y_1^*, \alpha_2^*) = \hat{L}(y_1^*, \alpha_1^*) = \hat{L}(y_2^*, \alpha_1^*) = \hat{L}(y_2^*, \alpha_2^*) = \hat{L}(y_1^*, \alpha_2^*)$. Because $\hat{L}$ is strictly concave in α, we have $\alpha_1^* = \alpha_2^*$. ■

Using the new Lagrangian function $\hat{L}$, **Algorithm-1** can be written as follows:

$$\dot{y} = -\Xi_y\left(\frac{\partial\hat{L}(y,\alpha)}{\partial y}\right) \tag{18}$$
$$\dot{\alpha} = \Xi_\alpha\left(\frac{\partial\hat{L}(y,\alpha)}{\partial \alpha}\right). \tag{19}$$

Let (y^*, α^*) be a saddle point of $\hat{L}(y,\alpha)$. Adopting the Lyapunov function

$$U(y,\alpha) = \sum_{i=1}^n \frac{1}{2\epsilon_{y_i}}(y_i - y_i^*)^2 + \sum_{i=1}^m \frac{1}{2\epsilon_{\alpha_i}}(\alpha_i - \alpha_i^*)^2 \tag{20}$$

following the methods in [25]–[29], and using the properties of $\hat{L}$ introduced in Proposition 8 and 10, we know that algorithm (18), (19) converges to the saddle point of $\hat{L}$ if $\ker(A^T) = 0$. Consequently, we know that **Algorithm-1** converges to the saddle point of L, which is an optimal point of (17).

2) Proof of Theorem 3, Theorem 5, and Theorem 7: Though the previous analysis for the general optimization problem could not be directly applied to prove Theorems 3, 5, and 7,[4] the ideas used in the proof are easily extended to prove those theorems. Since the proofs of the three theorems are very similar, here we only provide the detailed proof for Theorem 3.

Denote the Lagrangian function in (7) as $L(P^M, \omega, P, \lambda, \mu)$, where $P^M := \{P_j^M\}_{j\in\mathcal{N}}, \omega := \{\omega_j\}_{j\in\mathcal{N}}, P := \{P_{ij}\}_{(i,j)\in\mathcal{E}}, \lambda := \{\lambda_j^M\}_{j\in\mathcal{N}}, \mu := \{\mu_j\}_{j\in\mathcal{N}}$. Algorithm in (8) can be written as

$$\omega(t) = \arg\min_\omega \{L(P^M, \omega, P, \lambda, \mu)\} = \lambda$$
$$\dot{P} = -\Xi_P \frac{\partial L(P^M, \omega, P, \lambda, \mu)}{\partial P}$$
$$\dot{P}^M = -\Xi_{P^M} \frac{\partial L(P^M, \omega, P, \lambda, \mu)}{\partial P^M}$$
$$\dot{\lambda} = \Xi_\lambda \frac{\partial L(P^M, \omega, P, \lambda, \mu)}{\partial \lambda}$$
$$\dot{\mu} = \Xi_\mu \frac{\partial L(P^M, \omega, P, \lambda, \mu)}{\partial \mu}$$

where Ξ_P, Ξ_{P^M}, Ξ_λ, Ξ_μ are diagonal matrices standing for the stepsizes. Let $\hat{L}(P^M, P, \lambda, \mu) := L(P^M, \omega = \lambda, \lambda, \mu)$. Given the structure of OGC-1, we can obtain the following proposition about $\hat{L}$.

Proposition 11: $\hat{L}$ is strictly convex on P^M, strictly concave on λ, and linear on P and μ.

Moreover, we have the following Lemma about the saddle points of $\hat{L}$.

Proposition 12: Let (P^{M*}, ω^*) be the unique optimal point of OGC-1. Then, (P^M, P, λ, μ) is a saddle point of $\hat{L}$ if and only if $P^M = P^{M*}$, $\lambda = \omega^*$, $\mu_j = -\beta_j P_j^{M*} - \omega_j^*$, and $\sum_{k:j\to k} P_{jk} - \sum_{i:i\to j} P_{ij} = P_j^{M*} - P_j^L - D_j\omega_j^*$.

[4] This is because the corresponding As in optimization problems OGC-1 (6), OGC-2 (10), and OGC-3 (15) do not satisfy $\ker(A^T) = 0$.

