Some Problems in Demand Side Management

Lingwen Gan Libin Jiang Steven Low Ufuk Topcu Changhong Zhao
Engineering & Applied Science, Caltech

***Abstract*—We present a sample of problems in demand side management in future power systems and illustrate how they can be solved in a distributed manner using local information. First, we consider a set of users served by a single load-serving entity (LSE). The LSE procures capacity a day ahead. When random renewable energy is realized at delivery time, it manages user load through real-time demand response and purchases balancing power on the spot market to meet the aggregate demand. Hence optimal supply procurement by the LSE and the consumption decisions by the users must be coordinated over two timescales, a day ahead and in real time, in the presence of supply uncertainty. Moreover, they must be computed jointly by the LSE and the users since the necessary information is distributed among them. We present distributed algorithms to maximize expected social welfare. Instead of social welfare, the second problem is to coordinate electric vehicle charging to fill the valleys in aggregate electric demand profile, or track a given desired profile. We present synchronous and asynchronous algorithms and prove their convergence. Finally, we show how loads can use locally measured frequency deviations to adapt in real time their demand in response to a shortfall in supply. We design decentralized demand response mechanism that, together with the swing equation of the generators, jointly maximize disutility of demand rationing, in a decentralized manner.**

I. Introduction

There is a large literature on various forms of load side management in the electricity grid from the classical direct load control to the more recent real-time pricing [6], [5], [7]. Almost all demand response programs today target large industrial or commercial users, or, in the case of residential users, a small number of them, for two, among other, important reasons. First, demand side management is invoked rarely to mostly cope with a large correlated demand spike due to weather (e.g., during a few hottest days in summer) or a supply shortfall because of faults. Second, the lack of ubiquitous two-way communication in the current infrastructure prevents the participation of a large number of diverse users with heterogeneous and time-varying consumption requirements. Both reasons favor a simple and static mechanism involving a few large users that is sufficient to deal with the occasional need for load control, but both reasons are changing, because of renewable penetration and the deployment of a sensing, control, and two-way communication infrastructure. In this paper, we provide an overview of some of the demand side management algorithms that we have developed in [1]–[4]. They illustrate that it is possible to optimally control loads using decentralized algorithms.

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

II. Real-time demand response

A. Problem formulation

Consider a set $\mathcal{N}$ of N users that are served by a single load-serving entity (LSE). Without loss of generality, we assume each user $i \in \mathcal{N}$ operates a single appliance. Let q_i denote its energy consumption in the period of interest. An appliance i is characterized by

- a utility function $U_i(q_i)$ that quantifies the utility that user i obtains from using appliance i and consuming q_i amount of energy;
- consumption constraints: $\underline{q}_i \leq q_i \leq \overline{q}_i$.

For multi-period time-correlated case, see [1], [2].

The LSE procures energy for delivery in two steps. First, one day in advance, it procures "day-ahead" capacities P_d and pays $c_d(P_d)$ for the capacity. This entitles the LSE to purchase up to P_d amount of energy the following day at a price pre-determined by the day-ahead market. Let $P_o(t)$ denote the amount of the day-ahead energy that the LSE actually uses the following day and $c_o(P_o)$ denote its cost. The renewable energy is a nonnegative random variable P_r and we assume its cost is zero. At real time, the random variable P_r is realized and used to satisfy demand. The LSE satisfies any excess demand by using P_o from the day-ahead capacity. If there is still excess demand, the LSE purchases the balance P_b on the real-time energy market at a cost $c_b(P_b)$.

The real-time decisions (P_o, P_b) are made by the LSE so as to minimize its total cost. Let $Q := \sum_i q_i$ be the total demand and $\Delta(Q) := Q - P_r$ be the excess demand, i.e., in excess of the renewable generation. Given the excess demand $\Delta(Q(t))$ and the day-ahead capacity P_d, the LSE's decision that minimizes its total energy cost is

$$\begin{aligned} P_o^* &= [\Delta(Q)]_0^{P_d}, \\ P_b^* &= [\Delta(Q) - P_d]_+. \end{aligned}$$

The total supply cost that the LSE incurs is then a function only of P_d and Q and given by

$$c(Q, P_d; P_r) = c_d(P_d) + c_o\left([\Delta(Q)]_0^{P_d}\right) + c_b\left([\Delta(Q) - P_d]_+\right),$$

i.e., the total cost consists of the capacity cost c_d, the cost c_o of day-ahead energy, and the cost c_b of the real-time balancing energy.

