Optimal load-side control for frequency regulation in smart grids

Enrique Mallada, *Member, IEEE,* Changhong Zhao, *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 become increasingly necessary with the current acceleration of renewable penetration. 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. In particular, 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 the currently implemented frequency control), 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. We illustrate its effectiveness through simulations.

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 or AGC) 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]–[15] 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) [16]. Although these simulation studies and field trials are insightful, they fall short in predicting the (potential) behavior of a large-scale deployment of frequency control.

This has motivated the recent development of new analytic studies assessing the effect of distributed frequency control in power systems [17]–[23], and microgrids [24]–[26], which can be grouped into three main approaches. The first approach builds on *consensus algorithms* to provide efficiency guarantees to classical PI controllers [17], [25], [26]. It achieves efficiency but does not manage congestion, i.e., it does not enforce constraints such as thermal limits. The second approach *reverse engineers* the network dynamics as a primal-dual algorithm of an underlying optimization problem, and then add constraints and modifies the objective function while preserving the primal-dual interpretation of the network dynamics [20]–[23], [27]. It successfully achieves efficiency but it limits the type of operational constraints that can be satisfied. The third approach directly formulates an optimization problem that encodes *all operational constraints* and then designs a *hybrid algorithm* that combines the network dynamics with a subset of the primal-dual algorithm [19]. It is able to satisfy operational constraints, but the stability depends on the network parameters.

This work was supported by NSF NetSE grant CNS 0911041, NSF CPS grant CNS 1544771, ARPA-E grant DE-AR0000226, Southern California Edison, National Science Council of Taiwan R.O.C. grant NSC 103-3113-P-008-001, Caltech Resnick Institute, Johns Hopkins E2SHI Seed Grant, and Johns Hopkins WSE startup funds.

Enrique Mallada is with the Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD 21218 USA (e-mail:mallada@jhu.edu).

Changhong Zhao is with the National Renewable Energy Laboratory, Golden, CO 80401 USA (changhong.zhao@nrel.gov).

Steven Low is with the Department of Computing + Mathematical Sciences, California Institute of Technology, Pasadena, CA 91125 USA (e-mail: slow@caltech.edu).

Contributions of this work: In this paper we develop a method to jointly achieve primary + secondary frequency control, and congestion management (bringing line flows to within their limits), *in a distributed manner using controllable loads*. To contrast with the generation-side AGC frequency control, we refer to our solution as *load-side control*. To our knowledge, this method produces the *first* distributed controllers for demand response that are scalable and enforce required operational constraints for frequency regulation, such as restoring nominal frequency and preserving inter-area flows, while respecting line limits.

Our work builds on previous optimization-based approaches [19]–[23], [27]. The crux of our solution is the introduction of *virtual (line) flows* that can be used to implicitly constrain real flows without altering the primal-dual interpretation of the network dynamics. A virtual flow is a cyber quantity on each line that a controller computes based on information from its neighbors. In steady state, its value equals the actual line flows incident on that controller. This device allows us to impose arbitrary constraints on the (actual) line flows for congestion management and restoring inter-area flows, while retaining the ability to exploit network dynamics to help carry out the primal-dual algorithm.

Our contribution with respect to the existing literature is manifold. Unlike [18], [19], our global asymptotic stability result (Theorem 11 in Section IV) is independent of the controller gains, which is highly desirable for fully distributed deployments. Our results hold for arbitrary topologies and can impose inter-area constraints, thermal limits or *any linear equality or inequality constraint* in the line flows. Moreover, we provide a convergence analysis in the presence of unknown parameters (Section V) that is novel in the primal-dual literature and provides the necessary robustness for large scale distributed deployments. Finally, our framework can further extend to include intermediate buses without generators or loads – quite common in transmission networks [2, Chapter 9.3]– which are not considered by the existing literature, and to fully distribute non-local constraints like those imposed on inter-area flows (Section VI).

A preliminary version of this work has been presented in [28]. This paper extends [28] in several ways. First, the robustness study of our controllers with respect to uncertain load parameter (Section V) as well as the framework extensions (Section VI) are new. Second, we include detailed proofs that were omitted in [28] due to space constraints. Finally, we extend our simulations in Section VII to further illustrate the conservativeness of the uncertainty bounds of Theorem 15.

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 to its column vector, 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. For two vectors $a \in \mathbb{R}^{|S|}$ and $b \in \mathbb{R}^{|S'|}$ we define the column vector $x = (a, b) \in \mathbb{R}^{|S|+|S'|}$. Given a matrix A, we denote its transpose as A^T and use A_i to denote the ith 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 $\mathrm{diag}(a_i)_{i\in S}$. Similarly, for a sequence of matrices $\{A_h,\ h \in S\}$ we let $\mathrm{blockdiag}(A_h)_{h\in S}$ denote the block diagonal matrix. Finally, we use either $\mathbf{1}_n$ and $\mathbf{1}_{n\times m}$ ($\mathbf{0}_n$ and $\mathbf{0}_{n\times m}$) to denote the vector and matrix of all ones (zeros), or 1 (0) when its dimension is understood from the context.

For a function $f : \mathbb{R}^n \to \mathbb{R}^n$ we use $f'(x) := \frac{\partial}{\partial x} f(x)$ to denote the Jacobian and f^{-1} to denote its inverse. When $n = 1$, $f''(x)$ denotes the second derivative $\frac{\partial^2}{\partial x^2} f(x)$.

For given vectors $u \in \mathbb{R}^n$ and $a \in \mathbb{R}^n$, and set $S \subseteq \{1, \dots, n\}$, the operator $[a]^+_{u_S}$ is defined element-wise by

$$([a]^+_{u_S})_i = \begin{cases} [a_i]^+_{u_i}, & \text{if } i \in S, \\ a_i, & \text{if } i \notin S, \end{cases} \tag{1}$$

where $[a_i]^+_{u_i}$ is equal to a_i if either $a_i > 0$ or $u_i > 0$, and 0 otherwise. Whenever $u^*_S \geq 0$, the following relation holds:

$$(u_S - u^*_S)^T [a_S]^+_{u_S} \leq (u_S - u^*_S)^T a_S \tag{2}$$

since for any pair (u_i, a_i), with $i \in S$, that makes the projection active ($[a_i]^+_{u_i} = 0$) we must have $u_i \leq 0$ and $a_i \leq 0$, and therefore $(u_i - u^*_i)a_i \geq 0 = (u_i - u^*_i)^T [a_i]^+_{u_S}$.

A. Network Model

We consider a power network described by a directed graph $G(\mathcal{N}, \mathcal{E})$ where $\mathcal{N} = \{1, ..., |\mathcal{N}|\}$ is the set of buses, denoted by either i or j, and $\mathcal{E} \subset \mathcal{N} \times \mathcal{N}$ is the set of transmission lines denoted by either e or ij such that if $ij \in \mathcal{E}$, then $ji \notin \mathcal{E}$.

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. A generator bus may also have loads attached to it, but not otherwise. We assume the graph $G(\mathcal{N}, \mathcal{E})$ is connected, and make the following assumptions which are standard and well-justified for transmission networks [29]: (i) bus voltage magnitudes $|V_i|$ are constant for all $i \in \mathcal{N}$, (ii) lines $ij \in \mathcal{E}$ are lossless and characterized by their reactance $x_{ij} > 0$, and (iii) reactive power flows do not affect bus voltage phase angles and frequencies.

The evolution of the transmission network is therefore described by

$$\dot{\theta}_i = \omega_i \qquad i \in \mathcal{N} \tag{3a}$$

$$M_i \dot{\omega}_i = P^{in}_i - (d_i + D_i\omega_i) - \textstyle\sum_{j\in\mathcal{N}_i} B_{ij}(\theta_i - \theta_j) \qquad i \in \mathcal{G} \tag{3b}$$

$$0 = P^{in}_i - (d_i + D_i\omega_i) - \textstyle\sum_{j\in\mathcal{N}_i} B_{ij}(\theta_i - \theta_j) \qquad i \in \mathcal{L} \tag{3c}$$

where ω_i denotes the bus frequency, θ_i denotes the phase, d_i denotes an aggregate controllable load, $D_i\omega_i$ denotes the aggregate power consumption consumption due to uncontrollable frequency-sensitive loads and/or generators' damping. M_i denotes the generator's inertia, and P^{in}_i is the (constant) difference between mechanical power injected by a generator and the constant aggregate power consumed by loads. Notice that for load buses ($i \in \mathcal{L}$) $P^{in}_i < 0$. Finally, the line parameter $B_{ij} := 3\frac{|V_i||V_j|}{x_{ij}} \cos(\theta^0_i - \theta^0_j)$ represents the sensitivity of the power flow to phase variations. All variables θ_i, ω_i and d_i as

well as the parameter P_i^{in} are deviations from nominal values θ_i^0, ω^0, d_i^0, and $P_i^{in,0}$.

By letting the line power flows $P_e = P_{ij} := B_{ij}(\theta_i - \theta_j)$ for $e = ij \in \mathcal{E}$, we can equivalently rewrite (3) as

$$M_i\dot{\omega}_i = P_i^{in} - (d_i + D_i\omega_i) - \textstyle\sum_{e\in\mathcal{E}} C_{i,e}P_e \qquad i \in \mathcal{G} \quad (4a)$$

$$0 = P_i^{in} - (d_i + D_i\omega_i) - \textstyle\sum_{e\in\mathcal{E}} C_{i,e}P_e \qquad i \in \mathcal{L} \quad (4b)$$

$$\dot{P}_{ij} = B_{ij}(\omega_i - \omega_j) \qquad ij \in \mathcal{E} \quad (4c)$$

where $C_{i,e}$ are the elements of the incidence matrix $C \in \mathbb{R}^{|\mathcal{N}|\times|\mathcal{E}|}$ of the graph $G(\mathcal{N}, \mathcal{E})$ such that $C_{i,e} = -1$ if $e = ji \in \mathcal{E}$, $C_{i,e} = 1$ if $e = ij \in \mathcal{E}$ and $C_{i,e} = 0$ otherwise, and the line flows initial condition must satisfy $P_{ij}(0) = B_{ij}(\theta_i(0) - \theta_j(0))$.

Remark 1. *Equation* (3) *and* (4) *represent a linearized version of the nonlinear swing equations [2]. The dynamics are similar to the DC approximation but the assumptions do not require small nominal phase angle difference across each line; see [30, Section VII] for a first principles derivation of* (4)*. This model has been standard in the design of frequency controllers (see, e.g., [31] and [2, Chapter 11]) as it is able to capture the behavior of small frequency fluctuations due to supply demand imbalance. Extending the results of this work for nonlinear power flow models is subject of current research. We refer the reader to [32] for recent preliminary extensions on this direction.*

Remark 2. *Our model assumes that every bus has both $D_i > 0$ and controllable load d_i. While this is reasonable for load and generator buses (for $i \in \mathcal{G}$ the generator can implement $-d_i$), it is unreasonable for intermediate buses that have neither generators nor loads. This is addressed in Section VI. Our framework can also handle the case where d_i is present only for $i \in S \subset \mathcal{N}$. This case is omitted due to space constraints.*

B. Operational Constraints

We denote each control area using k and let $\mathcal{K} := \{1, \dots, |\mathcal{K}|\}$ be the set of all areas in the network. Let $\mathcal{N}_k \subseteq \mathcal{N}$ be set of buses in area k and $\mathcal{B}_k \subseteq \mathcal{E}$ the set of boundary edges of area k, i.e. $ij \in \mathcal{B}_k$ iff either $i \in \mathcal{N}_k$ or $j \in \mathcal{N}_k$, but not both.

Within each area, the AGC scheme seeks to restore the frequency to its nominal value and preserve a constant power transfer outside the area, i.e.

$$\textstyle\sum_{ij\in\mathcal{B}_k} \hat{C}_{k,ij}P_{ij} = \hat{P}_k \quad (5)$$

where $\hat{C}_{k,ij}$ is equal to 1, if $i \in \mathcal{N}_k$, -1, if $j \in \mathcal{N}_k$, and $\hat{P}_k$ is the net scheduled power injection of area k.[1]

By defining $\hat{C}_{k,ij}$ to be 0 whenever $ij \notin \mathcal{B}_k$ we can also relate $\hat{C} \in \mathbb{R}^{|\mathcal{K}|\times|\mathcal{E}|}$ with $C \in \mathbb{R}^{|\mathcal{N}|\times|\mathcal{E}|}$ using

$$\hat{C} = E_{\mathcal{K}}C \quad (6)$$

[1]The division of the power network into control areas is an artifact of the current control architecture. Our solution can achieve all the control objectives of primary and secondary frequency control without enforcing this constraint. Thus, we only include (5) to fit our solution within the existing architecture.

where $E_{\mathcal{K}} := [e_1 \ \dots \ e_{|\mathcal{K}|}]^T$ and $e_k \in \mathbb{R}^{|\mathcal{N}|}$, $k \in \mathcal{K}$, is a vector with elements $(e_k)_i = 1$ if $i \in \mathcal{N}_k$ and $(e_k)_i = 0$ otherwise.

Finally, the thermal limit constraints are usually given by

$$\underline{P} \le P \le \overline{P} \quad (7)$$

where $P := (P_e)_{e\in\mathcal{E}}$, $\overline{P} := (\overline{P}_e)_{e\in\mathcal{E}}$ and $\underline{P} := (\underline{P}_e)_{e\in\mathcal{E}}$ represent the line flow limits.

III. Optimal load-side control

Suppose the system (4) is in equilibrium, i.e. $\dot{\omega}_i = 0$ for all i and $\dot{P}_{ij} = 0$ for all ij, and at time 0, there is a disturbance, represented by a step change in the vector $P^{in} := (P_i^{in})_{i\in\mathcal{N}}$, that produces a power imbalance. Then, we are interested in designing a distributed control mechanism that rebalances the system, restores the frequency to its nominal value while 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 $c_i(d_i)$ to denote the cost or disutility of changing the load consumption by d_i. This allows us to formally describe our notion of efficiency in terms of the loads' power share. More precisely, we shall say that a load control is efficient if in equilibrium solves the Optimal Load Control (OLC) problem:

Problem 1 (OLC)**.**

$$\underset{d,\omega,\theta}{\text{minimize}} \quad \sum_{i\in\mathcal{N}} c_i(d_i) \quad (8a)$$

$$\text{subject to} \quad P^{in} - (d + D\omega) = L_B\theta \quad (8b)$$

$$\omega = 0 \quad (8c)$$

$$\hat{C}BC^T\theta = \hat{P} \quad (8d)$$

$$\underline{P} \le BC^T\theta \le \overline{P} \quad (8e)$$

where $d = (d_i)_{i\in\mathcal{N}}$, $\omega = (\omega_i)_{i\in\mathcal{N}}$, $\theta = (\theta_i)_{i\in\mathcal{N}}$, $D = \mathrm{diag}(D_i)_{i\in\mathcal{N}}$, $B = \mathrm{diag}(B_{ij})_{ij\in\mathcal{E}}$, $(BC^T\theta)_{ij} = B_{ij}(\theta_i - \theta_j)$ and $L_B := CBC^T$ is the B_{ij}-weighted Laplacian matrix.

Problem 1 embeds in a sole convex optimization problem the operational constraints of the three hierarchical layers of generation-side frequency control. Supply-demand balance (8b), which belongs to primary frequency control; frequency restoration (8c) and inter-area flow constraints (8d), which conform secondary frequency control; and thermal limits (8e) and efficient scheduling (8a), which are part of economic dispatch.

This sophisticated hierarchical approach evidences the incremental complexity required to enforce each additional constraint. Supply-demand balance (8b) is *intrinsically* enforced by the power system dynamics. This property was leveraged in [20] to develop a load-side primary frequency control. Frequency restoration (8c) and inter-area flow preservation (8d) require some notion of *integral control* to cancel their combined error (also known as area control error or ACE), and *communication* to send this (unique within each area) signal to each participating unit. However, the most challenging constraints are the thermal limits (8e) since they require a highly non-trivial coordination of resources –as illustrated in

Section VII (Figure 7)– which is currently implemented by solving an *optimization problem*.

