Optimal load-side control for frequency regulation in smart grids

Enrique Mallada, Member, IEEE, Changhong Zhao, Student Member, IEEE, and Steven Low, Fellow, IEEE

Abstract—Frequency control rebalances supply and demand while maintaining the network state within operational margins. It is implemented using fast ramping reserves that are expensive and wasteful, and which are expected to grow with the increasing penetration of renewables. The most promising solution to this problem is the use of demand response, i.e. load participation in frequency control. Yet it is still unclear how to efficiently integrate load participation without introducing instabilities and violating operational constraints.

In this paper we present a comprehensive load-side frequency control mechanism that can maintain the grid within operational constraints. Our controllers can rebalance supply and demand after disturbances, restore the frequency to its nominal value and preserve inter-area power flows. Furthermore, our controllers are distributed (unlike generation-side), can allocate load updates optimally, and can maintain line flows within thermal limits. We prove that such a distributed load-side control is globally asymptotically stable and robust to unknown load parameters. Simulations are used to illustrate the properties of our solution.

I. INTRODUCTION

Frequency control maintains the frequency of a power network at its nominal value when demand or supply fluctuates. It is traditionally implemented on the generation side and consists of three mechanisms that work in concert [1]–[3]. The primary frequency control, called the droop control and completely decentralized, operates on a timescale up to low tens of seconds and uses a governor to adjust, around a setpoint, the mechanical power input to a generator based on the local frequency deviation. The primary control can rebalance power and stabilize the frequency but does not restore the nominal frequency. The secondary frequency control (called automatic generation control) operates on a timescale up to a minute or so and adjusts the setpoints of governors in a control area in a centralized fashion to drive the frequency back to its nominal value and the inter-area power flows to their scheduled values. Finally, economic dispatch operates on a timescale of several minutes or up and schedules the output levels of generators that are online and the inter-area power flows. See [4], [5] for a recent hierarchical model of power systems and their stability analysis.

Load-side participation in frequency control offers many advantages, including faster response, lower fuel consumption and emission, and better localization of disturbances. The idea of using frequency adaptive loads dates back to [6] that advocates their large scale deployment to “assist or even replace turbine-governed systems and spinning reserve.” They also proposed to use spot prices to incentivize the users to adapt their consumption to the true cost of generation at the time of consumption. Remarkably it was emphasized back then that such frequency adaptive loads will “allow the system to accept more readily a stochastically fluctuating energy source, such as wind or solar generation” [6].

This is echoed recently in, e.g., [7]–[13] that argue for “grid-friendly” appliances, such as refrigerators, water or space heaters, ventilation systems, and air conditioners, as well as plug-in electric vehicles to help manage energy imbalance. Simulations in all these studies have consistently shown significant improvement in performance and reduction in the need for spinning reserves. A small scale project by the Pacific Northwest National Lab in 2006–2007 demonstrated the use of 200 residential appliances in primary frequency control that automatically reduce their consumption when the frequency of the household dropped below a threshold (59.95Hz) [14].

In spite of these simulation studies and field trials, it was not until very recently that analytic studies were developed on the (potential) behavior of the large-scale deployment of distributed frequency control. Some of the main examples of these studies focus on distributed secondary frequency control in power systems [5], [15]–[17], and microgrids [18]–[22]. However, a general solution is yet to be developed on how to rebalance supply and demand, restore nominal frequency, preserve inter-area flows and avoid thermal limit violations.

Another model was recently presented in [23] that formulates an optimal load control (OLC) problem where the objective is to minimize the aggregate disutility of tracking an operating point subject to power balance over the network. The main conclusion is that decentralized load-side primary frequency control, coupled with the power network dynamics, serves as a primal-dual algorithm to solve (the Lagrangian dual of) OLC. Like the droop control on the generation side, the scheme in [23] rebalances power and resynchronizes frequencies after a disturbance, but does not drive the system to a desirable operating point. Similar ideas since then have been developed to include AGC and governor dynamics [24] and to use load-side secondary frequency control to restore the frequency to its nominal value [25], [26].

This work was 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 103-3113-P-008-001, and Caltech Resnick Institute.

The authors are with the Department of Computing + Mathematical Sciences, California Institute of Technology, Pasadena, CA 91125 USA (email: {mallada, zhao, slow}@caltech.edu).
Contributions of this work: In this paper, we extend this framework to allow the system restore the desired operational constraints. We first modify the OLC problem to include the operational constraints in Section III. The crux of our contribution is the introduction of surrogate line flows that in equilibrium are equal to the real line flows. This allows us to derive a distributed solution that preserves the primal-dual interpretation of the network dynamics (Section IV) and guarantees global asymptotic stability (Section V).

A preliminary version of this work has been presented in [27]. This paper extends [27] in many ways: Firstly, we include several theorem proofs that were omitted in [27] due to space constraints. Secondly, we add a new section (Section VI) that illustrates the robustness of our controllers. That is, we show that stability is preserved under load parameters uncertainty (Theorem 17). Thirdly, we extend our numerical simulation section (Section VII) to further illustrate the conservativeness of the uncertainty bounds of Theorem 17.

Our contribution with respect to the existing literature is also manifold. Unlike [16], [17], our global asymptotic stability result (Theorem 10) is independent of the controller gains, which is highly desirable for fully distributed deployments. Our results hold for arbitrary topologies, in contrast with [17], [19], and can impose inter-area constraints and thermal limits. Finally, the results of Section VI are, to the best of our knowledge, novel among the literature of primal-dual algorithm convergence, and provide much desired robustness properties for large scale distributed deployments.

II. Preliminaries

Let \( \mathbb{R} \) be the set of real numbers and \( \mathbb{N} \) the set of natural numbers. Given a finite set \( S \subset \mathbb{N} \) we use \( |S| \) to denote its cardinality. For a set of scalar numbers \( a_i \in \mathbb{R}, i \in S, \) we denote \( a_S \) as the column vector of the \( a_i \) components, i.e. \( a_S := (a_i, i \in S) \in \mathbb{R}^{|S|}; \) we usually drop the subscript \( S \) when the set is known from the context. Similarly, for two vectors \( a, b \in \mathbb{R}^{|S|} \) and \( b \in \mathbb{R}^{|S|} \) we define the column vector \( x = (a, b) \in \mathbb{R}^{|S|+|S|} \). Given any matrix \( A \), we denote its transpose as \( A^T \) and use \( A_e \) to denote the \( i \)th row of \( A \). We will also use \( A_S \) to denote the sub matrix of \( A \) composed only of the rows \( A_i \) with \( i \in S \). The diagonal matrix of a sequence (\( a_i, i \in S \)) is represented by \( \text{diag}(a_i) \). Similarly, for a sequence of submatrices \( \{A_h, h \in S\} \) we let \( \text{blockdiag}(A_h) \) denote the block diagonal matrix. Finally, we use \( 1 \) (0) to denote the vector/matrix of all ones (zeros), where its dimension is understood from the context.

A. Network Model

We consider a power network described by a directed graph \( G(\mathcal{N}, \mathcal{E}) \) where \( \mathcal{N} = \{1, \ldots, |\mathcal{N}|\} \) is the set of buses and \( \mathcal{E} \subset \mathcal{N} \times \mathcal{N} \) is the set of transmission lines denoted by either \( e \) or \( ij \) such that \( i,j \in \mathcal{N} \), then \( ji \not\in \mathcal{E} \). We partition the buses \( \mathcal{N} = \mathcal{G} \cup \mathcal{L} \) and use \( \mathcal{G} \) and \( \mathcal{L} \) to denote the set of generator and load buses respectively. We assume the graph \( (\mathcal{N}, \mathcal{E}) \) is connected, and make the following assumptions which are well-justified for transmission networks [28]:

- Lines \( ij \in \mathcal{E} \) are lossless and characterized by their susceptances \( B_{ij} = B_{ji} > 0 \).
- Reactive power flows do not affect bus voltage phase angles and frequencies.

We further assume that the bus frequency \( \omega_i \) and line flows \( P_{ij} \) are close to schedule values \( \omega_0 \) and \( P_{ij}^0 \). In other words, \( P_{ij} = P_{ij}^0 + \delta P_{ij} \) and \( \omega_i = \omega_0 + \delta \omega_i \) where \( \delta P_{ij} \) and \( \delta \omega_i \) small enough; without loss of generality, we take here \( \omega_0 = 0 \). We refer the reader to [30] for a thorough motivation of the model. The evolution of the transmission network is then described by

\[
M_i \dot{\omega}_i = P_i^m - (d_i + \hat{d}_i) - \sum_{e \in \mathcal{E}} C_{i,e} P_e \quad i \in \mathcal{G} \quad (1a)
\]
\[
0 = P_i^m - (d_i + \hat{d}_i) - \sum_{e \in \mathcal{E}} C_{i,e} P_e \quad i \in \mathcal{L} \quad (1b)
\]
\[
\hat{P}_{ij} = B_{ij} (\omega_i - \omega_j) \quad ij \in \mathcal{E} \quad (1c)
\]
\[
\dot{\hat{d}}_i = D_i \omega_i \quad i \in \mathcal{N} \quad (1d)
\]

where \( d_i \) denotes an aggregate controllable load, \( \hat{d}_i = D_i \omega_i \) denotes an aggregate uncontrollable but frequency-sensitive load as well as damping loss at generator \( i \), \( M_i \) is the generator’s inertia, \( P_i^m \) is the mechanical power injected by a generator \( i \in \mathcal{G} \), and \( -P_i^m \) is the aggregate power consumed by constant loads for \( i \in \mathcal{L} \). Finally, \( C_{i,e} \) are the elements of the incidence matrix \( C \in \mathbb{R}^{N \times |\mathcal{E}|} \) of the graph \( G(\mathcal{N}, \mathcal{E}) \) such that \( C_{i,e} = -1 \) if \( e = ji \in \mathcal{E} \) and \( C_{i,e} = 1 \) if \( e = ij \in \mathcal{E} \) and \( C_{i,e} = 0 \) otherwise.

For notational convenience we will use whenever needed the vector form of (1), i.e.

\[
M_G \dot{\omega}_G = P_G^m - (d_G + \hat{d}_G) - C_G P
\]
\[
0 = P_L^m - (d_L + \hat{d}_L) - C_L P
\]
\[
\hat{P} = D_B \dot{\omega}
\]
\[
\dot{\bar{d}} = D \bar{d}
\]

where the matrices \( C_L \) and \( C_G \) are defined by splitting the rows of \( C \) between generator and load buses, i.e. \( C = [C_G^T \; C_L^T]^T \), \( D = \text{diag}(D_i)_{i \in \mathcal{N}}, D_B = \text{diag}(B_{ij})_{ij \in \mathcal{E}} \) and \( M_G = \text{diag}(M_i)_{i \in \mathcal{G}} \).

