

Dynamic Pricing of Power in Smart-Grid Networks

Qingsi Wang, Mingyan Liu and Rahul Jain

Abstract

In this paper we introduce the problem of dynamic pricing of power for smart-grid networks. This is studied within a network utility maximization (NUM) framework in a deterministic setting with a single provider, multiple users and a finite horizon. The provider produces power or buys power in a (deterministic) spot market, and determines a dynamic price to charge the users. The users then adjust their demand in response to the time-varying prices. This is typically categorized as the demand response problem, and we study a progression of related models by focusing on two aspects: 1) the characterization of the structure of the optimal dynamic prices in the Smart Grid and the optimal demand and supply under various interaction with a spot market; 2) a greedy approach to facilitate the solution process of the aggregate NUM problem and the optimality gap between the greedy solution and the optimal one.

I. INTRODUCTION

As the Smart Grid takes shape, new possibilities of efficient management of the electric power grid open up. One of these is pricing of electricity to consumers. Currently, temporal variations in the cost of electricity are hidden behind inflexible rate designs. This leads to inefficiencies with over-consumption during peak times, and under-consumption during off-peak times. It also makes the problem of matching demand and supply, both of which are uncertain, and affect grid stability, particularly acute. In the United States, only 60-100 hours of the year can account for 10-18 percent of the system peak load [1]. Meeting this critical peak load requires installation, operation and maintenance of expensive combustion gas-turbine generators, since these start almost instantaneously.

Dynamic pricing of electricity to consumers can remedy this problem by inducing consumers to switch off or defer some of the non-urgent, non-critical loads. For example, dish-washers in most households have a delayed start option. And yet, only a small fraction of consumers utilize this feature and use the appliance later at night at off-peak times. Dynamic pricing can make consumers sensitive to their time and amount of electricity consumption, thus potentially smoothing out peak-time system loads, and enhancing economic efficiency.

Dynamic pricing rate designs are receiving increased attention by state commissions, with the California Public Utilities Commission (CPUC) having set a deadline of 2011 for the state electric utilities to propose dynamic pricing rate structures. These are defined as an electric rate structure that reflects the actual wholesale market conditions. Dynamic pricing can take various forms such as *critical peak-pricing*, *time-of-use pricing* and *real-time pricing*. CPUC defines the real-time price as the rate linked to the actual price in the wholesale hourly electricity market [2]. Such prices can be communicated to the consumers over the *advanced metering infrastructure* (AMI), or a *smart-grid network* [3] and displayed on a *smart meter*.

Literature Overview. Problems related to the design and operation of a Smart Grid are receiving increasing attention. The problem of dynamic pricing to shape demand is typically called *demand response*. There exists

Q. Wang and M. Liu are with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109-2122, USA. Email: (qingsi, mingyan)@umich.edu

R. Jain is with the Department of Electrical Engineering and the Department of Industrial and Systems Engineering, University of Southern California, Los Angeles, USA. Email: rahul.jain@usc.edu

a large literature on demand response; the most relevant is [4]. This paper formulates the problem of optimal demand response in discrete time over a finite time horizon when users have some storage available, and the formulation can be categorized as a network utility maximization (NUM) problem (see [5], [6] and the references therein). Using Lagrangian duality, it establishes the existence of a solution, and gives a distributed algorithm based on gradient projection to compute this solution. The modeling of various appliances into the proposed optimization framework is also studied. In [7], a different approach is taken wherein demand is required to match supply, and users bid for load-shedding in an iterative supply function bidding mechanism, that was first proposed in [8]. Both competitive and Nash equilibrium analysis is provided but the focus is on a single instant, and without storage. In [9], the focus is on optimal power flow through the network, wherein existence of a zero duality gap solution is proved. In contrast to aforementioned work, particularly [4], in this paper we study a progression of related models by focusing on two aspects: 1) the characterization of the structure of the optimal dynamic prices in the Smart Grid and the optimal demand and supply under various interaction with a spot market; 2) a greedy approach to facilitate the solution process of the aggregate NUM and the optimality gap between the greedy solution and the optimal one.

Other works have focused on electricity markets. While [10] provides a stability analysis of the wholesale electricity market, [11] has focused on means for wind power generators to participate in the day-ahead electricity market. The difficulty in this is that wind-power generators cannot enter into binding contracts for a day-ahead power supply without assuming a huge risk. Thus, alternative market architectures are considered in [12], wherein entities called “aggregators” buy power from the wind power generators and participate in the day-ahead markets. The paper proposes optimal mechanisms that the aggregators can use to get the wind power generators to reveal their true distributions, and thus minimize the risk due to uncertainty in supply from the contracted wind power generators. A related paper is [13], wherein mechanisms for riskless dispatch by the generators are given for use with the current market architecture.

Results and organization of this paper. The system model is first presented in Section II, and three related problems are introduced. The first problem SYSTEM is a simplified special case of the main model with internalized production cost, which provides a preliminary framework for more realistic refinement. Its optimal solution is characterized in Section III based on a dual decomposition; moreover, a distributed algorithm to solve SYSTEM, which can be readily adapted for problems in the same family, is also proposed. In Section IV we extend the SYSTEM problem to the scenario where the provider may purchase in part from a spot market with an exogenous cost to narrow the demand-supply gap. This is referred to as the SYS_SPOT problem, and the interaction between the decision of purchase and the spot market price is investigated. We then consider the explicit effect of the production cost in a third problem in Section V, and the structure of its optimal control in the single-user scenario is studied. A greedy algorithm to approximate the solution to SYSTEM is presented in Section VI, where an upper bound on the optimality gap is established and heuristics on evaluating this gap are discussed. Numerical results of sample problems are illustrated in Section VII. We present an attempt of adopting distributed Newton’s method to problems in this work in Section VIII, and Section IX concludes the paper.

II. SYSTEM MODEL AND NOTATION

Consider a smart-grid network that consists of n users or households, and one provider. We consider a discrete-time model with a finite horizon t_f . A user i consumes $x_i(t)$ power at time step t , and the supply by the grid is denoted as $z_i(t)$. The excess supply/demand is charged to/discharged from the user’s battery with capacity B_i . The energy level/state of the battery before operation at time t is denoted by $y_i(t)$. We assume that the

total supply from the grid at time t is upper bounded by C_t , and the total demand of user i over time is upper bounded by D_i . For convenience we define the following vector forms of the above quantities: the demand of user i , $x_i = [x_i(0), x_i(1), \dots, x_i(t_f)]^T$; the supply of the grid $z_i = [z_i(0), z_i(1), \dots, z_i(t_f)]^T$; the state of user i 's battery, $y_i = [y_i(1), y_i(2), \dots, y_i(t_f + 1)]^T$; the battery capacities of users $B = [B_1, B_2, \dots, B_n]^T$; the supply constraint of the grid $C = [C_0, C_1, \dots, C_{t_f}]^T$; the demand constraint of users $D = [D_1, D_2, \dots, D_n]^T$. Also, $x = \text{vect}(x_1, \dots, x_n)$ with $\text{vect}(u_1, \dots, u_n)$ being the stacked vector from u_1 to u_n , and similarly define y and z . When user i draws power x_i over time, her satisfaction is measured by the utility function

$$U_i(x_i) = \sum_{t=0}^{t_f} V_i(x_i(t)),$$

where $V_i : \mathbf{R}_+ \rightarrow \mathbf{R}_+$ is a strictly increasing, strictly concave and twice differentiable function of $x_i(t)$. Denote by $\dot{y}_i(t)$ the change of battery state of user i at time t :

$$\dot{y}_i(t) = y_i(t+1) - y_i(t) = z_i(t) - x_i(t).$$

Assuming zero initial state, $y(0) = 0$, we then have

$$y_i(t) = y_i(0) + \sum_{s=0}^{t-1} \dot{y}_i(s) = \sum_{s=0}^{t-1} (z_i(s) - x_i(s)).$$