Unfortunately, the existing architecture is not suitable to solving our optimal load-side control problem since each layer is required to operate at a different timescale. One possible alternative could be to use a standard optimization algorithm to solve Problem 1 – since Problem 1 is convex, in principle it could be efficiently solved using one of the many optimization algorithms available [33]. However, unlike standard optimization problems where one usually has access to all the problem variables, here we can only modify the load demand d_i, while θ_i and ω_i react to these changes according to (3). This is a fundamental constraint that further requires that fast changes in d_i do not render the system unstable.

We overcome these restrictions in this paper by formulating an equivalent optimization problem whose primal-dual optimization algorithm embeds the line flow version of the swing equations (4). By doing so, instead of looking at the power network dynamics and their stability as a limitation to load control design, our controllers *cooperate with the network* in order to collectively solve Problem 1. Since our work cuts across several architectural layers, our solution can be interpreted as a *unified load-side control architecture for efficient primary and secondary frequency control with congestion management.*

Throughout this paper we make the following assumptions:

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

Assumption 2 (Strict feasibility in $\mathcal{D}$)**.** *The OLC problem* (8) *is feasible and there is at least one feasible (d, ω, θ) such that $d \in \operatorname{Int}\mathcal{D} := \Pi_{i=1}^{|\mathcal{N}|}\operatorname{Int}\mathcal{D}_i$.*

Assumptions 1 and 2 are sufficient for the optimality and convergence analysis of our controllers to hold. Assumption 1 ensures that the derived controllers are Lipschitz, continuously differentiable, and that the demand response never exceeds its capacity. Assumption 2 guarantees that, even in the presence of Assumption 1, the optimal solution of OLC is finite, and therefore allows us to use the KKT conditions [33, Ch. 5.2.3] to characterize it.

The next assumption will be used in Section V in order to guarantee convergence in the presence of uncertain parameters.

Assumption 3 (Lipschitz continuity of c_i')**.** *The functions $c_i'(\cdot)$ are Lipschitz continuous with Lipschitz constant $L > 0$.*

As we will show in Section V, Assumption 3 ensures enough responsiveness on the demand to guarantee convergence in the presence of uncertainty on D_i. This is critical for our robustness result (Theorem 15). Notice also that Assumption 3 implies that the domain of $\mathcal{D}_i = \mathbb{R}$ in Assumption 1. However, if the systems is designed with enough capacity such that $d_i^* \in [\underline{d}_i + \varepsilon, \overline{d}_i - \varepsilon]\ \forall i$, then one can always modify a cost function $c_i(\cdot)$ that satisfies Assumptions 1 and 2 for finite domains $\mathcal{D}_i = [\underline{d}_i, \overline{d}_i]$ and define a new cost function $\tilde{c}_i(\cdot)$ that satisfies Assumption 3 without modifying the optimal allocation d_i^*. More precisely, given $c_i(\cdot)$, define $\tilde{c}_i(\cdot)$ to be equal to $c_i(\cdot)$ inside $[\underline{d}_i + \varepsilon, \overline{d}_i - \varepsilon]$ and modify $\tilde{c}_i(\cdot)$ outside the subset so that Assumption 3 holds. It is easy to show then that the optimal solution will still be d_i^* and therefore we still get $\underline{d}_i \leq d_i^* \leq \overline{d}_i$.

A. Virtual Flows Reformulation

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 implicitly imposing the constraints (8c)-(8e) by using *virtual flows* instead of explicitly using ω_i and θ_i. This, together with an additional quadratic objective on ω_i and substituting $B_{ij}(\theta_i - \theta_j)$ with P_e in (8b), allows us to embed the network dynamics as part of the primal-dual algorithm while preserving all the desired constraints.

Problem 2 (VF-OLC)**.**

$$\underset{d,\omega,\phi,P}{\text{minimize}} \quad \sum_{i\in\mathcal{N}} c_i(d_i) + \frac{D_i{\omega_i}^2}{2} \tag{9a}$$

$$\text{subject to} \quad P^{in} - (d + D\omega) = CP \tag{9b}$$

$$P^{in} - d = L_B\phi \tag{9c}$$

$$\hat{C}BC^T\phi = \hat{P} \tag{9d}$$

$$\underline{P} \leq BC^T\phi \leq \overline{P} \tag{9e}$$

where $\phi = (\phi_i)_{i\in\mathcal{N}}$ represents the virtual phases and $(BC^T\phi)_{ij} = B_{ij}(\phi_i - \phi_j)$ is the corresponding *virtual flow* through line $ij \in \mathcal{E}$.

Although not clear at first sight, the constraint (9c) –together with (9b)– implicitly enforces that any optimal solution of VF-OLC $(d^*, \omega^*, \phi^*, P^*)$ will restore the frequency to its nominal value, i.e. $\omega_i^* = 0$. Similarly, constraint (9d) will impose (8d) and (9e) will impose (8e). As a result, the optimal solution of VF-OLC is not affected by the additional term $\sum_i D_i\omega_i^2/2$ in the objective and therefore the two problems (OLC and VF-OLC) are equivalent. This is formalized in Lemma 4 below.

We also highlight that VF-OLC uses P to represent line flows instead of the phase representation $BC^T\theta$ used in OLC. This selection is not arbitrary and it will be critical for embedding the network dynamics (4) –not (3)– as part of the primal-dual algorithm that solves VF-OLC (Theorem 6). A simple exercise shows that, if P is substituted with $BC^T\theta$ in (9), then the primal-dual algorithm no longer embeds (4), i.e., Theorem 6 no longer holds.

We use ν_i, λ_i and π_k to denote the Lagrange multipliers of constraints (9b), (9c) and (9d), and ρ_{ij}^+ and ρ_{ij}^- as multipliers of the right and left constraints of (9e), respectively. In order to make the presentation more compact sometimes we will use $x = (\phi, P) \in \mathbb{R}^{|\mathcal{N}|+|\mathcal{E}|}$ and $\sigma = (\lambda, \nu, \pi, \rho^+, \rho^-) \in \mathbb{R}^{2|\mathcal{N}|+|\mathcal{K}|+2|\mathcal{E}|}$, as well as $\rho := (\rho^+, \rho^-)$.

Using this notation we can write the Lagrangian of VF-OLC as

$$L(d, \omega, x, \sigma) = \textstyle\sum_{i\in\mathcal{N}} \left(c_i(d_i) + \frac{D_i{\omega_i}^2}{2}\right) + \nu^T(P^{in} - d - D\omega - CP) + \lambda^T(P^{in} - d - L_B\phi) + \pi^T(\hat{C}BC^T\phi - \hat{P})$$

$$
\begin{aligned}
&+\rho^{+T}(BC^T\phi-\overline{P})+\rho^{-T}(\underline{P}-BC^T\phi)\\
&=\sum_{i\in\mathcal{N}}c_i(d_i)-(\lambda_i+\nu_i)d_i+D_i\omega_i\,(\omega_i/2-\nu_i)+(\nu_i+\lambda_i)P_i^{in}\\
&\quad-P^TC^T\nu-\phi^T(L_B\lambda-CB\hat{C}^T\pi-CB(\rho^+-\rho^-))\\
&\quad-\pi^T\hat{P}-\rho^{+T}\overline{P}+\rho^{-T}\underline{P} \qquad (10)
\end{aligned}
$$

The next lemmas characterize the optimality conditions of VF-OLC and its equivalence with OLC. Their proofs can be found in the Appendix.

Lemma 3 (Optimality)**.** *Let $G(\mathcal{N},\mathcal{E})$ be a connected graph. Then $(d^*,\omega^*,\phi^*,P^*,\sigma^*{:=}(\lambda^*,\nu^*,\pi^*,\rho^{+*},\rho^{-*}))$ is a primal-dual optimal solution to VF-OLC if and only if $(d^*,\omega^*,\phi^*,P^*)$ is primal feasible, $\rho^{+*},\rho^{-*}\geq 0$,*

$$d_i^*=c_i'^{-1}(\nu_i^*+\lambda_i^*),\;\;\omega_i^*=\nu_i^*=\hat{\nu},\;\;i\in\mathcal{N}, \qquad (11)$$

where $c_i'^{-1}$ is the inverse of the derivative of c_i, $\hat{\nu}$ is some scalar,

$$\textstyle\sum_{j\in\mathcal{N}_i}B_{ij}(\lambda_j^*-\lambda_i^*)+C_iB(\hat{C}^T\pi^*+\rho^{+*}-\rho^{-*})=0 \qquad (12)$$

with C_i being the ith row of C, and

$$\rho_{ij}^{+*}(B_{ij}(\phi_i^*-\phi_j^*)-\overline{P}_{ij})=0,\quad ij\in\mathcal{E}, \qquad (13a)$$
$$\rho_{ij}^{-*}(\underline{P}_{ij}-B_{ij}(\phi_i^*-\phi_j^*))=0,\quad ij\in\mathcal{E}. \qquad (13b)$$

Moreover, d^*, ω^*, ν^* and λ^* are unique with $\hat{\nu}=0$.

Lemma 4 (OLC and VF-OLC Equivalence)**.** *Given any set of vectors $(d^*,\omega^*,\theta^*,\phi^*,P^*)$ satisfying $C^T\theta^*=C^T\phi^*$ and $L_B\theta^*=CP^*$. Then (d^*,ω^*,θ^*) is an optimal solution of OLC if and only $(d^*,\omega^*,\phi^*,P^*)$ is an optimal solution to VF-OLC.*

B. Distributed Optimal Load-side 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 [34]–[37].

Since by Lemma 4, VF-OLC provides the same optimal load schedule as OLC, we can solve VF-OLC instead. This will allow us to incorporate the network dynamics as part of an optimization algorithm that indirectly solves OLC.

To achieve this goal we first minimize (10) over d and ω which is achieved by setting $\omega_i=\nu_i$ and $d_i=c_i'^{-1}(\nu_i+\lambda_i)$ in (10). Thus, by letting $d_i(\cdot):=c_i'^{-1}(\cdot)$ we get

$$
\begin{aligned}
L(x,\sigma)&=\underset{d,\omega}{\text{minimize}}\;L\left(d,\omega,x{:=}(\phi,P),\sigma{:=}(\lambda,\nu,\pi,\rho^+,\rho^-)\right)\\
&=\Phi_{\mathcal{N}}(\lambda,\nu)-P^TC^T\nu-\pi^T\hat{P}-\rho^{+T}\overline{P}+\rho^{-T}\underline{P}\\
&\quad-\phi^T(L_B\lambda-CD_B\hat{C}^T\pi-CD_B(\rho^+-\rho^-)) \qquad (14)
\end{aligned}
$$

where $\Phi_S(\lambda_S,\nu_S):=\sum_{i\in S}\Phi_i(\lambda_i,\nu_i)$ and

$$
\begin{aligned}
\Phi_i(\lambda_i,\nu_i):=&c_i(d_i(\lambda_i+\nu_i))-(\lambda_i+\nu_i)d_i(\lambda_i+\nu_i)\\
&-D_i\frac{\nu_i^2}{2}+(\lambda_i+\nu_i)P_i^{in}. \qquad (15)
\end{aligned}
$$

The strict convexity of $L(d,\omega,x,\sigma)$ in (d,ω) and the fact that d and ω only appear in (9b) and (9c) gives rise to the following lemma whose proof is also in the Appendix.

Lemma 5 (Strict concavity of $L(x,\sigma)$ in (λ,ν))**.** *The function $\Phi_i(\lambda_i,\nu_i)$ in* (15) *is strictly concave. As a result, $L(x,\sigma)$ is strictly concave in (λ,ν).*

Next we reduce the Lagrangian $L(x,\sigma)$ by maximizing it for any ν_i with $i\in\mathcal{L}$. We let $y=(\lambda,\nu_{\mathcal{G}},\pi,\rho^+,\rho^-)$ and consider the Lagrangian

$$L(x,y)=\underset{\nu_i:i\in\mathcal{L}}{\text{maximize}}\;L(x,\sigma). \qquad (16)$$

Since $L(x,\sigma)$ is strictly concave in ν by Lemma 5, the minimizer of (16) is unique. Moreover, this also implies that $L(x,y)$ is strictly concave in $(\lambda,\nu_{\mathcal{G}})$.

Finally, the optimal load controllers can be then obtained by considering the primal-dual gradient law of $L(x,y)$ which is given by

$$\dot{y}=Y\left[\frac{\partial}{\partial y}L(x,y)^T\right]_\rho^+\;\;\text{and}\;\;\dot{x}=-X\frac{\partial}{\partial x}L(x,y)^T \qquad (17)$$

where X and Y are positive-definite diagonal matrices given by $X:=\text{diag}((\chi_e^P)_{e\in\mathcal{E}},(\chi_i^\phi)_{i\in\mathcal{N}})$ and $Y:=\text{diag}((\zeta_i^\nu)_{i\in\mathcal{G}},(\zeta_i^\lambda)_{i\in\mathcal{N}},(\zeta_k^\pi)_{k\in\mathcal{K}},(\zeta_e^{\rho^+})_{e\in\mathcal{E}},(\zeta_e^{\rho^-})_{e\in\mathcal{E}})$, and the projection $[\cdot]_\rho^+$ is defined as in (1) for $u=y$ and $u_S=\rho$. This projection ensures that the vector $\rho(t)=(\rho^+(t),\rho^-(t))$ remains within the positive orthant, that is $\rho^+(t)\geq 0$ and $\rho^-(t)\geq 0\;\forall t$.

Notice that although the frequency ω is no longer present in (17), the minimization in (14) requires $\omega_i=\nu_i$. As a result the Lagrange multipliers ν_i in (17) will play the role of the frequency ω_i in (4). The following theorem shows that this procedure indeed embeds the network dynamics as part of the primal-dual law (17).

Theorem 6 (Optimal Load-side Control)**.** *By setting $\zeta_i^\nu=M_i^{-1}$, $\chi_{ij}^P=B_{ij}$ and $\nu_i=\omega_i$, the primal-dual gradient law* (17) *is equivalent to the power network dynamics* (4) *together with*

$$\dot{\lambda}_i=\zeta_i^\lambda\Big(P_i^{in}-d_i-\textstyle\sum_{j\in\mathcal{N}_i}B_{ij}(\phi_i-\phi_j)\Big) \qquad (18a)$$
$$\dot{\pi}_k=\zeta_k^\pi\Big(\textstyle\sum_{ij\in\mathcal{B}_k}\hat{C}_{k,ij}B_{ij}(\phi_i-\phi_j)-\hat{P}_k\Big) \qquad (18b)$$
$$\dot{\rho}_{ij}^+=\zeta_{ij}^{\rho^+}\left[B_{ij}(\phi_i-\phi_j)-\overline{P}_{ij}\right]_{\rho_{ij}^+}^+ \qquad (18c)$$
$$\dot{\rho}_{ij}^-=\zeta^{\rho_{ij}^-}\left[\underline{P}_{ij}-B_{ij}(\phi_i-\phi_j)\right]_{\rho_{ij}^-}^+ \qquad (18d)$$
$$
\begin{aligned}
\dot{\phi}_i=\chi_i^\phi\Big(&\textstyle\sum_{j\in\mathcal{N}_i}B_{ij}(\lambda_i-\lambda_j)-\sum_{k\in\mathcal{K},e\in\mathcal{B}_k}C_{i,e}B_e\hat{C}_{k,e}\pi_k\\
&-\textstyle\sum_{e\in\mathcal{E}}C_{i,e}B_e(\rho_e^+-\rho_e^-)\Big) \qquad (18e)
\end{aligned}
$$
$$d_i=c_i'^{-1}(\lambda_i+\omega_i) \qquad (18f)$$

where (18a)*,* (18e) *and* (18f) *are for $i\in\mathcal{N}$,* (18b) *is for $k\in\mathcal{K}$, and* (18c) *and* (18d) *are for $ij\in\mathcal{E}$.*

Proof: By Lemma 5 and (14), $L(x,\sigma)$ is strictly concave in (λ,ν). Therefore, it follows that there exists a unique

$$\nu_{\mathcal{L}}^*(x,y)=\arg\max_{\nu_{\mathcal{L}}}L(x,\sigma). \qquad (19)$$

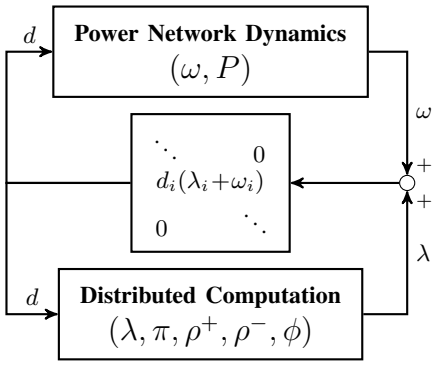


Fig. 1: Control architecture derived from OLC.