B. Operational Constraints

We denote each control area using \( k \) and let \( \mathcal{K} := \{1, \ldots, |\mathcal{K}|\} \) denote the set of areas. Within each area, the Automatic Generation Control (AGC) scheme seeks to restore the frequency to its nominal value as well as preserving a constant power transfer outside the area, i.e.

\[
\sum_{i \in \mathcal{N}_k} \sum_{e \in \mathcal{E}} C_{i,e} P_e = e_k^T \dot{C} P = \dot{P}_k, \; \forall k \in \mathcal{K}, \quad (2)
\]

where \( \mathcal{N}_k \subset \mathcal{N} \) is the set of buses of area \( k \in \mathcal{K}, e_k \in \mathbb{R}^{N_k} \), \( k \in \mathcal{K} \), is a vector with elements \( (e_k)_{i} = 1 \) if \( i \in \mathcal{N}_k \) and \( (e_k)_{i} = 0 \) otherwise, \( \dot{P}_k \) is the net scheduled power injection of area \( k \).

\[1\]The analysis can be extended to networks with constant R/X ratio [29].
Notice that if we define
\[ \hat{C} := E_k C \]  
with \( E_k := [e_1 \ldots e_{|k|}]^T \) and \( \hat{C} \in \mathbb{R}^{|K| \times |E|} \), then constraint (2) can be compactly expressed using
\[ \hat{C} P = \hat{P} \]  
where \( \hat{P} = (\hat{P}_e)_{e \in K} \in \mathbb{R}^{|K|} \). It is easy to see that \( \hat{C}_{k,e} (e = ij) \) is equal to 1 if \( ij \) is an inter-area line with \( i \in N_k \), -1 if \( ij \) is an inter-area line with \( j \in N_k \), and 0 otherwise.

Finally, the thermal limit constraints are usually given by
\[ P \leq \bar{P} \leq \check{P} \]  
where \( \check{P} := (\check{P}_e)_{e \in E} \) and \( P := (P_e)_{e \in E} \) represent the line flow limits; usually \( P = -\check{P} \) so that we get \( |P| \leq \check{P} \).

C. Efficient Load Control

Suppose the system (1) is in equilibrium, i.e. \( \dot{\omega}_i = 0 \) for all \( i \) and \( P_{ij} = 0 \) for all \( ij \), and at time 0, there is a disturbance, represented by a step change in the vector \( P^m := (P^m_i, i \in N) \), that produces a power imbalance. Then, we are interested in designing a distributed control mechanism that rebalances the system while preserving the frequency within its nominal value as well as maintaining the operational constraints of Section II-B. Furthermore, we would like this mechanism to produce an efficient allocation among all the users (or loads) that are willing to adapt.

We use \( c_i(d_i) \) and \( \frac{d_i^2}{2D_i} \) to denote the cost or disutility of changing the load consumption by \( d_i \) and \( \dot{d}_i \) respectively. This allows us to formally describe our notion of efficiency in terms of the loads' welfare. More precisely, we shall say that a load control \((d, \dot{d})\) is efficient if it solves the following problem.

**Problem 1 (WELFARE).**

\[
\begin{align*}
\text{minimize} & \quad \sum_{i \in N} c_i(d_i) + \frac{d_i^2}{2D_i} \\
\text{subject to} & \quad \text{operational constraints.}
\end{align*}
\]  

(6)

Problem 1 has been originally proposed in [30] for the case where the operational constraint is to balance supply and demand, i.e.
\[ \sum_{i \in N} (d_i + \dot{d}_i) = \sum_{i \in N} P^m_i. \]

It is shown in [30] that when
\[ d_i = c_i^{-1}(\dot{\omega}_i), \]
then (1) is a distributed primal-dual algorithm that solves (6) subject to (7).

Therefore, one can use problem (6)-(7) to forward engineering the desired controllers, by means of primal-dual decomposition, that can rebalance supply and demand. Like primary frequency control, the system (1) and (8) suffers from the disadvantage that the optimal solution of (6)-(7) may not recover the frequency to the nominal value and satisfy the additional operational constraints of Section II-B.

In the next section we shall see that a clever modification of (6)-(7) will be able to restore the nominal frequency while maintaining the interpretation of (1) as a component of the primal-dual algorithm that solves the modified optimization problem. An additional byproduct of the formulation is that we can also impose any type of linear equality and inequality constraint that the operator may require.

III. OPTIMAL LOAD-SIDE CONTROL

We now proceed to describe the optimization problem that will be used to derive the distributed controllers that achieve our goals. The crux of our solution comes from including additional constraints to Problem 1 that implicitly guarantee the desired operational constraints, yet still preserves the desired structure which allows the use of (1) as part of the optimization algorithm.

Thus, we will use the following modified version of Problem 1:

**Problem 2 (OLC).**

\[
\begin{align*}
\text{minimize} & \quad \sum_{i \in N} c_i(d_i) + \frac{d_i^2}{2D_i} \\
\text{subject to} & \quad P^m - (d + \dot{d}) = CP \quad \text{(9b)} \\
& \quad P^m - d = L_B v \quad \text{(9c)} \\
& \quad \hat{C} D_B C^T v = \hat{P} \quad \text{(9d)} \\
& \quad P \leq D_B C^T v \leq \check{P} \quad \text{(9e)}
\end{align*}
\]

where \( L_B := C D_B C^T \) is the \( B_{ij} \)-weighted Laplacian matrix.

Although not clear at first sight, the constraint (9c) implicitly enforces that any optimal solution of OLC \((d^*, \dot{d}^*, P^*, v^*)\) will restore the frequency to its nominal value, i.e. \( \dot{d}^*_i = D_i \dot{\omega}_i^* = 0 \). Similarly, we will use constraint (9d) to impose (2) (or equivalently (4)) and (9e) to impose (5).

Throughout this paper we make the following assumptions:

**Assumption 1** (Cost function). The cost function \( c_i(d_i) \) is \( \alpha \)-strongly convex and second order continuously differentiable \((c_i \in C^2 \text{ with } c''_i(d_i) \geq \alpha > 0 \) in the interior of its domain \( D_i := [d_i, \bar{d}_i] \subseteq \mathbb{R} \), such that \( c_i(d_i) \rightarrow +\infty \) whenever \( d_i \rightarrow \partial D_i \).

**Assumption 2** (Slater Condition). The OLC problem (9) is feasible and there is at least one feasible \((d, \dot{d}, P, v)\) such that \( d \in \text{Int } D := \Pi_{i \in N} D_i \) [31, Ch. 5.2.3].

The remainder of this section is devoted to understanding the properties of the optimal solutions of OLC. We will use \( \nu_i, \lambda_i, \pi_k \) as Lagrange multipliers of constraints (9b), (9c) and (9d), and \( \rho_{ij}^+ \) and \( \rho_{ij}^- \) as multipliers of the right and left constraints of (9e), respectively. In order to make the presentation more compact sometimes we will use \( x = (P, v) \in \mathbb{R}^{|E|} \) and \( \sigma = (\nu, \lambda, \pi, \rho^+, \rho^-) \in \mathbb{R}^{|E| + |N|} \), as well as \( \sigma_i = (\nu_i, \lambda_i) = (\pi_k) \) and \( \sigma_{ij} = (\rho_{ij}^+, \rho_{ij}^-) \). We will also use \( \rho := (\rho^+, \rho^-) \).
Next, we consider the dual function $D(\sigma)$ of the OLC problem.

$$D(\sigma) = \inf_{d, \hat{d}, x} L(d, \hat{d}, x, \sigma)$$

(10)

where

$$L(d, \hat{d}, x, \sigma) = \sum_{i \in \mathcal{N}} (c_i(d_i) + \frac{\hat{d}_i^2}{2D_i}) + \nu^T(P - (d + \hat{d})) - CP + \lambda^T(P - d - L_B \sigma) + \pi^T(\hat{C} \hat{B} C \nu - \tilde{P}) + \rho^T(D_B C \nu - \tilde{P}) + \rho^T(P - D_B C \nu)$$

$$= \sum_{i \in \mathcal{N}} (c_i(d_i) - (\lambda_i + \nu_i)d_i + \frac{\hat{d}_i^2}{2D_i} - \nu_i \hat{d}_i + (\nu_i + \lambda_i)P_{i}^m)$$

$$- \pi^T \tilde{P} - \rho^T \tilde{P} + \rho^T P$$

(11)

Since $c_i(d_i)$ and $\frac{\hat{d}_i^2}{2D_i}$ are radially unbounded, the minimization over $d$ and $\hat{d}$ in (10) is always finite for given $x$ and $\sigma$. However, whenever $C^T \nu \neq 0$ or $L_B \lambda - C_D C^T \pi - C_D B (\rho^+ - \rho^-) \neq 0$, one can modify $P$ or $\nu$ to arbitrarily decrease (11). Thus, the infimum is attained if and only if we have

$$C^T \nu = 0 \quad \text{and} \quad L_B \lambda - C_D C^T \pi - C_D B (\rho^+ - \rho^-) = 0.$$  

(12a) and (12b)

Moreover, the minimum value must satisfy

$$\frac{c_i'(d_i)}{\hat{d}_i} = \nu_i + \lambda_i \quad \text{and} \quad \frac{\hat{d}_i}{D_i} = \nu_i, \quad \forall i \in \mathcal{N}.$$  

(13)

Using (12) and (13) we can compute the dual function

$$D(\sigma) = \begin{cases} \Phi(\sigma), & \sigma \in \tilde{\mathcal{N}} \\ -\infty, & \text{otherwise,} \end{cases}$$

(14)

where

$$\tilde{\mathcal{N}} := \{ \sigma \in \mathbb{R}^{2|\mathcal{N}| + |\mathcal{K}| + 2|\mathcal{E}|} : (12a) \text{ and } (12b) \}$$

and the function $\Phi(\sigma)$ is decoupled in $\sigma_i = (\nu_i, \lambda_i), \sigma_k = (\pi_k)$ and $\sigma_{ij} = (\rho_{ij}^+, \rho_{ij}^-)$. That is,

$$\Phi(\sigma) = \sum_{i \in \mathcal{N}} \Phi_i(\sigma_i) + \sum_{k \in \mathcal{K}} \Phi_k(\sigma_k) + \sum_{ij \in \mathcal{E}} \Phi_{ij}(\sigma_{ij})$$

(15)

where $\Phi_k(\sigma_k) = -\pi_k \tilde{P}_k, \Phi_i(\sigma_i) = \rho_{ij}^- \tilde{P}_{ij} - \rho_{ij}^+ \tilde{P}_{ij}$ and

$$\Phi_i(\sigma_i) = c_i(d_i(\sigma_i)) + (\nu_i + \lambda_i)(P_{i}^m - d_i(\sigma_i)) - \frac{D_i}{2} \nu_i^2$$

(16)

with

$$d_i(\sigma_i) = c_i^{-1}(\nu_i + \lambda_i).$$

(17)

The dual problem of the OLC (DOLC) is then given by

**DOLC:**

maximize

$$\sum_{i \in \mathcal{N}} \Phi_i(\nu_i, \lambda_i) + \sum_{k \in \mathcal{K}} \Phi_k(\pi_k) + \sum_{ij \in \mathcal{E}} \Phi_{ij}(\rho_{ij})$$

subject to

(12a) and (12b)

$$\rho_{ij}^+ \geq 0, \quad \rho_{ij}^- \geq 0, \quad i, j \in \mathcal{E}$$

(18)

Clearly, DOLC is feasible (e.g. take $\sigma = 0$). Then, Assumption 2 implies dual optimal is attained.

