On Variability of Renewable Energy and Online Power Allocation

Chinwendu Enyioha, Sindri Magnússon, Kathryn Heal, Na Li, Carlo Fischione, and Vahid Tarokh

***Abstract*—As electric power system operators shift from conventional energy to renewable energy sources, power distribution systems will experience increasing fluctuations in supply. These fluctuations present the need to not only design online decentralized power allocation algorithms, but also characterize how effective they are given fast-changing consumer demand and generation. In this paper, we present an Online Decentralized Dual Descent (OD3) power allocation algorithm and determine (in the worst case) how much of observed social welfare can be explained by fluctuations in generation capacity and consumer demand. Convergence properties and performance guarantees of the OD3 algorithm are analyzed by characterizing the difference between the online decision and the optimal decision. We demonstrate validity and accuracy of the theoretical results in the paper through numerical experiments using real data.**

I. Introduction

As the quest to integrate more renewable energy sources into the power distribution grid continues, it is important to understand how systemic fluctuations both in renewable energy generation and consumer demand affect real-time power allocation policies and the social welfare of distribution systems. Despite the increasing popularity of renewable energy sources, they experience high variability in generation capacity [1, 2]. Though wind forecasting is generally used to reduce uncertainty in day ahead schedules for wind power generation, variability in the wind resource continues to pose challenges for the integration of wind power in forward electricity markets [3]. Similar conclusions can easily be made of solar energy. To better understand the impact of penetrating renewable energy into the grid, the market value of these renewable sources have also been analyzed; for instance, in [4], the authors characterized how the market value of renewable energy sources varies with grid integration. Their study found that the value of wind power fell from 110% of the average power price to about $50-80\%$ as wind penetration increased from zero to 30% of total electricity consumption. The outcome of the study on price fluctuation highlights the difficulty in integrating large-scale renewable energy into the distribution grid. Similar studies in [5, 6] noted the significant impact of variability of renewable energy sources on the operations of electricity markets.

This work was supported by the VR Chromos Project, NSF grant No. 1548204 and NSF CAREER 1553407.

C. Enyioha, K. Heal, N. Li, and V. Tarokh are with the School of Engineering and Applied Sciences, Harvard University, Cambridge, MA USA. (email: cenyioha@seas.harvard.edu; kathrynheal@g.harvard.edu; nali@seas.harvard.edu; vahid@seas.harvar.edu)

S. Magnússon and C. Fischione are with Electrical Engineering School, Access Linnaeus Center, KTH Royal Institute of Technology, Stockholm, Sweden. (e-mail: sindrim@kth.se; carlofi@kth.se)

The inherent variability of renewable energy sources easily propagates to distribution systems and poses a challenge to system operators who aim to coordinate or monitor the operation of electrical power system. Amongst others, an objective of system operators is to allocate power (usually limited in supply) to end users in a way that optimizes the aggregate social welfare of the system. Similar resource allocation problems have been studied widely in different contexts, including communication [7, 8] and power systems [9]. For example, in a system comprising N users where each user has a utility function $U_i(\cdot)$, the objective of the system operator is typically to maximize the aggregate social welfare subject to available system resource Q, mathematically formulated as:

$$\begin{aligned} \underset{\mathbf{q}_1,\dots,\mathbf{q}_N}{\text{maximize}} \quad & \sum_{i=1}^{N} U_i(\mathbf{q}_i) \\ \text{subject to} \quad & \sum_{i=1}^{N} \mathbf{q}_i = Q. \end{aligned} \tag{1}$$

In Problem (1), the task of the system operator is to determine how much power, $\mathbf{q}_i$, to allocate each user i to maximize the aggregate users' utility. Common schemes to solve Problem (1) in a distributed way (like descent, gradient-based algorithms), involve the system operator and users iteratively communicating and carrying out computations until they reach the optimal solution to trigger an event. A major challenge is that during the process, especially in the context of renewable energy generation, the system environment may change since the generation capacity and users' utility functions are dynamic. Furthermore, the procedure of iteratively carrying out computations until an optimal solution is attained is not only expensive in terms of communication overhead, but also costly in terms of the time it takes to coordinate, especially in power systems, given the need to track fast variability in renewable energy generation and consumer demands in real time [10, 11].

We propose an alternate approach to the aforementioned coordination scheme in which the decentralized power allocation algorithm is implemented in real-time: At each time-step, the system operator updates the coordination signal (which is usually able to be interpreted as the price signal), and users locally compute their optimal allocations based on the price signal and individual utility/need. With this approach, a number of interesting questions arise, including performance guarantees of the decentralized algorithm in real-time implementation. Since generation capacity and users' utility functions are constantly changing, the aggregate social welfare of the system will fluctuate and it is unclear how optimal the online decisions will be. In this paper, we focus

on investigating and characterizing performance guarantees of an Online Decentralized Dual Descent (OD3) power allocation algorithm. In particular, we assume the power supply and consumers' utilities fluctuate at the same time scale as iterations of our algorithm; we also derive a worst-case bound on the aggregate social welfare of users in the system. Recent efforts on dynamic optimization problems include [12, 13].

Although the literature on price volatility in electricity markets is rich, our work differs from existing literature. For instance, volatility of electricity prices in renewable energy markets was analyzed in [4–6]. Furthermore, volatility of electricity prices has also been studied in the context of spot market pricing, price-based Demand Response (DR) and wholesale market exposure [14–18]. The focus of the referenced works are on pricing dynamics in relation to the operation of electricity markets, in contrast, our focus is on characterizing the difference between real-time and optimal aggregate social welfare of a distribution system and deriving worst-case bounds on the difference using and OD3 algorithm and the coordination (price) signal as a tool.

Motivated by the aforementioned challenges, the main contribution of this work is to study a time-varying instance of the social welfare optimization problem, where the available power generation and utility functions of users change at time scales as fast as iterations of algorithms that solve them. This strategy is in contrast to the standard optimal power allocation paradigm, which assumes that the net power generation as well as users' utility functions are fixed and/or known *a priori* at each time step of a solution algorithm. In particular, we study how fast the optimal social welfare changes, as the available power in the system and utilities of users change. Our results, hence, show how observable changes in available power and users' utilities affect the social welfare. These results also give insights into how to regulate the power market; for example, incentivizing users to appropriately modify their utility functions, to ensure that the social welfare does not significantly deviate (and too quickly) from optimal. Next, we investigate how well the OD3 algorithm can track the time-varying optimal social welfare solutions. We provide bounds on the tracking performance that depend on how fast available supply and users' utility functions change over time. In [9], we investigated a special case of the OD3 algorithm – the static case, whereas in this paper we consider the more challenging *real-time* case. Preliminary results of this work appeared in [19].

The rest of the paper is organized as follows: Following notation and definitions, we introduce the system model as well as underlying assumptions, summarize the OD3 power allocation algorithm and state the main result in Section II. Sections III and IV respectively present a volatility analysis of the coordinating (price) signal and power allocation, and convergence analysis of our algorithm. We follow in Section V by briefly characterizing the choice of step-size in the presence of ramp constraints. In Section VI, we present a numerical simulation of our algorithm that demonstrates its performance and make our conclusions in Section VII.