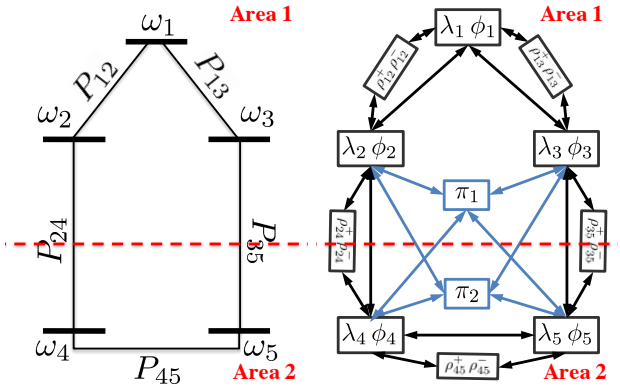


Fig. 2: Power network example (left) and the corresponding communication requirement to implement the distributed load control (18) (right).

By stationarity, $\nu^*_{\mathcal{L}}(x,y)$ must satisfy

$$\frac{\partial L}{\partial \nu_{\mathcal{L}}}(x,y,\nu^*_{\mathcal{L}}(x,y))^T = \frac{\partial \Phi_{\mathcal{L}}}{\partial \nu_{\mathcal{L}}}(\lambda_{\mathcal{L}},\nu^*_{\mathcal{L}}(x,y))^T - C_{\mathcal{L}}P \quad (20a)$$
$$= P^{in}_{\mathcal{L}} - D_{\mathcal{L}}\nu^*_{\mathcal{L}}(x,y) - d_{\mathcal{L}}(\lambda_{\mathcal{L}}+\nu^*_{\mathcal{L}}(x,y)) - C_{\mathcal{L}}P = 0 \quad (20b)$$

which is equivalent to (4b), i.e. $\nu^*_{\mathcal{L}}(x,y)$ implicitly satisfies (4b). Moreover, since $d_i(\cdot)$ is strictly increasing, there is a unique ω_i that satisfies (4b) with $d_i = d_i(\lambda_i+\omega_i)$ for fixed λ_i and P, which implies that $\omega_{\mathcal{L}} = \nu^*_{\mathcal{L}}(x,y)$.

We now iteratively apply the envelope theorem [38] on $L(x,y)$ defined in (16) to compute $\frac{\partial}{\partial x}L(x,y)$ and $\frac{\partial}{\partial y}L(x,y)$. For example, to compute $\frac{\partial}{\partial x}L(x,y)$ we use

$$\frac{\partial}{\partial x}L(x,y) = \frac{\partial}{\partial x}L(x,\sigma)\Big|_{\nu_{\mathcal{L}}=\nu^*_{\mathcal{L}}(x,y)} \quad (21a)$$
$$= \left(\frac{\partial}{\partial x}L(d,\omega,x,\sigma)\Big|_{(d,\omega)=(c'^{-1}(\lambda+\nu),\nu)}\right)\Big|_{\nu_{\mathcal{L}}=\nu^*_{\mathcal{L}}(x,y)}, \quad (21b)$$

where $c'^{-1}(\lambda+\nu) := (c_i'^{-1}(\lambda_i+\nu_i))_{i\in\mathcal{N}}$, which leads to

$$\frac{\partial}{\partial P}L(x,y)^T = -(C_{\mathcal{G}}^T\nu_{\mathcal{G}} + C_{\mathcal{L}}^T\nu^*_{\mathcal{L}}(x,y)) \quad (22a)$$
$$\frac{\partial}{\partial \phi}L(x,y)^T = -(L_B\lambda - CB(\hat{C}^T\pi + \rho^+ - \rho^-)) \quad (22b)$$

An analogous computation for $\frac{\partial}{\partial y}L(x,y)$ gives

$$\frac{\partial}{\partial \nu_{\mathcal{G}}}L(x,y)^T = P^{in} - (d_{\mathcal{G}}(\lambda_{\mathcal{G}}+\nu_{\mathcal{G}}) + D\nu_{\mathcal{G}}) - CP \quad (23a)$$
$$\frac{\partial}{\partial \lambda}L(x,y)^T = P^{in} - d(\lambda+\nu)|_{\nu_{\mathcal{L}}=\nu^*_{\mathcal{L}}(x,y)} - L_B\phi \quad (23b)$$
$$\frac{\partial}{\partial \pi}L(x,y)^T = \hat{C}BC^T\phi - \hat{P} \quad (23c)$$
$$\frac{\partial}{\partial \rho^+}L(x,y)^T = BC^T\phi - \overline{P} \quad (23d)$$
$$\frac{\partial}{\partial \rho^-}L(x,y)^T = \underline{P} - D_BC^T\phi \quad (23e)$$

where, for a set S, $d_S(\lambda_S+\nu_S) := (d_i(\lambda_i+\nu_i))_{i\in S}$ and $d(\lambda+\nu) = d_{\mathcal{N}}(\lambda_{\mathcal{N}}+\nu_{\mathcal{N}})$.

Finally, by setting $\nu_i = \omega_i$ and $\zeta_i^\nu = M_i^{-1}$ in (17), it follows that (23a) with (17) is equivalent to (4a). Since we have already shown that $\omega_{\mathcal{L}}$ must be equal to $\nu^*_{\mathcal{L}}(x,y)$, then, since by assumption $\chi^P_{ij} = B_{ij}$ for $ij \in \mathcal{E}$, an analogous argument shows that (22)-(23) with (17) is equivalent to (4) and (18). ■

Equations (4) and (18) 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 $L(x,y)$. We will show in the next section that this system does achieve optimality as intended.

We now illustrate the operation of our OLC algorithm using figures 1 and 2. Figure 1 shows the control architecture derived from our OLC problem. The cyber layer (lower block) is in charge of the distributed real-time computation of the cyber quantities λ_i, ϕ_i, $\rho_e^\pm$, and π_k . The the power network (upper block) evolves in parallel as a result of applying the load control $d_i(\lambda_i+\omega_i)$ and determines the frequency ω_i and line flows P_{ij}. Both ω_i and λ_i are required by the load at bus i to implement $d_i(\lambda_i+\omega_i)$.

Figure 2 illustrates the different agents and communication links required to implement the cyber layer. There are three types of agents: Bus agents (in charge of computing (λ_i,ϕ_i)), link agents (to compute $(\rho^+_{ij},\rho^-_{ij})$), and area agents (to compute π_k). If the net constant power injection P_i^{in} is available, the cyber layer only requires d_i as input to the bus agent to compute $(\lambda,\pi,\rho^+\rho^-,\phi)$. All the additional information required can be obtained from adjacent agents in the communication graph (right diagram in Figure 2). Figure 2 also illustrates how the only semi-centralized feature of our controller is the computation of π_k. This also affects the computation of ϕ_i at each bus i in the area boundary as they require to know the π_k of both corresponding adjacent areas (blue links in Figure 2). As mentioned before, this issue is solved in Section VI-B.

Remark 7 (Estimation of P_i^{in})**.** *One of the limitations of* (18) *is the need to know P_i^{in} for the bus agent to compute λ_i. If D_i is known, then P_i^{in} can be estimated from measurements of the net real-time bus injection $P_i^{in} - D_i\omega_i - d_i$. However, estimating D_i can also be challenging. This problem will be addressed in Section V where we propose a modified control scheme that can achieve the same equilibrium even with approximate estimates of D_i.*

Remark 8. *The procedure described in this section is indepen-*

dent of the constraints (8d)-(8e). *Therefore, such constraints can be generalized to arbitrary linear equality and inequality constraints on the line flows $BC^T\theta$. This property will be exploited in Section VI to further extend our framework.*

IV. Optimality and Convergence

In this section we will show that the system (4) and (18) 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 (4) and (18) is an optimal solution of (9), or equialently (8). 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 $(d(t), \omega(t), \phi(t), P(t), \lambda(t), \pi(t), \rho^+(t), \rho^-(t))$ converges to an equilibrium point of (4) and (18). Moreover, we will show that since $P(0) = BC^T\theta(0)$ (as shown in Section II-A), the line flows will converge to a point that satisfies (5) and (7).

Theorem 9 (Optimality)**.** *A point $p^* := (d^*, \omega^*, \phi^*, P^*, \lambda^*, \pi^*, \rho^{+*}, \rho^{-*})$ is an equilibrium point of* (4) *and* (18) *if and only if $(d^*, \omega^*, \theta^*)$ is an optimal solution of OLC and $(d^*, \omega^*, \phi^*, P^*, \lambda^*, \nu^*, \pi^*, \rho^{+*}, \rho^{-*})$ is a primal-dual optimal solution to the VF-OLC problem, with*

$$\omega^* = \nu^*,\ CP^* = L_B\theta^* \text{ and } C^T\theta^* = C^T\phi^*. \tag{24}$$

Proof: The proof of this theorem is a direct application of lemmas 3 and 4. Let p^* be an equilibrium of (4) and (18). Then, by definition of the projection $[\cdot]_\rho^+$ and (18c)-(18d), $\rho^{+*} \geq 0$ and $\rho^{-*} \geq 0$ and thus dual feasible.

Similarly, since $\dot\omega_i = 0$, $\dot\lambda_i = 0$, $\dot\pi_k = 0$, $\dot\rho_{ij}^+ = 0$ and $\dot\rho_{ij}^- = 0$, then (4a)-(4b) and (18a)-(18d) are equivalent to primal feasibility, i.e. $(d^*, \omega^*, \phi^*, P^*)$ is a feasible point of (9). Finally, by definition of (4) and (18), conditions (11), (12) and (13) are always satisfied by any equilibrium point. Thus we are under the conditions of Lemma 3 and therefore $(d^*, \omega^*, \phi^*, P^*, \lambda^*, \nu^*, \pi^*, \rho^{+*}, \rho^{-*})$ is primal-dual optimal of VF-OLC satisfying (24). Lemma 4 shows the remaining statement of the theorem. ∎

The rest of this section is devoted to showing that in fact for every initial condition $(\omega(0), \phi(0), P(0), \lambda(0), \pi(0), \rho^+(0), \rho^-(0))$, the system (4) and (18) converges to one such optimal solution. Furthermore, we will show that $P(t)$ converges to a P^* that satisfies (5) and (7).

Since we showed in Theorem 6 that (4) and (18) is equivalent to (17), we will provide our convergence result for (17). Our global convergence proof builds on recent results of [39] on global convergence of primal-dual algorithms for network flow control. Our proof extends [39] in the following aspects. Firstly, the Lagrangian $L(x, y)$ is not strictly concave in all of its variables. Secondly, the projection (1) introduces discontinuities in the vector field that prevents the use of the standard LaSalle's Invariance Principle [40].

We solve the latter issue using an invariance principle for Caratheodory systems [41]. We refer the reader to [42] for a detailed treatment that formalizes its application for primal-dual systems. The former issue is solved in Theorem 11 which makes use of the following additional lemma whose proof can be found in the Appendix.

Lemma 10 (Differentiability of $\nu_\mathcal{L}^*(x, y)$)**.** *Given any (x, y), the maximizer of* (16)*, $\nu_\mathcal{L}^*(x, y)$, is continuously differentiable provided $c_i(\cdot)$ is strongly convex. Furthermore, the derivative is given by*

$$\frac{\partial}{\partial x}\nu_\mathcal{L}^*(x,y) = [\overset{\phi}{0} \mid -(D_\mathcal{L} + \overset{P}{d'_\mathcal{L}})^{-1}C_\mathcal{L}]\ \ \nu_\mathcal{L} \tag{25}$$

$$\frac{\partial}{\partial y}\nu_\mathcal{L}^*(x,y) = [-(D_\mathcal{L} + \overset{\lambda_\mathcal{L}}{d'_\mathcal{L}})^{-1}d'_\mathcal{L} \mid \overset{\lambda_\mathcal{G}}{0} \mid \overset{\nu_\mathcal{G}}{0} \mid \overset{\pi}{0} \mid \overset{\rho}{0}]\ \ \nu_\mathcal{L} \tag{26}$$

where $D_S := \mathrm{diag}(D_i)_{i\in S}$, $d'_S = diag(d'_i)$, and $d'_i = d'_i(\lambda_i + \nu_i)$ for $i \in \mathcal{G}$ and $d'_i = d'_i(\lambda_i + \nu_i^*(x, y))$ for $i \in \mathcal{L}$.

We now present our main convergence result. Let E be the set of equilibrium points of (17)

$$E := \left\{(x, y) : \tfrac{\partial L}{\partial x}(x, y) = 0,\ \left[\tfrac{\partial L}{\partial y}(x, y)\right]_\rho^+ = 0\right\},$$

which by theorems 6 and 9 characterizes the set of optimal solutions of the OLC problem.

Theorem 11 (Global Convergence)**.** *The set E of equilibrium points of the primal-dual algorithm* (17) *is globally asymptotically stable. Furthermore, each individual trajectory converges to a point within E that is optimal with respect to the OLC problem.*

Proof: Following [39] we consider the candidate Lyapunov function

$$U(x, y) = \tfrac{1}{2}(x - x^*)^T X^{-1}(x - x^*) + \tfrac{1}{2}(y - y^*)^T Y^{-1}(y - y^*) \tag{27}$$

where $\left(x^* = (\phi^*, P^*), y^* = (\lambda^*, \nu_\mathcal{G}^*, \pi^*, \rho^{+*}, \rho^{-*})\right)$ is *any* equilibrium point of (17).

We divide the proof of this theorem in four steps:

Step 1: We first use the invariance principle for Caratheodory systems [41] to show that $(x(t), y(t))$ converges to the largest invariance set that satisfies $\dot U(x, y) \equiv 0$ between transitions of the projection $[\cdot]_\rho^+$, i.e.

$$(x(t), y(t)) \to \mathcal{M} \subseteq \{(x, y) : \dot U(x(t), y(t)) \equiv 0, t \in \mathbb{R}^+ \backslash \{t_k\}\} \tag{28}$$

where $\{t_k,\ k \in \mathbb{N}\}$ are the time instants when the projection changes between on and off.

Step 2: We show that any invariant trajectory $(x(t), y(t)) \in M$ must have $\lambda(t) \equiv \hat\lambda$ and $\nu(t) \equiv \hat\nu$ for some constant vectors $\hat\lambda$ and $\hat\nu$.

Step 3: We show that whenever $\lambda(t) \equiv \hat\lambda$ and $\nu(t) \equiv \hat\nu$, then the whole trajectory $(x(t), y(t))$ must be an equilibrium point, i.e. $M \subseteq E$.

Step 4: Finally, we show that even though the invariance principle guarantees only convergence to the set E. The convergence is always to some point within E, i.e. $(x(t), y(t)) \to (x^*, y^*) \in E$.

Proof of step 1: Differentiating U over time gives

$$\dot{U}(x,y) = -\frac{\partial L}{\partial x}(x,y)(x-x^*) + \left[\frac{\partial L}{\partial y}(x,y)\right]_\rho^+ (y-y^*) \quad (29)$$

$$\leq \frac{\partial L}{\partial x}(x,y)(x^*-x) + \frac{\partial L}{\partial y}(x,y)(y-y^*) \quad (30)$$

$$\leq L(x^*,y) - L(x,y) + L(x,y) - L(x,y^*) \quad (31)$$

$$= \underbrace{L(x^*,y) - L(x^*,y^*)}_{\leq 0} + \underbrace{L(x^*,y^*) - L(x,y^*)}_{\leq 0} \quad (32)$$

where (29) follows from (17) and (30) from (2). Step (31) follows from convexity (resp. concavity) of $L(x,y)$ in x (resp. y). Finally, equation (32) follows from the saddle property of the equilibrium point (x^*,y^*).