We make the following assumptions:

A1: The utility functions U_i are strictly concave, increasing, and continuously differentiable, and the cost functions c_d, c_o, c_b are convex, increasing, and continuously differentiable, with $c_d(0) = c_o(0) = c_b(0) = 0$.

A2: $c_b'(0) > c_o'(P_o)$ for all $P_o \geq 0$.

B. Optimal demand response

The welfare maximization reduces to the problem

$$\max_{P_d \geq 0} \left\{ -c_d(P_d) \ + \ E \max_{q \in [\underline{q}, \overline{q}]} W_1(q; P_d, P_r) \right\}, \qquad (1)$$

where the real-time welfare, given decision P_d and realization of P_r, is

$$W(q; P_d, P_r) := \sum_i U_i(q_i) - c_o\left([\Delta(Q)]_0^{P_d}\right) - c_b\left([\Delta(Q) - P_d]_+\right). \qquad (2)$$

The expectation E in (1) is taken with respect to P_r. The order of maximizations and expectation in (1) reflects the fact that the decision P_d must be made a day ahead based on the distribution of P_r, but the consumption decisions q should be made in real time after P_r is realized. Given P_d and a realization of P_r, $W(q; P_d, P_r)$ is a deterministic function of q. Hence our problem decomposes into two subproblems:

1) *Real-time demand response:* Optimize real-time welfare W_1 over consumptions q, given P_d, P_r:

$$\max_{q \in [\underline{q}, \overline{q}]} W(q; P_d, P_r) \ = \ \sum_i U_i(q_i) - c_o\left([\Delta(Q)]_0^{P_d}\right) - c_b\left([\Delta(Q) - P_d]_+\right) \qquad (3)$$

Let $q(P_d, P_r)$ denote an optimizer.

2) *Day-ahead capacity procurement:* maximize expected welfare over P_d:

$$\max_{P_d \geq 0} \{ \ -c_d(P_d) + EW_1(q(P_d, P_r); P_d, P_r) \ \}$$

We now consider each subproblem in turn.

For the real-time demand response subproblem, since P_d has been committed, the cost $c_d(P_d)$ has been given. Hence, (3) is equivalent to

$$\begin{aligned} \tilde{W}(P_d; P_r) \quad := \quad & \max_{q, y_o, y_b} \{\sum_i (\delta_i U_i(q_i)) - c_o(y_o) - c_b(y_b)\} \\ s.t. \quad & \underline{q}_i \leq q_i \leq \bar{q}_i, \forall i, \\ & 0 \leq y_o \leq P_d, y_b \geq 0, \\ & P_r + y_o + y_b \geq \sum_i q_i. \end{aligned}$$

Associate dual variables μ_1 and μ_2 with the last two constraints. Then a partial Lagrangian is

$$\begin{aligned} & \mathcal{L}(q, y_o, y_b; \mu_1, \mu_2) \\ = \quad & \sum_i U_i(q_i) - c_o(y_o) - c_b(y_b) + \mu_1(P_d - y_o) \\ & + \mu_2(P_r + y_o + y_b - \sum_i q_i). \end{aligned} \qquad (4)$$

Consequently, a primal-dual algorithm to solve problem (4) is as follows.

Algorithm 1: Given P_d, P_r, compute real-time consumption

Initially, user i sets $q_i^0 \in [\underline{q}_i, \bar{q}_i]$. The LSE lets $\mu_1^0 = \mu_2^0 = 0$, and $y_o^0 = y_b^0 = 0$. In iteration $k+1 = 1, 2, \ldots$, do the following.

1) Each user i computes

$$q_i^{q+1} \quad = \quad \left(q_i^k + \beta^k \cdot [u_i'(x_i^k) - \mu_2^k]\right)_{\underline{q}_i}^{\bar{q}_i},$$

where $\beta^k := \frac{1}{k+1}$ is the step size, and reports it to the LSE through a communication network. That is, the "price" posed to the users is μ_2^k.