Although $D(\sigma)$ is only finite on $\tilde{\mathcal{N}}, \Phi_i(\sigma_i), \Phi_k(\sigma_k)$ and $\Phi_{ij}(\sigma_{ij})$ are finite everywhere. Thus sometimes we use the extended version of the dual function (15) instead of $D(\sigma)$, knowing that $D(\sigma) = \Phi(\sigma)$ for $\sigma \in \tilde{\mathcal{N}}$. Given any $S \subset \mathcal{N}, K \subset \mathcal{K}$ or $U \subset \mathcal{E}$ we also define

$$\Phi_S(\sigma_S) := \sum_{i \in S} \Phi_i(\sigma_i), \quad \Phi_K(\sigma_K) := \sum_{k \in K} \Phi_k(\sigma_k)$$

and

$$\Phi_U(\sigma_U) := \sum_{ij \in U} \Phi_{ij}(\sigma_{ij})$$

such that $\Phi(\sigma) = \Phi_N(\sigma_N) + \Phi_K(\sigma_K) + \Phi_E(\sigma_E)$.

The following lemmas describe several properties of our optimization problem that we will use in latter sections.

**Lemma 3 (Strict concavity of $\Phi_S(\sigma_S)$).** Given any set $S \subseteq \mathcal{N}$, nonempty, the function $\Phi_S(\sigma_S)$ is the sum of strictly concave functions $\Phi_i(\sigma_i)$ and it is therefore strictly concave. Moreover, the (extended) dual function $\Phi(\sigma)$ is strictly concave with respect to $\sigma_N = (\nu, \lambda)$.

**Proof:** From the derivation of $\Phi_i(\sigma_i)$ it is easy to show that

$$\Phi_i(\sigma_i) = \min_{d_i, \hat{d}_i} L_i(d_i, \hat{d}_i, \sigma_i)$$

(19)

where

$$L_i(d_i, \hat{d}_i, \sigma_i) := c_i(d_i) + \frac{\hat{d}_i^2}{2D_i} + (\nu_i + \lambda_i)P_{i}^m - d_i - \nu_i \hat{d}_i.$$

Notice that $L_i(d_i, \hat{d}_i, \sigma_i)$ is linear in $\sigma_i$ and strictly convex in $(d_i, \hat{d}_i)$.

Let $d_i(\sigma_i)$ and $\hat{d}_i(\sigma_i)$ be the unique minimizer of (19). Then from (13) it follows that $d_i(\sigma_i) = d_i(\nu_i + \lambda_i) = c_i^{-1}(\nu_i + \lambda_i)$ and $\hat{d}_i(\sigma_i) = D_i \nu_i$.

We will first show that, given $\sigma_i \neq \sigma_{i,2}$, then

$$(d_i(\sigma_{i,1}), \hat{d}_i(\sigma_{i,1})) \neq (d_i(\sigma_{i,2}), \hat{d}_i(\sigma_{i,2})).$$

(20)

Suppose by contradiction that there is $\sigma_{i,1} \neq \sigma_{i,2}$ such that

$$(d_i(\sigma_{i,1}), \hat{d}_i(\sigma_{i,1})) = (d_i(\sigma_{i,2}), \hat{d}_i(\sigma_{i,2})).$$

Then by (13), $c_i^{-1}(\nu_{i,1} + \lambda_{i,1}) = c_i^{-1}(\nu_{i,2} + \lambda_{i,2})$ and $D_i \nu_{i,1} = D_i \nu_{i,2}. \text{ Thus, } \nu_{i,1} = \nu_{i,2} = \nu \text{ and } c_i^{-1}(\nu + \lambda_{i,1}) = c_i^{-1}(\nu + \lambda_{i,2}). \text{ But since } c_i(\cdot) \text{ is strictly convex, } c_i' \text{ and its inverse are strictly increasing which implies that } \lambda_{i,1} = \lambda_{i,2} = \lambda. \text{ Contradiction.}$

Finally, let $\theta \in [0, 1]$ and consider any two $\sigma_i \neq \sigma_{i,2}$.

Then, $\Phi_i(\theta \sigma_{i,1} + (1 - \theta)\sigma_{i,2}) = \min_{d_i, \hat{d}_i} L_i(d_i, \hat{d}_i, \theta \sigma_{i,1} + (1 - \theta)\sigma_{i,2})$

$$= \min_{d_i, \hat{d}_i} \theta L_i(d_i, \hat{d}_i, \sigma_{i,1}) + (1 - \theta)L_i(d_i, \hat{d}_i, \sigma_{i,2})$$

$$> \theta \min_{d_i, \hat{d}_i} L_i(d_i, \hat{d}_i, \sigma_{i,1}) + (1 - \theta) \min_{d_i, \hat{d}_i} L_i(d_i, \hat{d}_i, \sigma_{i,2})$$

$$= \theta \Phi_i(\sigma_{i,1}) + (1 - \theta) \Phi_i(\sigma_{i,2})$$

where the strict inequality follows from (20). Thus, $\Phi_i(\sigma_i)$ is strictly concave and by using the definition of strict concavity we get $\Phi_S(\sigma)$ is strictly concave $\forall S \subseteq \mathcal{N}$. ■
Lemma 4 (OLC Optimality). Given a connected graph $G(N, E)$, then there exists a scalar $\nu^*$ such that $(d^*, \hat{d}^*, x^*, \sigma^*)$ is a primal-dual optimal solution of OLC and DOLC if and only if $(d^*, \hat{d}^*, x^*)$ is primal feasible (satisfies (9b)-(9e)), $\sigma^*$ is dual feasible (satisfies (12) and (18)),
\[
d_i^* = D_i \nu_i^*, \quad d_i^* = \nu_i^* - \lambda_i^*, \quad \nu_i^* = \nu^*, \quad i \in N, \tag{21}
\]
and
\[
\rho_{ij}^{d^+}(B_{ij}(\nu_i^* - \nu_j^*) - \hat{P}_{ij}) = 0, \quad ij \in E, \quad \tag{22a}
\]
\[
\rho_{ij}^{d^-}(P_{ij} - B_{ij}(\nu_i^* - \nu_j^*)) = 0, \quad ij \in E. \quad \tag{22b}
\]
Moreover, $d^*$, $\hat{d}^*$, $\nu^*$ and $\lambda^*$ are unique with $\nu^* = 0$.

**Proof:** Assumptions 1 and 2 guarantee that the solution to the primal (OLC) is finite. Moreover, since by Assumption 2 there is a feasible $d \in \text{int} D$, then the Slater condition is satisfied [31] and there is zero duality gap.

Thus, since OLC only has linear equality constraints, we can use Karush-Kuhn-Tucker (KKT) conditions [31] to characterize the primal dual optimal solution. Thus $(d^*, \hat{d}^*, x^*, \sigma^*)$ is primal dual optimal if and only if we have:
(i) Primal feasibility: (9b)-(9e)
(ii) Dual feasibility: (12) and (18)
(iii) Stationarity:
\[
\frac{\partial}{\partial d} L(d^*, \hat{d}^*, x^*, \sigma^*) = 0, \quad \frac{\partial}{\partial d} L(d^*, \hat{d}^*, x^*, \sigma^*) = 0
\]
and
\[
\frac{\partial}{\partial x} L(d^*, \hat{d}^*, x^*, \sigma^*) = 0.
\]
(iv) Complementary slackness:
\[
\rho_{ij}^{d^+}(B_{ij}(\nu_i^* - \nu_j^*) - \hat{P}_{ij}) = 0, \quad ij \in E
\]
and
\[
\rho_{ij}^{d^-}(P_{ij} - B_{ij}(\nu_i^* - \nu_j^*)) = 0, \quad ij \in E.
\]
KKT conditions (i), (ii) and (iv) are already implicit by assumptions of the lemma. Furthermore, since the graph $G$ is connected then (12a) is equivalent to
\[
\nu_i^* = \nu^* \quad \forall i \in N.
\]
which is the third condition of (21).

Now, using (11), Stationarity (iii) is equivalent to (ii) and
\[
\frac{\partial L}{\partial d_i}(d^*, \hat{d}^*, x^*, \sigma^*) = c_i(d_i^*) - (\nu_i^* + \lambda_i^*) = 0 \tag{23a}
\]
\[
\frac{\partial L}{\partial d_i}(d^*, \hat{d}^*, x^*, \sigma^*) = \hat{d}_i^* - \nu_i^* = 0 \tag{23b}
\]
which are the remaining conditions of (21).

Since $c_i(d_i)$ and $\frac{\partial c_i}{\partial d_i}$ are strictly convex functions, by the same argument in the proof of Lemma 3 we get that $\nu_i^*$ and $\lambda_i^*$ are unique. To show $\nu^* = 0$ we use (i). Adding (9b) over $i \in N$ gives
\[
0 = \sum_{i \in N} \left( P_{ii}^m - (d_i^* + \hat{d}_i^*) - \sum_{e \in E} C_{ie} P_e \right)
\]
\[
= \sum_{i \in N} \left( P_{ii}^m - (d_i^* + \hat{d}_i^*) \right) - \sum_{e = e \in E} \left( C_{ie} P_e + C_{je} P_e \right)
\]
\[
= \sum_{i \in N} \left( P_{ii}^m - (d_i^* + \hat{d}_i^*) \right) \quad \tag{24}
\]
and similarly (9c) gives
\[
0 = \sum_{i \in N} P_i^m - d_i^* \tag{25}
\]
Thus, subtracting (24) from (25) gives
\[
0 = \sum_{i \in N} \hat{d}_i^* = \sum_{i \in N} D_i \nu^* = \nu^* \sum_{i \in N} D_i
\]
and since $D_i > 0 \forall i \in N$, it follows that $\nu^* = 0$. \hfill \blacksquare

**IV. DISTRIBUTED OPTIMAL LOAD CONTROL**

We now show how to leverage the power network dynamics to solve the OLC problem in a distributed way. Our solution is based on the classical primal dual optimization algorithm that has been of great use to design congestion control mechanisms in communication networks [32]-[35].