Therefore,

$$y_i = H(z_i - x_i),$$

where

$$H = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 1 & 1 & \ddots & 0 & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 & 0 \\ 1 & 1 & \cdots & 1 & 1 \end{bmatrix} \in \mathbf{R}^{(t_f+1) \times (t_f+1)}.$$

In addition, since the battery suffers wear and tear due to usage, we model maintenance and operational costs as

$$M(y_i) = \sum_{t=1}^{t_f+1} m(y_i(t)),$$

where $m : \mathbf{R}_+ \rightarrow \mathbf{R}_+$ is a strictly increasing, strictly convex and twice differentiable function of $y_i(t)$. The objective is to maximize the sum of individual utilities of both users and the provider or the social welfare. Formally, we consider the following related problems in this paper. The main model is given as follows.

$$\begin{aligned} & \text{maximize} && \sum_{i=1}^n (U_i(x_i) - M(y_i)) - P(z) \end{aligned} \tag{1}$$

$$\text{subject to} \quad y_i \preceq B_i \mathbf{1} \tag{2}$$

$$y_i = H(z_i - x_i) \tag{3}$$

$$\sum_{i=1}^n z_i \preceq C \quad (4)$$

$$\mathbf{1}^T x_i \leq D_i \quad (5)$$

$$x_i, y_i, z_i \geq 0, i = 1, 2, \dots, n \quad (6)$$

where $P : \mathbf{R}_+^{n(t_f+1)} \rightarrow \mathbf{R}_+$ is the production cost function to be specified later, and $\mathbf{1} = (1, 1, \dots, 1)^T$ with proper sizes. We begin with a simplified special case of this problem by assuming $P = 0$; this is denoted as **SYSTEM** (U, M, B, H, C, D) . The motivation for studying this degenerated model is twofold: 1) it models the scenario when the production cost is realized thus fixed *before* the distribution of energy or when the marginal cost is zero such as in the case of renewable energies; 2) more critically, it provides a preliminary framework that reveals the basic structure of optimal solutions and their algorithmic implementation that can be readily extended to more realistic models.

We then consider the setting wherein the grid may purchase power from an external spot market to narrow the demand-supply gap, while incurring an exogenous cost; this is formulated as the **SYS_SPOT** $(U, M, B, H, C, D; \sigma)$ problem as follows.

$$\text{maximize} \quad \sum_{i=1}^n (U_i(x_i) - M(y_i)) - \sigma^T \left[\sum_{i=1}^n z_i - C \right]^+ \quad (7)$$

$$\text{subject to} \quad y_i \preceq B_i \mathbf{1} \quad (8)$$

$$y_i = H(z_i - x_i) \quad (9)$$

$$\mathbf{1}^T x_i \leq D_i \quad (10)$$

$$x_i, y_i, z_i \geq 0, i = 1, 2, \dots, n \quad (11)$$

where $\sigma = [\sigma_0, \sigma_1, \dots, \sigma_{t_f}]$ is the price vector of unit power in the spot market, and $[\cdot]^+ = \max\{\cdot, 0\}$. Our main interest in this problem is the characterization of spot market prices that may incentivize the exogenous purchase. In particular, we show that there is a threshold price below which this purchase is justified, and it is closely related to the optimal solution to **SYSTEM**.

Our last model considers non-trivial production costs. We consider the linear cost in the original setting with multiple users, that is, $P(z) = \sigma^T d_z$, where $d_z = \sum_{i=1}^n z_i$ is the vector of aggregate demands, and σ can be either interpreted as the unit power generation cost or the spot market price when the provider produces nothing of its own and completely resorts to the spot market for supply. This model is denoted by **SYS2** (U, M, B, H, C, D) . We also consider a general convex cost in the single-user case that has an explicit control structure.

In the next section we start with the **SYSTEM** problem, and note that when constraint (4) is decoupled, the **SYSTEM** problem can be separated into several subproblems. This observation motivates us to consider the dual-decomposition based technique to solve the **SYSTEM** problem [5], [6], which we show next.

III. DUAL DECOMPOSITION OF SYSTEM AND THE PROPERTIES OF ITS SOLUTIONS

In this section the **SYSTEM** problem is analyzed based on a dual decomposition, which also prompts a distributed first-order algorithm to solve it, and the properties of both the primal and the dual optimal solutions (demand, supply and prices) are discussed.

A. Dual decomposition

We first reformulate **SYSTEM** (U, M, B, H, C, D) as

$$\text{maximize} \quad \sum_{i=1}^n (U_i(x_i) - M(y_i)) \quad (12)$$

$$\text{subject to} \quad \text{vect}(x, y, z) \in \mathcal{F} \quad (13)$$

$$\text{over} \quad \text{vect}(x, y, z) \in \mathcal{D} \quad (14)$$

where the domain and the feasible region are respectively given by

$$\mathcal{D} = \left\{ \text{vect}(x, y, z) \mid \begin{array}{l} x_i \succeq 0, \mathbf{1}^T x_i \leq D_i, \\ 0 \preceq y_i \preceq B_i \mathbf{1}, z_i \succeq 0, \forall i \end{array} \right\},$$

and

$$\mathcal{F} = \left\{ \text{vect}(x, y, z) \mid y_i = H(z_i - x_i), \sum z_i \preceq C, \forall i \right\}.$$

Note that $\text{relint } \mathcal{D} \cap \mathcal{F} \neq \emptyset$ when B, C and D are all positive, as we shall assume so. Indeed, let $\epsilon > 0$ be chosen later, and consider

$$\hat{x}_i(t) = \alpha \frac{D_i - \epsilon}{t_f + 1},$$

where $0 < \alpha \leq 1$ satisfying

$$\alpha \frac{\sum_{i=1}^n D_i}{t_f + 1} < \min_t C_t,$$

$\hat{z}_i(t) = \hat{x}_i(t) + \frac{\alpha\epsilon}{t_f+1}$ and $\hat{y}_i(t) = \sum_{s=0}^{t-1} (\hat{z}_i(s) - \hat{x}_i(s))$. Let $\epsilon < \min_i \{\min\{B_i, D_i\}\}$, and it can be verified that $\text{vect}(\hat{x}, \hat{y}, \hat{z}) \in \text{relint } \mathcal{D} \cap \mathcal{F}$. In addition, since this problem is convex and the inequality constraint is affine, Slater's condition is then satisfied, and hence, strong duality will hold. Because of the strict concavity of the objective function in x and y , there are unique optima for x, y and z . As pointed out before, the structure of the **SYSTEM** problem prompts a solution based on a dual decomposition. Its Lagrangian is given by

$$\begin{aligned} L &= \sum_{i=1}^n (U_i(x_i) - M(y_i)) + \sum_{i=1}^n \nu_i^T (H(z_i - x_i) - y_i) + \theta^T (C - \sum_{i=1}^n z_i) \\ &= \sum_{i=1}^n (U_i(x_i) - \nu_i^T H x_i) - \sum_{i=1}^n (M(y_i) + \nu_i^T y_i) - \sum_{i=1}^n (\theta - H^T \nu_i)^T z_i + \theta^T C. \end{aligned}$$