2) The LSE computes

$$\begin{aligned} \mu_1^{k+1} &= [\mu_1^k + \beta^k(y_o^k - P_d)]_+, \\ \mu_2^{k+1} &= [\mu_2^k + \beta^k(\sum_i q_i^k - P_r - y_o^k - y_b^k)]_+, \\ y_o^{k+1} &= [y_o^k + \beta^k(-c_o'(y_o^k) - \mu_1^k + \mu_2^k)]_0^{P_{max}}, \\ y_b^{k+1} &= [y_b^k + \beta^k(-c_b'(y_b^k) + \mu_2^k)]_0^{P_{max}}, \end{aligned}$$

where $P_{max} := \sum_i \bar{q}_i$. The LSE reports μ_2^{k+1} to active users.

It follows directly from convex optimization theory that

Theorem 1. *Algorithm 1 converges to the set of optimal solutions q^* and and dual variables μ^*.*

For the day-ahead capacity procurement subproblem, to decide P_d to maximize expected social welfare, the LSE solves

$$\max_{P_d \geq 0} \{E[\tilde{W}(P_d; P_r)] - c_d(P_d)\},$$

where $\tilde{W}(.)$ is defined in (4). The gradient of the objective function is $g(P_d) := E(\mu_1^*) - c_d'(P_d)$ (note that μ_1^* depends on P_d, P_r). A stochastic subgradient algorithm that converges to the set of optimal P_d is as follows.

Algorithm 2: Day-ahead energy

1) Initially, let $P_d^0 = 0$.
2) In step $m+1 = 1, 2, \ldots$, given a realization of P_r and δ (denoted by P_r^m, run Algorithm 1 to find μ_1^*, and denote it by μ_1^{*m}. Then, compute

$$P_d^{m+1} = \{P_d^m + \alpha^m[\mu_1^{*m} - c_d'(P_d^m)]\}_0^{P_{max}}$$

where $\alpha^m = 1/(m+1)$ is the step size.
Algorithm 2 can be run one day in advance by simulating the system (i.e., drawing samples of and P_r).

Theorem 2. *Algorithm 2 converges to a welfare-maximizing procurement P_d^* almost surely.*

III. Scheduling of EV charging

A. Problem Formulation

Consider a scenario where an electric utility negotiates with N electric vehicles (EVs) over T time slots of length ΔT on their charging profiles. The utility is assumed to know (precisely predict) the inelastic base electricity load profile (aggregated non-EV load) and aims to shape the aggregated charging profile of EVs to flatten the total load (base load plus EV load) profile. Each EV can charge after it plugs in and needs to be charged a specified amount of electricity by its deadline. For instance, an EV may plug in for charging at 9:00pm, specifying that it needs to be fully charged by 6:00am the next morning, or at least 80% full by 4:00am the next morning. In each time slot, the charging rate of an EV is a constant. Let $D(t)$ denote the base load in slot t, $r_n(t)$ denote the charging rate of EV n in slot t, $r_n := (r_n(1), \ldots, r_n(T))$ denote the charging profile of EV n, for $n \in \mathcal{N} := \{1, \ldots, N\}$ and $t \in \mathcal{T} := \{1, \ldots, T\}$. Our goal is to flatten the total load profile. This motivates the cost function

$$L(r) = L(r_1, \ldots, r_N) := \sum_{t=1}^{T} U\left(D(t) + \sum_{n=1}^{N} r_n(t)\right). \quad (5)$$

In (5), $r := (r_1, \ldots, r_N)$ denotes a charging profile of all EVs.

The charging profile r_n of EV n can take values in the interval $[0, \overline{r}_n]$ for some given $\overline{r}_n \succeq 0$, i.e.,

$$0 \leq r_n(t) \leq \overline{r}_n(t), \ n \in \mathcal{N}, \ t \in \mathcal{T}. \quad (6)$$

In order to impose arrival time and deadline constraints, $\overline{r}_n$ is considered to be time dependent with $\overline{r}_n(t) = 0$ for slots t before the arrival time and after the deadline of EV n. For each EV $n \in \mathcal{N}$, let B_n, $s_n(0)$, $s_n(T)$ and η_n denote its battery capacity, initial state of charge, final state of charge and charging efficiency respectively. The constraint that EV n needs to reach $s_n(T)$ state of charge by its deadline is captured by charging a pre-specified amount of energy over time

$$\eta_n \sum_{t \in \mathcal{T}} r_n(t)\Delta T = B_n(s_n(T) - s_n(0)), \ n \in \mathcal{N}. \quad (7)$$