Notation: Vectors and matrices are represented by boldface lower and upper case letters, respectively. We denote the set of real numbers by $\mathbb{R}$, a vector or matrix transpose as $(\cdot)^T$, and the L_2-norm of a vector by $||\cdot||$. The gradient of a function $f(\cdot)$ is denoted $\nabla f(\cdot)$, and $\langle\ \cdot,\cdot\rangle$ denotes the inner product of two vectors. We denote the vector of ones as $\mathbf{1}$, and respectively denote the set of positive and negative reals as $\mathbb{R}_+$ and $\mathbb{R}_-$.

Definitions

We say that a function $U:\mathbb{R}^R\to\mathbb{R}$ is *σ-strongly convex* if for all $\mathbf{q}_1,\mathbf{q}_2\in\mathbb{R}^R$ we have

$$\langle\nabla U(\mathbf{q}_1)-\nabla U(\mathbf{q}_2),\mathbf{q}_1-\mathbf{q}_2\rangle\geq\sigma\|\mathbf{q}_1-\mathbf{q}_2\|^2,\ \sigma>0.$$

A function $U(\mathbf{q})$ is strongly concave if $-U(\mathbf{q})$ is *strongly convex*, i.e., if for all $\mathbf{q}_1,\mathbf{q}_2\in\mathbb{R}^R$, we have

$$-\langle\nabla U(\mathbf{q}_1)-\nabla U(\mathbf{q}_2),\mathbf{q}_1-\mathbf{q}_2\rangle\geq\sigma\|\mathbf{q}_1-\mathbf{q}_2\|^2. \quad (2)$$

Given two vectors $\mathbf{q},\mathbf{r}\in\mathbb{R}^n$, we say $\mathbf{q}\leq\mathbf{r}$ if and only if $\mathbf{q}_i\leq\mathbf{r}_i$ for $i=1,\dots,n$.

A function $U(\cdot)$ is said to be *monotone decreasing*, if for all $\mathbf{q}_1\leq\mathbf{q}_2$, it holds that $U(\mathbf{q}_1)\geq U(\mathbf{q}_2)$.

II. Model and Algorithm

A. System Model

We consider a power distribution system comprising N users and R primary suppliers. We assume that some users may also locally generate power and sell to the system as secondary sources. The system operator's objective at time t is to solve a decentralized dynamic power allocation problem to maximize the aggregate social welfare of users in the system. We define the aggregate social welfare as the utility users gain from consuming or selling power $\mathbf{q}_i(t)$ at time t; that is, $\sum_{i=1}^N U_i^t(\mathbf{q}_i(t))$, where $U_i^t(\cdot)$ is the utility of user i at time t. We assume the entire system capacity available to the suppliers is the actual power generation, and that users have constantly changing power needs, resulting in a dynamic utility function. Let the convex function $C_j^t(\cdot)$ be the cost of generating power by supplier j at time t. Assuming that each user can choose from the R available power suppliers and let the allocation to the N users be $\mathbf{q}_1(t),\cdots,\mathbf{q}_N(t)$, where $\mathbf{q}_i(t)\in\mathbb{R}^R$ for $i=1,\dots,N$. The jth entry of the ith vector, $\mathbf{q}_i^j(t)$, represents the power allocation to the ith user from the jth supplier at time t. Furthermore, let the generation of the R primary suppliers at time t be denoted by the vector $Q(t)\in\mathbb{R}^R$, where Q_j, the jth entry of Q, represents the power capacity at the jth supplier. At time t, the objective of the system operator is to maximize the social welfare by solving the following

time-varying economic dispatch optimization program [20]:

$$\begin{aligned} \underset{\mathbf{q}_1(t),\ldots,\mathbf{q}_N(t)}{\text{maximize}} \quad & \sum_{i=1}^{M} U_i^t(\mathbf{q}_i(t)) - \sum_{j=M+1}^{N} C_j^t(\mathbf{q}_j(t)) \\ \text{subject to} \quad & \sum_{i=1}^{N} \mathbf{q}_i(t) = Q(t), \\ & \mathbf{q}_i(t) \in \mathcal{Q}_i. \end{aligned} \tag{3}$$

In Problem (3), for each user i, the set $\mathcal{Q}_i \subset \mathbb{R}_+$ if user i consumes power and $\mathcal{Q}_i \subset \mathbb{R}_-$ if user i generates power. Without loss of generality and for straightforwardness in analysis, this paper is focused on a special case of Problem (3) where the cost of generation is implicit and users are allowed to sell back power to the grid. Essentially, the system operator is focused on the efficiency and optimal operation of the system with a specific interest in maximizing the aggregate utility of users in the system. In particular, the problem we consider is:

$$\begin{aligned} \underset{\mathbf{q}_1(t),\ldots,\mathbf{q}_N(t)}{\text{maximize}} \quad & \sum_{i=1}^{N} U_i^t(\mathbf{q}_i(t)) \\ \text{subject to} \quad & \sum_{i=1}^{N} \mathbf{q}_i(t) = Q(t), \end{aligned} \tag{4}$$

where the constraint $\mathbf{q}_i(t) \in \mathcal{Q}_i$ is implicit, and the utility gained for power generation is thought of in terms of monetary profit. The system capacity at each time, $Q(t)$, is assumed to be the net power generated at time t. Thus, the equality constraint guarantees that power generation from both primary and secondary sources is not wasted. The economic dispatch problem in [20] is, for instance, a special case of Problem (4). In section V, we address the case with local ramp constraints.

Remark 1. *In practice, there will be local capacity constraints on the decision variables $\{\mathbf{q}_i(t)\}_{i=1}^N$. While we do not specifically model those capacity constraints, appropriate utility functions $U_i(\cdot)$ can be designed to approximate those constraints. Additionally, note that our formulation in Problem* (4) *is general in the sense that we allow $\{\mathbf{q}_i(t)\}_{i=1}^N$ to be negative, indicating the instances where users sell power back to suppliers. This means that each user i can be elastic suppliers as well.*

Problems like (4) are common in power systems, where the utility function may, for instance, represent deviation from desired power demand and capture time-varying utility in power consumed. To make the structure of the users' utility function in Problem (4) a bit more precise, we briefly present examples of such functions below:

Example 1. *Consider a power distribution system with one supplier. Let $s_i(t) \in \mathbb{R}$ and $q_i(t) \in \mathbb{R}$ respectively be the power demand and allocation of user i at time t. Let a_i be user i's utility if the allocation $q_i(t)$ satisfies the demand $s_i(t)$. If $q_i(t) \neq s_i(t)$ then user i suffers dis-utility. This can be modeled with the time-varying utility function*

$$U_i^t(q_i(t)) = a_i - (q_i(t) - s_i(t))^2, \tag{5}$$

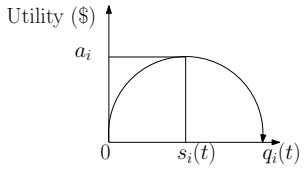


Fig. 1: A sample utility function showing saturation and diminishing marginal utility for excess consumption.

where the parameter $s_i(t)$ captures time-varying utility for user i. The quadratic function above is strongly concave in $\mathbf{q}_i(t)$ and is not uncommon in such social welfare optimization problems [8].

