Connecting Automatic Generation Control and Economic Dispatch from an Optimization View

Na Li, Lijun Chen, Changhong Zhao and Steven H. Low

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 so as 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.**

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 is to (i) maintain system nominal frequency, and (ii) let each area absorbs its own load changes so as to maintain the scheduled net interchanges between control areas [1], [2]. 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 individual 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, e.g., 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 (i) maintain nominal frequency and (ii) reach optimal power dispatch between different control areas (and/or different generator units) to balance supply and demand within the whole interconnected electricity system (and/or within the control area) to achieve economic efficiency. We call this modified AGC the economic AGC.

In order to keep the modification minimal and also 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 [3] on distributed load control, this work lends the promise to develop a modeling framework and solution approach for systematic design of distributed, low-complexity generation and load control to achieve system-wide efficiency and robustness.

There has been a large amount of work on AGC in the last few decades, including, e.g., stability and optimum parameter setting [4], optimal or adaptive controller design [5]–[7], decentralized control [8], [9], and multilevel or multi timescale control [10], [11]; see also [2] and the references therein for a thorough and up-to-date review on AGC. Most of these work focuses on improving the control performance of AGC, such as stability and transient dynamics, but not on improving the economic efficiency. References [12], [13] introduce approaches for AGC that also support an ED feature which operates at a slower time scale and interacts with AGC frequency stabilization function. For instance, reference [13] 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 control is decentralized, where each control area can update its generation based only on local information and communications with neighboring areas.

The paper is organized as follows. In Section II, we

[1] A similar approach has been used to design a decentralized optimal load control in our previous work [3].

introduce a dynamic power network model with AGC, the ED problem, and the objetive of the economic AGC. In Section III, we reverse-engineer the conventional AGC and in Section IV, we design an economic AGC scheme from the insight obtained by the reverse engineering. In Section V, we simulate and compare the convention AGC and the economic AGC. We conclude the paper in Section VI.

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 $\mathcal{N} = \{1, \cdots, n\}$ 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 non-dispatchable renewable generation. To simplify notation, all 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 the mechanical power input, and P_j^L 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:

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

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:

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

where

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

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 [3].

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 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 [1] for a detailed introduction of governor-turbine control.

ACE-based control: In the conventional AGC, 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 both 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 equations (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, i.e., $P_j^M = P_j^L$ for all $j \in \mathcal{N}$, and the frequency is returned to the nominal value, i.e., $\omega_j = 0$ 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, i.e., deviations in frequency and the net power interchanges with the 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 (i) maintain the nominal frequency and (ii) drive the mechanical power output $P_j^M, j \in \mathcal{N}$ to the optimum of the following ED problem:[2]

$$\begin{aligned} \min \quad & \sum_{j\in\mathcal{N}} C_j(P_j^M) && \text{(5a)}\\ \text{s.t.} \quad & P_j^M = P_j^L + \sum_{k:j\to k} P_{jk} - \sum_{i:i\to j} P_{ij},\ j\in\mathcal{N} && \text{(5b)}\\ \text{over} \quad & P_j^M, P_{ij}, j\in\mathcal{N}, (i,j)\in\mathcal{E}, \end{aligned}$$

where each generator at j incurs certain cost $C_j(P_j^M)$ when its power generation is P_j^M. Equation (5b) imposes power flow balanced at each bus. The cost function $C_j(\cdot)$ is assumed to be continuous, convex. We call this modified AGC as the economic AGC. In the following sections, we will show how to reverse and forward engineer the ACE-based AGC to design an economic AGC scheme.