Let
\[
L(x, \sigma) = \text{minimize}_{d, \hat{d}} L(d, \hat{d}, x, \sigma)
\]
\[
= L(d(\sigma), \hat{d}(\sigma), x, \sigma)
\]
\[
= \Phi(\sigma) - P^T C^T \nu
\]
\[
- \nu^T (L_B \lambda - CD_B \hat{C}^T \pi - CD_B (\rho^+ - \rho^-)) \tag{26}
\]
where $L(d, \hat{d}, x, \sigma)$ is defined as in (11), $d(\sigma) := (d_i(\sigma_i))$ and $\hat{d}(\sigma) := (d_i(\sigma_i))$ according to (21).

We then propose the following partial primal-dual law
\[
\dot{\nu}_i = \zeta_i^\nu (P_i^m - (d(\sigma) + D_i \nu_i) - C_i \hat{P}) \tag{27a}
\]
\[
0 = P_i^m - (d(\sigma) + D_i \nu_i) - C_i \hat{P} \tag{27b}
\]
\[
\dot{\lambda} = \zeta_i^\lambda (P_i^m - (\sigma - L_B v) \tag{27c}
\]
\[
\dot{\pi} = \zeta_i^\pi (C_D B C^T v - \hat{P}) \tag{27d}
\]
\[
\dot{\rho}^+ = \zeta_i^{\rho^+} (D_B C^T v + \hat{P}) \tag{27e}
\]
\[
\dot{\rho}^- = \zeta_i^{\rho^-} (P - D_B C^T v) \tag{27f}
\]
\[
\dot{\hat{P}} = \chi_i^P (C^T \nu) \tag{27g}
\]
\[
\dot{v} = \chi_i^v \left( L_B \lambda - C_D B \hat{C}^T \pi - C_D B (\rho^+ - \rho^-) \right) \tag{27h}
\]
where $\zeta_i^\nu = \text{diag}(\zeta_i^\nu)_{i \in G}$, $\zeta_i^\lambda = \text{diag}(\zeta_i^\lambda)_{i \in N}$, $\zeta_i^\pi = \text{diag}(\zeta_i^\pi)_{i \in E}$, $\zeta_i^{\rho^+} = \text{diag}(\zeta_i^{\rho^+})_{i \in E}$, $\zeta_i^{\rho^-} = \text{diag}(\zeta_i^{\rho^-})_{i \in E}$, $\chi_i^P = \text{diag}(\chi_i^P)_{i \in E}$ and $\chi_i^v = \text{diag}(\chi_i^v)_{i \in N}$.

The operator $[.]_+^*$ is a element-wise projection that maintains each element of the $u(t)$ within the positive orthant when $\hat{u} = [.]_+^*$, i.e., given any vector $a$ with same dimension as $u$ then $[a]_+^*$ is defined element-wise by
\[
[a]_+^* = \begin{cases} 
  a_e & \text{if } a_e > 0 \text{ or } u_e > 0, \\
  0 & \text{otherwise}. 
\end{cases}
\]

One property that will be used later is that given any constant vector $u^* \geq 0$, then
\[
(u - u^*)^T [a]_+^* \leq (u - u^*)^T a \tag{29}
\]
since for any pair $(u_e, a_e)$ that makes the projection active we must have by definition $u_e = 0$ and $a_e < 0$ and therefore
\[
(u_e - u_e^*) a_e = -u_e^* a_e \geq 0 = (u_e - u_e^*)^T [a]_+^*.
\]
The name of the dynamic law (27) comes from the fact that
\[
\frac{\partial}{\partial \nu} L(x, \sigma)^T = P^m - (d(\sigma) + D\nu) - CP \quad (30a)
\]
\[
\frac{\partial}{\partial \lambda} L(x, \sigma)^T = P^m - d(\sigma) - L_B v \quad (30b)
\]
\[
\frac{\partial}{\partial \pi} L(x, \sigma)^T = \tilde{C} D_B C^T v - \tilde{P} \quad (30c)
\]
\[
\frac{\partial}{\partial p^+} L(x, \sigma)^T = D_B C^T v - \bar{P} \quad (30d)
\]
\[
\frac{\partial}{\partial p^-} L(x, \sigma)^T = P - D_B C^T v \quad (30e)
\]
\[
\frac{\partial}{\partial v} L(x, \sigma)^T = -(C^T \nu) \quad (30f)
\]
\[
\frac{\partial}{\partial \nu} L(x, \sigma)^T = -(L_B \lambda - CD_B \dot{C} C^T \pi - CD_B (\rho^+ - \rho^-)) \quad (30g)
\]

Equations (27a), (27b) and (27g) show that dynamics (1) can be interpreted as a subset of the primal-dual dynamics described in (27) for the special case when \( \zeta' = M_i^{-1} \) and \( \chi_{ij}^\rho = B_{ij} \). Therefore, we can interpret the frequency \( \omega_i \) as the Lagrange multiplier \( \nu_i \).

This observation motivates us to propose a distributed load control scheme that is naturally decomposed into

**Power Network Dynamics:**
\[
\begin{align*}
\dot{\omega}_G &= M_G^{-1} (P_G^m - (d_G + \hat{d}_G) - C_G P) \quad (31a) \\
0 &= P_L^m - (d_L + \hat{d}_L) - C_L P \quad (31b) \\
\dot{\rho} &= D_B C^T \omega \quad (31c) \\
\hat{d} &= D \omega \quad (31d)
\end{align*}
\]

and

**Dynamic Load Control:**
\[
\begin{align*}
\dot{\lambda} &= \zeta^\lambda (P^m - d - L_B v) \quad (32a) \\
\dot{\pi} &= \zeta^\pi (C D_B C^T v - \bar{P}) \quad (32b) \\
\dot{\rho^+} &= \zeta^\rho^+ [D_B C^T v - \bar{P}]^+ \quad (32c) \\
\dot{\rho^-} &= \zeta^\rho^- [P - D_B C^T v]^+ \quad (32d) \\
\dot{\nu} &= \chi^\nu (L_B \lambda - CD_B \dot{C} C^T \pi - CD_B (\rho^+ - \rho^-)) \quad (32e) \\
\dot{d} &= e^{-1}(\omega + \lambda) \quad (32f)
\end{align*}
\]

Equations (31) and (32) show how the network dynamics can be complemented with dynamic load control such that the whole system amounts to a distributed primal-dual algorithm that tries to find a saddle point on \( L(x, \sigma) \). We will show in the next section that this system does achieve optimality as intended.

Figure 1 also shows the unusual control architecture derived from our OLC problem. Unlike traditional observer-based controller design architecture [36], our dynamic load control block does not try to estimate state of the network. Instead, it drives the network towards the desired state using a static feedback loop, i.e. \( d_i(\lambda_i + \omega_i) \).

**Remark 5.** One of the limitations of (32) is that in order to generate the Lagrange multipliers \( \lambda_i \), one needs to estimate \( P^m - d_i \), which is not easy since one cannot separate \( P^m \) from \( P^m - D_i \omega_i \) when one measures the power injection of a given bus without knowing \( D_i \). This problem will be addressed in Section VI where we propose a modified control scheme that can achieve the same equilibrium without needing to know \( D_i \) exactly.

V. OPTIMALITY AND CONVERGENCE

In this section we will show that the system (31)-(32) can efficiently rebalance supply and demand, restore the nominal frequency, and preserve inter-area flow schedules and thermal limits.

We will achieve this objective in two steps. Firstly, we will show that every equilibrium point of (31)-(32) is an optimal solution of (9). This guarantees that a stationary point of the system efficiently balances supply and demand and achieves zero frequency deviation.

Secondly, we will show that every trajectory \((\dot{d}(t), \hat{d}(t), P(t), v(t), \omega(t), \lambda(t), \pi(t), \rho^+(t), \rho^-(t))\) converges to an equilibrium point of (31)-(32). Moreover, the equilibrium point will satisfy (2) and (5).

**Theorem 6 (Optimality).** A point \( p^* = (d^*, \hat{d}^*, x^*, \sigma^*) \) is an equilibrium point of (31)-(32) if and only if it is a primal-dual optimal solution to the OLC problem.

**Proof:** The proof of this theorem is a direct application of Lemma 4. Let \((d^*, \hat{d}^*, x^*, \sigma^*)\) be an equilibrium point of (31)-(32). Then, by (31c) and (32c)-(32e), \( \sigma^* \) is dual feasible.

Similarly, since \( \omega_i = \hat{\omega}_i = 0, \hat{\lambda}_i = 0, \hat{\chi}_{ij}^p = 0 \) and \( \hat{\chi}_{ij}^- = 0 \), then (31a)-(31b) and (32a)-(32d) are equivalent to primal feasibility, i.e. \((d^*, \hat{d}^*, P^*, \nu^*)\) is a feasible point of (9). Finally, by definition of (31)-(32) conditions (21) and (22) are always satisfied by any equilibrium point. Thus we are under the conditions of Lemma 4 and therefore \( p^* = (d^*, \hat{d}^*, x^*, \sigma^*) \) is primal-dual optimal which also implies that \( \omega^* = 0 \).

**Remark 7.** Theorem 6 implies that every equilibrium solution of (31)-(32) is optimal with respect to OLC. However, it guarantees neither convergence to it nor that the line flows satisfy (2) and (5).

The rest of this section is devoted to showing that in fact for every initial condition \((P(0), v(0), \omega(0), \lambda(0), \pi(0), \rho^+=0, \rho^-=0)\).
\( \rho^+(0), \rho^-(0) \), the system (31)-(32) converges to one of such optimal solution. Furthermore, we will show that \( P(t) \) converges to a \( P^* \) that satisfies (2) and (5).

Since we showed in Section IV that (31)-(32) are just a special case of (27), we will provide our convergence result for (27). Our global convergence proof leverages the results of [37] on global convergence in network flow control. Unfortunately, the results presented there need further refinement due to the following reasons. Firstly, (27) is not a full primal-dual gradient law due to constraint (27b). Secondly, the dual function \( D(\sigma) \) is not strictly concave in all of its variables.

The next lemma shows that (27) can in fact be interpreted as a primal-dual gradient law with respect to a different Lagrangian.

**Lemma 8 (Primal-dual Gradient Law).** Let \( y = (\nu_L, \lambda, \pi, \rho)^2 \) and consider the reduced Lagrangian

\[
L(x, y) = \text{maximize}_{\nu_L} \ L(x, \sigma).
\]

Then, \( L(x, y) \) is concave in \( y \), convex in \( x \) and the dynamics (27) amount to

\[
\dot{y} = Y \left[ \frac{\partial}{\partial y} L(x, y)^T \right]^{\rho} \text{ and } \dot{x} = -X \frac{\partial}{\partial x} L(x, y)^T \tag{34}
\]

where the projection \([a]_P^+\) only acts in the \( \rho \) positions of \( a \), \( Y = \text{blockdiag}(\zeta^\mu, \zeta^\pi) \) and \( X = \text{blockdiag}(\lambda^\mu, \lambda^\pi) \).