Example 2. *In many cases, users have a choice between different power sources. For instance, if a user i has two suppliers, that is, $\mathbf{q}_i(t) \in \mathbb{R}^2$, the user may demand some power, $s_{i,1}(t)$, from conventional energy sources (e.g coal) and power, $s_{i,2}(t)$, from renewable (e.g solar, wind) energy sources. An example of a utility function that captures user i's options is*

$$U_i^t(\mathbf{q}_i(t)) = a_i - \left[(\mathbf{q}_i^1(t) - s_{i,1}(t))^2 + \delta(\mathbf{q}_i^2(t) - s_{i,2}(t))^2\right],$$

where $\delta > 0$ is a weighting parameter representing preferences on the power sources.

Since commonly used optimization algorithms are usually iterative, the time-varying nature of the users' local utility functions $U_i^t(\cdot)$ and capacity of primary power suppliers $Q(t)$ poses a challenge to solving Problem (4). Another challenge is that the utility functions $U_i^t(\cdot)$ are known locally by user i; hence, a solution to Problem (4) needs to be decentralized. We make some assumptions on Problem (4).

Assumption 1. *(Strong Concavity): We assume the users' utility functions, $U_i^t(\cdot)$, are strongly concave in the variable $\mathbf{q}_i^t$ with concavity parameter σ.*

Assumption 2. *(Lipschitz Gradients): The gradients of the utility function of each user i is Lipschitz continuous at each time-step. In other words, for all vector pairs $\mathbf{q}_1$ and $\mathbf{q}_2$, $\|\nabla U_i^t(\mathbf{q}_1) - \nabla U_i^t(\mathbf{q}_2)\| \leq L\|\mathbf{q}_1 - \mathbf{q}_2\|$, where $L < \infty$ is the Lipschitz constant.*

Assumptions 1 and 2 can easily be accomplished when σ and L are globally known. For example, users can be free to design their utility functions at each time, independent of the other users, so long as the concavity and Lipschitz parameters fall within the global parameters. Our objective is to present a distributed solution to Problem (4) that enables us characterize variation in the social welfare of the system at each time-step as presented in Theorem 1. We investigate the performance of the well-known dual descent algorithm for Problem (4). Decentralized algorithms that solve (4) can be achieved via duality theory. Let $\mathbf{p}(t) \in \mathbb{R}^R$ be the dual variable representing

the price charged by the R different suppliers at time t. Then, the dual problem of (4) is

$$\underset{\mathbf{p}}{\text{minimize}} \quad D^t(\mathbf{p}(t)) \tag{6}$$

where $D^t(\cdot)$ is the dual function and $D^t(\cdot)$ is:

$$D^t(\mathbf{p}(t)) = \max_{\mathbf{q}_i} \quad \mathcal{L}(\mathbf{q}, \mathbf{p}), \tag{7}$$

where

$$\mathcal{L}(\mathbf{q}, \mathbf{p}) = \left[\sum_{i=1}^{N} U_i^t(\mathbf{q}_i(\mathbf{p}(t))) - \mathbf{p}(t)^T \left(\sum_{i=1}^{N} \mathbf{q}_i(\mathbf{p}(t)) - Q(t)\right)\right],$$

and $\mathbf{q}_i(\mathbf{p}(t))$ is the power demand of user i based on price $\mathbf{p}(t)$ at time t. Based on (7), the respective local problem for each user i is to solve:

$$\mathbf{q}_i(\mathbf{p}(t)) = \arg\max_{\mathbf{q}_i} \left[U_i^t(\mathbf{q}_i) - \mathbf{p}^T \mathbf{q}_i\right]. \tag{8}$$

As expected, the optimal power demand at each time-step t for each user i depends on its utility as well as the coordinating signal (price) received from the system operator. The structure of the problem enables us to claim the following result:

Lemma 1. *(Strong Duality): Consider Problem* (4) *at time t, suppose Assumptions 1 and 2 hold, and let $\mathbf{p}^*(t)$ be the optimal solution to* (6)*. Then $\mathbf{q}(\mathbf{p}^*(t)) = \{\mathbf{q}_i(\mathbf{p}^*(t))\}_{i=1}^N$ (cf.* (8)*) is the optimal solution to* (4)*.*

Proof. Convexity of the problem coupled with the constraints $\sum_{i=1}^N \mathbf{q}_i(t) = Q(t)$, and $\mathbf{q}_i \in \mathbb{R}^R$ ensures that Problem (4) satisfies Slater's condition, yielding a zero duality gap [21, Chapter 5]. □

Proposition 1. *Consider Problem* (4) *and suppose Assumptions 1 and 2 hold, then the dual function* (7) *is strongly convex in $\mathbf{p}$ with parameter N/L, and its gradient is Lipschitz continuous with parameter $N\sigma$.*

Proof. See Appendix A for the proof. □

B. Algorithm

We solve Problem (4) in a decentralized manner via a dual descent algorithm with $\eta > 0$ as step size. Based on aggregate consumption at the previous time-step, the system operator determines and broadcasts the coordinating (price) signal $\mathbf{p}(t)$. Consumers then use the pricing information to compute their optimal allocation based on their current utility functions. The system operator gathers the total consumption and computes the next price. We summarize the OD3 procedure in Algorithm 1. Given Algorithm 1 to solve Problem (4), of interest is to understand and characterize how the online and optimal system decisions and social welfare of the system change with time, given the time-varying supply capacities and consumer utility functions. To prevent drastic variations in the users' utility functions and suppliers' capacities, we make the following assumptions:

Algorithm 1 An Online Decentralized Dual Descent (OD3) Algorithm for optimal power allocation

Initialization: Suppliers set initial price $\mathbf{p}(0)$; and let $\eta \in]0, \bar{\eta}]$ be given.
for $t = 0, \ldots$ **do**
  Suppliers broadcast $\mathbf{p}(t)$
  for Users $i = 1, \ldots, N$ **do**
    User i receives $\mathbf{p}(t)$ and solves (8)
  end for
  Suppliers measure $\sum_{i=1}^N \mathbf{q}_i(t) - Q(t)$ and compute next price
  $\mathbf{p}(t+1) = \mathbf{p}(t) - \eta(\sum_{i=1}^N \mathbf{q}_i(t) - Q(t))$
end for

Assumption 3. *Between successive time-steps the changes in suppliers' capacities is bounded by γ; that is,*

$$\|Q(t) - Q(t+1)\| \leq \gamma, \ \forall\ t. \tag{9}$$

Assumption 4. *We assume that each user i has an upper bound on how much its utility function changes between consecutive time steps; that is,*

$$\|\nabla U_i^{t+1}(\mathbf{q}_i) - \nabla U_i^t(\mathbf{q}_i)\| \leq \alpha, \quad \forall\ i, \mathbf{q}_i \text{ and } t. \tag{10}$$

Assumption 3 prevents drastic changes in the supply capacities between consecutive time-steps, ensuring a smoothness property in capacity at the supplier over time. It is also a pragmatic assumption in the real physical system, as generators are known to have ramp constraints [22, 23]. Assumption 4 ensures that the rate of change in user utility function over time is bounded – again to avoid drastic changes in power demand between consecutive time-steps for all users.