Proof: Since OGC-1 is strong convex on (P^M, ω), the optimal solution is unique. Then, by using the strong duality of OGC-1, it is straightforward to show the statement of the lemma. We omit the details here. ■

Now we are ready to proof Theorem 3. First, note that the algorithm in (8) is equivalent to the following algorithm:

$$\dot{P} = -\Xi_P \frac{\partial \hat{L}}{\partial P};\ \dot{P}^M = -\Xi_{P^M} \frac{\partial \hat{L}}{\partial P^M}$$
$$\dot{\lambda} = \Xi_\lambda \frac{\partial \hat{L}}{\partial \lambda}; \quad \dot{\mu} = \Xi_\mu \frac{\partial \hat{L}}{\partial \mu}. \tag{21}$$

Let $(P^{M*}, P^*, \lambda^*, \mu^*)$ be a saddle point of $\hat{L}$. Define a nonnegative function as

$$\begin{aligned} &U_{P^{M*},P^*,\lambda^*,\mu^*}(P^M, P, \lambda, \mu)\\ &= \frac{1}{2}(P^M - P^{M*})^T \Xi_{P^M}^{-1}(P^M - P^{M*})\\ &+ \frac{1}{2}(P - P^*)^T \Xi_P^{-1}(P - P^*)\\ &+ \frac{1}{2}(\lambda - \lambda^*)\Xi_\lambda^{-1}(\lambda - \lambda^*)\\ &+ \frac{1}{2}(\mu - \mu^*)\Xi_\mu^{-1}(\mu - \mu^*). \end{aligned} \tag{22}$$

Taking the derivative along the dynamics (21), we can show

$$\begin{aligned} \frac{\partial U}{\partial t} &\leq \hat{L}(P^{M*}, P^*, \lambda, \mu) - \hat{L}(P^{M*}, P^*, \lambda^*, \mu^*)\\ &\quad + \hat{L}(P^{M*}, P^*, \lambda^*, \mu^*) - \hat{L}(P^M, P, \lambda^*, \mu^*)\\ &\leq 0. \end{aligned} \tag{23}$$

For simplicity, we will denote (P^M, P, λ, μ) as z.

Lemma 13: $\partial U(z)/\partial t \leq 0$ for all z, and

$$\begin{aligned} &\left\{\hat{z} : \frac{\partial U(\hat{z})}{\partial t} = 0\right\} \subseteq \mathcal{Z}\\ &\triangleq \left\{\hat{z} : \hat{P}^M = P^{M*}, \hat{\lambda} = \lambda^*, \hat{L}(P^{M*},P,\alpha^*,\mu^*)\right.\\ &\quad \left. = \hat{L}(P^{M*},P,\alpha^*,\mu) = \hat{L}(P^{M*},P^*,\alpha^*,\mu^*)\right\}. \end{aligned}$$

Proof: Reference (23) has shown that $\partial U(z)/\partial t \leq 0$. To ensure $\partial U(\hat{z})/\partial t = 0$, we need that $\hat{L}(P^{M*}, P^*, \lambda, \mu) = \hat{L}(P^{M*}, P^*, \lambda^*, \mu^*)$ and $\hat{L}(P^{M*}, P^*, \lambda^*, \mu^*) - \hat{L}(P^M, P, \lambda^*, \mu^*)$. Because of the strictly convexity of $\hat{L}$ on P^M, strictly concavity of $\hat{L}$ on λ, and the separable structure of $\hat{L}$ on (P^M, P), and (λ, μ), respectively, $\hat{P}^M = P^{M*}$, $\hat{\lambda} = \lambda^*$. Then, we can conclude the statement of this lemma. ■

Using Proposition 12, Lemma 13, and the Lyapunov convergence theorem, we have the following convergence result:

Lemma 14: Any solution $(P^M(t), P(t), \lambda(t), \mu(t))$ of (21) for $t \geq 0$ asympotically approaches a nonempty, compact subset of the set of saddle points.

Proof: Equation (22) tells that $U(z) \geq 0$ for any z, and (23) tells that $U(z(t))$ is decreasing with time t and $U(z(t)) \leq U(z(0))$ for any $t \geq 0$. Because of the structure of $U(z)$ in (22), $z(t)$ is bounded for $t \geq 0$. By Lyapunov convergence theory [31], we know that $z(t) = (P^M(t), P(t), \lambda(t), \mu(t))$ converges to a nonempty invariant compact subset of $\mathcal{Z}$ (defined in Lemma 13). To ensure the subset is invariant, we have $\dot{P}^M = -\Xi_{P^M}(\partial\hat{L}(\hat{z})/\partial P^M) = 0$ and $\dot{\lambda} = \Xi_\lambda(\partial\hat{L}(\hat{z})/\partial P^M) = 0$, implying that $\hat{\mu} = -\beta_j P_j^{M*} - \omega_j^*$ and $\sum_{k:j\to k} P_{jk} - \sum_{i:i\to j} P_{ij} = P_j^{M*} - P_j^L - D_j\omega_j^*$. Therefore, we know $\hat{z}$ is a saddle point of $\hat{L}$ according to Proposition 12. ■

Now we are ready to conclude the convergence of the algorithm (21).