Moreover, under Assumption 1, any saddle point \((x^*, y^*)\) of \( L(x, y) \) is unique in \( \nu_L \) and \( \lambda \).

**Proof:**

By Lemma 3 and (26), \( L(x, \sigma) \) is strictly concave in \( \sigma^x = (\nu, \lambda) \). Therefore, it follows that there exists a unique

\[
\nu^*_L(x, y) = \text{arg max}_{\nu_L} L(x, \sigma). \tag{35}
\]

Moreover, by stationarity, \( \nu^*_L(x, y) \) must satisfy

\[
\frac{\partial L}{\partial \nu_L}(x, y, \nu^*_L(x, y)) = 0 \tag{36a}
\]

\[
= \frac{\partial L}{\partial \nu_L}(y, \nu^*_L(x, y)) - (C_L P)^T = 0 \tag{36b}
\]

which is equivalent to (27b), i.e. \( \nu^*_L(x, y) \) implicitly satisfies (27b).

We now apply the envelope theorem [38] on (33) to compute \( \frac{\partial}{\partial \xi} L(x, y) \) and \( \frac{\partial}{\partial y} L(x, y) \) using

\[
\frac{\partial}{\partial x} L(x, y) = \frac{\partial L}{\partial x} \bigg|_{\nu^*_L(x,y)} \tag{37a}
\]

\[
\frac{\partial}{\partial y} L(x, y) = \frac{\partial L}{\partial y} \bigg|_{\nu^*_L(x,y)} \tag{37b}
\]

Using equation (37), (34) only differs from (27a) and (27c)-(27h) on the locations where \( \nu_L \) must be substituted with \( \nu^*_L(x, y) \). However, since there is a unique \( \nu^*_L \) that satisfies (27b), then it follows that (34) and (27) are equivalent representations of the same system.

Finally, given any saddle-point of \( (x^*, y^*) \), \( (d^*, d^*, x^*, y^*, \nu^*_L) \) is a saddle point of \( L(d, d, x, \sigma) \). Therefore, using Lemma 4 the uniqueness properties follow.

We will also use the following lemma.

**Lemma 9 (Differentiability of \( \nu^*_L(x, y) \)).** Given any \((x, y)\), the maximizer of (33), \( \nu^*_L(x, y) \), is continuously differentiable provided \( c_i(\cdot) \) is strongly convex. Furthermore, the derivative is given by

\[
\frac{\partial}{\partial x} \nu^*_L(x, y) = \begin{bmatrix} -\frac{D_L + d'_L}{C_L} \nu_L \\
\nu_L \frac{\lambda_L \lambda_E}{\nu_L} \end{bmatrix} \tag{38}
\]

\[
\frac{\partial}{\partial y} \nu^*_L(x, y) = \begin{bmatrix} \nu_L \frac{\lambda_L}{\nu_L} \\
\nu_L \frac{\lambda_E}{\nu_L} \\
0 \end{bmatrix} \tag{39}
\]

where \( \nu^*_L \) is defined as:

\[
d_S = \begin{cases} \text{diag}(d'_L(\lambda_i + \nu^*_L(x, y)))_{i \in S} & \text{if } S \subseteq \mathcal{L} \\
\text{diag}(d'_L(\lambda_i + \nu^*_L(x, y)))_{i \in \mathcal{S}} & \text{if } S \subseteq \mathcal{G} \end{cases}
\]

**Proof:** We first notice that \( \nu^*_L(x, y), i \in \mathcal{L} \), depends only on \( \lambda_i \) and \( C_i := \sum_{e \in \mathcal{E}} C_i \). Which means that

\[
\frac{\partial \nu^*_L(x, y)}{\partial \lambda_i} = 0, \quad \frac{\partial \nu^*_L(x, y)}{\partial C_i} = 0, \quad \frac{\partial \nu^*_L(x, y)}{\partial \nu_L} = 0
\]

Now, by definition of \( \nu^*_L(x, y) \), for any \( i \in \mathcal{L} \) we have

\[
0 = \frac{\partial}{\partial \nu_L} \begin{bmatrix} L(x, y, \nu^*_L(x, y)) \\
\nu^*_L(x, y) \end{bmatrix} = \begin{bmatrix} P_i^m - (D_i \nu^*_L(x, y) + d_i \lambda_i + \nu^*_L(x, y)) - \sum_{e \in \mathcal{E}} C_{i,e} P_e \end{bmatrix} \tag{40}
\]

Therefore, if we fix \( P \) and take the total derivative of \( \frac{\partial}{\partial \nu_L} L(x, y, \nu^*_L(x, y)) \) with respect to \( \lambda_i \) we obtain

\[
0 = \frac{d}{d \lambda_i} \left( \frac{\partial}{\partial \nu_L} L(x, y, \nu^*_L(x, y)) \right) \tag{41}
\]

\[
= -(D_i + d'_L(\lambda_i + \nu^*_L(x, y))) \frac{\partial}{\partial \nu^*_L(x, y)} \tag{42}
\]

which here we used \( \nu^*_L(x, y) \) for short of \( \nu^*_L(x, y) \).

Now since by assumption \( c_i(\cdot) \) is strongly convex, i.e. \( c_i''(\cdot) \geq \alpha > 0 \), \( d'_L(\cdot) = \frac{1}{\alpha(\nu^*_L)} \leq \frac{1}{\alpha} < \infty \). Thus, \( D_i + d'_L(\lambda_i + \nu^*_L(x, y)) \) is finite and strictly positive, which implies that

\[
\frac{\partial}{\partial \lambda_i} \nu^*_L(x, y) = -\frac{d'_L(\lambda_i + \nu^*_L(x, y))}{(D_i + d'_L(\lambda_i + \nu^*_L(x, y)))}, \quad i \in \mathcal{L} \tag{43}
\]

Similarly, we obtain

\[
\frac{\partial}{\partial \nu_L} \nu^*_L(x, y) = \frac{1}{(D_i + d'_L(\lambda_i + \nu^*_L(x, y)))}, \quad i \in \mathcal{L} \tag{44}
\]

where \( C_i \) is the \( i \)th row of \( C \).

Finally, notice that whenever \( d'_L(\lambda_i + \nu^*_L(x, y)) \) exists, then \( \frac{\partial}{\partial \lambda_i} \nu^*_L \) and \( \frac{\partial}{\partial \nu_L} \nu^*_L \) also exists. \[\square\]

We now present our main convergence result. Let \( E \) be the set of equilibrium points of (27)

\[
E := \{(x, \sigma) : \frac{\partial L}{\partial x}(x, \sigma) = 0, \frac{\partial L}{\partial \sigma}(x, \sigma) = 0\}
\]

which by Theorem 6 is the set of optimal solutions of the OLC problem.
Now, differentiating with respect to time gives

\[ \frac{\partial}{\partial t} L(x, y^\gamma) = 0 \]

for all \( x \), then, the last calculation does not bring any additional information. This is the case of \( \frac{\partial}{\partial t} L(x, y^\gamma) \); see equation (30g). However, it is not the case for \( \frac{\partial}{\partial t} L(x, y^\gamma) \) and therefore, since \( \nu_\gamma \equiv 0 \), \( \frac{\partial}{\partial t} L(x, y^\gamma) = 0 \) and (30) implies that \( x(t) \) must satisfy \( C_L \nu_\gamma(x(t), y^\gamma) = 0 \) which implies that either \( \nu_\gamma(x(t), y^\gamma) = 0 \), when \( C_L \) is full row rank, or \( \nu_\gamma(x(t), y^\gamma) = 1 \alpha(t) \) where \( \alpha(t) \) is a time-varying scalar, when \( L = N \).

We now show that when \( L = N \) we also get \( \nu_\gamma(x(t), y^\gamma) = \hat{\nu}_\gamma \) for some constant vector \( \nu_\gamma \). Differentiating \( \nu_\gamma(x(t), y(t)) = 1 \alpha(t) \) with respect to time and using (38) we obtain

\[ (D_L + d_L')^{-1} C_L \dot{P}(t) = \dot{\alpha}(t) \]

which after left multiplying by \( 1^T (D_L + d_L') \) gives

\[ 1^T (D_L + d_L') 1 \alpha(t) = 0 \implies \dot{\alpha}(t) = 0. \]

Thus, in either case we obtain

\[ \nu_\gamma(x(t), y(t)) = \nu_\gamma(C_L P(t), \lambda_\gamma) = \hat{\nu}_\gamma \]

which implies that \( C_L P(t) \equiv C_L \hat{P} \) for some constant vector \( \hat{P} \).

Therefore, it follows that \( \nu_\gamma(x(t), y(t)) \) must satisfy

\[ \nu_\gamma(x(t), y(t)) = \nu_\gamma(\hat{x}, y(t)) \]

for some constant vector \( \hat{x} \).

Now, using (27b) with (51) we get

\[ P_{\nu_\gamma}^* - D_L \nu_\gamma(x(t), y(t)) = d_L \nu_\gamma(x(t), y(t)) + \lambda_\gamma(t) - C_L \hat{P} \equiv 0. \]

(52)

A similar argument using the fact that \( L(x^*, y) \equiv L(x^*, y^\gamma) \) gives

\[ \frac{\partial}{\partial y} L(x^*, y^\gamma) \left[ \frac{\partial}{\partial y} L(x^*, y^\gamma)^T \right]^+ \equiv 0. \]

(53)

Since the projection \( [.]^+ \) only acts on the \( \rho \) positions, (53) implies \( \frac{\partial}{\partial y} L(x^*, y^\gamma) \equiv 0 \), \( \frac{\partial}{\partial y} L(x^*, y^\gamma) \equiv 0 \) and \( \frac{\partial}{\partial y} L(x^*, y^\gamma) \equiv 0 \).

Now \( \frac{\partial}{\partial y} L(x^*, y^\gamma) \equiv 0 \) together with equation (30a) implies that

\[ P_{\nu_\gamma}^* - D_\nu L\nu_\gamma(t) - d_\nu L\nu_\gamma(t) + \lambda_\gamma(t) - C_\nu P^*_\nu - 0, \]

(54)

and \( \frac{\partial}{\partial y} L(x^*, y^\gamma) = 0 \) with (30b) implies

\[ P_{\nu_\gamma}^* - d_\nu L\nu_\gamma(t) + \lambda_\gamma(t) - C_\nu P^*_\nu \equiv 0, \]

(55)

\[ P_{\nu_\gamma}^* - d_\nu L\nu_\gamma(t) + \lambda_\gamma(t) - C_\nu P^*_\nu \equiv 0. \]

(56)

Using (54) and (55) together with the fact that \( d_\nu(t) \) is strictly increasing, we get \( \nu_\gamma(t) = \hat{\nu}_\gamma \) and \( \lambda_\gamma(t) \equiv \lambda(\gamma) \). Moreover, since \( P^* \) is primal optimal, Lemma 8 and Theorem 6 imply that \( \nu_\gamma(t) \equiv 0 \) and \( \lambda_\gamma(t) = \lambda_\gamma^* \). Finally, using now (52) together with (56), the same argumentation gives \( \nu_\gamma(x(t), y(t)) \equiv \hat{\nu}_\gamma \) and \( \lambda_\gamma(t) \equiv \lambda_\gamma^* \).