Therefore, since $U(x,y)$ is radially unbounded, the trajectories are bounded, and it follows from the invariance principle for Caratheodory systems [41] that $(x(t),y(t)) \to \mathcal{M}$, i.e. (28) holds. The steps 2 and 3 below basically characterize $\mathcal{M}$.

Proof of step 2: Notice that in order to have $\dot{U} \equiv 0$, both terms in (32) must be zero. In particular, we must have

$$L(x(t),y^*) \equiv L(x^*,y^*).$$

Now, differentiating with respect to time gives

$$0 \equiv \dot{\overline{L(x(t),y^*)}} \equiv \frac{\partial}{\partial x} L(x(t),y^*)\dot{x} \equiv -||\frac{\partial}{\partial x} L(x(t),y^*)||^2,$$

which implies that $\frac{\partial}{\partial x} L(x(t),y^*) \equiv 0$.

Therefore, the fact that $\nu_{\mathcal{G}}^* = 0$, $\frac{\partial}{\partial P} L(x(t),y^*) \equiv 0$, and (22a) holds, implies that $x(t)$ must satisfy $C_{\mathcal{L}}^T \nu_{\mathcal{L}}^*(x(t),y^*) \equiv 0$, which implies that either $\nu_{\mathcal{L}}^*(x(t),y^*) \equiv 0$ (when $C_{\mathcal{L}}$ is full row rank) or $\nu_{\mathcal{L}}^*(x(t),y^*) \equiv \mathbf{1}_n \alpha(t)$ (when $\mathcal{L} = \mathcal{N}$) where $\alpha(t)$ is a time-varying scalar.

We now show that when $\mathcal{L} = \mathcal{N}$ we get $\nu_{\mathcal{L}}^*(x(t),y^*) \equiv \hat{\nu}_{\mathcal{L}}$ for some constant vector $\hat{\nu}_{\mathcal{L}}$. Differentiating $\nu_{\mathcal{L}}^*(x(t),y^*) \equiv \mathbf{1}_n \alpha(t)$ with respect to time and using (25) we obtain

$$(D_{\mathcal{L}} + d'_{\mathcal{L}})^{-1} C_{\mathcal{L}} \dot{P}((t) \equiv \mathbf{1}_n \dot{\alpha}(t)$$

which after left multiplying by $\mathbf{1}_n^T (D_{\mathcal{L}} + d'_{\mathcal{L}})$ gives

$$\mathbf{1}_n^T (D_{\mathcal{L}} + d'_{\mathcal{L}}) \mathbf{1}_n \dot{\alpha}(t) \equiv 0 \implies \dot{\alpha}(t) \equiv 0.$$

Thus, in either case we obtain

$$\nu_{\mathcal{L}}^*(x(t),y^*) = \nu_{\mathcal{L}}^*(C_{\mathcal{L}} P(t), \lambda_{\mathcal{L}}^*) \equiv \hat{\nu}_{\mathcal{L}} \quad (33)$$

for some constant vector $\hat{\nu}_{\mathcal{L}}$, which implies that $C_{\mathcal{L}} P(t) \equiv C_{\mathcal{L}} \hat{P}$ for some constant vector $\hat{P}$.

Therefore, it follows that $\nu_{\mathcal{L}}^*(x(t),y(t))$ must satisfy

$$\nu_{\mathcal{L}}^*(x(t),y(t)) \equiv \nu_{\mathcal{L}}^*(\hat{x},y(t)) \quad (34)$$

for some constant vector $\hat{x}$.

Now, using (20) with (34) we get

$$P_{\mathcal{L}}^m - D_{\mathcal{L}} \nu_{\mathcal{L}}^*(\hat{x},y(t)) - d_{\mathcal{L}}(\lambda_{\mathcal{L}}(t) + \nu_{\mathcal{L}}^*(\hat{x},y(t))) - C_{\mathcal{L}} \hat{P} \equiv 0. \quad (35)$$

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

$$\frac{\partial}{\partial y} L(x^*,y) \left[\frac{\partial}{\partial y} L(x^*,y)^T\right]_\rho^+ \equiv 0. \quad (36)$$

Since the projection $[\cdot]_\rho^+$ only acts on the ρ positions, (36) implies $\frac{\partial}{\partial \nu_{\mathcal{G}}} L(x^*,y) \equiv 0$, $\frac{\partial}{\partial \lambda} L(x^*,y) \equiv 0$ and $\frac{\partial}{\partial \pi} L(x^*,y) \equiv 0$.

Now $\frac{\partial}{\partial \nu_{\mathcal{G}}} L(x^*,y) \equiv 0$ together with equation (23a) implies that

$$P_{\mathcal{G}}^m - D_{\mathcal{G}} \nu_{\mathcal{G}}(t) - d_{\mathcal{G}}(\lambda_{\mathcal{G}}(t) + \nu_{\mathcal{G}}(t)) - C_{\mathcal{G}} P^* \equiv 0, \quad (37)$$

and $\frac{\partial}{\partial \lambda} L(x^*,y) \equiv 0$ with (23b) implies

$$P_{\mathcal{G}}^m - d_{\mathcal{G}}(\lambda_{\mathcal{G}}(t) + \nu_{\mathcal{G}}(t)) - C_{\mathcal{G}} P^* \equiv 0 \quad (38)$$

$$P_{\mathcal{L}}^m - d_{\mathcal{L}}(\lambda_{\mathcal{L}}(t) + \nu_{\mathcal{L}}^*(x^*,y(t))) - C_{\mathcal{L}} P^* \equiv 0 \quad (39)$$

Using (37) and (38) together with the fact that $d_i(\cdot)$ is strictly increasing, we get $\nu_{\mathcal{G}}(t) \equiv \hat{\nu}_{\mathcal{G}}$ and $\lambda_{\mathcal{G}}(t) \equiv \hat{\lambda}_{\mathcal{G}}$, for constant vectors $\hat{\nu}_{\mathcal{G}}$ and $\hat{\lambda}_{\mathcal{G}}$. Moreover, since P^* is primal optimal, Lemma 6 and Theorem 9 imply that $\nu_{\mathcal{G}}(t) \equiv 0$ and $\lambda_{\mathcal{G}}(t) \equiv \lambda_{\mathcal{G}}^*$. Finally, now using (35) together with (39), the same argumentation gives $\nu_{\mathcal{L}}^*(x(t),y(t)) \equiv \hat{\nu}_{\mathcal{L}}$ and $\lambda_{\mathcal{L}}(t) \equiv \hat{\lambda}_{\mathcal{L}}$ for constant vectors $\hat{\nu}_{\mathcal{L}}$ and $\hat{\lambda}_{\mathcal{L}}$. This finishes step 2, i.e. $\lambda(t) \equiv \hat{\lambda}$ and $\nu(t) \equiv \hat{\nu}$.

Proof of step 3: Now, since $\dot{\lambda} \equiv 0$, it follows from (18a) that $C^T \phi(t) \equiv C^T \hat{\phi}$ for some constant vector $\hat{\phi}$ or equivalently $\phi(t) \equiv \hat{\phi} + \beta(t) \mathbf{1}_n$. Differentiating in time $\mathbf{1}_n^T (\chi^\phi)^{-1} \phi(t)$ gives $0 \equiv \mathbf{1}_n^T (\chi^\phi)^{-1} \dot{\phi} \equiv (\sum_{i \in \mathcal{N}} 1/\chi_i^\phi) \dot{\beta}$ which implies that $\beta(t) \equiv \hat{\beta}$ for constant scalar $\hat{\beta}$.

Suppose now that either $\dot{P} \neq 0$ or $\dot{\pi} \neq 0$. Since $C^T \phi(t) \equiv C^T \hat{\phi}$ and $\nu(t) \equiv \hat{\nu}$, $\dot{P}$ and $\dot{\pi}$ are constant. Thus, since the trajectories are bounded, we must have $\dot{P} \equiv 0$ and $\dot{\pi} \equiv 0$; otherwise $U(x,y)$ will grow unbounded (contradiction).

It remains to show that $\dot{\rho} \equiv 0$, i.e. $\dot{\rho}^+ \equiv \dot{\rho}^- \equiv 0$. Since $\phi(t) \equiv \hat{\phi}$, then the argument inside (18c) and (18d) is constant.

Now consider any ρ_e^+, $e = ij \in \mathcal{E}$. Then we have three cases: (i) $B_e(\hat{\phi}_i - \hat{\phi}_j) - \bar{P}_e > 0$, (ii) $B_e(\hat{\phi}_i - \hat{\phi}_j) - \bar{P}_e < 0$ and (iii) $B_e(\hat{\phi}_i - \hat{\phi}_j) - \bar{P}_e = 0$. 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 $\dot{\rho}_e^+ \equiv 0$, and case (iii) also implies $\dot{\rho}_e^+ \equiv 0$. An analogous argument gives $\dot{\rho}^- \equiv 0$. Thus, we have shown that $\mathcal{M} \subseteq E$.

Proof of step 4: We now use structure of $U(x,y)$ to achieve convergence to a single equilibrium. First, since $(x(t),y(t)) \to \mathcal{M}$ and $(x(t),y(t))$ is bounded, then there exists an infinite sequence $\{t_k\}$ such that $(x(t_k),y(t_k)) \to (\hat{x}^*,\hat{y}^*) \in \mathcal{M}$. We choose this specific $(\hat{x}^*,\hat{y}^*)$ in the definition of U. Now, it follows from (32) that $U(x(t),y(t))$ is non-increasing in t and therefore, since $U(x,y)$ is quadratic, it is lower bounded and thus $U(t) \to U^* = 0$ (by the choice of $(x^*,y^*) = (\hat{x}^*,\hat{y}^*)$). Finally, by continuity of $U(x,y)$, $(x(t),y(t)) \to (\hat{x}^*,\hat{y}^*)$.

Thus, it follows that $(x(t),y(t))$ converges to only one optimal solution within $\mathcal{M} \subseteq E$. ■

Finally, the following theorem shows that the system is able to restore the inter-area flows (5) and maintain the line flows within the thermal limits (7).

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* (5)*. Furthermore, if $P(0) = BC^T\theta^0$ for some $\theta^0 \in \mathbb{R}^{|\mathcal{N}|}$, then $P_{ij}^* = B_{ij}(\phi_i^* - \phi_j^*)$ and therefore* (7) *holds.*

Proof: By optimality, P^* and ϕ^* must satisfy

$$P^m - d^* = CP^* = L_B\phi^* = CBC^T\phi^* \tag{40}$$

Therefore using primal feasibility, (6) and (40) we have

$$\hat{P} = \hat{C}BC^T\phi^* = E_{\mathcal{K}}CBC^T\phi^* = E_{\mathcal{K}}CP^* = \hat{C}P^*$$

which is exactly (5).

Finally, to show that $P^*_{ij} = B_{ij}(\phi^*_i - \phi^*_j)$ we will use (4c). Integrating (4c) over time gives

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

Therefore, since $P(t) \to P^*$, we have $P^* = P(0) + BC^T\theta^*$ where θ^* is any finite vector satisfying $C^T\theta^* = \int_0^\infty C^T\nu(s)ds$.

Again by primal feasibility $CBC^T\phi^* = L_B\phi^* = CP^* = C(P(0) + BC^T\theta^*) = CBC^T(\theta^0 + \theta^*)$. Thus, we must have $\phi^* = (\theta^0 + \theta^*) + \alpha\mathbf{1}_n$ and it follows then that $P^* = BC^T(\theta^0 + \theta^*) = BC^T(\phi^* - \alpha\mathbf{1}_n) = BC^T\phi^*$. Therefore, since by primal feasibility $\underline{P} \le BC^T\phi^* \le \bar{P}$, then $\underline{P} \le P^* \le \bar{P}$. ∎

V. Convergence under uncertainty

In this section we discuss an important aspect of the implementation of the control law (18). We provide a modified control law that solves the problem raised in Remark 7, 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. (48)).

We propose an alternative mechanism to compute λ_i. Instead of (18a), we consider the following control law:

$$\dot{\lambda}_i = \zeta_i^\lambda\Big(M_i\dot{\omega}_i + a_i\omega_i + \sum_{e\in\mathcal{E}} C_{i,e}P_e - \sum_{j\in\mathcal{N}_i} B_{ij}(\phi_i - \phi_j)\Big) \tag{41a}$$

where $M_i := 0$ for $i \in \mathcal{L}$ and $a_i \in \mathbb{R}$ is a positive controller parameter that can be arbitrarily chosen. Notice that, while before D_i was an unknown quantity, M_i is usually known and a_i is a design parameter. Furthermore, while equation (41a) requires the knowledge of $\dot{\omega}_i$, this is only needed on generator buses and can therefore be measured from the generator's shaft angular acceleration using one of several existing mechanisms, see e.g. [43].

The parameter a_i plays the role of D_i. In fact, whenever $a_i = D_i$ then one can use (4a)-(4b) to show that (41a) is the same as (18a). More precisely, if we let $a_i = D_i + \delta a_i$, then using (4a)-(4b), (41a) becomes

$$\dot{\lambda} = \zeta_i^\lambda\Big(P_i^{in} - d_i + \delta a_i\omega_i - \textstyle\sum_{j\in\mathcal{N}_i} B_{ij}(\phi_i - \phi_j)\Big), \tag{42}$$

which is equal to (18a) when $\delta a_i = 0$. A simple equilibrium analysis shows that a_i does not affect the steady state behavior provided that $a_i \neq 0$ for some $i \in \mathcal{N}$. Thus, we focus in this section on studying the stability of our modified control law.

Using (42), we can express the new system using

$$\dot{x} = -X\frac{\partial}{\partial x}L(x,y)^T \tag{43a}$$

$$\dot{y} = Y\left[\frac{\partial}{\partial y}L(x,y)^T + g(x,y)\right]_\rho^+ \tag{43b}$$

where $g(x,y) := [(\delta A_{\mathcal{L}}\nu^*_{\mathcal{L}})^T\ (\delta A_{\mathcal{G}}\nu_{\mathcal{G}})^T\ \overset{(\nu_{\mathcal{G}},\pi,\rho)}{0}\ ]^T$, (44)

with $\nu^*_{\mathcal{L}} := \nu^*_{\mathcal{L}}(x,y)$ and $\delta A_S := \mathrm{diag}(\delta a_i)_{i\in S}$.

The system (43) is no longer a primal-dual algorithm. The main result of this section shows, that provided that a_i does not depart significantly from D_i (see (48)), convergence to the optimal solution is preserved.

To show this result, we provide a novel convergence proof that makes use of the following lemmas whose proofs can be found in the Appendix.

Lemma 13 (Second order derivatives of $L(x,y)$)**.** *Whenever Lemma 10 holds, then we have*

$$\frac{\partial^2}{\partial x^2}L(x,y) = \begin{matrix}\phi\\P\end{matrix}\overset{\begin{matrix}\phi & P\end{matrix}}{\begin{bmatrix}0 & 0\\ 0 & C_{\mathcal{L}}^T(D_{\mathcal{L}} + d'_{\mathcal{L}})^{-1}C_{\mathcal{L}}\end{bmatrix}} \quad \textit{and} \tag{45}$$

$$\frac{\partial^2}{\partial y^2}L(x,y) = -\overset{\begin{matrix}\lambda_{\mathcal{L}} & \lambda_{\mathcal{G}} & \nu_{\mathcal{G}} & (\pi,\rho)\end{matrix}}{\begin{bmatrix} D_{\mathcal{L}}(D_{\mathcal{L}} + d'_{\mathcal{L}})^{-1}d'_{\mathcal{L}} & 0 & 0 & 0\\ 0 & d'_{\mathcal{G}} & d'_{\mathcal{G}} & 0\\ 0 & d'_{\mathcal{G}} & (D_{\mathcal{G}} + d'_{\mathcal{G}}) & 0\\ 0 & 0 & 0 & 0\end{bmatrix}}\begin{matrix}\lambda_{\mathcal{L}}\\ \lambda_{\mathcal{G}}\\ \nu_{\mathcal{G}}\\ (\pi,\rho)\end{matrix} \tag{46}$$

with $\frac{\partial^2}{\partial x^2}L(x,y) \succeq 0$ *and* $\frac{\partial^2}{\partial y^2}L(x,y) \preceq 0$.