Remark 1. *In the conventional ACE-based AGC, if a bus j denotes a control area, then the corresponding generation change P_j^C in (4) is allocated to generator units within this area 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 [14] for detailed description. Thus if the net loads fluctuate fast and dramatically due to the large penetration of renewable energy, this 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. In fact, a system manager 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 and the generality of our results, 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,

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

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 limit. □

Consider the following optimization problem:

OGC-1

$$\begin{aligned} \min \quad & \sum_{j\in\mathcal{N}} C_j(P_j^M) + \sum_{j\in\mathcal{N}} \frac{D_j}{2}|\omega_j|^2 && \text{(6a)}\\ \text{s.t.} \quad & P_j^M = P_j^L + D_j\omega_j + \sum_{k:j\to k} P_{jk} - \sum_{i:i\to j} P_{ij} && \text{(6b)}\\ & P_j^M = P_j^L && \text{(6c)}\\ \text{over} \quad & \omega_j, P_j^M, P_{ij},\ j\in\mathcal{N}, (i,j)\in\mathcal{E}, \end{aligned}$$

where equation (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,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 the information on the objective function and the constraints. However, in practice these information is unknown. Moreover, even if we know an optimum, we cannot just set the system to the optimum. As 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 important implication 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 $C_j(P_j^M) = \frac{\beta_j}{2}(P_j^M)^2$ where $\beta_j > 0$:

[2] Because all the variables denote the deviations in this paper, it may be not straightforward to interpret this ED problem, e.g., how this problem is connected with the slower timescale ED problem using the absolute value of each variable instead of the deviated value? This problem can be seen as revising energy dispatch, because of the load changes, over the nominal values that are determined by the slower time-scale ED problem that is usually operated by ISOs or generating companies.

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

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

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

$$\omega_j = \lambda_j \tag{7a}$$
$$\dot{P}_{ij} = \epsilon_{P_{ij}}(\lambda_i - \lambda_j) \tag{7b}$$
$$\dot{P}_j^M = -\epsilon_{P_j}(\beta_j P_j^M + \lambda_j + \mu_j) \tag{7c}$$
$$\dot{\lambda}_j = \epsilon_{\lambda_j}\left(P_j^M - P_j^L - D_j\omega_j - \sum_{k:j\to k} P_{jk} + \sum_{i:i\to j} P_{ij} \right) \tag{7d}$$
$$\dot{\mu_j} = \epsilon_{\mu_j}\left(P_j^M - P_j^L\right), \tag{7e}$$

where $\epsilon_{P_{ij}}$, ϵ_{P_j}, ϵ_{λ_j} and ϵ_{μ_j} are positive stepsizes. Note that equation (7a) solves $\max_{w_j} \frac{D_j}{2} w_j^2 - \lambda_j D_j w_j$ rather than follows the primal gradient algorithm with respect to w_j; hence the algorithm (7) is called a "partial" primal-dual gradient algorithm. See the Appendix for a description of the general form of partial primal-dual gradient algorithm and its convergence.

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

$$\begin{aligned} \omega_j &= \lambda_j \\ P_j^C &= K_j M_j \left(\lambda_j - \frac{1}{\epsilon_{\mu_j} M_j}\mu_j \right), \end{aligned}$$

the partial primal-dual gradient algorithm (7) becomes:

$$\dot{\omega}_j = -\frac{1}{M_j}\left(D_j\omega_j - P_j^M + P_j^L + \sum_{k:j\to k} P_{jk} - \sum_{i:i\to j} P_{ij} \right) \tag{8a}$$
$$\dot{P}_{ij} = \epsilon_{P_{ij}}(\omega_i - \omega_j) \tag{8b}$$
$$\dot{P}_j^M = -\epsilon_{P_j}\beta_j\left(P_j^M - \frac{\epsilon_{\mu_j}}{K_j\beta_j}P_j^C + \frac{1+\epsilon_{\mu_j}M_j}{\beta_j}\omega_j \right) \tag{8c}$$
$$\dot{P}_j^C = -K_j\left(D_j\omega_j + \sum_{k:j\to k} P_{jk} - \sum_{i:i\to j} P_{ij} \right). \tag{8d}$$