Theorem 15: Any solution $(P^M(t), P(t), \lambda(t), \mu(t))$ of (21) for $t \geq 0$ asympotically converges to a saddle point $(P^{M*}, P^*, \lambda^*, \mu^*)$. The saddle point $(P^{M*}, P^*, \lambda^*, \mu^*)$ may depend on the initial point $(P^M(0), P(0), \lambda(0), \mu(0))$.

Proof: The proof of Lemma 14 shows that $\{z(t)\}_{t\geq 0}$ is a bounded sequences; therefore, we know that there exists a subsequence $\{z(t_j) = (P^M(t_j), P(t_j), \lambda(t_j), \mu(t_j))\}$ that converges to a point $z^\infty = (P_\infty^M, P_\infty, \lambda_\infty, \mu_\infty)$. This implies that

$$\lim_{t_j\to\infty} U_{z^\infty}(z(t_j)) = 0. \tag{24}$$

As shown in Lemma 14, z^∞ is a saddle point of $\hat{L}$. Therefore, Lemma 13, Lemma 14, and their proof tell that

$$\lim_{t\to\infty} U_{z_\infty}(z(t)) = u \tag{25}$$

for some constant u. Since $\{z(t_j)\}$ is a subsequence of $\{z(t)\}$, (24) tells that $u = 0$. Therefore, we can conclude that $z(t)$ converges to z_∞. ■

The rest of the proof follows the exact same argument of the analysis for the general optimization problem. Thus, we omit the details here to avoid duplication.

ACKNOWLEDGMENT

The authors would like to thank Dr. S. Low of the California Institute of Technology for the helpful and constructive discussions.

REFERENCES

[1] N. Li, C. Zhao, L. Chen, and S. Low, "Connecting automatic generation control and economic dispatch from an optimization view," in *Proc. Amer. Control Conf.*, 2014, pp. 735–740.

[2] A. Bergen and V. Vittal, *Power Systems Analysis*, 2nd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 1999.

[3] P. Kumar *et al.*, "Recent philosophies of automatic generation control strategies in power systems," *IEEE Trans. Power Syst.*, vol. 20, no. 1, pp. 346–357, Feb. 2005.

[4] C. Zhao, U. Topcu, N. Li, and S. Low, "Design and stability of load-side primary frequency control in power systems," *IEEE Trans. Autom. Control*, vol. 59, no. 5, pp. 1177–1189, May 2014.

[5] J. Nanda and B. Kaul, "Automatic generation control of an interconnected power system," in *Proc. Inst. Elect. Eng.*, 1978, vol. 125, no. 5, pp. 385–390.

[6] O. I. Elgerd and C. Fosha, "The megawatt frequency control problem: A new approach via optimal control theory," *IEEE Trans. Power App. Syst.*, vol. PAS-89, no. 4, pp. 563–577, Apr. 1970.

[7] M. Aldeen and H. Trinh, "Load-frequency control of interconnected power systems via constrained feedback control schemes," *Comput. Elect. Eng.*, vol. 20, no. 1, pp. 71–88, 1994.

[8] C.-T. Pan and C.-M. Liaw, "An adaptive controller for power system load-frequency control," *IEEE Trans. Power Syst.*, vol. 4, no. 1, pp. 122–128, Feb. 1989.

[9] M. S. Ćalović, "Automatic generation control: Decentralized area-wise optimal solution," *Elect. Power Syst. Res.*, vol. 7, no. 2, pp. 115–139, 1984.

[10] M. Zribi, M. Al-Rashed, and M. Alrifai, "Adaptive decentralized load frequency control of multi-area power systems," *Int. J. Elect. Power Energy Syst.*, vol. 27, no. 8, pp. 575–583, 2005.

[11] N. Bengiamin and W. Chan, "Multilevel load-frequency control of interconnected power systems," *Proc. Inst. Elect. Eng.*, vol. 125, no. 6, pp. 521–526, 1978.

[12] U. Ozguner, "Near-optimal control of composite systems: The multi time-scale approach," *IEEE Trans. Autom. Control*, vol. AC-24, no. 4, pp. 652–655, Aug. 1979.

[13] D. B. Eidson and M. D. Ilic, "Advanced generation control with economic dispatch," in *Proc. 34th IEEE Conf. Dec. Control*, 1995, vol. 4, pp. 3450–3458.

[14] M. Ilic and C.-N. Yu, "Minimal system regulation and its value in a changing industry," in *Proc. IEEE Int. Conf. Control Appl.*, 1996, pp. 442–449.

[15] X. Zhang and A. Papachristodoulou, "A real-time control framework for smart power networks with star topology," in *Proc. IEEE Amer. Control Conf.*, 2013, pp. 5062–5067.

[16] F. Dörfler, J. Simpson-Porco, and F. Bullo, "Breaking the hierarchy: Distributed control and economic optimality in microgrids," *IEEE Trans. Control Netw.*, to be published.