Lemma 14 (Partial derivatives of $g(x,y)$)**.** *Whenever Lemma 10 holds, then*

$$\frac{\partial}{\partial x}g(x,y) = \overset{\begin{matrix}\phi & P\end{matrix}}{\begin{bmatrix}0 & -\delta A_{\mathcal{L}}(D_{\mathcal{L}} + d'_{\mathcal{L}})^{-1}C_{\mathcal{L}}\\ 0 & 0\end{bmatrix}}\begin{matrix}\lambda_{\mathcal{L}}\\ (\lambda_{\mathcal{G}},\nu_{\mathcal{G}},\pi,\rho)\end{matrix}$$

$$\frac{\partial}{\partial y}g(x,y) = \overset{\begin{matrix}\lambda_{\mathcal{L}} & \lambda_{\mathcal{G}} & \nu_{\mathcal{G}} & (\pi,\rho)\end{matrix}}{\begin{bmatrix}-\delta A_{\mathcal{L}}(D_{\mathcal{L}} + d'_{\mathcal{L}})^{-1}d'_{\mathcal{L}} & 0 & 0 & 0\\ 0 & \delta A_{\mathcal{G}} & 0 & 0\\ 0 & 0 & 0 & 0\end{bmatrix}}\begin{matrix}\lambda_{\mathcal{L}}\\ \lambda_{\mathcal{G}}\\ (\nu_{\mathcal{G}},\pi,\rho)\end{matrix}$$

Unfortunately, the conditions of Theorem 11 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. Therefore the sub-matrix of (46) corresponding to the states λ and $\nu_{\mathcal{G}}$ can become arbitrarily close to singular which makes the system non-robust to perturbations of the form of (44).

This problem is solved by using Assumption 3 of Section III which ensures that $d'_i(\lambda_i + \nu_i)$ is uniformly bounded away from zero. More precisely, using Assumptions 1 and 3 we can show that $\alpha \le c''_i \le L$ which implies

$$\underline{d}' := 1/L \quad \le \quad d'_i = 1/c''_i \quad \le \quad \bar{d}' := 1/\alpha. \tag{47}$$

Theorem 15 (Global convergence of perturbed system)**.** *Whenever assumptions 1, 2 and 3 hold. The system* (43) *converges to a point in the optimal set E for every initial condition whenever*

$$\delta a_i \in 2(\ \underline{d}' - \sqrt{\underline{d}'^2 + \underline{d}'D_{\min}},\ \underline{d}' + \sqrt{\underline{d}'^2 + \underline{d}'D_{\min}}\ ). \tag{48}$$

where $D_{\min} := \min_{i\in\mathcal{N}} D_i$.

Proof: We prove this theorem in three steps:

Step 1: We first show that under the dynamics (43), the time derivative of (27) is upper-bounded by

$$\dot{U}(z) \leq \int_0^1 (z-z^*)^T [H(z(s))](z-z^*)ds \tag{49}$$

where $z = (x, y)$, $z^* = (x^*, y^*)$, $z(s) = z^* + s(z - z^*)$, and $H(z)$ is given by (57).

Step 2: We then show that under the assumption (48) $H(z) \preceq 0$, and that for any $\tilde{z} = (\tilde{\phi}, \tilde{P}, \tilde{\lambda}, \tilde{\nu}_{\mathcal{G}}, \tilde{\pi}, \tilde{\rho}) \in \mathbb{R}^{2|\mathcal{N}|+3|\mathcal{E}|+|\mathcal{G}|+|\mathcal{K}|}$, we have

$$\tilde{z}^T H(z)\tilde{z}=0,\ \forall z \iff \tilde{z} \in \{\tilde{z} \in \mathbb{R}^{\mathrm{Z}} : \tilde{\lambda}=0, \tilde{\nu}_{\mathcal{G}}=0, C_{\mathcal{L}}\tilde{P}=0\} \tag{50}$$

where $\mathrm{Z} = 2|\mathcal{N}| + 3|\mathcal{E}| + |\mathcal{G}| + |\mathcal{K}|$.

Step 3: We finally use (50) and the invariance principle for Caratheodory systems [41] to show that $\nu(t) \equiv 0$ and $\lambda(t) \equiv \lambda^*$.

The rest of the proof follows from **steps 3 and 4** of Theorem 11.

We use $z = (x, y)$ and compactly express (43) using

$$\dot{z} = Z[f(z)]_\rho^+ \tag{51}$$

where $Z = \mathrm{blockdiag}(X, Y)$ and

$$f(z) := \begin{bmatrix} -\frac{\partial}{\partial x}L(x,y)^T \\ \frac{\partial}{\partial y}L(x,y)^T + g(x,y) \end{bmatrix}.$$

Similarly, (27) becomes $U(z) = \frac{1}{2}(z - z^*)^T Z^{-1}(z - z^*)$.

Proof of step 1: We now recompute $\dot{U}(z)$ differenlty, i.e.

$$\dot{U}(z) = \tfrac{1}{2}((z-z^*)^T[f(z)]_\rho^+ + [f(z)]_\rho^{+T}(z-z^*)) \tag{52}$$
$$\leq \tfrac{1}{2}((z-z^*)^T f(z) + f(z)^T(z-z^*)) = (z-z^*)^T f(z) \tag{53}$$
$$= \textstyle\int_0^1 (z-z^*)^T[\frac{\partial}{\partial z}f(z(s))](z-z^*)ds + (z-z^*)^T f(z^*) \tag{54}$$
$$\leq \tfrac{1}{2}\textstyle\int_0^1 (z-z^*)^T\left[\frac{\partial}{\partial z}f(z(s))^T + \frac{\partial}{\partial z}f(z(s))\right](z-z^*)ds \tag{55}$$
$$= \textstyle\int_0^1 (z-z^*)^T[H(z(s))](z-z^*)ds \tag{56}$$

where (52) follows from (51), (53) from (2), and (54) form the fact that $f(z) - f(z^*) = \int_0^1 \frac{\partial}{\partial z}f(z(s))(z - z^*)ds$, where

$$\frac{\partial}{\partial z}f(z) = \begin{bmatrix} -\frac{\partial^2}{\partial x^2}L(x,y) & -\frac{\partial^2}{\partial x \partial y}L(x,y) \\ \frac{\partial^2}{\partial x \partial y}L(x,y)^T & \frac{\partial^2}{\partial y^2}L(x,y) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \frac{\partial}{\partial x}g(x,y) & \frac{\partial}{\partial y}g(x,y) \end{bmatrix}.$$

Finally, (55) 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 (49) can be expressed as

$$H(z) = \frac{1}{2}\left[\frac{\partial}{\partial z}f(z)^T + \frac{\partial}{\partial z}f(z)\right] = \begin{bmatrix} -\frac{\partial^2}{\partial x^2}L(x,y) & 0 \\ 0 & \frac{\partial^2}{\partial y^2}L(x,y) \end{bmatrix} + \begin{bmatrix} 0 & \frac{1}{2}\frac{\partial}{\partial x}g(x,y)^T \\ \frac{1}{2}\frac{\partial}{\partial x}g(x,y) & \frac{1}{2}\left(\frac{\partial}{\partial y}g(x,y)^T + \frac{\partial}{\partial y}g(x,y)\right) \end{bmatrix}$$

which using lemmas 13 and 14 becomes

$$H(z) = \begin{array}{c} \begin{matrix} \phi & (P,\lambda_{\mathcal{L}}) & (\lambda_{\mathcal{G}},\nu_{\mathcal{G}}) & (\pi,\rho) \end{matrix} \\ \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & H_{P,\lambda_{\mathcal{L}}}(z) & 0 & 0 \\ 0 & 0 & H_{\lambda_{\mathcal{G}},\nu_{\mathcal{G}}}(z) & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{array} \begin{matrix} \phi \\ (P,\lambda_{\mathcal{L}}) \\ (\lambda_{\mathcal{G}},\nu_{\mathcal{G}}) \\ (\pi,\rho) \end{matrix} \tag{57}$$

where

$$H_{P,\lambda_{\mathcal{L}}}(z) = \begin{bmatrix} -C_{\mathcal{L}}^T(D_{\mathcal{L}} + d'_{\mathcal{L}})^{-1}C_{\mathcal{L}} & -\frac{1}{2}C_{\mathcal{L}}^T(D_{\mathcal{L}} + d'_{\mathcal{L}})^{-1}\delta A_{\mathcal{L}} \\ -\frac{1}{2}\delta A_{\mathcal{L}}(D_{\mathcal{L}} + d'_{\mathcal{L}})^{-1}C_{\mathcal{L}} & -(D_{\mathcal{L}} + \delta A_{\mathcal{L}})(D_{\mathcal{L}} + d'_{\mathcal{L}})^{-1}d'_{\mathcal{L}} \end{bmatrix}$$

and $\quad H_{\lambda_{\mathcal{G}},\nu_{\mathcal{G}}}(z) = \begin{bmatrix} -d'_{\mathcal{G}} & \frac{1}{2}\delta A_{\mathcal{G}} - d'_{\mathcal{G}} \\ \frac{1}{2}\delta A_{\mathcal{G}} - d'_{\mathcal{G}} & -(d'_{\mathcal{G}} + D_{\mathcal{G}}) \end{bmatrix}.$

It will prove useful in the next step to rewrite $H_{P,\lambda_L}(z)$ using

$$H_{P,\lambda_L}(z) = \tilde{C}^T \tilde{D}^{\frac{1}{2}}(z)\tilde{H}(z)\tilde{D}^{\frac{1}{2}}(z)\tilde{C} \tag{58}$$

where

$$\tilde{C} = \mathrm{blockdiag}(C_{\mathcal{L}}, I),\ \tilde{D}(z) = \mathrm{blockdiag}(D_{\mathcal{L}} + d'_{\mathcal{L}}, D_{\mathcal{L}} + d'_{\mathcal{L}})^{-1}$$

and $\quad \tilde{H}(z) = \begin{bmatrix} -I & -\frac{1}{2}\delta A_{\mathcal{L}} \\ -\frac{1}{2}\delta A_{\mathcal{L}} & -(D_{\mathcal{L}} + \delta A_{\mathcal{L}})d'_{\mathcal{L}} \end{bmatrix}.$

Notice that since $\tilde{D}(z) \succ 0$, $\tilde{D}^{\frac{1}{2}}(z)$ in (58) always exists.

Proof of step 2: To show that $H(z) \preceq 0$ and (50) holds, it is enough to show that

$$\tilde{H}(z) \prec 0 \quad \text{and} \quad H_{\lambda_{\mathcal{G}},\nu_{\mathcal{G}}}(z) \prec 0, \quad \forall z. \tag{59}$$

To see this, assume for now that (59) holds. Then, using (58) it follows that $H_{P,\lambda_{\mathcal{L}}}(z) \preceq 0$, which implies by (57) and $H_{\lambda_{\mathcal{G}},\nu_{\mathcal{G}}}(z) \prec 0$ that $H(z) \preceq 0$. Moreover, $\tilde{z}^T H(z)\tilde{z} = 0\ \forall z$ if and only if

$$[\,\tilde{P}^T\ \tilde{\lambda}_{\mathcal{L}}^T\,]H_{P,\lambda_{\mathcal{L}}}(z)[\,\tilde{P}^T\ \tilde{\lambda}_{\mathcal{L}}^T\,]^T = 0 \tag{60}$$

and

$$[\,\tilde{\lambda}_{\mathcal{G}}^T\ \tilde{\nu}_{\mathcal{G}}^T\,]H_{\lambda_{\mathcal{G}},\nu_{\mathcal{G}}}(z)[\,\tilde{\lambda}_{\mathcal{G}}^T\ \tilde{\nu}_{\mathcal{G}}^T\,]^T = 0. \tag{61}$$

Therefore using (58) it follows that (60) and $\tilde{H}(z) \prec 0\ \forall z$ implies that $C_{\mathcal{L}}\tilde{P} = 0$ and $\tilde{\lambda}_{\mathcal{L}} = 0$. Similarly, $H_{\lambda_{\mathcal{G}},\nu_{\mathcal{G}}}(z) \prec 0$ $\forall z$ and (61) implies $\tilde{\lambda}_{\mathcal{G}} = \tilde{\nu}_{\mathcal{G}} = 0$. This finishes the proof of (50). It remains to show that (59) holds whenever (48) holds.

Proof of $\tilde{H}(z) \prec 0$*:* By definition of negative definite matrices, $\tilde{H}(z) \prec 0$ if and only if all the roots of the characteristic polynomials

$$p_i(\mu_i) = (\mu_i + 1)(\mu_i + (D_i + \delta a_i)d'_i) - \delta a_i^2/4 = \mu_i^2 + (1 + (D_i + \delta a_i)d'_i)\mu_i + (D_i + \delta a_i)d'_i - \delta a_i^2/4$$

are negative for every $i \in \mathcal{L}$ and $\forall z$ (recall d'_i depends on z).

Thus, applying Ruth-Hurwitz stability criterion we get the following necessary and sufficient condition:

$$\delta a_i^2 - 4(D_i + \delta a_i)d'_i < 0 \tag{62a}$$
$$1 + (D_i + \delta a_i)d'_i > 0 \tag{62b}$$

for every $i \in \mathcal{L}$.

Now, equation (62a) can be equivalently rewritten as:

$$2(d_i' - \sqrt{d_i'(d_i'+D_i)}) < \delta a_i < 2(d_i' + \sqrt{d_i'(d_i'+D_i)}). \quad (63)$$

Since $d_i' \in [\underline{d}', \bar{d}']$, $D_i \geq D_{\min}$ and the function $x - \sqrt{x(x+y)}$ is decreasing in both x and y for $x, y \geq 0$, then

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

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

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

Therefore, (62a) holds whenever δa_i satisifes (48).

Finally, (62b) holds whenever $\delta a_i > -\frac{1}{d_i'} - D_i$ which in particular holds if $\delta a_i > -D_{\min}$. The following calculation shows that $2(\underline{d}' - \sqrt{\underline{d}'(\underline{d}'+D_{\min})}) > -D_{\min}$ which implies that (62b) holds under condition (48):

$$2(\underline{d}' - \sqrt{\underline{d}'(\underline{d}'+D_{\min})}) > -D_{\min} \iff$$
$$\underline{d}'(\underline{d}'+D_{\min}) < \underline{d}'^2 + \frac{D_{\min}^2}{4} + \underline{d}' D_{\min} \iff 0 < \frac{D_{\min}^2}{4}.$$

Therefore (62) holds whenever (48) holds.

Proof of $H_{\nu_\mathcal{G},\lambda_\mathcal{G}}(z) \prec 0$: Similarly, we can show that *all* the eigenvalues of $H_{\nu_\mathcal{G},\lambda_\mathcal{G}}(z)$ are the roots of the polynomials

$$\begin{aligned} p_i(\mu_i) &= (\mu_i + D_i + d_i')(\mu_i + d_i') - (\frac{\delta a_i}{2} - d_i')^2 \\ &= \mu_i^2 + (D_i + 2d_i')\mu_i + (D_i + \delta a_i)d_i' - \frac{\delta a_i^2}{4} \end{aligned}$$

which, since $D_i + 2d_i' > 0$, are negative if and only if (62a) is satisfied $\forall i \in \mathcal{G}$. Therefore, (48) also guarantees that $H_{\nu_\mathcal{G},\lambda_\mathcal{G}} \prec 0$.

Proof of step 3: Since by Step 2 $H(z) \preceq 0 \; \forall z$, (49) implies that $\dot{U} \leq 0$ whenever (48) holds. Thus, we are left to apply again the invariance principle for Caratheodory systems [41] and characterize its invariant set $\mathcal{M}$ (28).