Define the charging rate sum

$$R_n := B_n(s_n(T) - s_n(0))/(\eta_n \Delta T)$$

for $n \in \mathcal{N}$. Then, the constraint in (7) can be written as

$$\sum_{t=1}^{T} r_n(t) = R_n, \ n \in \mathcal{N}. \quad (8)$$

Definition 1. *Let $U : \mathbb{R} \to \mathbb{R}$ be strictly convex. A charging profile $r = (r_1, \ldots, r_N)$ is*

1) *feasible, if it satisfies the constraints* (6) *and* (8)*;*
2) *optimal, if it solves the optimal charging (OC) problem*

$$\textbf{OC}\begin{cases} \min\limits_{r_1,\ldots,r_N} & \sum_{t=1}^{T} U\left(D(t) + \sum_{n=1}^{N} r_n(t)\right) \\ \text{s.t.} & 0 \leq r_n(t) \leq \overline{r}_n(t), \ t \in \mathcal{T}, n \in \mathcal{N}; \\ & \sum_{t=1}^{T} r_n(t) = R_n, \ n \in \mathcal{N}. \end{cases}$$

3) *valley-filling, if there exists $A \in \mathbb{R}$ such that*

$$\sum_{n \in \mathcal{N}} r_n(t) = [A - D(t)]^+, \ t \in \mathcal{T}.$$

Remark 1. *Optimality of a charging profile r is independent of the choice of the utility function U (proved in Theorem 4). That is, if r is optimal with respect to a strictly convex utility function, then it is optimal with respect to any other strictly convex utility function. Therefore, we can choose $U(x) = x^2$ without loss of generality, and see that optimal charging profiles minimize the l_2 norm of the total load profile. Since the l_1 norm is a constant for all feasible r due to* (8)*, minimizing the l_2 norm "flattens" the total load profile.*

Remark 2. *If the objective is to track a given load profile G rather than to flatten the total load, we can change the objective function to*

$$\sum_{t=1}^{T} U\left(D(t) + \sum_{n=1}^{N} r_n(t) - G(t)\right)$$

without affecting the results [3]. For ease of presentation, we focus on the objective function in (5)*.*

B. Optimal Charging Profile

Theorem 3. *If a feasible charging profile r is valley-filling, then it is optimal.*

Valley-filling is our intuitive notion of optimality. However, it may not be always achievable. For example, the "valley" in inelastic base load may be so deep that even if all EVs charge at their maximum rate, the valley still cannot be completely filled. Besides, EVs may have stringent deadlines such that the potential for shifting the load over time to yield valley-filling is limited. The notion of optimality in Definition 1 takes care of these cases and agrees with the intuitive notion of optimality when valley-filling is achievable.

Since the objective function U depends on r only through its aggregate R_r, optimal charging profile are clearly nonunique. Let $\mathcal{O}$ be the set of all optimal points.

Theorem 4. *The set $\mathcal{O}$ of optimal charging profiles does not depend on the choice of U. That is, if r^* is optimal with respect to a strictly convex utility function, then r^* is also optimal with respect to any other strictly convex utility function.*

We now propose a decentralized algorithm for computing optimal charging profiles as the solution to the optimal control problem OC. By decentralized, we mean that EVs choose their own charging profiles, instead of being instructed by a centralized infrastructure. The utility only uses control signals, e.g. prices, to guide EVs' decisions. We assume that all EVs are available for negotiation at the beginning of the scheduling horizon (even though they are not necessarily available for charging as reflected by time-varying $\overline{r}_n$). Figure 1 shows the information exchange between the utility and the EVs for the implementation of this algorithm. Given the "price" profile broadcast by the utility, each EV chooses its charging profile independently, and reports back to the utility. The utility guides EVs' decision-making by altering the "price" profile. We assume U' is Lipschitz with the Lipschitz constant $\beta > 0$, i.e.,

$$|U'(x) - U'(y)| \leq \beta|x - y|$$

for all x, y.