We have obtained so far that \( \nu(t) \) and \( \lambda(t) \) are constant.

Now, since \( \lambda \equiv 0 \), it follows from (27c) that \( C^T \nu(t) \equiv C^T \hat{\nu} \) for some constant vector \( \hat{\nu} \) or equivalently \( \nu(t) \equiv \hat{\nu} + \beta(t) \).

Differentiating in time \( 1^T (x^\gamma)^{-1} \nu(t) \) gives \( 0 \equiv 1^T (x^\gamma)^{-1} \dot{\nu} \equiv (\sum_{i \in N} \frac{\partial}{\partial s_i}) \beta \) which implies that \( \beta(t) \equiv \beta^* \).
Suppose now that either $\dot{P} \neq 0$ or $\pi \neq 0$. Since $C^T \nu(t) = C^T \hat{v}$ and $\nu(t) \equiv \hat{v}$, $P$ and $\pi$ are constant. Thus, since the trajectories are bounded, we must have $\dot{P} \equiv 0$ and $\pi \equiv 0$; otherwise $U(x,y)$ will grow unbounded (contradiction).

It remains to show that $\rho = 0$, i.e. $\rho^+ = \rho^- = 0$. Since $v(t) \equiv \hat{v}$, then the argument inside (27e) and (27f) is constant.

Now consider any $\rho^+_e$, $e \in E$. Then we have three cases: (i) $a_e(\hat{v}_i - \hat{v}_j) - \hat{P}_e > 0$, (ii) $a_e(\hat{v}_i - \hat{v}_j) - \hat{P}_e < 0$ and (iii) $a_e(\hat{v}_i - \hat{v}_j) - \hat{P}_e = 0$. Case (i) implies $\rho^+_e(t) \to +\infty$ which cannot happen since the trajectories are bounded. Case (ii) implies that $\rho^+_e(t) \equiv 0$ which implies that $\rho^+_e \equiv 0$, and case (iii) also implies $\rho^+_e \equiv 0$. An analogous argument gives $\rho^- \equiv 0$.

Thus, we have shown that $M \subseteq E$. Unfortunately, since there is an affine space of equilibria $(x^*, \sigma^*)$ in $E$, the invariance principle for hybrid automata [39] does not guarantee that $(x(t), \sigma(t))$ will converge to one specific $(x^*, \sigma^*)$ value.

Fortunately, we can use structure of $U(x, y)$ as in [23] to achieve convergence to a single equilibrium. Since $(x(t), \sigma(t)) \to M$ and $(x(t), \sigma(t))$ are bounded, then there exists an infinite sequence of time values $\{t_k\}$ such that $(x(t_k), \sigma(t_k)) \to (\hat{x}^*, \hat{y}^*, \hat{v}^*_e) \in M$. Thus, using this specific equilibrium $(\hat{x}^*, \hat{y}^*)$ in the definition of $U(x, y)$, it follows that $U(x(t_k), y(t_k)) \to 0$, which by continuity of $U(x, y)$ implies that $(x(t), y(t)) \to (\hat{x}^*, \hat{y}^*)$.

Thus, it follows that $(x(t), \sigma(t))$ converges to only one optimal solution within $M \subseteq E$.

**Remark 11.** Theorem 10 uses the version the Invariance Principle for hybrid automata presented in [39] which requires the hybrid automaton to be deterministic, nonblocking and continuous with respect to initial conditions (see [39, Theorem IV.11]). While we neither define nor discuss these assumptions here, we stress that our system does satisfy all three of them. In particular, continuity follows from [40] and the fact that the dynamics (without the projection) are Lipschitz.

Finally, the following theorem shows that under mild conditions the system is able to restore the inter-area flows (2) and maintain the line flows within the thermal limits (5).

**Theorem 12** (Inter-area Constraints and Thermal Limits). Given any primal-dual optimal solution $(x^*, \sigma^*) \in E$, the optimal line flow vector $P^*$ satisfies (2). Furthermore, if $P(0) = D’B^T \theta^0$, then $P_{ij}^* = B_{ij}(v^*_i - v^*_j)$ and therefore (5) holds.

**Remark 13.** Theorem 12 shows that the system (31)-(32) can impose the constraints (2) and (5) for $P$ by constraining $v$. In other words, the constraints on $v$ indirectly guarantee (2) and (5).

**Proof:** By optimality, $P^*$ and $v^*$ must satisfy

$$P^m - d^* = CP^* = L_B v^* = CDB^T \nu^*$$

Therefore using primal feasibility, (3) and (57) we have

$$\dot{P} = \dot{\chi}DB^T \nu^* = E_k CD_B^T \nu^*$$

which is exactly (4).

Finally, to show that $P_{ij}^* = B_{ij}(v^*_i - v^*_j)$ we will use (31c). Integrating (31c) over time gives

$$P(t) - P(0) = \int_0^t DB^T \nu(s) ds.$$  

Therefore, since $P(t) \to P^*$, we have

$$P^* = P(0) + DB^T \theta^*$$

where $\theta^*$ is any finite vector satisfying $C^T \theta^* = \int_0^\infty C^T \nu(s) ds$.

Again by primal feasibility

$$CD_B^T \nu^* = L_B v^* = CP^* = C(P(0) + DB^T \theta^*) = CD_B^T (\theta^0 + \theta^*).$$

Thus, we must have $v^* = (\theta^0 + \theta^*) + \alpha1$ and it follows then that $P^* = DB^T (\theta^0 + \theta^*) = DB^T (v^* - \alpha1) = DB^T v^*$. Therefore, since by primal feasibility $P \leq DB^T v^* \leq \bar{P}$, then $\lim P(t) \leq P^*$.

**Examples 14.** The assumption of Theorem 12 of having $P(0) = DB^T \theta^0$ is equivalent to substituting (31) with

$$\omega_\theta = M_\omega^{-1}(P_m^D - (d_G - D_G \omega_\theta)) - C_G DB^T \theta$$

$$0 = P_m^L - (d_L + C_L \omega_L) - C_L DB^T \theta$$

$$\dot{\theta} = \omega$$

which is the linearization of the power network using phases instead of line flows as states. Therefore, this assumption can be done without loss of generality.

**VI. CONVERGENCE UNDER UNCERTAINTY**

In this section we discuss an important aspect of the implementation of the control law (32). We provide a modified control law that solves the problem raised in Remark 5, i.e. that does not require knowledge of $D_i$. We show that the new control law still converges to the same equilibrium provided the estimation error of $D_i$ is small enough (c.f. (71)).

We propose an alternative mechanism to compute $\lambda_i$. Instead of (32), we consider the following control law:

**Dynamic Load Control (2):**

$$\dot{\lambda} = \lambda \chi^e (M_\omega + B_\omega + CP - L_B v)$$

$$\dot{\pi} = \pi \chi^e (\tilde{C} DB^T \nu - \hat{P})$$

$$\dot{\rho}^+ = \rho^+ [DB^T \nu \pi^+ - \rho^+]$$

$$\dot{\rho}^- = \rho^- [P - DB^T \nu \rho^-]$$

$$\dot{v} = \chi^e (L_B \lambda - CD_B \pi - CD_B (\rho^+ - \rho^-))$$

$$\dot{d} = e^{-1} (\omega + \lambda)$$

where $M = \text{diag}(M_i)_{i \in N}$ with $M_i = 0$ for $i \in \mathcal{L}$, and $B = \text{diag}(b_i)_{i \in N}$.

Notice that the only difference between (32) and (59) is that we substitute (32a) with (59a) where now we only need to estimate $M_i$ for the generators; which is usually known.

The parameter $b_i$ plays the role of $D_i$. In fact, whenever $b_i = D_i$ then one can use (31a) and (31b) to show that (59a)
Lemma 9 holds, then we have
\[
\dot{\lambda} = \zeta^\lambda(M\dot{\omega} + B\dot{\omega} + CP - LBv)
= \zeta^\lambda(M\dot{\omega} + D\dot{\omega} + \delta B\omega + CP - LBv)
= \zeta^\lambda(P^m - d - CP + \delta B\omega + CP - LBv).
\]
(60)
which is equal to (32a) when \(\delta b_i = 0\).

Using (60), we can express the system (31) and (59) by
\[
\dot{x} = -X \frac{\partial}{\partial x} L(x, y)^T
\]
\[
y = Y \left[ \frac{\partial}{\partial y} L(x, y)^T + g(x, y) \right] + \rho
\]
where
\[
g(x, y) := \begin{bmatrix} 0 \\ \delta B\nu^\varphi \\ \lambda C \\ \lambda C \\ \pi \\ \rho \end{bmatrix}
\]
with matrix \(\delta B_S := \text{diag}(\delta b_i)_{i \in S}\).

The main result of this section shows that under certain conditions on \(b_i\), convergence to the optimal solution is preserved despite the fact that (31) with (59) is no longer a primal dual algorithm. The basic intuition behind this result is that when one uses \(b_i\) instead of \(D_i\), the system dynamics are no longer a primal-dual law, yet provided \(b_i\) does not distant too much form \(D_i\), the convergence properties are preserved.

To show this result, we provide a novel convergence proof that make use of the following lemmas.

**Lemma 15** (Second order derivatives of \(L(x, y)\)). Whenever Lemma 9 holds, then we have
\[
\frac{\partial^2 L}{\partial x^2}(x, y) = \begin{bmatrix} C^T_x(D_L + d_L^2)^{-1}C_L & 0 \\ 0 & 0 \end{bmatrix}
\]
(63)
and
\[
\frac{\partial^2 L}{\partial y^2}(x, y) = -\begin{bmatrix} (D^*_L + d^*_L)^{-1} & 0 & 0 & 0 \\ 0 & 0 & D_L(D_L + d_L^2)^{-1}d_L^2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}
\]
(64)
with \(\frac{\partial^2 L}{\partial y^2} L(x, y) \geq 0\) and \(\frac{\partial^2 L}{\partial x^2} L(x, y) \leq 0\).

**Proof:** Using the Envelope Theorem [38] in (33) we have
\[
\frac{\partial L}{\partial x}(x, y) = \frac{\partial L}{\partial x}(x, y, \nu^\varphi(x, y))
\]
which implies that
\[
\frac{\partial^2 L}{\partial x^2}(x, y) = \frac{\partial L}{\partial x} \left[ \frac{\partial L}{\partial x}(x, y, \nu^\varphi(x, y)) \right]
= \frac{\partial^2 L}{\partial x\partial \nu^\varphi}(x, y) + \frac{\partial^2 L}{\partial x\partial \nu_L}(x, y, \nu^\varphi(x, y)) \frac{\partial \nu^\varphi(x, y)}{\partial \nu_L}
= \frac{\partial^2 L}{\partial x\partial \nu_L}(x, y, \nu^\varphi(x, y)) \frac{\partial \nu^\varphi(x, y)}{\partial \nu_L}.
\]
(65)
where the last step follows from \(L(x, \sigma)\) being linear in \(x\).