Let $\delta z = (z(t) - z^*)$ and similarly define $\delta P = (P(t) - P^*)$, $\delta\lambda_\mathcal{L} = \lambda_\mathcal{L}(t) - \lambda_\mathcal{L}^*$, $\delta\lambda_\mathcal{G} = \lambda_\mathcal{G}(t) - \lambda_\mathcal{G}^*$ and $\delta\nu_\mathcal{G} = \nu_\mathcal{G}(t) - \nu_\mathcal{G}^*$. Then since $\dot{U} \equiv 0$ iff $\delta z^T H(z) \delta z \equiv 0$, then it follows from (50) that $z(t) \in \mathcal{M}$ if and only if $C_\mathcal{L}\delta P \equiv 0$, $\delta\lambda \equiv 0$ and $\delta\nu_\mathcal{G} \equiv 0$.

This implies that $C_\mathcal{L} P(t) \equiv C_\mathcal{L} P^*$, $\lambda(t) \equiv \lambda^*$ and $\nu_\mathcal{G}(t) \equiv \nu_\mathcal{G}^* = 0$, which in particular also implies that $\nu_\mathcal{L}^*(x(t), y(t)) = \nu_\mathcal{L}^*(C_\mathcal{L} P(t), \lambda(t)) \equiv \nu_\mathcal{L}^*(C_\mathcal{L} P^*, \lambda^*) = 0$. Therefore we have shown that $z(t) \in \mathcal{M}$ if and only if $\lambda(t) \equiv \lambda^*$ and $\nu(t) \equiv 0$ which finalizes Step 3.

As mentioned before, the rest of the proof follows from **steps 2 and 3** of Theorem 11. ∎

VI. Framework Extensions

In this section we extend the proposed framework to derive controllers that enhance the solution described before. More precisely, we will show how we can modify our controllers in order to account for buses that have zero power injection (Section VI-A) and how to fully distribute the implementation of the inter-area flow constraints (Section VI-B). Although in principle both extensions could be combined, we present them separately to simplify presentation.

A. Zero Power Injection Buses

We now show how our design framework can be extended to include buses with zero power injection. Let $\mathcal{Z}$ be the set of buses that have neither generators nor loads. Thus, we consider a power network whose dynamics are described by

$$\dot{\theta}_{\mathcal{G}\cup\mathcal{L}} = \omega_{\mathcal{G}\cup\mathcal{L}} \quad (64a)$$
$$M_\mathcal{G}\dot{\omega}_\mathcal{G} = P_\mathcal{G}^{in} - (d_\mathcal{G} + D_\mathcal{G}\omega_\mathcal{G}) - L_{B,(\mathcal{G},\mathcal{N})}\theta \quad (64b)$$
$$0 = P_\mathcal{L}^{in} - (d_\mathcal{L} + D_\mathcal{L}\omega_\mathcal{L}) - L_{B,(\mathcal{L},\mathcal{N})}\theta \quad (64c)$$
$$0 = -L_{B,(\mathcal{Z},\mathcal{N})}\theta \quad (64d)$$

where $L_{B,(S,S')}$ is the sub-matrix of L_B consisting of the rows in S and columns in S'.

We will use Kron reduction to eliminate (64d). Equation (64d) implies that the (θ_i, $i \in \mathcal{Z}$) is uniquely determined by the buses adjacent to $\mathcal{Z}$, i.e. $\theta_\mathcal{Z} = L_{B,(\mathcal{Z},\mathcal{Z})}^{-1} L_{B,(\mathcal{Z},\mathcal{G}\cup\mathcal{L})}\theta_{\mathcal{G}\cup\mathcal{L}}$. Thus we can rewrite (64) using only $\theta_{\mathcal{G}\cup\mathcal{L}}$ which gives

$$\dot{\theta}_{\mathcal{G}\cup\mathcal{L}} = \omega_{\mathcal{G}\cup\mathcal{L}} \quad (65a)$$
$$M_\mathcal{G}\dot{\omega}_\mathcal{G} = P_\mathcal{G}^{in} - (d_\mathcal{G} + D_\mathcal{G}\omega_\mathcal{G}) - L_{B,(\mathcal{G},\mathcal{G}\cup\mathcal{L})}^\sharp\theta_{\mathcal{G}\cup\mathcal{L}} \quad (65b)$$
$$0 = P_\mathcal{L}^{in} - (d_\mathcal{L} + D_\mathcal{L}\omega_\mathcal{L}) - L_{B,(\mathcal{L},\mathcal{G}\cup\mathcal{L})}^\sharp\theta_{\mathcal{G}\cup\mathcal{L}} \quad (65c)$$

where $L_B^\sharp = L_{B,(\mathcal{G}\cup\mathcal{L},\mathcal{G}\cup\mathcal{L})} - L_{B,(\mathcal{G}\cup\mathcal{L},\mathcal{Z})} L_{B,(\mathcal{Z},\mathcal{Z})}^{-1} L_{B,(\mathcal{Z},\mathcal{G}\cup\mathcal{L})}$ is the Schur complement of L_B after removing the rows and columns corresponding to $\mathcal{Z}$. The matrix $L_B^\sharp$ is also a Laplacian of a reduced graph $G(\mathcal{G}\cup\mathcal{L}, \mathcal{E}^\sharp)$ and therefore it can be expressed as $L_B^\sharp = C^{\sharp T} B^\sharp C^{\sharp T}$ where $C^\sharp$ is the incidence matrix of $G(\mathcal{G}\cup\mathcal{L}, \mathcal{E}^\sharp)$ and $B^\sharp = \mathrm{diag}(B_{ij}^\sharp)_{ij\in\mathcal{E}^\sharp}$ are the line susceptances of the reduced network.

This reduction allows to use (65) (which is equivalent to (3)) to also model networks that contain buses with zero power injection. The only caveat is that some of line flows of the vector $BC^T\theta$ are no longer present in $B^\sharp C^{\sharp T}\theta_{\mathcal{G}\cup\mathcal{L}}$ – when a bus is eliminated using Kron reduction, its adjacent lines B_e, $e \in \mathcal{E}$, are substituted by an equivalent clique with new line impedances $B_{e'}^\sharp$, $e' \in \mathcal{E}^\sharp$. As a result, some of the constraints (8d)-(8e) would no longer have a physical meaning if we directly substitute $BC^T\theta$ with $B^\sharp C^{\sharp T}\theta_{\mathcal{G}\cup\mathcal{L}}$ in (3).

We overcome this issue by showing that each original $B_{ij}(\theta_i - \theta_j)$ in $G(\mathcal{N}, \mathcal{E})$ can be replaced by a linear combination of line flows $B_{i'j'}^\sharp(\theta_{i'} - \theta_{j'})$ of the reduced network $G(\mathcal{G}\cup\mathcal{L}, \mathcal{E}^\sharp)$.

For any θ satisfying (64d) we have

$$L_B\theta = \begin{bmatrix} q_{\mathcal{G}\cup\mathcal{L}} \\ \mathbf{0}_{|\mathcal{Z}|} \end{bmatrix} = \begin{bmatrix} L_B^\sharp\theta_{\mathcal{G}\cup\mathcal{L}} \\ \mathbf{0}_{|\mathcal{Z}|} \end{bmatrix}.$$

Thus it follows that

$$\begin{aligned} BC^T\theta &= BC^T L_B^\dagger \begin{bmatrix} q_{\mathcal{G}\cup\mathcal{L}} \\ \mathbf{0}_{|\mathcal{Z}|} \end{bmatrix} = BC^T L_{B,(\mathcal{N},\mathcal{G}\cup\mathcal{L})}^\dagger q_{\mathcal{G}\cup\mathcal{L}} \\ &= BC^T L_{B,(\mathcal{N},\mathcal{G}\cup\mathcal{L})}^\dagger C^\sharp B^\sharp C^{\sharp T}\theta_{\mathcal{G}\cup\mathcal{L}} := A^\sharp B^\sharp C^{\sharp T}\theta_{\mathcal{G}\cup\mathcal{L}} \end{aligned} \quad (66)$$

where $L_B^\dagger$ is the pseudo-inverse of L_B.

Therefore, by substituting $BC^T\theta$ with $A^\sharp B^\sharp C^{\sharp T}\theta_{\mathcal{G}\cup\mathcal{L}}$ in (8) and repeating the procedure of Section III we obtained a modified version of (18) in which (18a)-(18e) becomes

$$\dot{\lambda}_i = \zeta_i^\lambda \Big(P_i^{in} - d_i - \textstyle\sum_{j\in\mathcal{N}_i^\sharp} B_{ij}^\sharp(\phi_i - \phi_j) \Big) \quad (67a)$$

$$\dot{\pi}_k = \zeta_k^{\pi}\left(\sum_{e\in\mathcal{B}_k, ij\in\mathcal{E}^\sharp} \hat{C}_{k,e}A^\sharp_{e,ij}B^\sharp_{ij}(\phi_i-\phi_j) - \hat{P}_k\right) \quad (67b)$$

$$\dot{\rho}_e^+ = \zeta_e^{\rho^+}\left[\sum_{ij\in\mathcal{E}^\sharp} A^\sharp_{e,ij}B^\sharp_{ij}(\phi_i-\phi_j) - \overline{P_e}\right]^+_{\rho_e^+} \quad (67c)$$

$$\dot{\rho}_e^- = \zeta_e^{\rho^-}\left[\underline{P_e} - \sum_{ij\in\mathcal{E}^\sharp} A^\sharp_{e,ij}B^\sharp_{ij}(\phi_i-\phi_j)\right]^+_{\rho_e^-} \quad (67d)$$

$$\dot{\phi}_i = \chi_i^{\phi}\Big(\sum_{j\in\mathcal{N}_i^\sharp} B^\sharp_{ij}(\lambda_i-\lambda_j) - \sum_{e\in\mathcal{E}} C^\sharp_{i,e}B^\sharp_e \sum_{e'\in\mathcal{E},\, k\in\mathcal{K}} A^\sharp_{e,e'}\hat{C}_{k,e'}\pi_k$$
$$-\sum_{e\in\mathcal{E}} C^\sharp_{i,e}B^\sharp_e \sum_{e'\in\mathcal{E}} A^\sharp_{e,e'}(\rho^+_{e'}-\rho^-_{e'})\Big) \quad (67e)$$

where (67a) and (67e) are for $i\in\mathcal{G}\cup\mathcal{L}$, (67b) is for $k\in\mathcal{K}$, and (67c) and (67d) are for the original lines $e\in\mathcal{E}$.

It can be shown that the analysis described in Sections IV and V still holds under this extension.

Remark 16. *The only additional overhead incurred by the proposed extension is the need for communication between buses that are adjacent on the graph $G(\mathcal{G}\cup\mathcal{L},\mathcal{E}^\sharp)$ and were not adjacent in $G(\mathcal{N},\mathcal{E})$ (see Figure 3 for an illustration).*

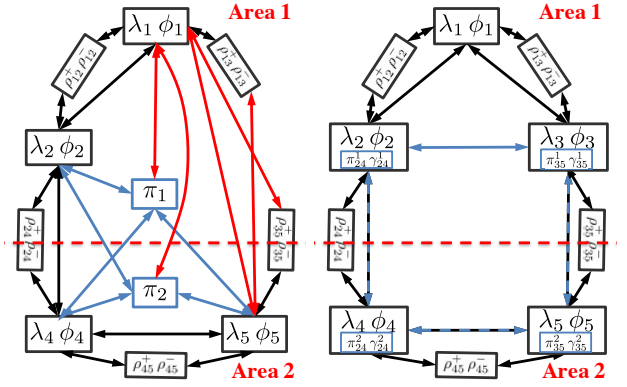


Fig. 3: Communication requirements for the power network in Fig. 2. Left side for the case when bus 3 has no injection (Section VI-A), and right side for the distributed inter-area flow constraint formulation (Section VI-B).

B. Distributed Inter-area Flow Constraints

We now show how we can fully distribute the implementation of the inter-area flow constraints. The procedure is analogous to Section VI-A and therefore we will only limit to describe what are the modifications that need to be done to (8) in order to obtain controllers that are fully distributed.

We define for each area k an additional graph $G(\mathcal{B}_k,\mathcal{E}^k)$ where we associate each boundary edge $e\in\mathcal{B}_k$ with a node and define *undirected* edges $\{e,e'\}\in\mathcal{E}^k$ that describe the communication links between e and e'. Using this formulation, we decompose equation (5) for each k into $|\mathcal{B}_k|$ equations

$$\hat{C}_{k,e}P_e - \frac{\hat{P}_k}{|\mathcal{B}_k|} = \sum_{e':\{e,e'\}\in\mathcal{B}_k}(\gamma_e-\gamma_{e'}),\quad e\in\mathcal{B}_k,\, k\in\mathcal{K} \quad (68)$$

where γ_e is a new primal variable that aims to guarantee indirectly (5). In fact, it is easy to see by summing (68) over $e\in\mathcal{B}_k$ that

$$\sum_{e\in\mathcal{B}_k}\left(\hat{C}_{k,e}P_e - \frac{\hat{P}_k}{|\mathcal{B}_k|}\right) = \sum_{e\in\mathcal{B}_k}\sum_{e':\{e,e'\}\in\mathcal{B}_k}(\gamma_e^k-\gamma_{e'}^k) = 0.$$

which is equal to (5).

Therefore, since whenever (5) holds, one can find a set of γ_e satisfying (68), then we can substitute (8d) with (68). If we let π_{ij}^k be the Lagrange multiplier associated with (68), then by replacing (18b) and (18e) with

$$\dot{\pi}_{ij}^k = \zeta_{k,ij}^{\pi}\Big(\hat{C}_{k,ij}B_{ij}(\phi_i-\phi_j) - \frac{\hat{P}_k}{|\mathcal{B}_k|} - \sum_{e:\{ij,e\}\in\mathcal{B}_k}\gamma_{ij}^k-\gamma_e^k\Big) \quad (69a)$$

$$\dot{\gamma}_{ij}^k = \chi_{k,ij}^{\gamma}\Big(\sum_{e:\{ij,e\}\in\mathcal{B}_k}\pi_{ij}^k-\pi_e^k\Big) \quad (69b)$$

$$\dot{\phi}_i = \chi_i^{\phi}\Big(\sum_{j\in\mathcal{N}_i} B_{ij}(\lambda_i-\lambda_j) - \sum_{k\in\mathcal{K},\, e\in\mathcal{B}_k} C_{i,e}B_e\hat{C}_{k,e}\pi_e^k$$
$$-\sum_{e\in\mathcal{E}} C_{i,e}B_e(\rho_e^+-\rho_e^-)\Big) \quad (69c)$$

we can distribute the implementation of the inter-area flow constraint. Figure 3 shows how the communication requirements are modified by this change. In particular, since π_e^k and γ_e^k can be co-located and computed together with λ_i and ϕ_i, where i denotes the bus of area k adjacent to the tie-line e, many of the communication links used for λ_i and ϕ_i can be reused. It can be shown that the *additional* communication links required to implement the distributed version of the inter-area flow constraints is *at most* $|\mathcal{B}_k|-1$ per area, while for the centralized solution this number is always $2|\mathcal{B}_k|$ per area. Finally, if each boundary bus has only one incident boundary edge, i.e. if $\sum_{k\in\mathcal{K},\, e\in\mathcal{B}_k}C_{i,e}B_e\hat{C}_{k,e}\pi_e^k$ has at most one term, the convergence results of sections IV and V extend to this case.

VII. Numerical Illustrations

We now illustrate the behavior of our control scheme. We consider the widely used IEEE 39 bus system, shown in Figure 4, 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 [44] data set. Each bus is assumed to have a controllable load with $\mathcal{D}_i=[-d_{\max},d_{\max}]$, with $d_{\max}=1$p.u. on a 100MVA base with $c_i(\cdot)$ and the corresponding $d_i(\cdot)=c_i'^{-1}(\cdot)$ as shown in Figure 5.

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 $\chi_i^{\phi}=\zeta_i^{\lambda}=\zeta_k^{\pi}=\zeta_e^{\rho^+}=\zeta_e^{\rho^-}=1$, for all $i\in\mathcal{N}$, $k\in\mathcal{K}$ 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^{in} are corrected so that they initially add up to zero by evenly distributing the mismatch among the load buses. $\hat{P}$ is obtained from the starting stationary condition. We initially set $\overline{P}$ and $\underline{P}$ so that they are not binding.