Let $\nu = \text{vect}(\nu_1, \dots, \nu_n)$. The dual function is then given by

$$\begin{aligned} g(\nu, \theta) &= \max_{\text{vect}(x, y, z) \in \mathcal{D}} L(x, y, z, \nu, \theta) \\ &= \sum_{i=1}^n \max_{\substack{x_i \succeq 0, \\ \mathbf{1}^T x_i \leq D_i}} W_i(x_i; \nu_i) - \sum_{i=1}^n \min_{0 \preceq y_i \preceq B_i \mathbf{1}} R_i(y_i; \nu_i) - \sum_{i=1}^n \min_{z_i \succeq 0} (\theta - H^T \nu_i)^T z_i + \theta^T C, \end{aligned}$$

where

$$W_i(x_i; \nu_i) = U_i(x_i) - \nu_i^T H x_i,$$

and

$$R_i(y_i; \nu_i) = M(y_i) + \nu_i^T y_i,$$

and we introduce

$$\begin{aligned} \text{USER}_i : \text{maximize} \quad & W_i(x_i; \nu_i) \\ \text{subject to} \quad & x_i \succeq 0, \mathbf{1}^T x_i \leq D_i \end{aligned} \quad (15)$$

and

$$\begin{aligned} \text{USER}'_i : \text{minimize} \quad & R_i(y_i; \nu_i) \\ \text{subject to} \quad & 0 \preceq y_i \preceq B_i \mathbf{1} \end{aligned} \quad (16)$$

as two subproblems for user i . Since the minimization with respect to z_i is a linear problem, which is readily solved, the dual problem is given by DUAL_SYS (U, M, B, H, C, D):

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^n (W_i^*(\nu_i) - R_i^*(\nu_i)) + \theta^T C \\ \text{subject to} \quad & \theta - H^T \nu_i \succeq 0, \theta \succeq 0, i = 1, 2, \dots, n \end{aligned}$$

where $W_i^*(\nu_i)$ and $R_i^*(\nu_i)$ are the optimal values of problem (15) and (16) for a given ν_i . Given now this dual decomposition structure, a distributive algorithm can be readily proposed to solve SYSTEM, and this is widely studied in the literature, see e.g. [5], [4]. We report our version of the algorithm in Appendix A and show its convergence.

B. Properties of solutions to SYSTEM and DUAL_SYS

Denote by $x^* = \text{vect}(x_1^*, \dots, x_n^*)$ the optimum for x in SYSTEM; ν^* is an optimal solution for ν of DUAL_SYS, and similarly define y^* , z^* and θ^* , respectively. We first note that $H^T \nu_i$ can be interpreted as the price vector that the provider charges from user i for each unit of consumed energy, in light of subproblem USER_i, given that it is component-wise non-negative. Indeed, the following result establishes this premise.

Proposition 1: $H^T \nu_i^* \succeq 0$ for all i .

Proof: The proof can be found in Appendix B. ■

Remark 1: Since $z_i(t)$ is the total amount of energy distributed to user i at time t , under the interpretation of $H^T \nu_i$ being a price vector it is as if the energy buffered in battery were only charged when it is later consumed while incurring a *storage cost before* the consumption. This obviously is not what actually happens in practice, but serves as an interesting interpretation of the result. We do note that if there is no battery buffering in the model, this interpretation is consistent with the notion of shadow price in the literature (see e.g. [14], [15]), i.e., entries of $H^T \nu_i$ are shadow prices of power.

Hence, $\theta^*(t) = \max_i (H^T \nu_i^*)_t$ for all t . Moreover, 1) if we assume that $x_i^* \in \text{int } \mathcal{X}_i$ for all i , where \mathcal{X}_i is the feasible set of (15), i.e., $\mathcal{X}_i = \{x_i \mid x_i \succeq 0, \mathbf{1}^T x_i \leq D_i\}$, $H^T \nu_i^*$ is then uniquely determined for each i ; 2) alternatively, if we assume that $x_i^* \in \text{int } \mathcal{X}_i$ for at least one i , say i_k , and $z_{i_k}^* \succ 0$, $H^T \nu_{i_k}^*$ is then uniquely determined and from the minimization with respect to z_{i_k} we get that $\theta^* = H^T \nu_{i_k}^*$. In either case, θ^* is also uniquely determined.

In Section IV, we will consider the situation where there is a gap between the users demand and the power generation capacity of the grid at some time instants. At those instants, the grid may purchase power from the spot market to narrow the demand-supply gap. We will show therein that the above unique θ^* is a “lower bound” on the price of unit power in the spot market above which the grid decides to make no purchase.

Also, we can determine the closed form of y^* given ν^* based on (16). Since the objective and the constraint are both separable in time, we conclude

$$y_i^*(t) = [m'^{-1}(-\nu_i^*(t-1))]_0^{B_i}$$

for all i and $t > 0$, where $[\cdot]_a^b$ is the projection operator onto $[a, b]$, i.e., $[x]_a^b := \min\{\max\{x, a\}, b\}$. Let $\rho_i^* := H^T \nu_i^*$, and $\dot{\rho}_i^*(t) := \rho_i^*(t) - \rho_i^*(t-1) = -\nu_i^*(t-1)$. We then obtain

$$y_i^*(t) = [m'^{-1}(\dot{\rho}_i^*(t))]_0^{B_i}.$$

If we assume that $0 < y_i^*(t) < B_i$, the above result can be rewritten as

$$\dot{\rho}_i^*(t) = m'(y_i^*(t)).$$

Therefore, at optimality the marginal cost of reallocating the power purchase from $t-1$ to t is equal to the marginal cost of storing the energy in battery after purchasing in $t-1$, which is intuitively appealing. When the demand is known and fixed a priori, some structural results on the optimal power generation and battery scheduling in a single user setting can be found in [16]. In Section V, we will also discuss in detail the explicit structure of optimal control in the single-user problem when a non-zero production cost is incorporated in our main model.

IV. SPOT MARKET

In this section, we consider the SYS_SPOT problem. As before, we reformulate this problem and its domain and the feasible region are the same as in the SYSTEM problem. Also, there are unique optima for x , y and z because of the strict concavity of the objective function in x and y .

Denote by $z^*(\sigma) = (z_1^*(\sigma), \dots, z_n^*(\sigma))$ the optimal solution for z of SYS_SPOT given σ . Define

$$\mathcal{S} = \left\{ \sigma \mid \sum_{i=1}^n z_i^*(\sigma) \preceq C \right\}.$$

In the following, we characterize \mathcal{S} using the solutions of the SYSTEM problem and its dual problem DUAL_SYS. To avoid ambiguity, symbols without hat are variables in SYS_SPOT, and others are for the SYSTEM problem and its dual. We assume that either of the two cases below is true:

Assumption 1:

- a) $x_i^* \in \text{int } \mathcal{X}_i$ for all i , or
- b) $x_i^* \in \text{int } \mathcal{X}_i$ for at least one i , say i_k , and $z_{i_k}^* \succ 0$.

Recall that x_i^* denotes the optimal solution to the SYS_SPOT problem for x_i , and \mathcal{X}_i is the feasible set of (15). In the following proposition, we show that \mathcal{S} can be characterized with a threshold of prices.

Proposition 2: $\sigma \in \mathcal{S}$ if and only if $\sigma \succeq \hat{\theta}^*$, noting that $\hat{\theta}^*$ is the optimal dual variable in DUAL_SYS (U, M, B, H, C, D) .