Theorem 1. *(Main Result) Suppose Algorithm 1 is used to solve Problem 4, suppose Assumptions 3 and 4 hold, and that each $U_i^t(\cdot)$ is Lipschitz with parameter L'. Let $\mathbf{q}_i(t)$ and $\mathbf{q}_i^*(t)$ respectively be the power allocation obtained from Algorithm 1 and the optimal power allocation at time t. The difference between the aggregate online social welfare (given fluctuations in the system) and aggregate optimal social welfare is bounded by*

$$\left\|\sum_{i=1}^{N} U_i^t(\mathbf{q}_i(t)) - \sum_{i=1}^{N} U_i^t(\mathbf{q}_i^*(t))\right\| \leq W, \tag{11}$$

where

$$W = NL'\left[\frac{c^t}{\sigma}||\mathbf{p}(0) - \mathbf{p}^*(0)|| + \frac{L^2}{\sigma^2}\left(\frac{\gamma}{N} + \frac{\alpha}{\sigma}\right)\right],$$

$$c = \left(1 - \frac{2\eta\sigma N}{(1+\sigma L)}\right)^{1/2} \quad \text{and} \quad 0 < \eta \leq \frac{2L}{N(1+L\sigma)}.$$

Furthermore, $\mathbf{p}(0)$ and $\mathbf{p}^*(0)$ are respectively the initial online and optimal prices, γ and α are respectively the bounds on variations of the suppliers' generation and users' utility functions.

The proof of Theorem 1 is given in Section IV-B, and based

on a series of lemmas and propositions developed in the succeeding sections of the paper.

Remark 2. *As can be observed, the upper bound on difference between the aggregate optimal and online social welfare of the system depends not only on the structure of the users' utility function, but also on how much it changes; this, of course, is in addition to the magnitude of generation fluctuation. Also, the bound comprises a fixed term that depends on these fluctuations and a transient term that decays as the OD3 algorithm attempts to closely track the optimal decisions.*

Note that for Theorem 1, we assume that the utility functions themselves (and not their gradients as in Assumption 2) are Lipschitz continuous.

III. Volatility Analysis

In this section we derive bounds on changes in the optimal price $\mathbf{p}^*(t)$ between consecutive time-steps, based on Assumptions 3 and 4.

Theorem 2. *(Volatility of the Optimal Price): Consider Problem* (4) *with Dual Problem* (6)*. Suppose Assumptions 2 – 4 hold; then*

$$\|\mathbf{p}^*(t) - \mathbf{p}^*(t+1)\| \leq \frac{L^2}{\sigma}\left(\frac{\gamma}{N} + \frac{\alpha}{\sigma}\right). \tag{12}$$

Proof. Given that the Dual Problem (6) is unconstrained, convex and differentiable, for all t

$$\nabla D^t(\mathbf{p}^\star(t)) = \sum_{i=1}^{N} [\nabla U_i^t]^{-1}(\mathbf{p}^*(t)) - Q(t) = 0.$$

In particular, let $\Gamma_t(\mathbf{p}) = \sum_{i=1}^{n} [\nabla U_i^t]^{-1}(\mathbf{p})$; then $\mathbf{p}^*(t) = \Gamma_t^{-1}(Q(t))$. And by Lemma 4 in Appendix E, the inverse function of the gradient exists. Using the triangle inequality, it follows that

$$\begin{aligned}\|\mathbf{p}^*(t) - \mathbf{p}^*(t+1)\| &= ||\Gamma_t^{-1}(Q(t)) - \Gamma_{t+1}^{-1}(Q(t+1))|| \\ &\leq ||\Gamma_t^{-1}(Q(t)) - \Gamma_t^{-1}(Q(t+1))|| \\ &\quad + ||\Gamma_t^{-1}(Q(t+1)) - \Gamma_{t+1}^{-1}(Q(t+1))|| \\ &\leq \frac{L^2\gamma}{\sigma N} + \frac{\alpha L^2}{\sigma^2} = \frac{L^2}{\sigma}\left(\frac{\gamma}{N} + \frac{\alpha}{\sigma}\right).\end{aligned}$$

□

Theorem 2 above implies that price fluctuation is higher when the changes in supply capacity γ and user utility functions α are high. Observe that though the system operator's decisions are optimal between successive time-steps, the deviation in pricing between consecutive time-steps is non-zero. This is because (both the cost function and constraint set of) the particular social welfare problem being solved by the system operator changes at each time-step. These changes, as noted earlier, are due to the variation in supply capacity and changes in utility functions of the consumers (respectively assumed to be upper-bounded by γ and α). Furthermore, the presence of more users in the system, captured by N (in the first term of the bound) dampens any significant price changes observed between consecutive time-steps caused by the time-varying supplier capacities. In addition, the more concave the form of the utility function of users (for $\sigma > 1$), the lower the volatility in the price signal between consecutive time-steps.

Given that the power allocation to any user $\mathbf{q}_i(\mathbf{p}(t))$ at any time t depends on the price at that time, a consequence of Theorem 2 are bounds on the change in the optimal primal variables between consecutive time-steps.

Corollary 1. *Consider Problem* (4)*, and given Assumptions 1 and 2, the optimal power allocation between consecutive iterates, satisfies*

$$\|\mathbf{q}_i^*(t+1) - \mathbf{q}_i^*(t)\| \leq \frac{L^2}{\sigma^2}\left(\frac{\gamma}{N} + \frac{\alpha}{\sigma}\right) + \frac{\alpha}{\sigma}.$$

Proof. See Appendix B for a proof. □

Similar to the bound in Theorem 2, we find in Corollary 1 that the magnitude of changes in power allocation between consecutive time steps is proportional to the changes experienced in supply capacities and consumers' utility functions. Furthermore, as highlighted earlier, when no fluctuations are recorded in consumers' utility functions, the size of the user base mitigates the changes in power demand between successive time steps. The bounds derived on the changes in optimal price and power allocation to each user show that so long as the supply capacity and consumer utility functions are not fixed, the optimal operating point of the distribution system will fluctuate between time instances t. Further, one can also observe from the bounds above that if the optimal operating point is fluctuating, the fluctuation is bounded.

IV. Convergence Analysis of Algorithm 1

In this section, we analyze convergence of Algorithm 1 and its performance with regards to tracking the optimal solution (in section IV-A). In section IV-B, we present a proof of the main result of the paper based on the results established in section IV-A.

A. Convergence and Tracking

In Lemma 2, we show that by appropriately choosing a step-size η for Algorithm 1, the coordinating/price signal update (cf line 7 of Algorithm 1) is a descent direction.

Lemma 2. *For every step-size η such that $0 < \eta \leq 2L/(N(1+L\sigma))$, and $\mathbf{p} \in \mathbb{R}^R$, the following holds*

$$||\mathbf{p}^*(t) - (\mathbf{p} - \eta\nabla D_t(\mathbf{p}))||^2 \leq c^2||\mathbf{p}^*(t) - \mathbf{p}||^2, \tag{13}$$

where c is as defined in Theorem 1.