If we set $\epsilon_{P_{ij}} = B_{ij}$, $\epsilon_{\mu_j} = \frac{R_jK_j}{1-R_jK_jM_j}$, $\beta_j = \frac{R_j}{1-R_jK_jM_j}$, and $\epsilon_{P_j} = \frac{1}{\beta_j T_j}$, then the partial primal-dual algorithm (8) 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 V, 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 work to characterize the range of B_j which guarantees the convergence of the algorithm. Nonetheless, the algorithm in (8) 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 above chosen ϵ_{λ_j}, ϵ_{μ_j}, $\epsilon_{P_{ij}}$ and ϵ_{Pj}, the partial primal-dual gradient algorithm (8) (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: Please see the Appendix for the convergence of the partial primal-dual gradient algorithm. □

Remark 2. *We have made an equivalence transformation in the above: from algorithm (7) to algorithm (8). 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 can not be directly observed or estimated, while P_j^C can be estimated/calculated based on the observable variables ω_j and P_{ij}. As the control should be based on observable or estimable variables, the power system implements algorithm (8) instead of (7) for the ACE-based AGC.*

The above reverse-engineering, i.e., 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 (i) keeps the frequency deviation to 0, i.e., $\omega_j = 0$ for all $j \in \mathcal{N}$, and

(ii) achieves economic efficiency, i.e., the mechanical power generation solves the ED problem (5).

Consider the following optimization problem:

OGC-2

$$\min \sum_{j\in\mathcal{N}} C_j(P_j^M) + \sum_{j\in\mathcal{N}} \frac{D_j}{2}|\omega_j|^2 \quad (9a)$$

$$\text{s.t. } P_j^M = P_j^L + D_j\omega_j + \sum_{k:j\to k} P_{jk} - \sum_{i:i\to j} P_{ij} \quad (9b)$$

$$P_j^M = P_j^L + \sum_{k:j\to k}\gamma_{jk} - \sum_{i:i\to j}\gamma_{ij} \quad (9c)$$

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

where γ_{ij} are auxiliary variables introduced to facilitate the algorithm design. As will become clear later, the reason to include constraint (9c) is in order to keep $\omega_j = 0$ for all $j \in \mathcal{N}$ and to derive an implementable control algorithm, similar to equations (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 (9b) and (9c) 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 (9c) 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 can derive the following partial prime-dual algorithm solving OGC-2:

$$\omega_j = \lambda_j \quad (10a)$$

$$\dot{P}_{i,j} = \epsilon_{P_{ij}}(\lambda_i - \lambda_j) \quad (10b)$$

$$\dot{P}_j^M = -\epsilon_{P_j}(C_j'(P_j^M) + \lambda_j + \mu_j) \quad (10c)$$

$$\dot{\gamma}_{ij} = \epsilon_{\gamma_{ij}}(\mu_i - \mu_j) \quad (10d)$$

$$\dot{\lambda}_j = \epsilon_{\lambda_j}\left(P_j^M - P_j^L - D_j\omega_j - \sum_{k:j\to k} P_{jk} + \sum_{i:i\to j} P_{ij}\right) \quad (10e)$$

$$\dot{\mu}_j = \epsilon_{\mu_j}\left(P_j^M - P_j^L - \sum_{k:j\to k}\gamma_{jk} + \sum_{i:i\to j}\gamma_{ij}\right), \quad (10f)$$