Proof: 1) Necessity. First of all, the Lagrangian of SYS_SPOT is given by $L(x, y, z, \nu, \mu)$,

$$\begin{aligned} L &= \sum_{i=1}^n (U_i(x_i) - M(y_i)) - \sigma^T \left[\sum_{i=1}^n z_i - C \right]^+ + \sum_{i=1}^n \nu_i^T (H(z_i - x_i) - y_i) \\ &= \sum_{i=1}^n (U_i(x_i) - \nu_i^T H x_i) - \sum_{i=1}^n (M(y_i) + \nu_i^T y_i) - \sigma^T \left[\sum_{i=1}^n z_i - C \right]^+ + \sum_{i=1}^n \nu_i^T H z_i. \end{aligned}$$

Group the terms in a similar way as in DUAL_SYS, and the dual problem is

$$\begin{aligned} & \text{minimize} && \sum_{i=1}^n (W_i^*(\nu_i) - R_i^*(\nu_i)) - G^*(\nu; \sigma) \\ & \text{subject to} && \mu_i + \nu_i \succeq 0 \\ & && \mu_i \succeq 0, i = 1, 2, \dots, n \end{aligned}$$

where W_i^* and R_i^* are the respective optimal values of problems (15) and (16) for a given ν_i as in DUAL_SYS, and $G^*(\nu; \sigma)$ is the optimal value of

$$\begin{aligned} & \text{minimize} && \sigma^T \left[\sum_{i=1}^n z_i - C \right]^+ - \sum_{i=1}^n \nu_i^T H z_i \\ & \text{subject to} && z_i \succeq 0, i = 1, 2, \dots, n \end{aligned} \tag{17}$$

Because of strong duality¹, the optimal solution (x^*, y^*, z^*) to SYS_SPOT is also the maximizer of the Lagrangian $L(x, y, z, \nu^*)$ where ν^* is an optimal solution of the dual. Therefore, z^* is an optimal solution of (17) when $\nu = \nu^*$. Since $\sum_{i=1}^n z_i^*$ is finite, we must have $\sigma \succeq H^T \nu_i^*$ for all i . On the other hand, provided $\sum_{i=1}^n z_i^* \preceq C$, we can obtain the same solution (x^*, y^*, z^*) of SYS_SPOT by solving the SYSTEM problem with the same parameter set (U, M, B, H, C, D) . Recalling the properties of solutions to SYSTEM in Section III-B, under assumption a) we have ν_i^* and $\hat{\nu}_i^*$ are uniquely determined and $\nu_i^* = \hat{\nu}_i^*$ for all i . In addition, $\hat{\theta}^*(t) = \max_i (H^T \hat{\nu}_i^*)_t$ for all t . Under assumption b) we have $\nu_{i_k}^* = \hat{\nu}_{i_k}^*$ and $\hat{\theta}^* = H^T \hat{\nu}_{i_k}^*$. In either case, it follows that $\sigma \succeq \hat{\theta}^*$.

2) Sufficiency. We only provide an outline of the proof here. Our strategy is to show that the optimal solution $(\hat{x}^*, \hat{y}^*, \hat{z}^*)$ to the SYSTEM problem with optimal dual variables $(\hat{\nu}^*, \hat{\theta}^*)$ also satisfies the KKT conditions of SYS_SPOT when $\sigma \succeq \hat{\theta}^*$, and then because of strong duality and convexity of SYS_SPOT, it is also the optimal solution of SYS_SPOT with optimal dual variables $\nu^* = \hat{\nu}^*$, given the spot market price σ . To this end, note that the KKT conditions of SYS_SPOT only differ from those of SYSTEM in the dual feasibility and the stationarity conditions because of the replacement of $\hat{\theta}$ with σ and the non-differentiable term in the objective function. Recall that $\hat{\theta}^*(t) = \max_i (H^T \hat{\nu}_i^*)_t$ for all t , and we have $\sigma \succeq H^T \hat{\nu}_i^*$ if $\sigma \succeq \hat{\theta}^*$. Let $\partial_{z_i} L$ denote the subdifferential of the Lagrangian of SYS_SPOT with respect to z_i for all i . Hence, when $\sigma \succeq \hat{\theta}^*$ and $\nu_i = \hat{\nu}_i^*$, $0 \in \partial_{z_i} L$ for all $z_i = \hat{z}_i^*$, and hence combining other optimality conditions, we obtain $(\hat{x}^*, \hat{y}^*, \hat{z}^*, \hat{\nu}^*)$ is a primal-dual optimal solution given $\sigma \succeq \hat{\theta}^*$. ■

V. SYS2 AND THE STRUCTURE OF OPTIMAL CONTROL FOR A SINGLE USER

Under some circumstances, we have an explicit form of the solution to the SYS2 problem, which is given by the following proposition.

Proposition 3: At time t , $\sum_{i=1}^n z_i^*(t) = C_t$ if $(H^T \nu_i^*)_t > \sigma_t$ for some i , and $z_i^*(t) = 0$ if $(H^T \nu_i^*)_t < \sigma_t$.

Proof: The results follow from the KKT conditions. Firstly, the dual feasibility and the complementary slackness specify that

$$\theta \succeq 0, \theta^T \left(C - \sum_{i=1}^n z_i \right) = 0.$$

¹It can be shown that the reformulated SYS_SPOT problem is convex and Slater's condition is satisfied.

for all t . From this, using the complementary slackness conditions (24), we can conclude that either $P_t(0) = \rho^*(t) + \theta_0^*(t)$, or $P_t(C) = \rho^*(t) - \theta_1^*(t)$, or $P_t(z^*(t)) = \rho^*(t)$. We can summarize this as

$$z^*(t) = [P_t^{-1}(\rho^*(t))]_0^C. \quad (26)$$

Similarly, from KKT conditions (19), we get

$$V'(x^*(t)) = \rho^*(t) + (\lambda_1^*(t) - \lambda_0^*(t)),$$

which together with the complementary slackness conditions (23), implies that

$$x^*(t) = [V'^{-1}(\rho^*(t))]_0^D. \quad (27)$$

From (26) and (27), we conclude that $z^*(t)$ is non-decreasing and $x^*(t)$ is non-increasing in $\rho^*(t)$ for each t . Moreover, $x^*(t+1) \leq x^*(t)$ whenever $\rho^*(t+1) \geq \rho^*(t)$. In the following, we assume that $P_t = P$, and consequently $z^*(t+1) \geq z^*(t)$ whenever $\rho^*(t+1) \geq \rho^*(t)$. Using (25) and the complementary slackness conditions (22), we also conclude that if $y(t) > 0$ for all $t \in \mathcal{T} = \{\tau_1, \tau_1 + 1, \dots, \tau_2\}$, $\rho(t)$ is then strictly increasing in t over \mathcal{T} .

Combing the results obtained above, we conclude the following only possible dynamics of the optimal control, and the detailed analysis can be found in Appendix D.

- (i) $x^*(0) > z^*(0)$ when $y_0 > 0$, and x^* is non-increasing, y^* decreasing, z^* non-decreasing in t . Either $t_{x=z} = t_{y=0} \leq t_f$ (i.e., $x^*(t) = z^*(t)$ and $y^*(t) = 0$ simultaneously) and from then on the control and the state maintain fixed until $t = t_f$, or $x^*(t) > z^*(t)$ and $y^*(t) > 0$ until the time horizon.
- (ii) $x^*(0) = z^*(0)$ when $y_0 = 0$, and the control and the state maintain fixed until $t = t_f$.

In addition, once $x^*(t) = z^*(t) = u$, we have $u = [\min\{C, D, \tilde{u}\}]^+$ where \tilde{u} is such that $V'(\tilde{u}) = P'(\tilde{u})$.

To conclude, we sketch the dynamics of the optimal control and the state of case (i) in Figure 1. Note that the change of x^* , y^* and z^* are not necessarily linear.

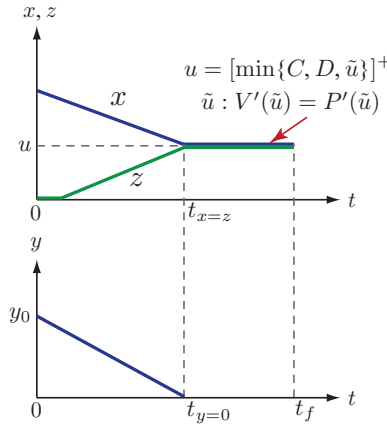


Fig. 1. The optimal policy for a single user

VI. A GREEDY APPROACH

In this section, we consider a greedy approach to approximating the solution to the SYSTEM problem. The need for a greedy approach can be motivated by the following considerations:

- 1) depending on the definition of time scale of a step, the convergence time may exceed the time step;

- 2) some parameters such as B , C and D may be stochastic within the time horizon, and we can only observe their realizations a posteriori.