[17] M. Andreasson, D. V. Dimarogonas, H. Sandberg, and K. H. Johansson, "Distributed pi-control with applications to power systems frequency control," in *Proc. IEEE Amer. Control Conf.*, 2014, pp. 3183–3188.

[18] M. Andreasson, D. V. Dimarogonas, K. H. Johansson, and H. Sandberg, "Distributed vs. centralized power systems frequency control," in *Proc. IEEE Eur. Control Conf.*, 2013, pp. 3524–3529.

[19] E. Mallada and S. H. Low, "Distributed frequency-preserving optimal load control," presented at the IFAC World Congr., Cape Town, South Africa, 2014.

[20] M. Bürger, C. De Persis, and S. Trip, "An internal model approach to (optimal) frequency regulation in power grids," in *Proc. 21st Int. Symp. Math. Theory Netw. Syst.*, 2014, pp. 577–583.

[21] E. Mojica-Nava, C. A. Macana, and N. Quijano, "Dynamic population games for optimal dispatch on hierarchical microgrid control," *IEEE Trans. Syst., Man, Cybern.: Syst.*, vol. 44, no. 3, pp. 306–317, Mar. 2014.

[22] R. Mudumbai, S. Dasgupta, and B. B. Cho, "Distributed control for optimal economic dispatch of a network of heterogeneous power generators," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 1750–1760, Nov. 2012.

[23] S. T. Cady and A. Dominguez-Garcia, "Distributed generation control of small-footprint power systems," in *Proc. IEEE North Amer. Power Symp.*, 2012, pp. 1–6.

[24] P. Kundur, *Power System Stability and Control*. New York, NY, USA: McGraw-Hill Education, 1994.

[25] K. J. Arrow, L. Hurwicz, and H. Uzawa, *Studies in Linear and Non-Linear Programming*. Stanford, CA, USA: Stanford Univ. Press, 1958.

[26] J. T. Wen and M. Arcak, "A unifying passivity framework for network flow control," *IEEE Trans. Autom. Control*, vol. 49, no. 2, pp. 162–174, Feb. 2004.

[27] A. Nedić and A. Ozdaglar, "Subgradient methods for saddle-point problems," *J. Optimiz. Theory Appl.*, vol. 142, no. 1, pp. 205–228, 2009.

[28] D. Feijer and F. Paganini, "Stability of primal-dual gradient dynamics and applications to network optimization," *Automatica*, vol. 46, no. 12, pp. 1974–1981, 2010.

[29] S. You and L. Chen, "Reverse and forward engineering of frequency control in power networks," presented at the IEEE Conf. Dec. Control, Los Angeles, CA, USA, Los Angeles, CA, USA, 2014.

[30] D. Bertsekas, *Nonlinear Programming*, 2nd ed. Belmont, MA, USA: Athena Scientific, 2008.

[31] H. K. Khalil, Ed., *Nonlinear Systems*, 3rd ed. Englewood Cliffs, NJ, USA: Prentice-Hall, 2002.

Na Li (M'09) received the B.S. degree in mathematics and applied mathematics from Zhejiang University, Zhejiang, China, and the Ph.D. degree in control and dynamical systems from the California Institute of Technology, Pasadena, CA, USA, in 2013.

Currently, she is an Assistant Professor in the School of Engineering and Applied Sciences at Harvard University, Cambridge, MA, USA. She was a Postdoctoral Associate with the Laboratory for Information and Decision Systems at Massachusetts Institute of Technology, Cambridge. Her research lies in the design, analysis, optimization, and control of distributed network systems, with particular applications toward power networks and systems biology/physiology.

Prof. Li entered the Best Student Paper Award Finalist at the 2011 IEEE Conference on Decision and Control.

Changhong Zhao (M'15) received the B.Eng. degree in automation from Tsinghua University, Beijing, China, in 2010, and is currently pursuing the Ph.D. degree in electrical engineering at the California Institute of Technology, Pasadena, CA, USA.

His research is on control and optimization of power systems, with a particular focus on power system dynamics and stability, frequency and voltage regulation, and distributed load-side control.

Lijun Chen (M'05) received the Ph.D. degree from the California Institute of Technology, Pasadena, CA, USA, in 2007.

Currently, he is an Assistant Professor of Computer Science and Telecommunications at the University of Colorado at Boulder. His current research interests are in optimization and control of networked systems, distributed optimization and control, convex relaxation and parsimonious solutions, as well as game theory and its engineering applications.

Dr. Chen was a co-recipient of the Best Paper Award at the IEEE International Conference on Mobile Ad-hoc and Sensor Systems in 2007. He is currently an editor of IEEE TRANSACTIONS ON COMMUNICATIONS.