Let $\epsilon_{\lambda_j} = \frac{1}{M_j}$, $\epsilon_{P_{ij}} = B_{ij}$, $\epsilon_{\mu_j} = \frac{R_jK_j}{1-R_jK_jM_j}$ and $\epsilon_{P_j} = \frac{1-R_jK_jM_j}{T_jR_j}$ as in Section III. By using linear transformation $\omega_j = \lambda_j$ and $P_j^C = K_jM_j\left(\lambda_j - \frac{1}{\epsilon_{\mu_j}M_j}\mu_j\right)$, the partial primal-dual gradient algorithm (10) becomes:

$$\dot{\omega}_j = -\frac{1}{M_j}\left(D_j\omega_j - P_j^M + P_j^L + \sum_{k:j\to k} P_{jk} - \sum_{i:i\to j} P_{ij}\right) \quad (11a)$$

$$\dot{P}_{ij} = B_{ij}(\omega_i - \omega_j) \quad (11b)$$

$$\dot{P}_j^M = -\frac{1}{T_j}\left(\frac{1-R_jK_jM_j}{R_j}C_j'(P_j^M) - P_j^C + \frac{1}{R_j}\omega_j\right) \quad (11c)$$

$$\dot{P}_j^C = -K_j\left(D_j\omega_j + \sum_{k:j\to k}(P_{jk} - \gamma_{jk}) - \sum_{i:i\to j}(P_{ij} - \gamma_{ij})\right) \quad (11d)$$

$$\dot{\gamma}_{ij} = \epsilon_{\gamma_{ij}}\left(\left(M_i\omega_i - \frac{P_i^C}{K_i}\right)\epsilon_{\mu_i} - \left(M_j\omega_j - \frac{P_j^C}{K_j}\right)\epsilon_{\mu_j}\right). \quad (11e)$$

Compared with algorithm (8) (i.e., the power system dynamics with the ACE-based AGC), the difference in algorithm (11) is the new variables γ_{ij} and the marginal cost $C_j'(\cdot)$ in the generation control (11c). Note that γ_{ij} can be calculated based on the observable/measurable variables. So, the above algorithm is implementable. However, it might be not practical to add 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 (11d, 11e) by the following dynamics:

$$\dot{P}_j^C = -K_j\left(D_j\omega_j + \sum_{k:j\to k}(P_{jk} - \gamma_j + \gamma_k) - \sum_{i:i\to j}(P_{ij} - \gamma_i + \gamma_j)\right) \quad (12a)$$

$$\dot{\gamma}_i = \epsilon_\gamma\left(\left(M_i\omega_i - \frac{P_i^C}{K_i}\right)\epsilon_{\mu_i}\right). \quad (12b)$$

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 (11a–11c, 12a–12b) 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 $P_j^{C*} = \frac{1-R_jK_jM_j}{R_j}C_j'(P_j^{M*})$.*

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 (11a–11c, 12a–12b) as an economic AGC for the power system. By comparing with the ACE-based AGC in (1)-(4) and the economic AGC in (11a–11c, 12a–12b), 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 a local communication about the new local auxiliary variable γ_j based on (12a–12b) can improve the economic efficiency of AGC.

Remark 3. *We can actually derive a simpler and yet implementable algorithm without introducing variable $\gamma_{ij}, (i, j) \in \mathcal{E}$ (or γ_i, $i \in \mathcal{N}$). However, in order to have minimal modification to the existing conventional AGC and also keep the resulting control decentralized, we choose to derive the algorithm (11) and (12).*

V. Case Study

Consider a small 4-area interconnected system, as shown in Figure 1. The values of the generator and transmission line parameters are shown in Table II and I. Notice that though 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. For each area, the generation cost takes on the form of $C_i(P_{M_i}) = a_i P_{M_i}^2$ where a is randomly drawn from $[1, 2]$.