Thus, we consider the following sequence of single-step optimization problems as approximations to the SYSTEM problem. We denote each by SYS_GREEDY $(V, B, C_t, D; t)$,

$$\text{maximize} \quad \sum_{i=1}^n V_i(x_i(t)) \quad (28)$$

$$\text{subject to} \quad \sum_{i=1}^n x_i(t) \leq C_t \quad (29)$$

$$x_i(t) + \sum_{s=0}^{t-1} x_i^*(s) \leq D_i \quad (30)$$

$$x_i(t) \geq 0, i = 1, 2, \dots, n \quad (31)$$

where $(x_i^*(s), i = 1, \dots, n)$ is the optimal solution to SYS_GREEDY $(V, B, C_s, D; s)$. Note that because of the myopic nature of each single-step optimization, it does not consider utilizing battery to store energy for future use, and thus the dimension of the greedy formulation reduces to that of the space of x .

There are a number of interesting problems to investigate. One of them is whether there is a speed advantage of SYS_GREEDY over SYSTEM, and the other is the optimality gap between them. We will present numerical results on the convergence rate of SYS_GREEDY in Section VII, and in the rest of this section we will characterize the solution of SYS_GREEDY and subsequently the optimality gap.

In the following presentation including Section VII, we limit our discussion to the case where the utility functions of users at a given time step are identical, i.e., $V_i = V$. We denote by x_i^* and $x_i^{g,*}$ the optimal solutions of SYSTEM and SYS_GREEDY, respectively. Define $\chi_i^{*,g} = \mathbf{1}^T x_i^{g,*}$ and $\chi_i^* = \mathbf{1}^T x_i^*$, and denote by $\bar{\chi}_i$, \mathcal{N}_0 and \mathcal{N}_1 the output of the following procedure.

Procedure GetChi(B, C, D)

Initialize $\mathcal{N}_0 \leftarrow \{1, 2, \dots, n\}$, $\mathcal{N}_1 \leftarrow \emptyset$, $\text{prev_}\mathcal{N}_0 \leftarrow \mathcal{N}_0$

repeat

$\text{prev_}\mathcal{N}_0 \leftarrow \mathcal{N}_0$

for $i \in \mathcal{N}_0$ **do**

if $(\mathbf{1}^T C - \sum_{j \in \mathcal{N}_1} D_j) / |\mathcal{N}_0| < D_i$ **then**

$\bar{\chi}_i \leftarrow (\mathbf{1}^T C - \sum_{j \in \mathcal{N}_1} D_j) / |\mathcal{N}_0|$

else

$\bar{\chi}_i \leftarrow D_i$

$\mathcal{N}_0 \leftarrow \mathcal{N}_0 - \{i\}$, $\mathcal{N}_1 \leftarrow \mathcal{N}_1 \cup \{i\}$

end if

end for

until $\mathcal{N}_0 = \text{prev_}\mathcal{N}_0$ or $\mathcal{N}_0 \leftarrow \emptyset$

return $\bar{\chi}_i, i = 1, \dots, n, \mathcal{N}_0, \mathcal{N}_1$

The above procedure can be literally described as a best-effort balanced water-filling process. That is, it attempts to fill each water tank (user) the same amount of water (energy), subject to their capacity (demand) constraints, and whenever smaller ones of them overflow, the excess goes to larger tanks in equal amounts. The following proposition asserts that this procedure in fact results in the optimal total per-user amount of energy.

Proposition 4: $\chi_i^{*,g} = \chi_i^* = \bar{\chi}_i$ for all i .

Proof: That $\chi_i^{*,g} = \bar{\chi}_i$ for all i is straightforward, and we prove that $\chi_i^* = \bar{\chi}_i$ for all i in Appendix C. ■

Remark 2: The balanced property of optimal solutions is a consequence of the concavity and the user-independence of utility function that we assumed. For user-dependent utility functions, we conjecture that the solution may be in a scaled form of balanced power distribution.

Let p^* and $p^{g,*}$ be the optimal value of SYSTEM and the sum of optimal values of SYS_GREEDY, respectively. Note that $x_i(t) = \chi_i^* (\frac{C_t}{\mathbf{1}^T C})$, $i = 1, \dots, n$, $t = 0, \dots, t_f$, are feasible solutions to the sequence of greedy optimization. The optimality gap is then upper bounded by

$$p^* - p^{g,*} \leq (t_f + 1) \sum_{i=1}^n V\left(\frac{\chi_i^*}{t_f + 1}\right) - \sum_{t=0}^{t_f} \sum_{i=1}^n V\left(\chi_i^* \frac{C_t}{\mathbf{1}^T C}\right). \quad (32)$$

Eqn. (32) suggests that the optimality gap would shrink as the power generation capacity C is more balanced among time steps. Moreover, observe that the loss of optimality of the greedy approach results from the myopic optimization that overlooks the use of energy buffer (i.e., battery) to balance energy dispatch between steps. We thus conjecture that the vector C in a roughly ascending order could yield a smaller optimality gap.

VII. NUMERICAL ILLUSTRATION

In this section, we illustrate the demand sensitivity to spot market price that we studied in Section IV, using two sample problems that are specified in Table I. We also empirically show the advantage in the convergence rate of the proposed greedy approach compared to the SYSTEM problem, when the respective optimization is performed using the same distributed algorithm based on gradient projection.

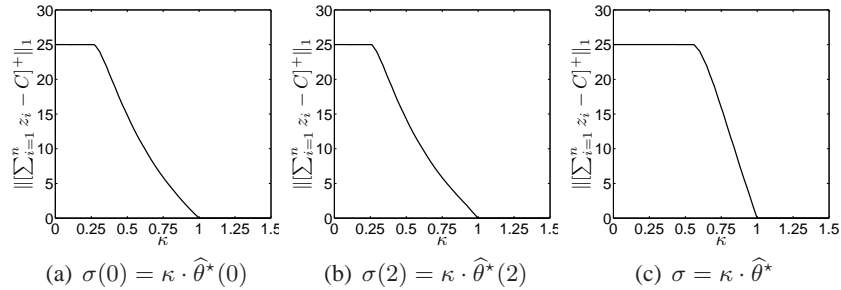


Fig. 2. Demand sensitivity to spot market price for the sample problem $n = 10, t_f = 4$.

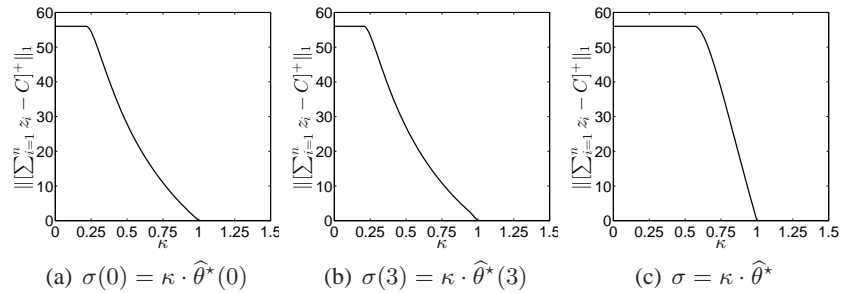


Fig. 3. Demand sensitivity to spot market price for the sample problem $n = 20, t_f = 6$.

As shown in Proposition 2, $\hat{\theta}^*$ assumes the threshold value of spot market price σ , and in Figure 2 and 3 we tune σ in three different ways to demonstrate the behaviors of optimal demand. The tests are performed for both sample problems and we elaborate the one with $n = 10, t_f = 4$ for example. In the first scenario, we

set $\sigma(t) = \hat{\theta}^*(t)$ for $t > 0$ and $\sigma(0) = \kappa \cdot \hat{\theta}^*(0)$, where κ is a tunable parameter. We scale the third entry of σ in the second scenario and the entire price vector in the third one. In Figure 2 and 3, we plot the curves of $\|\sum_{i=1}^n z_i - C\|_1$ versus κ , respectively. As can be seen in the figures, the demand increases as κ decreases when $\kappa < 1$ until hitting the demand constraint, and remains unchanged when $\kappa \geq 1$, as suggested by Proposition 2.