Algorithm 3:
Given scheduling horizon $\mathcal{T}$, base load profile D, the number

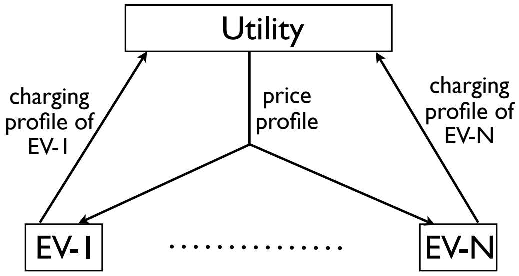

Fig. 1: Schematic view of the information flow patterns between the utility and the EVs. Given the "price" profile, the EVs choose their charging profiles independently. The utility guides EVs' decision making by altering the "price" profile based on total demand profile.

N of EVs, charging rate sum R_n and charging rate upper bound $\overline{r}_n$ for EV $n \in \mathcal{N}$, pick a step size γ satisfying

$$0 < \gamma < \frac{1}{N\beta}.$$

1) Initialize the "price" profile and the charging profile as

$$p^0(t) := U'(D(t)), \; r_n^0(t) := 0$$

for $t \in \mathcal{T}$ and $n \in \mathcal{N}$, $k \leftarrow 0$.

2) The utility broadcasts γp^k to all EVs.

3) Each EV $n \in \mathcal{N}$ calculates a new charging profile r_n^{k+1} as the solution to the following optimization problem

$$\begin{aligned} \min_{r_n} \quad & \sum_{t \in \mathcal{T}} \gamma p^k(t) r_n(t) + \frac{1}{2}\left(r_n(t) - r_n^k(t)\right)^2 \quad (9) \\ \text{s.t.} \quad & 0 \le r_n(t) \le \overline{r}_n(t), \; t \in \mathcal{T}; \\ & \sum_{t \in \mathcal{T}} r_n(t) = R_n, \end{aligned}$$

and reports r_n^{k+1} to the utility.

4) The utility collects charging profiles r_n^{k+1} from the EVs, and updates the "price" as

$$p^{k+1}(t) := U'\left(D(t) + \sum_{n=1}^{N} r_n^{k+1}(t)\right) \quad (10)$$

for $t \in \mathcal{T}$.

5) Iterate step (2)–(4) until convergence, and return r_n for all n in the last iteration as the charging schedule.

Algorithm 3 is in fact a gradient projection method for the optimal charging problem in Definition 1. We now establish its convergence to the set $\mathcal{O}$ of optimal charging profiles. Let the superscript k for each variable denote its respective value in iteration k. Let $R^k := \sum_{n \in \mathcal{N}} r_n^k$ denote the aggregated charging profile in iteration k.

Theorem 5. *Charging profiles converge to optimal charging profiles, i.e.,* $r^k \to \mathcal{O}$ *as* $k \to \infty$. *Furthermore, optimal charging profiles have the same aggregate charging profile* R^{opt}, *and aggregate charging profiles converge to it, i.e.,* $R^k \to R^{\text{opt}}$ *as* $k \to \infty$.

IV. Optimal load control

A. Problem formulation

Let $\mathbb{R}$ denote the set of real numbers and $\mathbb{C}$ denote the set of complex numbers. A variable without a subscript usually denotes a vector with appropriate components, e.g., $d := (d_l, l \in \mathcal{L}(j))$, $\omega := (\omega_j, j \in \mathcal{V})$, $P := (P_{ij}, (i,j) \in E)$. For a matrix A, A^t denotes its transpose and A^* its complex conjugate transposed.

1) Transmission network model: The transmission network is described by a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ where $\mathcal{V} = \{1, ..., N\}$ is the set of buses and $\mathcal{E}$ is the set of transmission lines connecting the buses. We adopt the following assumptions [1]

- The lines $(i,j) \in \mathcal{E}$ are lossless and characterized by reactances x_{ij}.
- The bus voltage magnitudes $|V_i|$ are constant.
- Reactive power is ignored.

We assume that $\mathcal{E}$ is directed, with an arbitrary orientation, so that $(j,i) \notin \mathcal{E}$ if $(i,j) \in \mathcal{E}$. We use (i,j) and $i \to j$ interchangeably to denote a link in $\mathcal{E}$. We also assume without loss of generality that $\mathcal{G}$ is connected. To simplify notation, we assume all variables represent deviations from their nominal (operating) values and are in per unit.