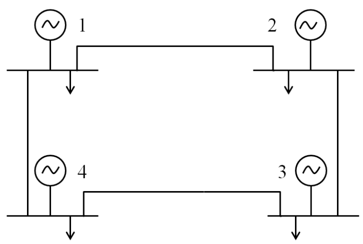

Fig. 1: A 4-area interconnected system

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

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

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

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

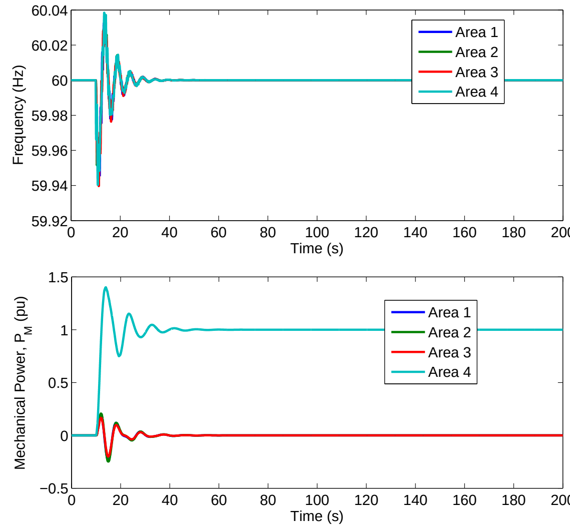

Fig. 2: The ACE-based AGC

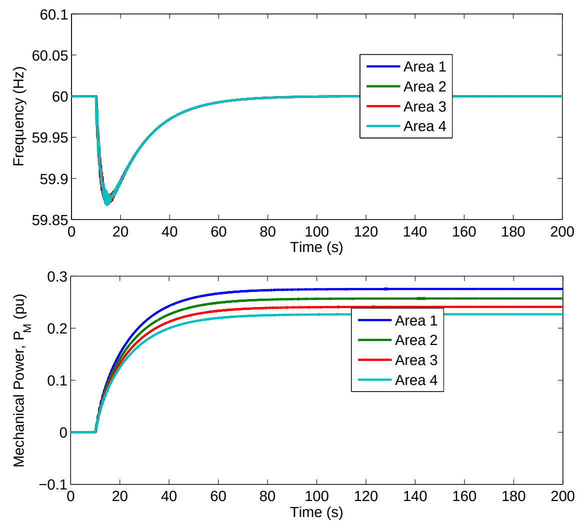

Fig. 3: The economic AGC

Simulations results show that our proposed AGC scheme works well even in these non-ideal, practical systems.

At time $t = 10$s, a step change of load occurs at area 4 where $P_4^L = 1$ pu. Figure 2 shows the dynamics of the frequencies and mechanical power outputs for the 4 areas using ACE-based AGC (1)–(4). Figure 3 shows the dynamics of the frequencies and mechanical power outputs for the 4 areas using the economic AGC (11a–11c, 12a–12b). Figure 4 compares the total generation costs using the ACE-based AGC and the economic AGC with the minimal generation cost of the ED problem (5). We see that the economic AGC does not only track the optimal value of the ED problem but also smooths out the frequency dynamics.

VI. 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

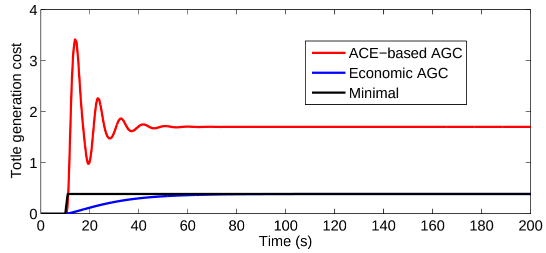

Fig. 4: The generation cost

the ED into the AGC and achieves economic efficiency. Combined with the previous work [3] on distributed load control, this work lends the promise to develop a modeling framework and solution approach for systematic design of distributed, low-complexity generation and load control to achieve system-wide efficiency and robustness.

ACKNOWLEDGMENT

This work is supported by NSF NetSE grant CNS 0911041, ARPA-E grant DE-AR0000226, Southern California Edison, National Science Council of Taiwan, R.O.C. grant, NSC 101- 3113-P-008-001, the Caltech Resnick Institute, and the Okawa Foundation.

REFERENCES

[1] A. Bergen and V. Vittal. *Power Systems Analysis*. Prentice Hall, 2 edition, 1999.

[2] P Kumar, Dwarka P Kothari, et al. Recent philosophies of automatic generation control strategies in power systems. *Power Systems, IEEE Transactions on*, 20(1):346–357, 2005.