TABLE I
PARAMETERS OF THE TWO SAMPLE PROBLEMS IN SECTION VII.

	B	C	D
$n = 10, t_f = 4$	$[10, \dots, 10]^T$	$[8, 7, \dots, 4]^T$	$[1, 2, \dots, 10]^T$
$n = 20, t_f = 6$	$[10, \dots, 10]^T$	$[10, 9, \dots, 4]^T$	$[0.5, 1, \dots, 10]^T$
$V(a) = \log(1 + a)$ and $m(b) = \frac{1}{2}b^2$			

TABLE II
TOTAL RUNTIME (ITERATIONS) OF SAMPLE PROBLEMS.

	SYSTEM	SYS_GREEDY
$n = 10, t_f = 4$	25.0959s(100)	0.0206s(210)
$n = 20, t_f = 6$	68.0778s(158)	0.0316s(197)

In Table II, we report the respective convergence rates of SYSTEM and SYS_GREEDY. Though these records are machine-dependent, it demonstrates the significant comparative advantage of the latter in speed, and the greedy approach can be used as a good approximation to SYSTEM when the optimality gap bounded by (32) is acceptable.

VIII. A DISTRIBUTED PRIMAL-DUAL ALGORITHM: NEWTON'S METHOD

In this section, we present a distributed second-order algorithm, namely Newton's method, to solve the SYSTEM problem. We first reformulate SYSTEM and approximate it with an equality-constrained problem, and apply Newton's method on the primal problem. Theory and application of distributed Newton's method in network optimization can be found in [17], [18], [19] and etc. Instead of solving the dual problem where updating Lagrange dual variables can be interpreted as dynamic pricing, this algorithm works on the primal side and the solving process can be regarded as dynamic resource slicing [5]. In this section, we assume that $-U_i$ is self-concordant [20] for all i for the proof of convergence. A convex function $f : \mathbf{R} \rightarrow \mathbf{R}$ is self-concordant if $|f'''(x)| \leq 2f''(x)^{3/2}$ for all $x \in \text{dom } f$, and self-concordant functions include linear, convex quadratic functions, negative logarithm and etc.

We first introduce non-negative slack variables to three inequality constraints, that is,

$$\mathbf{1}^T x_i + p_i = D_i, \quad y_i + q_i = B_i \mathbf{1}, \quad \sum_{i=1}^n z_i + r = C,$$

and using the logarithmic barrier function for non-negative constraints, we approximate our original problem with an equality constrained problem as follows. Let $p = \text{vect}(p_1, \dots, p_n)$ and $q = \text{vect}(q_1, \dots, q_n)$. To simplify our notation, we also define $w = \text{vect}(x, y, z, p, q, r)$, that is, for instance, $w_i = x_i$ for $1 \leq i \leq n$ and $w_i(t) = x_i(t)$ and etc. We also define the logarithm for vector $u = (u_1, \dots, u_m) \in \mathbf{R}^m$ to be

$$\log u = \sum_t \log u_t.$$

Let

$$f(w) = -\sum_{i=1}^n U_i(w_i) + \sum_{i=n+1}^{2n} M(w_i) - \tau \sum_{i=1}^{5n+1} \log w_i,$$

and f is also self-concordant for $\tau \geq 1$. Define

$$A = \begin{bmatrix} \Xi & \mathbf{0} & \mathbf{0} & I_n & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & I_{n(t_f+1)} & \mathbf{0} & \mathbf{0} & I_{n(t_f+1)} & \mathbf{0} \\ \Gamma & I_{n(t_f+1)} & -\Gamma & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & S & \mathbf{0} & \mathbf{0} & I_{t_f+1} \end{bmatrix}$$

where $I(m)$ is the $m \times m$ identity matrix,

$$\Xi = \begin{bmatrix} \mathbf{1}^T & & \\ & \ddots & \\ & & \mathbf{1}^T \end{bmatrix}, \Gamma = \begin{bmatrix} H & & \\ & \ddots & \\ & & H \end{bmatrix},$$

and $S = [I_{t_f+1}, \dots, I_{t_f+1}]$, and also define $b = \text{vect}(D, B_1 \mathbf{1}, \dots, B_n \mathbf{1}, \mathbf{0}, C)$. The original problem can be then approximated as

$$\begin{aligned} & \text{minimize} && f(w) \\ & \text{subject to} && Aw = b \end{aligned}$$

It can be shown that as τ approaches zero, the optimum of the above approximating problem converges to that of the original one.

Like the gradient projection method, the Newton's method solves the above problem through an iterated update procedure. In each iteration, we have

$$w^{k+1} = w^k + \gamma^m \alpha^k \Delta w^k,$$

where α^k is the step-size, Δw^k is the Newton direction, γ is the discounting factor and

$$m = \min\{m' \in \mathbf{Z}_+ \mid w^k + \gamma^{m'} \alpha^k \Delta w^k \succeq 0\}.$$

Let $H^k = \nabla^2 f(w^k)$ and $g_k = \nabla f(x^k)$. The Newton direction is then given by the solution to the following KKT system,

$$\begin{bmatrix} H_k & A^T \\ A & \mathbf{0} \end{bmatrix} \begin{bmatrix} \Delta w^k \\ v^k \end{bmatrix} = - \begin{bmatrix} g_k \\ \mathbf{0} \end{bmatrix}$$

where v^k are dual variables associated with the equality constraint and $\mathbf{0}$ is null matrices with proper dimensions, and it yields

$$\begin{aligned} \Delta w^k &= -H_k^{-1}(g_k + A^T v^k), \\ (AH_k^{-1}A^T)v^k &= -AH_k^{-1}g_k. \end{aligned}$$

Note that f is separable for each $w_i(t)$, and hence g_k only depends on local information with terms

$$\frac{\partial f}{\partial w_i(t)} = \begin{cases} -V'_i(x_i(t)) - \frac{\tau}{x_i(t)}, & w_i(t) = x_i(t) \\ m'(y_i(t)) - \frac{\tau}{y_i(t)}, & w_i(t) = y_i(t) \\ -\frac{\tau}{u(t)} & w_i(t) = u(t), \\ & u = z_i, p_i, q_i, r \end{cases}$$

and H_k is a diagonal matrix with terms on the diagonal

$$\frac{\partial^2 f}{\partial w_i(t)^2} = \begin{cases} -V''_i(x_i(t)) + \frac{\tau}{x_i(t)^2}, & w_i(t) = x_i(t) \\ m''_i(y_i(t)) + \frac{\tau}{y_i(t)^2}, & w_i(t) = y_i(t) \\ \frac{\tau}{u(t)^2} & w_i(t) = u(t), \\ & u = z_i, p_i, q_i, r \end{cases}$$

Therefore, Δw^k can be computed from collected data distributedly given the vector v^k . Define the Newton decrement $\lambda(w^k)$ as

$$\lambda(w^k) = (\Delta w_k^T H_k \Delta w_k)^{1/2},$$

and use the step-size [18]

$$\alpha^k = \begin{cases} \frac{c}{\lambda(x^k)+1}, & \text{if } \lambda(x^k) \geq \frac{1}{4} \\ 1, & \text{o.w.} \end{cases}$$

where c is some positive scalar that satisfies $5/6 < c < 1$. The description of algorithm is then as follows.