Proof. Choose any $\mathbf{p} \in \mathbb{R}^R$, and let $r_t := \|\mathbf{p}^*(t) - \mathbf{p}\|$. Then

$$\begin{aligned}||\mathbf{p}^*(t) - (\mathbf{p} - \eta\nabla D_t(\mathbf{p}))||^2 &= r_t^2 - 2\eta\langle \mathbf{p}^*(t) - \mathbf{p}, \nabla D_t(\mathbf{p})\rangle \\ &\quad + \eta^2||\nabla D_t(\mathbf{p})||^2 \\ &\leq \left(1 - \frac{2\eta N\sigma}{1+L\sigma}\right) r_t^2 \\ &\quad + \eta\left(\eta - \frac{2L}{N(1+L\sigma)}\right)\|\nabla D_t(\mathbf{p})\|^2,\end{aligned}$$

where we have used the fact that $D(\cdot)$ is N/L-strongly convex, has $N\sigma-$Lipschitz continuous gradients and satisfies [24, Theorem 2.1.12]:

$$\langle \nabla D(\mathbf{p}) - \nabla D(\mathbf{p}^*(t)), \mathbf{p} - \mathbf{p}^* \rangle \geq \|\mathbf{p} - \mathbf{p}^*(t)\|^2 + \|\nabla D(\mathbf{p}) - \nabla D(\mathbf{p}^*(t))\|^2, \quad (14)$$

and that at the optimal $\mathbf{p}^*(t)$, $\nabla D_t(\mathbf{p}^*(t)) = 0$. And the last inequality in the proof comes from the fact that $D(\cdot)$ ($N\sigma-$Lipschitz continuous gradients) satisfies

$$0 \leq D(\mathbf{p}^*(t)) - D(\mathbf{p}) - \langle \nabla D(\mathbf{p}), \mathbf{p}^*(t) - \mathbf{p} \rangle \leq \frac{N\sigma}{2} \|\mathbf{p} - \mathbf{p}^*(t)\|^2.$$

If the step-size η is chosen such that $0 < \eta < 2L/N(1+L\sigma)$, one always obtains $\eta \left(\frac{2L}{N(1+\sigma L)} - \eta \right) ||\nabla D_t(\mathbf{p})||^2 \geq 0$; thence,

$$||\mathbf{p}^*(t) - (\mathbf{p} - \eta \nabla D_t(\mathbf{p}))||^2 \leq \left(1 - \frac{2\eta\sigma N}{(1+\sigma L)}\right)^2 ||\mathbf{p}^*(t) - \mathbf{p}||^2,$$

and it follows that, for each iterate $\mathbf{p}$ obtained from the OD3 Algorithm 1, the bound in (13) holds. □

The convergence of the OD3 algorithm in Lemma 2 exploits strong convexity of the dual function, $D(\cdot)$, derived from the structure of Problem (4). Choice of η is important for convergence of the algorithm in the static case, and tracking of the optimal solutions in the dynamic case. If η is small, the variable $\mathbf{p}(t)$ is updated in a conservative way, resulting in slow convergence. The step size that yields the optimal convergence can easily be obtained by minimizing the parameter c with respect to η. Based on Lemma 2, we present upper bounds on the real-time difference between the optimal price and allocation, and prices and allocations generated by our OD3 algorithm.

Theorem 3. *(Tracking the Optimal Dual Variable): Consider the system Problem* (4) *and its Dual Problem* (6)*. Regardless of the initial price* $\mathbf{p}(0)$*, the difference between the optimal price* $\mathbf{p}^*(t+1)$ *and the price iterate of Algorithm 1 at each* t *is*

$$\|\mathbf{p}(t+1) - \mathbf{p}^*(t+1)\| \leq \frac{b}{1-c} + c^t \left(\|\mathbf{p}(0) - \mathbf{p}^*(0)\| - \frac{b}{1-c} \right), \quad (15)$$

where from Theorem 2 and Lemma 2 respectively,

$$b = L^2 \left(\frac{\gamma}{\sigma N} + \frac{\alpha}{\sigma^2} \right) \quad \text{and} \quad c = \left(1 - \frac{2\eta\sigma N}{(1+\sigma L)}\right)^{\frac{1}{2}},$$

again for $0 < \eta \leq 2L/(N(1+L\sigma))$.

The bound in Eq. (15) has two components: The first term is a constant term which does not vanish over time and depends on the magnitude of the system fluctuation in capacity and changes in user utility functions. The second term of the bound is a transient term (similar to the dual gradient algorithm), which converges to 0 as $t \to \infty$.

Proof. See Appendix C for proof. □

Remark 3. *We reiterate that all the bounds presented in this paper represent the worst-case. For example, for each user* i *and at each time* t*, the actual Lipschitz constant of* $\nabla U_i^t(\cdot)$ *and concavity parameter of* $U_i^t(\cdot)$*, the change in supply capacity or utility function between consecutive time-steps may be less than the parameters used in the bounds.*

Theorem 4. *(Tracking the Optimal Primal Variable): Consider the System Problem* (4) *and its Dual Problem* (6)*. Regardless of the initial unit of power allocation per user* $\mathbf{q}_i(0)$*, the gap between the optimal allocation* $\mathbf{q}_i^*(t+1)$ *and the allocation generated by Algorithm 1 at each* t *is*

$$\|\mathbf{q}_i(t+1) - \mathbf{q}_i^*(t+1)\| \leq \frac{c^t}{\sigma} \|\mathbf{p}(0) - \mathbf{p}^*(0)\| + \frac{L^2}{\sigma^2} \left(\frac{\gamma}{N} + \frac{\alpha}{\sigma} \right),$$

where the constant c *comes from Theorem 3.*

Proof. See Appendix D for proof. □

Remark 4. *As in Theorem 3, the bound in Theorem 4 has two terms – a transient term that converges to* 0 *as* $t \to \infty$*, and a term that depends on fluctuations in user utility functions and supplier capacity in the distribution system, which does not vanish over time, so long as the systemic fluctuations exist.*

B. Proof of Main Result (Theorem 1)

We can now prove Theorem 1), by using Lemma 2, and Theorems 3 and 4.

Proof. To begin, since each $U_i^t(\cdot)$ are L'-Lipschitz continuous, for any two vectors $\mathbf{q}_i(t), \mathbf{q}_i^*(t) \in \mathbb{R}^R$, we have

$$||U_i^t(\mathbf{q}_i(t)) - U_i^t(\mathbf{q}_i^*(t))|| \leq L' ||\mathbf{q}_i(t) - \mathbf{q}_i^*(t)||.$$

Summing both sides and using the triangle inequality,

$$\left|\left| \sum U_i^t(\mathbf{q}_i(t)) - \sum U_i^t(\mathbf{q}_i^*(t)) \right|\right| \leq \sum ||U_i^t(\mathbf{q}_i(t)) - U_i^t(\mathbf{q}_i^*(t))|| \leq NL' ||\mathbf{q}_i(t) - \mathbf{q}_i^*(t)||.$$

From Theorem 4 we can further bound the Right-Hand-Side (RHS) of the above expression; that is,

$$NL' ||\mathbf{q}(t) - \mathbf{q}^*(t)|| \leq NL' \left[\frac{c^t}{\sigma} ||\mathbf{p}(0) - \mathbf{p}^*(0)|| + \frac{L^2}{\sigma^2} \left(\frac{\gamma}{N} + \frac{\alpha}{\sigma} \right) \right],$$

where

$$c = \left(1 - \frac{2\eta\sigma N}{1+\sigma L}\right)^{1/2}.$$

Since $c < 1$ (from Lemma 2), the first term of the bound above goes to zero as t goes to infinity, and the second term is a constant term. Hence, the deviation of the aggregate online social welfare (computed using the OD3 algorithm), from the aggregate optimal social welfare is bounded as proved. □

Having proved the main result of the paper, next, we consider the original Problem 4 in the presence of ramp constraints in the power allocation. In particular, the question we address next is: how well does Algorithm 1 perform if

the difference in allocation between consecutive time-steps is constrained? In the next section, we briefly highlight the importance of the problem.