[3] C. Zhao, U. Topcu, N. Li, and S. Low. Power system dynamics as primal-dual algorithm for optimal load control. In *arXiv:1305.0585*, 2012.

[4] J Nanda and BL Kaul. Automatic generation control of an interconnected power system. In *Proceedings of the Institution of Electrical Engineers*, volume 125, pages 385–390, 1978.

[5] O. I. Elgerd and C. Fosha. The megawatt frequency control problem: A new approach via optimal control theory. *IEEE Transactions on Power Apparatus and Systems*, 89(4):563–577, 1970.

[6] M Aldeen and H Trinh. Load-frequency control of interconnected power systems via constrained feedback control schemes. *Computers & electrical engineering*, 20(1):71–88, 1994.

[7] C-T Pan and C-M Liaw. An adaptive controller for power system load-frequency control. *Power Systems, IEEE Transactions on*, 4(1):122–128, 1989.

[8] Milan S Ćalović. Automatic generation control: Decentralized area-wise optimal solution. *Electric power systems research*, 7(2):115–139, 1984.

[9] M Zribi, M Al-Rashed, and M Alrifai. Adaptive decentralized load frequency control of multi-area power systems. *International Journal of Electrical Power & Energy Systems*, 27(8):575–583, 2005.

[10] NN Bengiamin and WC Chan. Multilevel load-frequency control of interconnected power systems. *Electrical Engineers, Proceedings of the Institution of*, 125(6):521–526, 1978.

[11] Umit Ozguner. Near-optimal control of composite systems: The multi time-scale approach. *Automatic Control, IEEE Transactions on*, 24(4):652–655, 1979.

[12] D Brian Eidson and Marija D Ilic. Advanced generation control with economic dispatch. In *Decision and Control, Proceedings of the 34th IEEE Conference on*, volume 4, pages 3450–3458, 1995.

[13] MD Ilic and Chien-Ning Yu. Minimal system regulation and its value in a changing industry. In *Control Applications, Proceedings of the 1996 IEEE International Conference on*, pages 442–449, 1996.

[14] Prabha Kundur. *Power system stability and control*. Tata McGraw-Hill Education, 1994.

[15] D.P. Bertsekas. *Nonlinear programming, 2nd edition*. Athena Scientific Belmont, MA, 2008.

[16] H. K. Khalil, editor. *Nonlinear Systems, 3rd Edition*. Prentice Hall, 2002.

APPENDIX

Consider the following optimization problem:

$$\begin{aligned} \min_{x,y} \quad & f(x) + g(y) \\ \text{s.t.} \quad & Ax + By = C, \end{aligned} \tag{13}$$

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

The Lagrangian function of this optimization problem is given by:

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

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

The partial primal-dual gradient algorithm is given as follows:

$$\begin{aligned} \textbf{Algorithm-1:} \quad x(t) &= \min_x \left\{ f(x) + \lambda^T Ax \right\} \\ \dot{y} &= -\Xi_y \left(\frac{\partial g(y)}{\partial y} + B^T \lambda \right) \\ \dot{\lambda} &= \Xi_\lambda (Ax + By - C) \end{aligned}$$

where $\Xi_y = diag(\epsilon_{y_i})$ and $\Xi_\lambda = diag(\epsilon_{\lambda_j})$.

In the following we will study the convergence of this algorithm.