Now, by definition of \(\nu^\varphi(x, y)\) it follows that
\[
\frac{\partial L}{\partial \nu_L}(x, y, \nu^\varphi(x, y)) = 0.
\]
(66)

Differentiating (66) with respect to \(x\) gives
\[
0 = \frac{\partial^2 L}{\partial \nu_L \partial x}(x, y, \nu^\varphi(x, y)) + \frac{\partial^2 L}{\partial \nu_L \partial \nu^\varphi}(x, y, \nu^\varphi(x, y)) \frac{\partial \nu^\varphi(x, y)}{\partial \nu_L}
\]
and therefore
\[
\frac{\partial^2 L}{\partial x \partial \nu^\varphi}(x, y) = \left[ \frac{\partial^2 L}{\partial \nu_L \partial x}(x, y, \nu^\varphi(x, y)) \right]^T
= -\frac{\partial \nu^\varphi(x, y)}{\partial x} \frac{\partial^2 L}{\partial \nu^\varphi \partial \nu_L}(x, y, \nu^\varphi(x, y)).
\]
(67)
Substituting (67) into (65) gives
\[
\frac{\partial^2 L}{\partial x^2}(x, y) = -\frac{\partial \nu^\varphi(x, y)}{\partial x} \frac{\partial^2 L}{\partial x \partial \nu^\varphi}(x, y) \frac{\partial \nu^\varphi(x, y)}{\partial \nu_L} \frac{\partial \nu^\varphi(x, y)}{\partial \nu_L}.
\]
(68)

It follows from (36) and (16) that
\[
\frac{\partial^2 L}{\partial \nu^\varphi}(x, y, \nu^\varphi(x, y)) = \frac{\partial^2 \Phi^\varphi}{\partial \nu^\varphi}(\nu^\varphi(x, y), \lambda_L)
= -(D_L + d_L^2).
\]
(69)
Therefore, substituting (38) and (69) into (68) gives (63).

A similar calculation using (39) gives (64).

**Lemma 16** (Partial derivatives of \(g(x, y)\)). Whenever Lemma 9 holds, then
\[
\frac{\partial g(x, y)}{\partial x} = \begin{bmatrix} P & v \\ 0 & 0 \end{bmatrix}
\]
\[
\frac{\partial g(x, y)}{\partial y} = \begin{bmatrix} \nu^\varphi & \lambda^\varphi & \lambda C & \pi & \rho \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}
\]
Proof: By definition of $g(x, y)$ we have
\[
\frac{\partial}{\partial y} g(x, y) = \frac{\partial}{\partial y} \left[ \begin{array}{c}
0 \\
\delta B_C \nu_G \\
\delta B_C \nu_G^2(x, y) \\
U(x, y) = 0 \\
\frac{\partial}{\partial y} \end{array} \right] = \left[ \begin{array}{c}
\frac{\partial}{\partial y} 0 \\
\delta B_C \frac{\partial}{\partial y} \nu_G \\
\delta B_C \frac{\partial}{\partial y} \nu_G^2(x, y) \\
U(x, y) = 0 \\
\frac{\partial}{\partial y} \end{array} \right].
\]

Thus, using Lemma 9 we obtain \( \frac{\partial}{\partial y} g(x, y) \). A similar computation gives \( \frac{\partial}{\partial x} g(x, y) \).

Unfortunately, the conditions of Theorem 10 will not suffice to guarantee convergence of the perturbed system. The main difficulty is that \( d_i'(\lambda_i + \nu_i) > 0 \) can become arbitrarily close to zero and therefore the block of the Jacobian of (34) corresponding to load bus states can become arbitrarily close to singular.

Thus, we require the following additional assumption.

**Assumption 3** (Lipschitz continuity of \( c'_i(\cdot) \)). The marginal cost functions \( c'_i(\cdot) \) are Lipschitz continuous with Lipschitz constant \( L > 0 \). Notice that assumption 3 implies that the domain of \( D_i = \mathbb{R} \) in Assumption 1. However, one can always design a cost function \( c_i \) such that the optimal solution of (9) lies within predefined bounds \( d_i^l \leq d_i^* \leq d_i^u \), provided that Assumption 2 is satisfied for such bounds.

Using now assumptions 1 and 3 we can show that \( \alpha \leq c''_i \leq L \) which implies
\[
d_i' := \frac{1}{L} \leq d_i' = \frac{1}{c''_i} \leq \tilde{d} := \frac{1}{\alpha}. \tag{70}
\]

**Theorem 17** (Global convergence of perturbed system). Under assumptions 1, 2 and 3 hold. The system (61) converges to a point in the optimal set \( E \) for every initial condition whenever
\[
\delta b_i \in 2( \delta \tilde{d} - \sqrt{\delta^2 + \delta^2 D_{\text{min}}, \delta \tilde{d}' + \sqrt{\delta^2 + \delta^2 D_{\text{min}}}} ). \tag{71}
\]

Proof: When \( \delta B = 0 \), (61) reduces to (34) which we have shown to converge using the function (44) as a Lyapunov function. Here, we will show that under the assumptions of this theorem, \( U \) is still a Lyapunov function of (61).

Let \( z = (x, y) \) and define \( f(z) \) to be
\[
f(z) = \left[ \begin{array}{c}
- \frac{\partial}{\partial y} L(x, y) \\
\frac{\partial}{\partial y} L(x, y) \end{array} \right] T + g(x, y).
\]

Thus, using this notation, (61) becomes
\[
\dot{z} = Z[f(z)]^+ . \tag{72}
\]

We can therefore rewrite (44) as
\[
U(z) = \frac{1}{2} (z - z^*)^T Z^{-1} (z - z^*) \tag{73}
\]

and recompute \( \dot{U}(z) \) in a different way
\[
\dot{U}(z) = \frac{1}{2} ((z - z^*)^T [f(z)]^+ + [f(z)]^T (z - z^*)) \tag{74}
\]
\[
\leq \frac{1}{2} ((z - z^*)^T f(z) + f(z)^T (z - z^*)) = (z - z^*)^T f(z) \tag{75}
\]
\[
\leq \int_0^1 (z - z^*)^T \left[ \frac{\partial}{\partial z} f(z(s)) \right] (z - z^*) ds + (z - z^*)^T f(z) \tag{76}
\]
\[
\leq \int_0^1 (z - z^*)^T \left[ \frac{\partial}{\partial z} f(z(s)) \right] + (z - z^*)^T f(z) \tag{77}
\]
\[
= \int_0^1 (z - z^*)^T H(z(s))(z - z^*) ds \tag{78}
\]
where (74) follows from (72), (75) from (29), and (76) form the fact that
\[
f(z) - f(z^*) = \int_0^1 \frac{\partial}{\partial z} f(z(s))(z - z^*) ds,
\]
where the vector
\[
z(s) = (x(s), y(s)) := (z - z^*) s + z^*
\]
and
\[
\frac{\partial}{\partial z} f(z) = \left[ \begin{array}{c}
- \frac{\partial}{\partial x} L(x, y) \\
\frac{\partial}{\partial y} L(x, y) \end{array} \right] T + \left[ \begin{array}{c}
0 \\
\frac{\partial}{\partial y} g(x, y)
\end{array} \right].
\]

Finally, (77) follows from the fact that either \( f_i(z^*) = 0 \), or \( z_i - z_i^* = z_i \geq 0 \) and \( f_i(z^*) < 0 \), which implies \( (z - z^*)^T f(z^*) \leq 0 \).

Therefore, \( H(z) \) in (78) can be expressed as
\[
H(z) = \frac{1}{2} \left[ \frac{\partial}{\partial z} f(z)^T + \frac{\partial}{\partial z} f(z) \right]
\]
\[
= \left[ \begin{array}{c}
- \frac{\partial}{\partial x} L(x, y) \\
\frac{\partial}{\partial y} L(x, y)
\end{array} \right] T + \left[ \begin{array}{c}
0 \\
\frac{\partial}{\partial y} g(x, y) \end{array} \right].
\]

We will now show that, under the assumptions of the theorem, \( H(z) \) is negative semidefinite for all \( z \).

Using lemmas 15 and 16, \( H(z) \) becomes equal to (79). Thus, we can see that the perturbation in the primal-dual dynamics affects the coupling between \( P \) and \( \lambda_C \) as well as \( \nu_G \) and \( \lambda_G \). Moreover, we can show that \( H(z) \) is negative semidefinite provided that the matrices
\[
H_{P, \lambda_C}(z) = \left[ \begin{array}{cc}
- C_T (D_C + d_C^T)^{-1} C_L & - \frac{1}{2} C_T (D_C + d_C^T)^{-1} \delta B_C \\
- \frac{1}{2} \delta B_C (D_C + d_C^T)^{-1} C_L & - (D_C + \delta B_C)(D_C + d_C^T)^{-1} d_C^T
\end{array} \right]
\]
and
\[
H_{\nu_G, \lambda_C}(z) = \left[ \begin{array}{cc}
- (d_G' + D_G) \frac{1}{2} \delta B_G - d_G' & - \frac{1}{2} \delta B_G - d_G' \\
\frac{1}{2} \delta B_G - d_G' & -d_G'
\end{array} \right]
\]
are negative semidefinite.
A simple algebraic manipulation makes $H_{P,\lambda}(z)$ equal to

$$H_{P,\lambda}(z) = \tilde{C}^T \tilde{D} \tilde{H} \tilde{D}^T \tilde{C} \tag{80}$$

where

$$\tilde{C} = \begin{bmatrix} C_L & 0 \\ 0 & I \end{bmatrix}, \quad \tilde{D} = \begin{bmatrix} (D_L + d'_L)^{-1} & 0 \\ 0 & (D_L + d'_L)^{-1} \end{bmatrix}$$

and

$$\tilde{H} = \begin{bmatrix} -I & -\frac{1}{2} \delta B_L \\ -\frac{1}{2} \delta B_L & -(D_L + \delta B_L)d'_L \end{bmatrix}$$

Notice that $\tilde{D} \succ 0$ implies that $\tilde{D}^T$ exists. Thus, a sufficient condition for $H_{P,\lambda}(z) \preceq 0$ is $\tilde{H} \preceq 0$.