V. The case with Ramp Constraints

In Problem (4), no ramp constraints were imposed on the power allocation decisions between consecutive time-steps. Ramp constraint (on both generation and allocation) in solving decision and optimization problems are important because the ramp rate[1] of generators are not infinite [22, 23]. In this section, we briefly analyze the OD3 Algorithm 1; in particular, we refine the choice of step-size in the context of the ramp rate constraints by considering the following problem:

$$\begin{aligned} \underset{\{\mathbf{q}_i(t)\}_{i=1}^N}{\text{maximize}} \quad & \sum_{i=1}^N U_i^t(\mathbf{q}_i(t)) \\ \text{subject to} \quad & \|\mathbf{q}_i(t)-\mathbf{q}_i(t-1)\| \leq \hat{q} \;\forall\; i, \\ & \sum_{i=1}^N \mathbf{q}_i(t) = Q(t). \end{aligned} \tag{16}$$

We would show that one can satisfy the ramp rate constraint by appropriately tuning the step-size η of Algorithm 1. To demonstrate this, we present the following result:

Lemma 3. *Suppose Problem* (16) *is to be solved by Algorithm 1. If the step-size in Algorithm 1 is chosen such that*

$$0 \leq \eta \leq \min\left(\frac{2L}{N(1+L\sigma)}, \frac{\hat{q}}{N\alpha\kappa}\right), \tag{17}$$

where $\kappa = c^2\|\mathbf{p}^*(t)-\mathbf{p}\|^2$ *(the RHS of Equation* (13)*),* $\hat{q}$ *is the upper bound on the ramp-rate between consecutive iterates, and* N *is the number of users, then the power allocation iterates* $\{\mathbf{q}_i(t)\}_{i=1}^N$ *produced by Algorithm 1 satisfy the ramp-rate constraints for Problem* (16)*.*

Proof. Recall that the primal update (local at each node i) is expressed as:

$$\mathbf{q}_i(\mathbf{p}(t)) = \arg\max_{\mathbf{q}_i} \left[U_i^t(\mathbf{q}_i) - \mathbf{p}^T\mathbf{q}_i\right],$$

and by Lemma 4 in Appendix E,

$$\begin{aligned} \|\mathbf{q}_i(t)-\mathbf{q}_i(t+1)\| &= \|(\nabla U_i^t)^{-1}(\mathbf{p}) - (\nabla U_i^{t+1})^{-1}(\mathbf{p})\| \\ &\leq \frac{\alpha}{\sigma}\|\mathbf{p}(t)-\mathbf{p}(t+1)\| \\ &= \frac{\alpha}{\sigma}\|\eta\nabla D^t(\mathbf{p}(t))\|. \end{aligned}$$

By Proposition 1, $\nabla D^t(\cdot)$ is Lipschitz continuous with parameter $N\sigma$. Hence, using that $\nabla D^t(\mathbf{p}^*(t)) = 0$ we have

$$\|\nabla D^t(\mathbf{p}(t))\| \leq N\sigma\|\mathbf{p}(t)-\mathbf{p}^*(t)\| \leq N\sigma\kappa,$$

which implies that $\alpha/\sigma\|\eta\nabla D^t(\mathbf{p}(t))\| \leq \alpha\eta N\kappa$. Hence, for $\alpha\eta N\kappa \leq \hat{q}$ to hold, it is necessary that $\eta \leq \hat{q}/(N\alpha\kappa)$. In Lemma 2, we established that the step-size η in Algorithm 1 needed to satisfy $\eta \leq 2L/(N(1+L\sigma))$; wherefore, (17) holds. □

To parse Lemma 3, recall that the allocation to each user $\mathbf{q}_i(t)$ depends on the prices $\mathbf{p}(t)$, which in turn depend on the rate at which the gradient of the dual function $\nabla D(\cdot)$ changes with respect to the price and the step-size η. This means that given $\hat{q}$ in Problem (16), the step-size η can be chosen small enough to simultaneously result in a contraction paramer c (in Lemma 2), and satisfy $\|\mathbf{q}_i(t)-\mathbf{q}_i(t+1)\| \leq \hat{q}$ for each user in the online allocation policy.

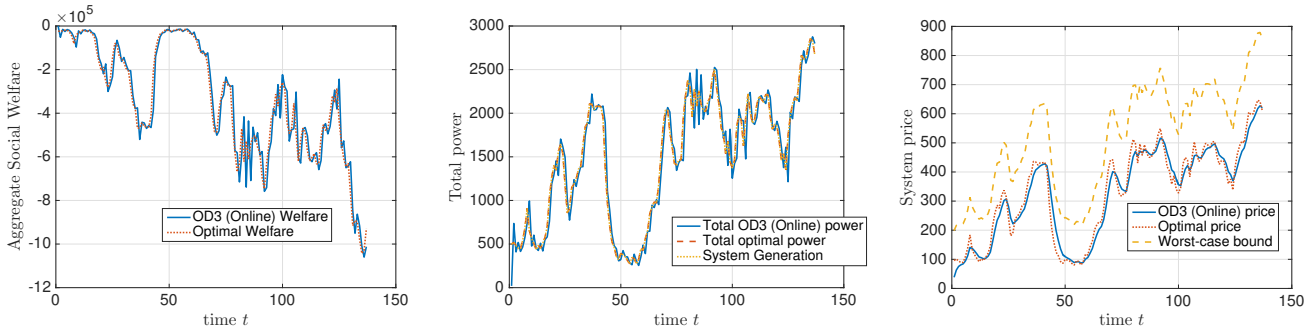
VI. Numerical Illustration

In this section, we demonstrate and validate the theoretical results developed in the preceding sections of the paper using real power generation data from the *IESO system operator Canada* Independent Electricity System Operator [25]. We illustrate the performance of the OD3 power allocation algorithm on Problem (4). In the convergence analysis of the OD3 algorithm and bounds derived, a critical element of the transient term that vanishes to zero as $t \to \infty$ is the difference between the initial online and optimal price. The farther away $\mathbf{p}(0)$ is from $\mathbf{p}^*(0)$, the longer it takes for the Algorithm 1 to converge. To show effectiveness of Algorithm 1, in the numerical experiments we will present two cases – the first in which the initial online price $\mathbf{p}(0) = 0$ is close to initial optimal price $\mathbf{p}^*(0)$; and a second case in which the initial online price $\mathbf{p}(0) = 1000$ is far away (or arbitrary chosen) from the optimal price $\mathbf{p}^*(0)$.