The dynamics at bus i with a generator is modeled by the swing equation

$$M_j \dot{\omega}_j \;=\; P_j^m - P_j^e$$

where ω_i is the frequency deviation from its nominal value, M_i is the inertia constant of the generator, P_i^m is the deviation in mechanical power injection to bus i from its nominal value, and P_i^e is the deviation in electric power from its nominal value. Each bus may have two types of loads, *frequency-sensitive* (e.g. motor-type) loads and *frequency-insensitive* (but controllable) loads. The total change in frequency-sensitive loads at bus i when the frequency deviation is ω_i is $\hat{d}_j := D_j \omega_j$ where D_j is the damping constant. Let $\mathcal{L}(j)$ denote the set of frequency-insensitive, controllable loads at bus j, and $(d_l, l \in \mathcal{L}(j))$ denote the deviations (from their nominal values) of frequency-insensitive loads on bus j. Then the electric power P_j^e is the sum of all frequency-sensitive loads, frequency-insensitive loads, and power flows from bus i to other buses

$$P_j^e \;=\; D_j \omega_i + \sum_{l \in \mathcal{L}(j)} d_l + \sum_{j \to k} P_{jk} - \sum_{i \to j} P_{ij}$$

Here P_{ij} is the deviation (from its nominal value) of branch flow from bus i to bus j. Our goal is to control the frequency-insensitive loads d_l in response to disturbances P_i^m in generation power. The swing equation can thus be rewritten as

$$\dot{\omega}_j = -\frac{1}{M_j}\left(\sum_{l \in \mathcal{L}(j)} d_l + D_j \omega_j - P_j^m + P_j^{\text{out}} - P_j^{\text{in}}\right), \quad (11)$$

[1]This is similar to the standard DC approximation except that we do not assume the phase angle difference is small across each link.

where $P_j^{\text{out}} := \sum_{j\to k} P_{jk}$ and $P_j^{\text{in}} := \sum_{i\to j} P_{ij}$ are total branch power flows out and into bus j, respectively.

We assume that the branch flows follow the dynamics

$$\dot{P}_{ij} = B_{ij}\,\omega^0\,(\omega_i - \omega_j)\,, \tag{12}$$

where ω^0 is the common nominal frequency on which the per-unit convention is based, and

$$B_{ij} := \frac{|V_i||V_j|}{x_{ij}}\cos\left(\theta_i^0 - \theta_j^0\right). \tag{13}$$

The dynamic model (12)–(13) is motivated by the following model of deviations in branch flows P_{ij} when the deviations are small [8] [9, Chapter 11]:

$$P_{ij} = B_{ij}(\theta_i - \theta_j) \tag{14}$$

where θ_i are the phase angle deviations of the bus voltages, i.e., the voltage phasors are $V_i := |V_i|e^{j(\theta_i^0+\theta_i)}$ with the nominal phase angles θ_i^0. While the model (14) assumes that the differences $\theta_i - \theta_j$ of the deviations are small, it does not assume the differences $\theta_i^0 - \theta_j^0$ of their nominal values are small.

In summary, the dynamic model of the transmission network is specified by (11)–(13). In steady state, the mechanical power deviations P_i^m are equal to the electric power deviations P_i^e, so $\dot{\omega}_i = 0$ and $\dot{P}_{ik} = 0$.

2) Optimal load control: Suppose a step change $P^m := (P_1^m, ..., P_N^m)$ in generation is injected to the N buses.[2] How should the frequency-insensitive loads $d := (d_l,\ l \in \mathcal{L}(i), i = 1, \ldots, N)$ in the network be reduced (or increased) in real time in a way that (i) balances the generation shortfall (or surplus), (ii) resynchronizes the bus frequencies, and (iii) minimizes the aggregate disutility of load control? We now formulate this as an optimal load control (OLC) problem.