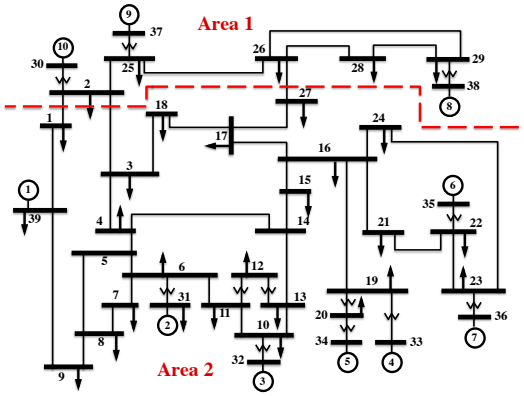


Fig. 4: IEEE 39 bus system: New England

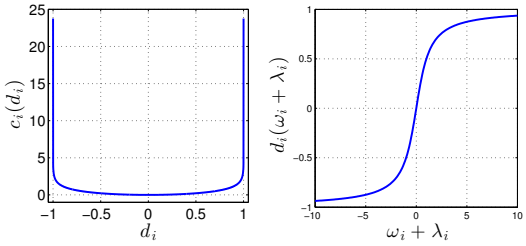


Fig. 5: Disutility $c_i(d_i)$ and corresponding load function $d_i(\omega_i+\lambda_i) = c_i'^{-1}(\omega_i+\lambda_i)$

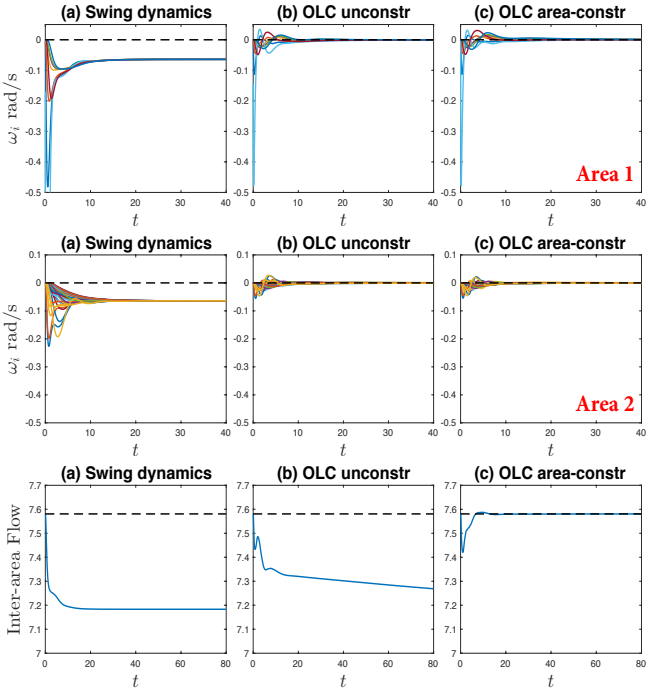


Fig. 6: Areas Frequencies and Aggregate Inter-area Flow

We simulate the OLC-system as well as the swing dynamics (4) without load control ($d_i = 0$), after introducing a perturbation at bus 29 of $P_{29}^{in} = -2$p.u.. In some scenarios we disable a few of the OLC constraints. This is achieved by fixing the corresponding Lagrange multiplier to be zero.

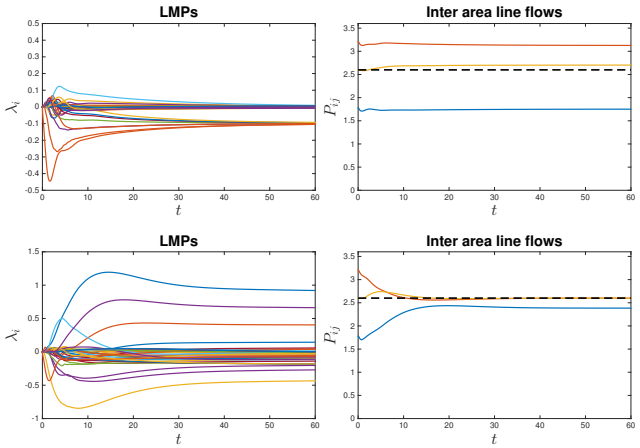


Fig. 7: LMPs and inter area line flows: without thermal limits (top), with thermal limits (bottom)

Figure 6 shows the evolution of the bus frequencies and the inter-area flow 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 (a), the OLC controllers can jointly rebalance the power as well as recovering the nominal frequency (b and c). The frequency stabilization when using OLC seems to be similar or even better than the swing dynamics. Figure 6 shows that, interestingly, even in the case where the inter-area flow constraint is not active (b) the inter-area flow takes longer to settle to the new value. This has a smoothing effect that makes the transition of the power flows to the new steady-state less sudden.

Now, we illustrate the action of the thermal constraints by adding a constraint of $\overline{P}_e = 2.6$p.u. and $\underline{P}_e = -2.6$p.u. to the tie lines between areas. Figure 7 (top) shows the values of the multipliers λ_i, that correspond to the Locational Marginal Prices (LMPs), and the line flows of the tie lines for the same scenario displayed in Figure 6 (c), i.e. without thermal 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 (bottom of Figure 7), we can see that the system converges to a new operating point that satisfies our constraints.

Finally, we show the conservativeness of the bound obtained in Theorem 15. We simulate the perturbed system (4), (41a) and (18b)-(18f) under the same conditions as in Figure 7 (top), i.e., without enforcing thermal limits. We set the scalars a_is such that the corresponding δa_is are homogeneous for every bus i. We also do not enforce Assumption 3 and use the same d_i as described in Figure 5. This implies that (48) in Theorem 15 is not satisfied (because $\underline{d}' = 0$).

Figure 8 shows the evolution of the frequency ω_i and LMPs λ_i for different values of δa_i belonging to $\{-0.4, -0.21, -0.2, -0.19, 0.0\}$. Since $D_i = 0.2$ at all the buses, then $\delta a_i = -0.2$ is the threshold that makes a_i go from positive to negative as δa_i decreases. Even though condition (48) is not satisfied for any δa_i, our simulations show that

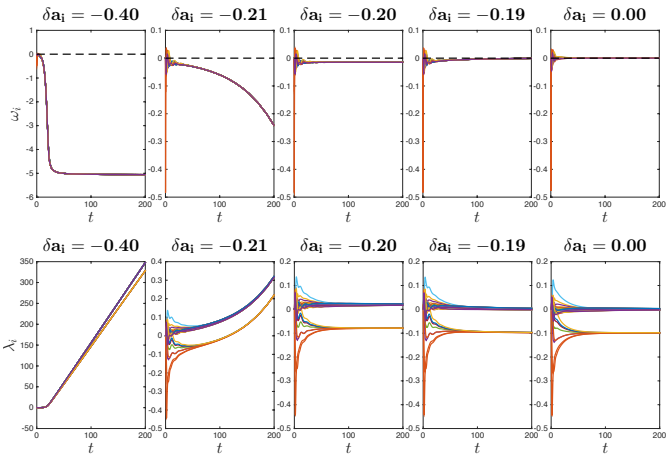


Fig. 8: Frequency and Location Marginal Prices evolution for homogeneous perturbation $\delta a_i \in \{-0.4, -0.21, -0.2, -0.19, 0\}$

the system converges whenever $a_i \geq 0$ ($\delta a_i \geq -0.2$). The case when $\delta a_i = -0.2$ is of special interest. Here, the system converges, yet the nominal frequency is not restored. This is because the terms $\delta a_i \omega_i$ (42) are equal to the terms $D_i \omega_i$ in (4a)-(4b). Thus $\dot{\omega}_i$ and $\dot{\lambda}_i$ can be made simultaneously zero with nonzero ω_i^*. Fortunately, this can only happen when $a_i = 0 \ \forall i$ which can be avoided since a_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.

Appendix

A. Proof of Lemma 3

Proof: Assumptions 1 and 2 guarantee that the solution to the primal (OLC) is finite. Moreover, since by Assumption 2 there is a feasible $d \in \operatorname{Int} \mathcal{D}$, then the Slater condition is satisfied [33] and there is zero duality gap.

Thus, since OLC only has linear equality constraints, we can use Karush-Kuhn-Tucker (KKT) conditions [33] to characterize the primal dual optimal solution. Thus $(d^*, \omega^*, P^*, \phi^*, \sigma^*)$ is primal dual optimal if and only if we have:

(i) Primal and dual feasibility: (9b)-(9e) and $\rho^{+*}, \rho^{-*} \geq 0$.
(ii) Stationarity:
$\frac{\partial}{\partial d} L(d^*, \omega^*, x^*, \sigma^*) = 0$, $\frac{\partial}{\partial \omega} L(d^*, \omega^*, x^*, \sigma^*) = 0$ and $\frac{\partial}{\partial x} L(d^*, \omega^*, x^*, \sigma^*) = 0$.
(iii) Complementary slackness:

$$\rho_{ij}^{+*}(B_{ij}(\phi_i^* - \phi_j^*) - \bar{P}_{ij}) = 0, \qquad ij \in \mathcal{E};$$
$$\rho_{ij}^{-*}(\underline{P}_{ij} - B_{ij}(\phi_i^* - \phi_j^*)) = 0, \qquad ij \in \mathcal{E}.$$

KKT conditions (i) and (iii) are already implicit by assumptions of the lemma.

The stationarity condition (ii) is given by

$$\frac{\partial L}{\partial d_i}(d^*, \omega^*, P^*, \phi^*, \sigma^*) = c_i'(d_i^*) - (\nu_i^* + \lambda_i^*) = 0 \quad (70a)$$
$$\frac{\partial L}{\partial \omega_i}(d^*, \omega^*, P^*, \phi^*, \sigma^*) = D_i(\omega_i^* - \nu_i^*) = 0 \quad (70b)$$
$$\frac{\partial L}{\partial P_{ij}}(d^*, \omega^*, P^*, \phi^*, \sigma^*) = \nu_j^* - \nu_i^* = 0 \quad (70c)$$
$$\frac{\partial L}{\partial \phi_i}(d^*, \omega^*, P^*, \phi^*, \sigma^*) = \textstyle\sum_{j \in \mathcal{N}_i} B_{ij}(\lambda_j^* - \lambda_i^*)$$
$$+ \textstyle\sum_{e \in \mathcal{E}} C_{i,e} B_e (\sum_{k \in \mathcal{K}} \hat{C}_{k,e} \pi_k^* + \rho_e^{+*} - \rho_e^{-*}) = 0 \quad (70d)$$

Since $D_i > 0$ equation (70b) implies $\nu_i^* = \omega_i^*$. Thus, (70b) and (70a) amount to the first and second conditions of (11). Furthermore, since the graph G is connected then (70c) is equivalent to $\nu_i^* = \hat{\nu} \ \forall i \in \mathcal{N}$ which is the third condition of (11).

Since $c_i(d_i)$ and $\frac{D_i \omega_i^2}{2}$ are strictly convex functions, it is easy to show that ν_i^* and λ_i^* are unique. To show $\hat{\nu} = 0$ we use (i). Adding (9b) over $i \in \mathcal{N}$ gives

$$\begin{aligned} 0 &= \textstyle\sum_{i \in \mathcal{N}} \left(P_i^{in} - (d_i^* + D_i \omega_i^*) - \sum_{e \in \mathcal{E}} C_{ie} P_e\right) \\ &= \textstyle\sum_{i \in \mathcal{N}} \left(P_i^{in} - (d_i^* + D_i \omega_i^*)\right) - \sum_{e = ij \in \mathcal{E}} \left(C_{ie} P_e + C_{je} P_e\right) \\ &= \textstyle\sum_{i \in \mathcal{N}} \left(P_i^{in} - (d_i^* + D_i \omega_i^*)\right) \end{aligned} \quad (71)$$

and similarly (9c) gives

$$0 = \textstyle\sum_{i \in \mathcal{N}} P_i^{in} - d_i^* \quad (72)$$

Thus, subtracting (71) from (72) gives

$$0 = \textstyle\sum_{i \in \mathcal{N}} D_i \omega_i^* = \sum_{i \in \mathcal{N}} D_i \nu_i^* = \hat{\nu} \sum_{i \in \mathcal{N}} D_i$$

and since $D_i > 0 \ \forall i \in \mathcal{N}$, it follows that $\hat{\nu} = 0$. ∎

B. Proof of Lemma 4

Proof: Let $(d^* \omega^* = 0, \theta^*)$ be an optimal solution of OLC. Then, by letting $\phi^* = \theta^*$ and $P^* = BC^T \theta^*$, it follows that $(d^*, \omega^* = 0, \phi^*, P^*)$ is a feasible solution of VF-OLC. Suppose that $(d^*, \omega^*, \phi^*, P^*)$ is not optimal with respect to VF-OLC, then the solution $(\hat{d}^*, \hat{\omega}^*, \hat{\phi}^*, \hat{P}^*)$ of VF-OLC has strictly lower cost $\sum_{i \in \mathcal{N}} c_i(\hat{d}_i^*) + \frac{D_i \hat{\omega}_i^{*2}}{2} < \sum_{i \in \mathcal{N}} c_i(d_i^*)$. By Lemma 3 we have that $\hat{\omega}^* = 0$. Then, it follows that by setting $\hat{\theta}^* = \hat{\phi}^*$, $(\hat{d}^*, \hat{\omega}^*, \hat{\theta}^*)$ is a feasible solution of OLC with strictly lower cost than the supposedly optimal $(d^*, \omega^*, \theta^*)$. Contradiction. Therefore $(\hat{d}^*, \hat{\omega}^*, \hat{\phi}^*, \hat{P}^*)$ is an optimal solution of VF-OLC. The converse is shown analogously. ∎

C. Proof of Lemma 5

Proof: A straightforward differentiation shows that the Hessian of $\Phi_i(\nu_i, \lambda_i)$ is given by

$$\frac{\partial^2}{\partial(\nu_i, \lambda_i)^2} \Phi_i(\nu_i, \lambda_i) = \begin{matrix} \\ \nu_i \\ \lambda_i \end{matrix} \overset{\begin{matrix} \nu_i & \lambda_i \end{matrix}}{\begin{bmatrix} -(d_i' + D_i) & -d_i' \\ -d_i' & -d_i' \end{bmatrix}} \quad (73)$$

where d_i' is short for $d_i'(\lambda_i+\nu_i)$ and denotes de derivative of $d_i(\cdot)=c_i^{'-1}(\cdot)$ with respect to its argument.

Since c_i is strictly convex $d_i'>0$. Thus, since $D_i>0$, (73) is negative definite which implies that $\Phi_i(\nu_i,\lambda_i)$ is strictly concave. Finally, it follows from (14) that $L(x,\sigma)$ is strictly concave in (ν,λ). ∎

D. Proof of Lemma 10

Proof: We first notice that $\nu_i^*(x,y)$, $i\in\mathcal{L}$, depends only on λ_i and $C_iP:=\sum_{e\in E}C_{i,e}P_e$. Which means that $\frac{\partial}{\partial\phi}\nu_{\mathcal{L}}^*(x,y)=0$, $\frac{\partial}{\partial\nu_{\mathcal{G}}}\nu_{\mathcal{L}}^*(x,y)=0$, $\frac{\partial}{\partial\pi}\nu_{\mathcal{L}}^*(x,y)=0$, $\frac{\partial}{\partial\rho}\nu_{\mathcal{L}}^*(x,y)=0$ and $\frac{\partial}{\partial\lambda_{\mathcal{L}}}\nu_{\mathcal{L}}^*(x,y)$ is diagonal.