For the simulation, we consider a system comprising $N = 10$ users and one supplier, that is $R = 1$. The number of suppliers do not affect the theoretical bounds, so the use of one supplier in the experiments is without loss of generality. We refer interested readers to Example 2, where we presented an example of a utility function for a user with the ability to select from multiple suppliers. For each user $i = 1, \ldots, 10$, we use a utility function of the form presented in Example 1; $U_i^t(q_i(t)) = -(q_i(t) - s_i(t))^2$, where the demand $s_i(t)$ takes on different values at each time-step; thus resulting in a time-varying utility function. This choice of utility function $U_i^t(\cdot)$ is strongly concave (in this case with concavity parameter $\sigma = 2$), and satisfies Assumption 1. The *real data* on power supply was obtained from [25], where generation data from biofuel, wind and solar sources obtained in $5-$minute intervals was used. The data captures real-world variability and fluctuation experienced in renewable energy generation systems. In Figure 2, we illustrate the performance of the OD3 Algorithm on Problem (4) and compare the aggregate social welfare, allocation and price computed by the OD3 Algorithm in real time with their respective optimal counterparts in the case where the initial online price $p(0)$ is close to the initial optimal price $p^*(0)$, given the fluctuating system dynamics.

From the plots in Figure 2, one can observe that the online decisions of the OD3 algorithm are close to the optimal decisions at each time-step, despite the fluctuating consumers' utility functions and supply generation. The system generation

[1]Ramp rate is generally understood to be the speed at which generators can increase generation.



(a) Aggregate Social Welfare (b) Optimal Power Allocation (c) Optimal Price

Fig. 2: Plots 2a, 2b and 2c respectively show aggregate social welfare, aggregate optimal allocation and optimal price from using Algorithm 1 over time. Plot 2a shows the social welfare of the system computed by the OD3 algorithm in real-time and the optimal social welfare over time. Plot 2b shows the the total power allocated to users in real-time in comparison to the optimal total power and the total power supply available over time. The plot 2c illustrates the prices at each time-step generated by the real-time OD3 algorithm and the optimal price, as well as the theoretial bounds derived from our analysis.

and total optimal power allocation at each time are identical, as observed in Figure 2b. The online decisions are not equal to the optimal decisions over time because the parameters of the problem change at the same time scale as iterations of the OD3 algorithm. Further, the significant variations experienced in the generation capacity (noticed between time-steps 75 and 120) propagate to the online allocation and price (where relatively significant spikes in the power allocated and price per unit power are observed). In Figure 2c, we plot the optimal and OD3 prices, as well as the worst-case bounds derived from our analysis.

In Figure 3, we consider the case the initial online price $p(0)$ is chosen such that it is far away from the initial optimal price $p^*(0)$ and present iterates of the OD3 Algorithm, again comparing the online aggregate social welfare, aggregate power allocation and price to their respective optimal values. For this case, $p(0)$=1000 was chosen as the initial online price. Despite the large initial online price, the plots in Figure 3 indicate that the aggregate online social welfare (in Figure 3a), the aggregate online power allocation (in Figure 3b) and online price (in Figure 3c), all closely track their respective optimal counterparts. In Figure 3c, we plot the optimal and OD3 prices, as well as the worst-case bounds derived from our analysis. As expected, the system generation and total optimal power allocation in Figure 3b overlap at each time. This indicates that the OD3 algorithm is able to shrink the difference between the optimal decisions and the online decisions over time, except for the system fluctuations that do remain present over time. To lessen the volatility experienced in the first few time steps of the algorithm resulting from choosing an 'improper' initial price, a system operator may use other sources of power to allow more flexibility in pricing before the coordination stabilizes. As in Figure 2, we can also observe that in Figure 3, the spikes resulting from the inconsistent generation and dynamic consumer utility functions propagate to the coordinating signal and aggregate social welfare of the system.

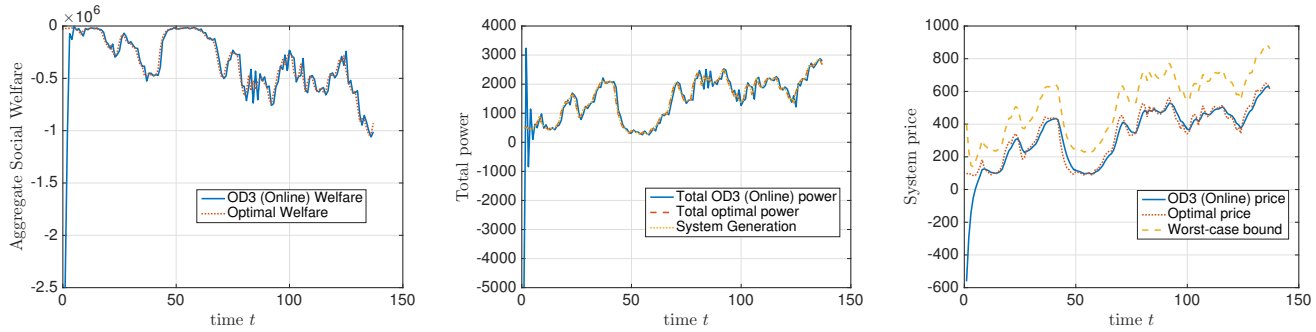
VII. Conclusions

In this paper, we considered the problem of allocating electric power to users in a power system, where the utility functions of the users and generation capacities of the suppliers vary with time. We assumed the changes in the supplier capacities and user utility functions change at the same time-scale as iterations in the algorithm, and presented a worst-case analysis of the OD3 algorithm. In particular, we investigated and characterized performance guarantees of the OD3 power allocation algorithm by deriving bounds on the difference between the aggregate real-time and optimal social welfare of the system as a function of systemic fluctuations. The OD3 algorithm uses a one-way message passing protocol between the users and the suppliers in which a system operator broadcasts a coordinating signal (price) and the users locally compute their power allocation based on the price received and their utility function. The results obtained were illustrated using real data [25]. We derived worst-case bounds on the fluctuation of the aggregate social welfare within a particular power system. It is important to notice that these worst-case bounds depend on parameters that define each user's utility functions and generation fluctuations, as well as on the number of users in the system. A natural question of interest is how a system operator may influence or incentivize users to, for instance, choose utility functions with parameters that help better optimize the aggregate social welfare and mitigate the aggregate system volatility. The ability to adjust these parameters in an altruistic way may have an effect on the amount of fluctuation experienced in the distribution system.