Define

$$q(\lambda) \triangleq \min_x \left\{ f(x) + \lambda^T Ax \right\}$$

$$\hat{L}(y, \lambda) \triangleq q(\lambda) + g(y) + \lambda^T (By - C).$$

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

Proposition 6. *$q(\lambda)$ is a concave function and its gradient is given as $\frac{\partial q(\lambda)}{\partial \lambda} = Ax$. If $ker(A^T) = 0$, then $q(\lambda)$ is a strictly concave function. As a consequence, given any y, there is a unique maximizer for $\max_\lambda \hat{L}(y, \lambda)$.*

Proof: This proposition follows directly from Proposition 6.1.1 in [15]. □

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

Lemma 7. *If (x^*, y^*, λ^*) is a saddle point of L, then (y^*, λ^*) is a saddle point of $\hat{L}$ and $x^* = argmin_x \left\{ f(x) + (\lambda^*)^T Ax \right\}$. Moreover, if (y^*, λ^*) is a saddle point of $\hat{L}$, then (x^*, y^*, λ^*) is a saddle point of L where $x^* = argmin_x \left\{ f(x) + (\lambda^*)^T Ax \right\}$.*

Proof: The proof is straightforward by comparing the first order conditions of saddles points for both L and $\hat{L}$.

Notice that convexity of f, g, and concavity of q implies that those first order conditions are necessary and sufficient conditions for saddle points. □

In the following, we will assume that $ker(A^T) = 0$. Thus $q(\lambda)$ and $\hat{L}$ are stricitly concave on λ.

Now we study the convergenc of **Algorithm-1**. With $\hat{L}(y, \lambda)$, **Algorithm-1** can be written as follows:

$$\dot{y} = -\Xi_y \left(\frac{\partial \hat{L}(y,\lambda)}{\partial y} \right) \quad (14)$$

$$\dot{\lambda} = \Xi_\lambda \left(\frac{\partial \hat{L}(y,\lambda)}{\partial \lambda} \right) \quad (15)$$

Let (y^*, λ^*) be a saddle point of $\hat{L}(y, \lambda)$. Define a nonegative function as:

$$\begin{aligned} U(y,\lambda) &= \frac{1}{2} \begin{bmatrix} y - y^* \\ \lambda - \lambda^* \end{bmatrix}^T \begin{bmatrix} \Xi_y^{-1} & \\ & \Xi_\lambda^{-1} \end{bmatrix} \begin{bmatrix} y - y^* \\ \lambda - \lambda^* \end{bmatrix} \\ &= \sum_{i=1}^n \frac{1}{2\epsilon_{y_i}} (y_i - y_i^*)^2 + \sum_{i=1}^m \frac{1}{2\epsilon_{\lambda_i}} (\lambda_i - \lambda_i^*)^2 \quad (16) \end{aligned}$$

Notice that $U \geq 0$ for any (y, λ). The derivative of U along the trajectory defined in (14,15) is given as:

$$\begin{aligned} \frac{\partial U}{\partial t} &= -\frac{\partial \hat{L}(y,\lambda)}{\partial y}^T (y - y^*) + \frac{\partial \hat{L}(y,\lambda)}{\partial \lambda} (\lambda - \lambda^*) \\ &\leq -\hat{L}(y,\lambda) + \hat{L}(y^*,\lambda) + \hat{L}(y,\lambda) - \hat{L}(y,\lambda^*) \quad (17) \\ &= \hat{L}(y^*,\lambda) - \hat{L}(y^*,\lambda^*) + \hat{L}(y^*,\lambda^*) - \hat{L}(y,\lambda^*) \\ &\leq 0 \end{aligned}$$

where the first equality comes from (14,15,16), the first inequality follows from the strictly concavity of $\hat{L}$ in λ and convexity of $\hat{L}$ in y and last inequality comes from that (y^*, λ^*) is a saddle point of $\hat{L}$. Therefore U is actually a Lyapunov function of (14,15). For simplicity, we will denote (y, λ) as z.