We will require however that

$$\tilde{H} \prec 0, \tag{81}$$

which can only happen if the polynomials

$$p_i(\mu_i) = (\mu_i + 1)(\mu_i + (D_i + \delta b_i)d'_i) - \frac{\delta b_i^2}{4}$$

$$= \mu_i^2 + (1 + (D_i + \delta b_i)d'_i)\mu_i + (D_i + \delta b_i)d'_i - \frac{\delta b_i^2}{4}, \quad i \in \mathcal{L}$$

have roots with negative real parts. Thus, applying Ruth-Hurwitz stability we get the following necessary and sufficient condition:

$$\delta b_i^2 - 4(D_i + \delta b_i)d'_i < 0 \tag{82a}$$

and

$$1 + (D_i + \delta b_i)d'_i > 0 \tag{82b}$$

for every $i \in \mathcal{L}$. Thus provided (82) holds, we can guarantee that $H_{P,\lambda}(z) \preceq 0$.

Similarly, we can show that all the eigenvalues of $H_{\nu G,\lambda}(z)$ are the roots of the polynomials

$$p_i(\mu_i) = (\mu_i + D_i + d'_i)(\mu_i + d'_i) - \frac{(\delta b_i}{2} - d'_i)^2$$

$$= \mu_i^2 + (D_i + 2d'_i)\mu_i + (D_i + \delta b_i)d'_i - \frac{\delta b_i^2}{4}$$

which, since $D_i + 2d'_i > 0$, have negative real part if and only if (82a) is satisfied for each $i \in G$. Therefore, (82a) guarantees that $H_{\nu G,\lambda} < 0$.

Now, equation (82a) can be equivalently rewritten as:

$$2(d'_i - \sqrt{d'_i(d'_i + D_i)}) < \delta b_i < 2(d'_i + \sqrt{d'_i(d'_i + D_i)}). \tag{83}$$

Since $d'_i \in [d'_i, \bar{d}_i]$, $D_i \geq D_{\text{min}}$ and the function

$$x - \sqrt{x(x + y)}$$

is decreasing in both $x$ and $y$ for $x, y \geq 0$, then

$$2(d'_i - \sqrt{d'_i(d'_i + D_i)}) \leq 2(d'_i - \sqrt{d'_i(d'_i + D_{\text{min}})}).$$

Similarly, since $x + \sqrt{x(x + y)}$ is increasing for $x, y \geq 0$,

$$2(d'_i + \sqrt{d'_i(d'_i + D_i)}) \geq 2(d'_i + \sqrt{d'_i(d'_i + D_{\text{min}})}).$$

Therefore, (82a) holds whenever $\delta b_i$ satisfies (71). Finally, (82b) holds whenever $\delta b_i > -\frac{1}{d'_i} - D_i$ which in particular holds if

$$\delta b_i > -D_{\text{min}}.$$
so that

\[ i \omega_i + \lambda_i \]

\[ i \]

VII. NUMERICAL ILLUSTRATIONS

We now illustrate the behavior of our control scheme. We consider the widely used IEEE 39 bus system, shown in Figure 2, to test our scheme. We assume that the system has two independent control areas that are connected through lines (1, 2), (2, 3) and (26, 27). The network parameters as well as the initial stationary point (pre fault state) were obtained from the Power System Toolbox [41] data set.

Each bus is assumed to have a controllable load with \( D_i = [-d_{max}, d_{max}] \), with \( d_{max} = 1 \) p.u. on a 100MVA base and disutility function

\[
c_i(d_i) = \int_0^{d_i} \tan\left(\frac{\pi}{2d_{max}} s\right) ds = -\frac{2d_{max}}{\pi} \ln(|\cos(\frac{\pi}{2d_{max}} d_i)|).
\]

Thus, \( d_i(\sigma_i) = \frac{1}{\pi} (\omega_i + \lambda_i) = \frac{2d_{max}}{\pi} \arctan(\omega_i + \lambda_i) \). See Figure 3 for an illustration of both \( c_i(d_i) \) and \( d_i(\sigma_i) \).

![Fig. 2: IEEE 39 bus system: New England](image)

![Fig. 3: Disutility \( c_i(d_i) \) and load function \( d_i(\omega_i + \lambda_i) \)](image)

Throughout the simulations we assume that the aggregate generator damping and load frequency sensitivity parameter \( D_i = 0.2 \) \( \forall i \in \mathcal{N} \) and \( \chi_i^e = \zeta_i^e = \zeta_i^m = \zeta_i^e \) \( \forall i \in \mathcal{N}, k \in \mathcal{K} \) and \( e \in \mathcal{E} \). These parameter values do not affect convergence, but in general they will affect the convergence rate. The values of \( P^m \) are corrected so that they initially add up to zero by evenly distributing the mismatch among the load buses. \( \tilde{P} \) is obtained from the starting stationary condition. We initially set \( \tilde{P} \) and \( P \) so that they are not binding.

We simulate the OLC-system as well as the swing dynamics (31) without load control (\( d_i = 0 \)), after introducing a perturbation at bus 29 of \( P_{29} = -2 \) p.u.. Figures 4 and 5 show the evolution of the bus frequencies for the uncontrolled swing dynamics (a), the OLC system without inter-area constraints (b), and the OLC with area constraints (c).

It can be seen that while the swing dynamics alone fail to recover the nominal frequency, the OLC controllers can jointly rebalance the power as well as recovering the nominal frequency. The convergence of OLC seems to be similar or even better than the swing dynamics, as shown in Figures 4 and 5.

![Fig. 4: Frequency evolution: Area 1](image)

![Fig. 5: Frequency evolution: Area 2](image)

Now, we illustrate the action of the thermal constraints by adding a constraint of \( P_e = 2.6 \) p.u. and \( P_e = -2.6 \) p.u. to the tie lines between areas. Figure 6 shows the values of the multipliers \( \lambda_i \), that correspond to the Locational Marginal Prices (LMPs), and the line flows of the tie lines for the same scenario displayed in Figures 4c and 5c, i.e. without thermal limits.

![Fig. 6: LMPs and inter area lines flows: no thermal limits](image)

![Fig. 7: LMPs and inter area lines flows: with thermal limits](image)
limits. It can be seen that neither the initial condition, nor the new steady state satisfy the thermal limit (shown by a dashed line). However, once we add thermal limits to our OLC scheme, we can see in Figure 7 that the system converges to a new operating point that satisfies our constraints.

Finally, we show the conservativeness of the bound obtained in Theorem 17. We simulate the system (31) and (59) under the same conditions as in Figure 6. We set $B_i$ such that the corresponding $\delta b$'s are homogeneous for every bus $i$. We also do not impose the bounds (70) on $d_i(\cdot)$ and use instead $d_i$ as described in Figure 3. This last assumption actually implies that the interval in (71) is empty since $d' = 0$.

![Fig. 8: Bus frequency evolution for homogeneous perturbation $\delta b_i \in \{-0.4, -0.21, -0.2, -0.19, 0.0\}$](image)

![Fig. 9: Location Marginal Prices evolution for homogeneous perturbation $\delta b_i \in \{-0.4, -0.21, -0.2, -0.19, 0.0\}$](image)

Figures 8 and 9 show the evolution of the frequency $\omega_i$ and LMPs $\lambda_i$ for different values of $\delta b_i$ belonging to $\{-0.4, -0.21, -0.2, -0.19, 0.0\}$. Since $D_i = 0.2$ at all the buses, then $\delta b_i = -0.2$ is the threshold that makes $B_i$ go from positive to negative as $\delta b_i$ decreases.

Despite condition (71) is not satisfied for any $\delta b_i$, our simulations show that the system converges whenever $B_i \geq 0$ ($\delta b_i \geq -0.2$). The case when $\delta b_i = -0.2$ is of special interest. Here, the system converges, yet the nominal frequency is not restored. This is because the terms $\delta b_i\omega_i$ (60) are equal to the terms $D_i\omega_i$ in (31a)-(31b). Thus $\omega_i$ and $\lambda_i$ can be made simultaneously zero with nonzero $\omega_i^*$. Fortunately, this can only happen when $B_i = 0$ which can be avoided since $B_i$ is a designed parameter.

**VIII. Concluding Remarks**

This paper studies the problem of restoring the power balance and operational constraints of a power network after a disturbance by dynamically adapting the loads. We show that provided communication is allowed among neighboring buses, it is possible to rebalance the power mismatch, restore the nominal frequency, and maintain inter-area flows and thermal limits. Our distributed solution converges for every initial condition and is robust to parameter uncertainty. Several numerical simulations verify our findings and provide new insight on the conservativeness of the theoretical sufficient condition.

Several future directions must be explored to assess the full potential of this control scheme. Nonlinear versions of the network as well as additional generator and load dynamics need to be included in the analysis. A thorough study that considers delays as well as implementations where only a subset of the network buses implements our controllers needs to be undertaken in order to assess robustness and the feasibility of an incremental deployment of the proposed scheme.

**References**


Enrique Mallada (S’09-M’13) received the ingeniero en telecommunications degree from Universidad ORT, Uruguay, and the Ph.D. degree in electrical and computer engineering with a minor in applied mathematics from Cornell University, Ithaca, NY, in 2005 and 2014, respectively. He is currently a Post-doctoral Fellow in the Center for the Mathematics of Information (CMI) at the California Institute of Technology, Pasadena, CA. His research interests include networks, control, nonlinear dynamics and optimization, with applications to power and information systems.

Dr. Mallada was a recipient of the Organization of American States (OAS) Academic Scholarship from 2008 to 2010, the Jacobs Fellowship from Cornell University in 2011, and the Cornell ECE Director’s Thesis Research Award in 2014.

Changhong Zhao (S’12) is a PhD candidate in electrical engineering at California Institute of Technology. His research is focused on dynamics, stability analysis and distributed control and optimization of power networks, with applications in frequency and voltage regulations, load-side ancillary services, and demand response. Before coming to Caltech, he received the B.Eng. degree in automatic control from Tsinghua University, Beijing, China, in 2010. He received the B.S. degree in automatic control from Tsinghua University, Beijing, China, in 2010.

Steven Low (F’08) received the B.S. degree from Cornell University, Ithaca, NY, USA, and the Ph.D. degree from the University of California, Berkeley, USA, both in electrical engineering.

He is a Professor of the Computing and Mathematical Sciences and Electrical Engineering Departments at the California Institute of Technology, Pasadena, CA, USA. Before that, he was with AT&T Bell Laboratories, Murray Hill, NJ, USA, and the University of Melbourne, Australia. He is a Senior Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (and the mentor for the annual JSAC series on Smart Grid), a Senior Editor of the IEEE TRANSACTIONS ON CONTROL OF NETWORK SYSTEMS, a Steering Committee Member of the IEEE TRANSACTIONS ON NETWORK SCIENCE AND ENGINEERING, and on the editorial board of NOW Foundations and Trends in Networking, and in Power Systems. He also served on the editorial boards of IEEE/ACM TRANSACTIONS ON NETWORKING, IEEE TRANSACTIONS ON AUTOMATIC CONTROL, ACM Computing Surveys, Computer Networks Journal.