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

$$\begin{aligned}0=\frac{\partial}{\partial\nu_i}L(x,y,\nu_{\mathcal{L}}^*(x,y))=&P_i^{in}-D_i\nu_i^*(x,y)\\&-d_i(\lambda_i+\nu_i^*(x,y))-\textstyle\sum_{e\in E}C_{i,e}P_e\end{aligned}\tag{74}$$

Therefore, if we fix P and take the total derivative of $\frac{\partial}{\partial\nu_i}L(x,y,\nu_{\mathcal{L}}^*(x,y))$ with respect to λ_i we obtain

$$0=\frac{d}{d\lambda_i}\left(\frac{\partial}{\partial\nu_i}L(x,y,\nu_{\mathcal{L}}^*(x,y))\right)\tag{75}$$

$$=-(D_i+d_i'(\lambda_i+\nu_i^*))\frac{\partial}{\partial\lambda_i}\nu_i^*-d_i'(\lambda_i+\nu_i^*)\tag{76}$$

where here we used ν_i^* for short of $\nu_i^*(x,y)$.

Now since by assumption $c_i(\cdot)$ is strongly convex, i.e. $c_i''(\cdot)\geq\alpha$, $d_i'(\cdot)=\frac{1}{c_i''(\cdot)}\leq\frac{1}{\alpha}<\infty$. Thus, (D_i+d_i') is finite and strictly positive, which implies that

$$\frac{\partial}{\partial\lambda_i}\nu_i^*(x,y)=-\frac{d_i'(\lambda_i+\nu_i^*(x,y))}{(\,D_i+d_i'(\lambda_i+\nu_i^*(x,y))\,)},\quad i\in\mathcal{L}.$$

Similarly, we obtain

$$\frac{\partial}{\partial P}\nu_i^*(x,y)=-\frac{1}{(\,D_i+d_i'(\lambda_i+\nu_i^*(x,y))\,)}C_i,\quad i\in\mathcal{L}.$$

where C_i is the ith row of C.

Finally, notice that whenever $d_i'(\lambda_i+\nu_i^*)$ exists, then $\frac{\partial}{\partial x}\nu_i^*$ and $\frac{\partial}{\partial y}\nu_i^*$ also exists. ∎

E. Proof of Lemma 13

Proof: Using the Envelope Theorem [38] in (16) we have

$$\frac{\partial L}{\partial x}(x,y)=\frac{\partial L}{\partial x}(x,y,\nu_{\mathcal{L}}^*(x,y))$$

which implies that

$$\begin{aligned}\frac{\partial^2L}{\partial x^2}(x,y)&=\frac{\partial}{\partial x}\left[\frac{\partial L}{\partial x}(x,y,\nu_{\mathcal{L}}^*(x,y))\right]\\&=\frac{\partial^2L}{\partial x^2}(x,y,\nu_{\mathcal{L}}^*(x,y))+\frac{\partial^2L}{\partial x\partial\nu_{\mathcal{L}}}(x,y,\nu_{\mathcal{L}}^*(x,y))\frac{\partial}{\partial x}\nu_{\mathcal{L}}^*(x,y)\\&=\frac{\partial^2L}{\partial x\partial\nu_{\mathcal{L}}}(x,y,\nu_{\mathcal{L}}^*(x,y))\frac{\partial}{\partial x}\nu_{\mathcal{L}}^*(x,y).\end{aligned}\tag{77}$$

where the last step follows from $L(x,\sigma)$ being linear in x.

Now, by definition of $\nu_{\mathcal{L}}^*(x,y)$ it follows that

$$\frac{\partial L}{\partial\nu_{\mathcal{L}}}(x,y,\nu_{\mathcal{L}}^*(x,y))=0.\tag{78}$$

Differentiating (78) with respect to x gives

$$0=\frac{\partial^2L}{\partial\nu_{\mathcal{L}}\partial x}(x,y,\nu_{\mathcal{L}}^*(x,y))+\frac{\partial^2L}{\partial\nu_{\mathcal{L}}^2}(x,y,\nu_{\mathcal{L}}^*(x,y))\frac{\partial}{\partial x}\nu_{\mathcal{L}}^*(x,y)$$

and therefore

$$\begin{aligned}\frac{\partial^2L}{\partial x\partial\nu_{\mathcal{L}}}(x,y,\nu_{\mathcal{L}}^*(x,y))&=\left[\frac{\partial^2L}{\partial\nu_{\mathcal{L}}\partial x}(x,y,\nu_{\mathcal{L}}^*(x,y))\right]^T\\&=-\frac{\partial}{\partial x}\nu_{\mathcal{L}}^*(x,y)^T\frac{\partial^2L}{\partial\nu_{\mathcal{L}}^2}(x,y,\nu_{\mathcal{L}}^*(x,y)).\end{aligned}\tag{79}$$

Substituting (79) into (77) gives

$$\frac{\partial^2L}{\partial x^2}(x,y)=-\frac{\partial}{\partial x}\nu_{\mathcal{L}}^*(x,y)^T\frac{\partial^2L}{\partial\nu_{\mathcal{L}}^2}(x,y,\nu_{\mathcal{L}}^*(x,y))\frac{\partial}{\partial x}\nu_{\mathcal{L}}^*(x,y).\tag{80}$$

It follows from (20) and (15) that

$$\frac{\partial^2L}{\partial\nu_{\mathcal{L}}^2}(x,y,\nu_{\mathcal{L}}^*(x,y))=\frac{\partial^2\Phi_{\mathcal{L}}}{\partial\nu_{\mathcal{L}}^2}(\nu_{\mathcal{L}}^*(x,y),\lambda_{\mathcal{L}})=-(D_{\mathcal{L}}+d_{\mathcal{L}}').\tag{81}$$

Therefore, substituting (25) and (81) into (80) gives (45).

A similar calculation using (26) gives (46). ∎

F. Proof of Lemma 14

Proof: By definition of $g(x,y)$ we have

$$\frac{\partial}{\partial x}g(x,y)=\;[(\delta A_{\mathcal{L}}\overset{\lambda_{\mathcal{L}}}{\frac{\partial}{\partial x}}\nu_{\mathcal{L}}^*)^T\;(\delta A_{\mathcal{G}}\overset{\lambda_{\mathcal{G}}}{\frac{\partial}{\partial x}}\nu_{\mathcal{G}})^T\quad\overset{(\nu_{\mathcal{G}},\,\pi,\,\rho)}{\frac{\partial}{\partial x}0}\;]^T.$$

Thus, using Lemma 10 we obtain $\frac{\partial}{\partial x}g(x,y)$. A similar computation gives $\frac{\partial}{\partial y}g(x,y)$. ∎

Acknowledgment

The authors would like to thank the associate editor and the anonymous reviewers whose valuable comments and suggestions considerably improved the manuscript.

References

[1] A. J. Wood, B. F. Wollenberg, and G. B. Sheble, "Power generation, operation and control," *John Wiley&Sons*, 1996.

[2] A. R. Bergen and V. Vittal, *Power Systems Analysis*, 2nd ed. Prentice Hall, 2000.

[3] J. Machowski, J. Bialek, and J. R. Bumby, *Power System Dynamics and Stability*. John Wiley & Sons, Oct. 1997.

[4] M. D. Ilic, "From Hierarchical to Open Access Electric Power Systems," *Proceedings of the IEEE*, vol. 95, no. 5, pp. 1060–1084, 2007.

[5] A. K. Bejestani, A. Annaswamy, and T. Samad, "A Hierarchical Transactive Control Architecture for Renewables Integration in Smart Grids: Analytical Modeling and Stability," *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 2054–2065, Jul. 2014.

[6] F. C. Schweppe, R. D. Tabors *et al.*, "Homeostatic Utility Control," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-99, no. 3, pp. 1151–1163, May 1980.

[7] D. Trudnowski, M. Donnelly, and E. Lightner, "Power-system frequency and stability control using decentralized intelligent loads," in *Transmission and Distribution Conference and Exhibition, 2005/2006 IEEE PES*, May 2006, pp. 1453–1459.

[8] N. Lu and D. Hammerstrom, "Design considerations for frequency responsive grid friendlytm appliances," in *Transmission and Distribution Conference and Exhibition, 2005/2006 IEEE PES*, May 2006, pp. 647–652.

[9] J. A. Short, D. G. Infield, and L. L. Freris, "Stabilization of Grid Frequency Through Dynamic Demand Control," *IEEE Transactions on Power Systems*, vol. 22, no. 3, pp. 1284–1293, 2007.

[10] M. Donnelly, D. Harvey *et al.*, "Frequency and stability control using decentralized intelligent loads: Benefits and pitfalls," in *Power and Energy Society General Meeting, 2010 IEEE*, July 2010, pp. 1–6.

[11] A. Brooks, E. Lu *et al.*, "Demand Dispatch," *IEEE Power and Energy Magazine*, vol. 8, no. 3, pp. 20–29, 2010.

[12] D. Callaway and I. Hiskens, "Achieving controllability of electric loads," in *Proceedings of the IEEE*. IEEE, 2011, pp. 184–199.

[13] A. Molina-Garcia, F. Bouffard, and D. S. Kirschen, "Decentralized Demand-Side Contribution to Primary Frequency Control," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 411–419, 2011.

[14] Y. Lin, P. Barooah *et al.*, "Experimental Evaluation of Frequency Regulation From Commercial Building HVAC Systems," *IEEE Transactions on Smart Grid*, vol. 6, no. 2, pp. 776–783, Mar. 2015.

[15] S. P. Meyn, P. Barooah *et al.*, "Ancillary service to the grid using intelligent deferrable loads," *IEEE Transactions on Automatic Control*, vol. 60, no. 11, pp. 2847–2862, 2015.

[16] D. Hammerstrom, J. Brous *et al.*, "Pacific Northwest GridWise Testbed Demonstration Projects, Part II: Grid Friendly Appliance Project," Pacific Northwest National Lab, Tech. Rep., Oct. 2007.

[17] M. Andreasson, D. V. Dimarogonas *et al.*, "Distributed Control of Networked Dynamical Systems: Static Feedback, Integral Action and Consensus," *IEEE Transactions on Automatic Control*, vol. 59, no. 7, pp. 1750–1764, 2014.

[18] ——, "Distributed vs. centralized power systems frequency control," in *2013 European Control Conference (ECC)*. IEEE, 2013, pp. 3524–3529.

[19] X. Zhang and A. Papachristodoulou, "A real-time control framework for smart power networks: Design methodology and stability," *Automatica*, vol. 58, pp. 43–50, Aug. 2015.

[20] C. Zhao, U. Topcu *et al.*, "Design and Stability of Load-Side Primary Frequency Control in Power Systems," *IEEE Transactions on Automatic Control*, vol. 59, no. 5, pp. 1177–1189, 2014.

[21] N. Li, L. Chen *et al.*, "Connecting automatic generation control and economic dispatch from an optimization view," *American Control Conference (ACC)*, pp. 735–740, 2014.

[22] E. Mallada and S. H. Low, "Distributed frequency-preserving optimal load control," in *IFAC World Congress*, 2014, pp. 5411–5418.

[23] S. You and L. Chen, "Reverse and Forward Engineering of Frequency Control in Power Networks," in *Conference on Decision and Control*, 2014.

[24] W. Liu, W. Gu *et al.*, "Decentralized Multi-Agent System-Based Cooperative Frequency Control for Autonomous Microgrids With Communication Constraints," *IEEE Transactions on Sustainable Energy*, vol. 5, no. 2, pp. 446–456, Apr. 2014.

[25] Q. Shafiee, J. M. Guerrero, and J. C. Vasquez, "Distributed Secondary Control for Islanded Microgrids - A Novel Approach," *IEEE Transactions on Power Electronics*, vol. 29, no. 2, pp. 1018–1031, 2014.

[26] F. Dörfler, J. Simpson-Porco, and F. Bullo, "Breaking the Hierarchy: Distributed Control & Economic Optimality in Microgrids," *IEEE Transactions on Control of Network Systems*, vol. 3, no. 3, pp. 241–253, 2016.

[27] C. Zhao, U. Topcu, and S. H. Low, "Optimal Load Control via Frequency Measurement and Neighborhood Area Communication," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 3576–3587, 2013.

[28] E. Mallada, C. Zhao, and S. Low, "Optimal load-side control for frequency regulation in smart grids," in *2014 52nd Annual Allerton Conference on Communication, Control, and Computing*. IEEE, 2014, pp. 731–738.

[29] P. Kundur, *Power System Stability And Control*. McGraw-Hill, 1994.

[30] C. Zhao, U. Topcu *et al.*, "Power System Dynamics as Primal-Dual Algorithm for Optimal Load Control," *ArXiv e-prints*, May 2013.

[31] F. deMello, R. Mills, and W. B'Rells, "Automatic Generation Control Part I - Process Modeling," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-92, no. 2, pp. 710–715, Jan. 1973.

[32] C. Zhao, E. Mallada, and S. H. Low, "Distributed generator and load-side secondary frequency control in power networks," *2015 49th Annual Conference on Information Sciences and Systems (CISS)*, pp. 1–6, 2015.

[33] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, Mar. 2004.

[34] F. P. Kelly, A. K. Maulloo, and D. K. H. Tan, "Rate Control for Communication Networks: Shadow Prices, Proportional Fairness and Stability," *The Journal of the Operational Research Society*, vol. 49, no. 3, pp. 237–252, Mar. 1998.

[35] S. H. Low and D. E. Lapsley, "Optimization flow control. I. Basic algorithm and convergence," *IEEE/ACM Transactions on Networking*, vol. 7, no. 6, pp. 861–874, 1999.

[36] R. Srikant, *The Mathematics of Internet Congestion Control*, ser. Systems & Control: Foundations & Applications. Boston, MA: Springer Science & Business Media, Dec. 2012.

[37] D. P. Palomar and M. Chiang, "A tutorial on decomposition methods for network utility maximization," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 8, pp. 1439–1451, Aug. 2006.

[38] A. Mas-Colell, M. D. Whinston, and J. R. Green, *Microeconomic Theory*. Oxford University Press, USA, 1995.

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

[40] H. K. Khalil, *Nonlinear systems*, 3rd ed. Prentice Hall, 2002.

[41] A. Bacciotti and F. Ceragioli, "Nonpathological Lyapunov functions and discontinuous Carathéodory systems," *Automatica*, 2006.

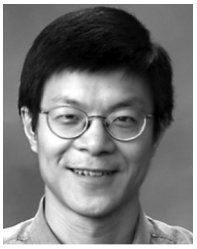
[42] A. Cherukuri, E. Mallada, and J. Cortés, "Convergence of caratheodory solutions for primal-dual dynamics in constrained concave optimization," in *SIAM conference on control and its applications*, July 2015.

[43] A. Tilli and M. Montanari, "A Low-Noise Estimator of Angular Speed and Acceleration from Shaft Encoder Measurements," *AUTOMATIKA: Journal for Control, Measurement, Electronics, Computing and Communications*, vol. 42., no. 3-4, pp. 169–176, Dec. 2001.

[44] J. H. Chow and K. W. Cheung, "A toolbox for power system dynamics and control engineering education and research," *IEEE Transactions on Power Systems*, vol. 7, no. 4, pp. 1559–1564, 1992.

Enrique Mallada (S'09-M'13) is an Assistant Professor of Electrical and Computer Engineering at Johns Hopkins University. Prior to joining Hopkins in 2016, he was a Post-Doctoral Fellow in the Center for the Mathematics of Information at California Institute of Technology from 2014 to 2016. He received his Ingeniero en Telecomunicaciones degree from Universidad ORT, Uruguay, in 2005 and his Ph.D. degree in Electrical and Computer Engineering with a minor in Applied Mathematics from Cornell University in 2014. Dr. Mallada was awarded the ECE Director's PhD Thesis Research Award for his dissertation in 2014, and the Cornell University Jacobs Fellowship in 2011. His research interests lie in the areas of control, dynamical systems and optimization, with applications to engineering networks such as power systems and the Internet.

Changhong Zhao (S'12-M'15) received the B. Eng. degree in Automation from Tsinghua University, and the PhD degree in Electrical Engineering from California Institute of Technology. He is currently a Research Engineer with the Power Systems Engineering Center at National Renewable Energy Laboratory, Golden, CO, USA. His research interests include power system dynamics and stability, frequency and voltage regulation, and control and optimization of distributed energy resources. He was a recipient of the Caltech Demetriades-Tsafka-Kokkalis PhD Thesis Prize.



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 TRANSACTIONS ON CONTROL OF NETWORK SYSTEMS, 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*, as well as that of the *Journal of Sustainable Energy, Grids and Networks*.