Second-Order Algorithm for Power Distribution

User i 's algorithm At iterations $k = 1, 2, \dots$,

- 1) Receives from the grid Δw^k and $\lambda(w^k)$, and updates α^k and w^{k+1} .
- 2) Computes the associated terms in g_{k+1} and H_{k+1} , and communicates them to the grid.

Grid's algorithm At iterations $k = 1, 2, \dots$,

- 1) Receives from each user g_k and H_k .
 - 2) Computes v^k , Δw_k and $\lambda(w^k)$, and broadcast them to all users.
-

Note that our implementation is different from that in [17] and [18], where the computation of v^k is centralized at the grid, due to our introduction of partial computational capacity of the grid. The self-concordance based convergence analysis is similar to [18] where τ is assumed greater than 1 to preserve the self-concordance of f , and the analysis is omitted for brevity².

IX. CONCLUDING REMARKS

In this paper, we have studied key features of optimal dynamic prices in smart-grid networks. Moreover, an attempt of adopting distributed Newton's method to the model in this paper, namely the SYSTEM problem, is also presented, which is however mainly designed to solve the primal problem and is categorized in the domain of dynamic resource slicing instead of dynamic pricing. In this work, the explicit control structure for time-invariant power production and energy buffering costs has been studied, and continuation of this effort under more general

²For the simplicity of presentation, we implicitly assume that the error in computing v^k is negligible. In [18], the authors assume the H_k -quadratic norm of error is upper bounded.

cost functions is one direction of our future research. On the other hand, deviating from the deterministic setting and investigating the optimal pricing with stochastic demand and supply is also one ongoing research project.

APPENDIX A

A DUAL-DECOMPOSITION BASED DISTRIBUTED ALGORITHM

Let

$$Q(\nu, \theta) = \sum_{i=1}^n (W_i^*(\nu_i) - R_i^*(\nu_i)) + \theta^T C,$$

and

$$\mathcal{E} = \left\{ \text{vect}(\mu, \nu, \theta) \mid \theta - H^T \nu_i \succeq 0, \theta \succeq 0, \forall i \right\}.$$

Define $x_i^*(\nu_i)$ and $y_i^*(\nu_i)$ ³ as the respective unique maximizers of (15) and (16). Let $Q^k = Q(\nu^k, \theta^k)$. We solve DUAL_SYS using the gradient projection method, and the updating rule in the k th iteration is given by

$$\begin{aligned} \tilde{\nu}_i^{k+1} &= \nu_i^k - \alpha^k \nabla_{\nu_i} Q^k = \nu_i^k + \alpha^k (H x_i^*(\nu_i^k) + y_i^*(\nu_i^k)) \\ \tilde{\theta}^{k+1} &= \theta^k - \alpha^k \nabla_{\theta} Q^k = \theta^k - \alpha^k C \end{aligned}$$

and

$$\text{vect}(\nu^{k+1}, \theta^{k+1}) = \left[\text{vect}(\tilde{\nu}^{k+1}, \tilde{\theta}^{k+1}) \right]_{\mathcal{E}},$$

where α^k is a diminishing step-size satisfying

$$\sum_{k=1}^{\infty} \alpha^k = \infty, \sum_{k=1}^{\infty} (\alpha^k)^2 < \infty,$$

e.g., $\alpha^k = \frac{1+\Gamma}{k+\Gamma}$ with Γ a fixed non-negative number, and $[x_0]_{\mathcal{X}}$ is the projection of x_0 onto a set \mathcal{X} , or defined as the following constrained quadratic program

$$\begin{aligned} &\text{minimize} && \frac{1}{2} \|x - x_0\|_2^2 \\ &\text{subject to} && x \in \mathcal{X} \end{aligned}$$

Because all gradients are bounded with diminishing step-size α^k , the gradient projection method converges for the dual problem [21], [22], [5]. Since $x_i^*(\nu_i)$ and $y_i^*(\nu_i)$ are the respective unique maximizers of (15) and (16) and the strong duality holds, $x_i^*(\nu_i^*)$ and $y_i^*(\nu_i^*)$ are primal optimal, where ν_i^* are the solutions of the dual problem. To summarize, we have the following:

First-Order Algorithm for Power Distribution

User i 's algorithm At iterations $k = 1, 2, \dots$,

- 1) Receives from the grid the dual variables ν_i^k
- 2) Computes $x_i^*(\nu_i^k)$ and $y_i^*(\nu_i^k)$, and communicates them to the grid

Grid's algorithm At iterations $k = 1, 2, \dots$,

- 1) Receives from each user $x_i^*(\nu_i^k)$ and $y_i^*(\nu_i^k)$ for all i
- 2) Computes the dual variable ν_i^{k+1} and θ^{k+1} for all i , and communicates them accordingly to each user

³With a slight abuse of notation, we will use $x(t)$ to denote the t th entry of x , and $x(\nu)$ as a function of ν .

The algorithm would terminate when the following stopping criterion is satisfied:

$$\|\text{vect}(\nu^k, \theta^k) - \text{vect}(\nu^{k-1}, \theta^{k-1})\| \leq \epsilon,$$

where ϵ is a predefined constant.

APPENDIX B

PROOF OF PROPOSITION 1

We prove by contradiction. Assume that there exists at least one i such that $(H^T \nu_i^*)_t = \sum_{s=t}^{t_f} \nu_i^*(s) < 0$ for some t . Let $I = \arg \min_t (H^T \nu_i^*)_t$. We first note that $x^* \neq 0$ and further more there exists $t_0 \in I$ such that $x_i^*(t_0) > 0$. Otherwise, one can switch the value of $x_i^*(t_0)$ with that of some positive entry in x_i^* and obtain an improved value of $W_i(x_i; \nu_i^*)$, which is a contradiction to the fact that x_i^* is the maximizer of $W_i(x_i; \nu_i^*)$. Therefore,

- 1) if there exists $t < t_0$ such that $(H^T \nu_i^*)_t > (H^T \nu_i^*)_{t_0}$, let $t_1 = \max\{t : (H^T \nu_i^*)_t > (H^T \nu_i^*)_{t_0}, t < t_0\}$ and we then have $\nu_i^*(t_1) > 0$; otherwise $(H^T \nu_i^*)_{t_1} < (H^T \nu_i^*)_{t_0}$. Moreover, $\nu_i^*(t) = 0$ for $t_1 < t < t_0$ if any. We then set

$$\nu_i^*(t_0) \leftarrow \nu_i^*(t_0) + \epsilon$$

and

$$\nu_i^*(t_1) \leftarrow \nu_i^*(t_1) - \epsilon$$

where $0 < \epsilon < \min\{\nu_i^*(t_1), -(H^T \nu_i^*)_{t_0}\}$, where “ $y \leftarrow x$ ” signifies assigning the value of x to y . By doing so, the value of W_i is strictly decreased while that of R_i is non-decreasing and the feasibility constraints are maintained. We thus obtain an improved dual objective value, which leads to a contradiction.

- 2) If $(H^T \nu_i^*)_t = (H^T \nu_i^*)_{t_0}$ for all $t < t_0$ if any, we set

$$\nu_i^*(t_0) \leftarrow \nu_i^*(t_0) + \epsilon$$

where $0 < \epsilon < -(H^T \nu_i^*)_{t_0}$, and the same argument applies as in 1).