The disturbance P^m in generation causes a nonzero frequency deviation ω_i. This incurs a cost to frequency-sensitive loads and suppose this cost is $\frac{1}{2D_i}\hat{d}_i^2$ in total at bus i. Suppose the frequency-insensitive load l is to be changed by an amount d_l and this will incur a cost (disutility) of $c_l(d_l)$. We assume $-\infty < \underline{d}_l \le d_l \le \overline{d}_l < \infty$. Our goal is to minimize the total cost over $(d, \hat{d})$ while balancing generation and load across the network:

OLC

$$\min_{\underline{d}\le d\le\overline{d},\hat{d}} \quad \sum_{j=1}^{N}\left(\sum_{l\in\mathcal{L}(j)} c_l(d_l) + \frac{1}{2D_j}\hat{d}_j^2\right) \tag{15}$$

$$\text{subject to} \quad \sum_{j=1}^{N}\left(\sum_{l\in\mathcal{L}(j)} d_l + \hat{d}_j\right) = \sum_{j=1}^{N} P_j^m \tag{16}$$

Remark 3. *Note that* (16) *does not require balance of generation and load at each individual bus, but only balance across the entire network. This is less restrictive and offers more opportunity to minimize costs. Additional constraints can be imposed if it is desirable that certain buses balance their own supply and demand, e.g., for economic or regulatory reasons.*

We make the following assumptions:

C0: OLC is feasible.

C1: The cost functions c_l are strictly convex and twice continuously differentiable on $\left[\underline{d}_l, \overline{d}_l\right]$.

B. Load control and swing dynamics as primal-dual solution

The objective function of the dual problem of OLC is

$$\sum_{j=1}^{N}\Phi_j(\nu) := \sum_{j=1}^{N} \min_{\underline{d}_j\le d_j\le\overline{d}_j,\hat{d}_j}\Bigg(\sum_{l\in\mathcal{L}(j)}(c_l(d_l) - \nu d_l) + \left(\frac{1}{2D_j}\hat{d}_j^2 - \nu\hat{d}_j\right) + \nu P_j^m\Bigg).$$

Hence,

$$\Phi_j(\nu) := \sum_{l\in\mathcal{L}(j)}(c_l(d_l(\nu)) - \nu d_l(\nu)) - \frac{1}{2}D_j\nu^2 + \nu P_j^m, \tag{17}$$

where

$$d_l(\nu) := \left[c_l^{'-1}(\nu)\right]_{\underline{d}_l}^{\overline{d}_l}. \tag{18}$$

This objective function has a scalar variable ν and is not separable across buses j. Its direct solution hence requires coordination across all buses. A *distributed* version of the dual problem where each bus j optimizes its own variable ν_j that are constrained to be equal at optimality is the following.

DOLC

$$\max_{\nu_j} \quad \Phi(\nu) := \sum_{j=1}^{N}\Phi_j(\nu_j)$$
$$\text{subject to} \quad \nu_i = \nu_j \qquad \text{for all } (i,j)\in\mathcal{E}.$$

Theorem 6. 1) *DOLC has a unique optimal solution ν^* with $\nu_i^* = \nu_j^* = \nu^*$.*[3]
2) *OLC has a unique optimal solution $(d^*, \hat{d}^*)$ where $d_l^* = d_l^*(\nu^*)$ is given by* (18) *and $\hat{d}_l^* = D_i\nu^*$.*
3) *There is no duality gap.*

Instead of solving OLC directly, Theorem 6 suggests solving its dual DOLC and recovering the unique optimal solution $(d^*, \hat{d}^*)$ of the primal problem OLC from the unique dual optimal ν^*. To derive a distributed solution for DOLC, consider its Lagrangian

$$L(\nu, \pi) := \sum_{j=1}^{N}\Phi_j(\nu_j) - \sum_{i\to j}\pi_{ij}(\nu_i - \nu_j) \tag{19}$$

where ν is the (vector) variable for DOLC and π is the associated dual variable for the dual of DOLC. Hence π_{ij}, for all $(i,j)\in\mathcal{E}$, measure the cost of not synchronizing

[2]If there is no generator at bus i, then $P_i^m = 0$.

[3]We abuse notation and use ν^* to denote both the vector and the common value of its components.

the variables ν_i and ν_j across buses i and j. A primal-dual algorithm for DOLC takes the form (using (17)–(18))

$$\begin{aligned}\dot{\nu}_j &= \gamma_j \frac{\partial L}{\partial \nu_j}(\nu(t), \pi(t)) \\ &= -\gamma_i \left(\sum_{l\in\mathcal{L}(j)} d_l(\nu_j) + D_j \nu_j - P_j^m + \pi_j^{\text{out}} - \pi_j^{\text{in}} \right),\end{aligned} \tag{20}$$

$$\dot{\pi}_{ij} = -\xi_{ij} \frac{\partial L}{\partial \pi_{ij}}(\nu(t), \pi(t)) \;=\; \xi_{ij}(\nu_i - \nu_j) \tag{21}$$

where $\gamma_i > 0$, $\xi_{ij} > 0$ are step sizes, $\pi_j^{\text{out}} := \sum_{k:j\to k} \pi_{jk}$, and $\pi_j^{\text{in}} := \sum_{i:i\to j} \pi_{ij}$.