Lemma 8. *$\frac{\partial U(z)}{\partial t} \leq 0$ for all z, and $\left\{ \hat{z} : \frac{\partial U(\hat{z})}{\partial t} = 0 \right\} = \left\{ \hat{z} : \hat{\lambda} = \lambda^*, \hat{L}(\hat{y}, \lambda^*) = \hat{L}(y^*, \lambda^*) \right\}$.*

Proof: (17) has shown that $\frac{\partial U(z)}{\partial t} \leq 0$. To ensure $\frac{\partial U(\hat{z})}{\partial t} = 0$, we need that $\hat{L}(y^*, \hat{\lambda}) = \hat{L}(y^*, \lambda^*) = \hat{L}(\hat{y}, \lambda^*)$, which implies that $\hat{\lambda} = \lambda^*$ because $\hat{L}$ is strictly concave in λ and (y^*, λ^*) is a saddle point. Thus we can conclude the lemma. □

Lemma 9. *Given any two saddle points (y_1^*, λ_1^*), (y_2^*, λ_2^*), we have $\lambda_1^* = \lambda_2^*$, and $\hat{L}(y_1^*, \lambda_1^*) = \hat{L}(y_2^*, \lambda_2^*)$. Any solution $(y(t), \lambda(t))$ of (14,15) for $t \geq 0$ asympotically approaches to a nonempty, compact subset of the set of saddle points.*

Proof: (16) tells that $U(z) \geq 0$ for any z, and (17) 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 (16), $z(t) = (y(t), \lambda(t))$ is bounded for $t \geq 0$. By Lyapunov convergence theory [16] , $z(t) = (y(t), \lambda(t))$ converges to a nonempty invariant compact subset of $\left\{ \hat{z} : \frac{\partial U(\hat{z})}{\partial t} = 0 \right\} = \left\{ \hat{z} : \hat{\lambda} = \lambda^*, \hat{L}(\hat{y}, \lambda^*) = \hat{L}(y^*, \lambda^*) \right\}$. To ensure the subset is invariant, we have $\dot{\lambda} = \frac{\partial \hat{L}(\hat{Z})}{\partial \lambda} = 0$ which implies that such $\hat{z}$ is a saddle point of $\hat{L}$. □

Theorem 10. *Any solution $(y(t), \lambda(t))$ of (14,15) for $t \geq 0$ asympotically converges to a saddle point (y^*, λ^*). The saddle point (y^*, λ^*) may depend on the initial point $(y(0), \lambda(0))$.*

Proof: The proof of Lemma 9 show that $\{z(t)\}_{t \geq 0}$ is a bounded sequences, therefore, we know that there exists a subsequence $\{z(t_j) = (y(t_j), \lambda(t_j))\}$ converges to a point $z^\infty = (y^\infty, \lambda^\infty)$. This implies that:

$$\lim_{t_j \to \infty} \sum_{i=1}^n \frac{1}{2\epsilon_{y_i}} (y_i(t_j) - y_i^\infty)^2 + \sum_{i=1}^m \frac{1}{2\epsilon_{\lambda_i}} (\lambda_i(t_j) - \lambda_i^\infty)^2 = 0. \quad (18)$$

As shown in Lemma 9, $z^\infty = (y^\infty, \lambda^\infty)$ is a saddle point of $\hat{L}$. Therefore Lemma 8,9 tells that:

$$\begin{aligned} &\lim_{t \to \infty} U(y(t) - y^\infty, \lambda(t) - \lambda^\infty) \\ &= \lim_{t \to \infty} \sum_{i=1}^n \frac{1}{2\epsilon_{y_i}} (y_i(t) - y_i^\infty)^2 + \sum_{i=1}^m \frac{1}{2\epsilon_{\lambda_i}} (\lambda_i(t) - \lambda_i^\infty)^2 \\ &= u \end{aligned}$$

for some constant u. Since $\{z(t_j) = (y(t_j), \lambda(t_j))\}$ is a subsequence of $\{z(t)\}$, (18) tells that $u = 0$. Therefore, we can conclude that $(y(t), \lambda(t))$ converges to $(y^\infty, \lambda^\infty)$. □

The above proof for the general partial primal-dual gradient algorithm can be easily extended to prove Theorem 3 and 5. We omit the details here due to the space limitation.