APPENDIX C

PROOF OF PROPOSITION 4

Note first that $\sum_{i=1}^n \chi_i^* = \sum_{i=1}^n \bar{\chi}_i = \min\{\mathbf{1}^T C, \mathbf{1}^T D\}$. The proof of this claim is straightforward and we thus omit it here for brevity. Without loss of generality, we assume that D is sorted in an ascending order, and therefore $\bar{\chi}_i$ is non-decreasing in i as the procedure suggests. Moreover, if $\bar{\chi}_j < \bar{\chi}_i$ for some $j < i$, then $\bar{\chi}_j = D_j$. Assume for contradiction that $\chi_i^* < \bar{\chi}_i \leq D_i$ for some i , and there then exists at least one $j \neq i$ such that $\chi_j^* > \bar{\chi}_j$. Thus,

- 1) if $j < i$, then $\bar{\chi}_j \leq \bar{\chi}_i$. If $\bar{\chi}_j < \bar{\chi}_i$, we have $\chi_j^* > \bar{\chi}_j = D_j$, which is infeasible. Therefore, $\bar{\chi}_j = \bar{\chi}_i$ and consequently $\chi_j^* > \chi_i^*$. Hence, there exists some t such that $x_j^*(t) > x_i^*(t)$. Consider an ϵ such that $0 < \epsilon \leq \min\{D_i - \chi_i^*, (x_j^*(t) - x_i^*(t))/2\}$ and set $x_j^*(t) \leftarrow x_j^*(t) - \epsilon$ and $x_i^*(t) \leftarrow x_i^*(t) + \epsilon$. We then obtain a feasible but improved objective value, which is a contradiction.
- 2) If $j > i$, then $\bar{\chi}_j \geq \bar{\chi}_i$. Hence $\chi_j^* > \chi_i^*$ and the same argument applies.

The above argument also applies symmetrically to the case when $\bar{\chi}_i < \chi_i^* \leq D_i$ for some i . Therefore, $\chi_i^* = \bar{\chi}_i$ for all i .

APPENDIX D

DYNAMICS OF THE SINGLE-USER OPTIMAL CONTROL

We rule out the suboptimal cases as follows.

1. If $x^*(0) < z^*(0)$, then $y^*(1) > 0$ and $\rho^*(1) = \rho^*(0) + M'(y^*(1)) + \mu^*(1) > \rho^*(0)$. Hence, $z^*(1) \geq z^*(0)$ and $x^*(1) \leq x^*(0)$, and the battery keeps charging until either fully charged before the time horizon and maintain the state, or the time horizon is reached; however, this is clearly suboptimal.
2. If $x^*(0) > z^*(0)$, which is only feasible when $y_0 > 0$, let us consider the sequence of events $x^*(t) \leq z^*(t)$ and $y^*(t) = 0$. Let $t_{x \leq z}$ be the first time step when $x^*(t) \leq z^*(t)$ or be infinity if it does not happen before the horizon, and similarly we define other quantities.
 - 2a) If $x^*(t) \leq z^*(t)$ happens before $y^*(t) = 0$, i.e., $\tilde{t} := t_{x \leq z} < t_{y=0}$, then $y^*(\tilde{t} + 1) \geq y^*(\tilde{t}) > 0$ and thus $\rho^*(\tilde{t} + 1) = \rho^*(\tilde{t}) + M'(y^*(\tilde{t} + 1)) + \mu^*(\tilde{t} + 1) > \rho^*(\tilde{t})$. Hence, $z^*(\tilde{t} + 1) \geq z^*(\tilde{t})$ and $x^*(\tilde{t} + 1) \leq x^*(\tilde{t})$, and again we are in case 1, which is suboptimal. For the boundary case that $t_{x \leq z} = t_f$, the suboptimality is also immediate, and we omit the discussion of boundary cases in the following.
 - 2b) If $t_{x \leq z} > t_{y=0} =: \tilde{t}$, i.e., $y^*(\tilde{t}) = 0$ while $x^*(\tilde{t}) > z^*(\tilde{t})$, this however is infeasible.
 - 2c) If $t_{x \leq z} = t_{y=0} =: \tilde{t}$, then $y^*(\tilde{t} + 1) > 0$ and $\rho^*(\tilde{t} + 1) > \rho^*(\tilde{t})$. Hence, $x^*(\tilde{t} + 1) < z^*(\tilde{t} + 1)$, and we are in case (2a), which is suboptimal.
3. If $x^*(0) = z^*(0)$ and $y_0 > 0$, then this is clearly suboptimal.

REFERENCES

- [1] A. Faruqui, R. Hledik, and J. Tsoukalis, "The power of dynamic pricing," *The Electricity Journal*, vol. 22, pp. 42–56, April 2009.
- [2] California Public Utilities Commission, Docket #: A.06-03-005. http://docs.cpuc.ca.gov/PUBLISHED/NEWS_RELEASE/85953.htm.
- [3] *The Smart Grid: An Introduction*. The US Department of Energy, 2008.
- [4] N. Li, L. Chen, and S. H. Low, "Optimal Demand Response Based on Utility Maximization in Power Networks," in *Proceedings of the 2011 IEEE Power & Energy Society General Meeting*, 2011.
- [5] D. P. Palomar and M. Chiang, "A Tutorial on Decomposition Methods for Network Utility Maximization," *IEEE J. Sel. Areas Commun.*, vol. 24, pp. 1439–1451, 2006.
- [6] M. Chiang, S. H. Low, A. R. Calderbank, and J. C. Doyle, "Layering as Optimization Decomposition: A Mathematical Theory of Network Architectures," *Proceedings of The IEEE*, 2007.
- [7] L. Chen, N. Li, S. Low, and J. Doyle, "Two Market Models for Demand Response in Power Networks," in *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on*, pp. 397–402, oct. 2010.
- [8] P. D. Klemperer and M. A. Meyer, "Supply Function Equilibria in Oligopoly under Uncertainty," *Econometrica*, vol. 57, pp. 1243–1277, November 1989.
- [9] J. Lavaei and S. H. Low, "Zero Duality Gap in Optimal Power Flow Problem," *IEEE Transactions on Power Systems*, 2011.
- [10] M. Roozbehani, M. Dahleh, and S. Mitter, "On the Stability of Wholesale Electricity Markets under Real-time Pricing," in *Decision and Control (CDC), 2010 49th IEEE Conference on*, pp. 1911–1918, dec. 2010.
- [11] E. Bitar, R. Rajagopal, P. Khargonekar, K. Poolla, and P. Varaiya, "Bringing Wind Energy to Market," *IEEE Transactions on Power Systems*, pp. 1–11, 2011.
- [12] W. Tang and R. Jain, "Stochastic Resource Auctions for Renewable Energy Integration," *working paper*, 2011.
- [13] P. Varaiya, F. Wu, and J. Bialek, "Smart Operation of Smart Grid: Risk-Limiting Dispatch," *Proceedings of the IEEE*, vol. 99, pp. 40–57, jan. 2011.
- [14] F. Kelly, "Charging and Rate Control for Elastic Traffic," *European Transactions on Telecommunications*, vol. 8, no. 1, pp. 33–37, 1997.
- [15] 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, 1998.
- [16] K. Chandy, S. Low, U. Topcu, and H. Xu, "A Simple Optimal Power Flow Model with Energy Storage," in *Decision and Control (CDC), 2010 49th IEEE Conference on*, pp. 1051–1057, dec. 2010.
- [17] A. Jadbabaie, A. Ozdaglar, and M. Zargham, "A Distributed Newton Method for Network Optimization," in *Decision and Control (CDC), Proceedings of the 48th IEEE Conference on*, pp. 2736–2741, dec. 2009.

- [18] E. Wei, A. Ozdaglar, and A. Jadbabaie, "A Distributed Newton Method for Network Utility Maximization," in *Decision and Control (CDC), Proceedings of the 49th IEEE Conference on*, pp. 1816 –1821, dec. 2010.
- [19] D. P. Bertsekas, "Centralized and Distributed Newton Methods for Network Optimization and Extensions," *Lab. for Information and Decision Systems Report LIDS-P-2866, MIT*, April, 2011.
- [20] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.
- [21] D. P. Bertsekas, *Nonlinear Programming*. Athena Scientific, 1999.
- [22] D. P. Bertsekas, A. Nedić, and A. E. Ozdaglar, *Convex Analysis and Optimization*. Athena Scientific, 2003.