It is then remarkable that (20)–(21) become identical to (11)–(12), if we identify ν with frequency deviations and π with branch flows

$$\nu_j(t) = \omega_j(t), \qquad \pi_{ij}(t) = P_{ij}(t),$$

and step sizes γ_i, ξ_{ij} with system parameters

$$\gamma_j = M_j^{-1}, \qquad \xi_{ij} = B_{ij} w^0.$$

For convenience, we collect the system dynamics and load control as

$$\begin{aligned}\dot{\omega}_j &= -\frac{1}{M_j}\Bigg(\sum_{l\in\mathcal{L}(j)} d_l(t) + \hat{d}_j(t) - P_j^m \qquad (22)\\ &\qquad + P_j^{\text{out}}(t) - P_j^{\text{in}}(t) \Bigg) \\ \dot{P}_{ij} &= B_{ij}\, \omega^0 \left(\omega_i(t) - \omega_j(t)\right) \qquad (23)\\ \hat{d}_j(\omega_j(t)) &= D_j \omega_j(t) \qquad (24)\\ d_l(\omega_j(t)) &= \left[c_l^{'-1}(\omega_j(t)) \right]_{\underline{d}_l}^{\overline{d}_l} \qquad \text{for all } l \in \mathcal{L}(j), (25)\end{aligned}$$

where $P_j^{\text{out}}(t) := \sum_{j\to k} P_{jk}(t)$ and $P_j^{\text{in}}(t) := \sum_{i\to j} P_{ij}(t)$ are total branch power flows out and into bus j, ω^0 is the common nominal frequency, and B_{ij} are given by (13). The dynamics (22)–(24) are automatically carried out by the power system while the local control (25) need to be implemented at each frequency-insensitive load. Let $(d(t), \hat{d}(t), \omega(t), P(t))$ denote a trajectory generated by the load control and the swing dynamics (22)–(25).

Theorem 7. *Any trajectory $(d(t), \hat{d}(t), \omega(t), P(t))$ converges to a limit $(d^*, \hat{d}^*, \omega^*, P^*)$ such that*

1) *$(d^*, \hat{d}^*)$ is the unique vector of optimal load control for OLC;*
2) *ω^* is the unique vector of optimal frequency deviations for DOLC;*
3) *P^* is a vector of optimal branch flows for the dual of DOLC.*

References

[1] L. Jiang and S. Low, "Energy procurement and real-time demand response with uncertain renewable energy," in *Proceedings of Allerton Conference on Communication, Control and Computing*, 2011.

[2] L. Jiang and S. Low, "Multi-period optimal procurement and demand responses in the presence of uncertain supply," in *Proceedings of IEEE Conference on Decision and Control*, 2011.

[3] L. Gan, U. Topcu and S. Low, "Optimal decentralized protocols for electric vehicle charging," in *IEEE Transactions on Power System*, to appear, 2012.

[4] C. Zhao, U. Topcu and S. Low, "Swing dynamics as primal-dual algorithm for optimal load control," sumbitted to *IEEE SmartGridComm*, 2012.

[5] M. H. Albadi and E. F. El-Saadany, "Demand response in electricity markets: An overview," in *Proceedings of the IEEE Power Engineering Society General Meeting*, 2007.

[6] C. W. Gellings and J. H. Chamberlin, "Demand-Side Management: Concepts and Methods," Fairmont Press, 1988.

[7] M. D. Ilic, "Dynamic monitoring and decision systems for enabling sustainable energy services," in *Proceedings of the IEEE*, vol. 99, no. 1, pp. 58–79, 2011.

[8] A. J. Wood and B. F. Wollenberg, "Power Generation, Operation, and Control," 2nd edition, John Wiley & Sons, Inc., 1996.

[9] A. R. Bergen and V. Vittal, "Power Systems Analysis," 2nd edition, Prentice Hall, 2000.