Appendix A
Proof of Proposition 1

Proof. Let L be the global Lipschitz continuity parameter for all users i and time t. Consider the decomposed dual function $D(\mathbf{p}) = \sum_i D_i(\mathbf{p})$, where

$$D_i^t(\mathbf{p}) = U_i^t(\mathbf{q}_i(\mathbf{p}(t))) - \mathbf{p}^T\mathbf{q}_i(\mathbf{p}(t)) + \frac{1}{N}\mathbf{p}(t)^T Q(t),$$



(a) Aggregate Social Welfare (b) Optimal Power Allocation (c) Optimal Price

Fig. 3: Plots 3a, 3b and 3c respectively show aggregate social welfare, aggregate optimal allocation and optimal price from using Algorithm 1 over time. Plot 3a shows the social welfare of the system computed by the OD3 algorithm in real-time and the optimal social welfare over time. Plot 3b shows the the total power allocated to users in real-time in comparison to the optimal total power and the total power supply available over time. The plot 3c illustrates the prices at each time-step generated by the real-time OD3 algorithm and the optimal price, as well as the theoretical bounds derived from our analysis.

with gradient

$$\nabla D_i(\mathbf{p}(t)) = Q(t)/N - \mathbf{q}_i(\mathbf{p}(t)) = Q(t)/N - (\nabla U_i^t)^{-1}(\mathbf{p}(t))$$

where $\mathbf{q}_i(\mathbf{p})$ is defined in (8). To prove the Proposition, we will show that $D_i(\mathbf{p})$ is $1/L-$convex. From [24, Theorem 2.1.9] $D_i(\mathbf{p})$ is $1/L-$convex if and only if for all $\mathbf{p}_1$, $\mathbf{p}_2$,

$$\begin{aligned}\frac{1}{L}\|\mathbf{p}_2-\mathbf{p}_1\|^2 &\leq \langle \mathbf{p}_2-\mathbf{p}_1, \nabla D_i(\mathbf{p}_2)-\nabla D_i(\mathbf{p}_1)\rangle \\ &= \left\langle \mathbf{p}_2-\mathbf{p}_1, \left(-(\nabla U_i)^{-1}(\mathbf{p}_2)\right)-\left(-(\nabla U_i)^{-1}(\mathbf{p}_1)\right)\right\rangle,\end{aligned} \tag{18}$$

where the equality comes from taking the gradient of the dual function and using the optimal solutions (8). The inverse $(\nabla U)^{-1}$ exists since U is strongly concave. Hence, ∇U_i is bijective and we can choose vectors $\mathbf{q}_1, \mathbf{q}_2 \in \mathbb{R}^R$ such that $\nabla U_i(\mathbf{q}_1) = \mathbf{p}_1$ and $\nabla U_i(\mathbf{q}_2) = \mathbf{p}_2$. Since U_i is $L-$smooth (from Assumption 1), we have that [24, Theorem 2.1.5]:

$$-\langle \nabla U_i(\mathbf{q}_2)-\nabla U_i(\mathbf{q}_1), \mathbf{q}_2-\mathbf{q}_1\rangle \geq \frac{1}{L}\|\nabla U_i(\mathbf{q}_2)-\nabla U_i(\mathbf{q}_1)\|^2,$$

or by using $\nabla U(\mathbf{q}_1) = \mathbf{p}_1$ and $\nabla U(\mathbf{q}_2) = \mathbf{p}_2$, we get

$$\left\langle \mathbf{p}_2-\mathbf{p}_1, \left(-(\nabla U_i)^{-1}(\mathbf{p}_2)\right)-\left(-(\nabla U_i)^{-1}(\mathbf{p}_1)\right)\right\rangle \geq \frac{1}{L}\|\mathbf{p}_2-\mathbf{p}_1\|^2.$$

Hence, (18) holds and we can conclude that $D_i(\mathbf{p})$ is $1/L$-convex; therefore, $\sum_{i=1}^N D_i(\cdot)$ is N/L strongly convex. □

Appendix B
Proof of Corollary 1

Proof. Using the fact that $\mathbf{q}_i^*(t) = (\nabla U_i^t)^{-1}(\mathbf{p}^*(t))$, $\nabla U_i^t(\cdot)$ is $1/\sigma$-Lipschitz continuous, and the bound in Theorem 2, via the triangle inequality one obtains

$$\begin{aligned}\|\mathbf{q}_i^*(t+1)-\mathbf{q}_i^*(t)\| &\leq \|[\nabla U_i^{t+1}]^{-1}(\mathbf{p}^*(t+1)) - [\nabla U_i^t]^{-1}(\mathbf{p}^*(t))\| \\ &\leq \|[\nabla U_i^{t+1}]^{-1}(\mathbf{p}^*(t+1)) \\ &\quad -[\nabla U_i^t]^{-1}(\mathbf{p}^*(t+1))\| \\ &\quad + \|[\nabla U_i^t]^{-1}(\mathbf{p}^*(t+1))-[\nabla U_i^t]^{-1}(\mathbf{p}^*(t))\| \\ &\leq \frac{1}{\sigma}\|\mathbf{p}^*(t)-\mathbf{p}^*(t+1)\|+\frac{\alpha}{\sigma} \\ &\leq \frac{L^2}{\sigma^2}\left(\frac{\gamma}{N}+\frac{\alpha}{\sigma}\right)+\frac{\alpha}{\sigma}\end{aligned}$$

where we have used that $[\nabla U_i^t]^{-1}$ is $1/\sigma$-Lipschitz continuous, see Lemma 4-b), and Eq. (25) in Lemma (5) (in Appendix E) together with Theorem (2). □

Appendix C
Proof of Theorem 3

Proof. Using the triangle inequality, we can express the LHS of (15) as

$$\|\mathbf{p}(t+1)-\mathbf{p}^*(t+1)\| \leq \|\mathbf{p}(t+1)-\mathbf{p}^*(t)\|+\|\mathbf{p}^*(t)-\mathbf{p}^*(t+1)\|.$$

Observe that bounds for the first and second summands of the RHS above have been respectively derived in Lemma 2 and Theorem 2. Since $\|\mathbf{p}^*(t)-\mathbf{p}^*(t+1)\| \leq b$ (from Theorem 2), one obtains $\|\mathbf{p}(t+1)-\mathbf{p}^*(t+1)\| \leq \|\mathbf{p}(t+1)-\mathbf{p}^*(t)\|+b$, which further simplifies into $\|\mathbf{p}(t+1)-\mathbf{p}^*(t+1)\| \leq c\|\mathbf{p}(t)-\mathbf{p}^*(t)\|+b$. Further simplification results in

$$\begin{aligned}\|\mathbf{p}(t+1)-\mathbf{p}^*(t+1)\| &\leq c^t\|\mathbf{p}(0)-\mathbf{p}^*(0)\|+\sum_{i=0}^{t} c^i b \\ &= c^t\|\mathbf{p}(0)-\mathbf{p}^*(0)\|+\frac{1-c^t}{1-c}b \\ &= \frac{b}{1-c}+c^t\left(\|\mathbf{p}(0)-\mathbf{p}^*(0)\|-\frac{b}{1-c}\right),\end{aligned} \tag{19}$$

which concludes the proof. □

APPENDIX D
PROOF OF THEOREM 4

Proof. For simplicity in notation, let $\mathcal{Q} = \|\mathbf{q}_i(t+1) - \mathbf{q}_i^*(t+1)\|$. Via the triangle inequality, we can express the LHS of (15) as

$$\begin{aligned}\mathcal{Q}=&\|(\nabla U_i^{t+1})^{-1}(\mathbf{p}(t+1)) - (\nabla U_i^{t+1})^{-1}(\mathbf{p}^*(t+1))\| \\ &\leq \|(\nabla U_i^{t+1})^{-1}(\mathbf{p}(t+1)) - (\nabla U_i^{t+1})^{-1}(\mathbf{p}^*(t))\| \\ &+ \ \|(\nabla U_i^{t+1})^{-1}(\mathbf{p}^*(t)) - (\nabla U_i^{t+1})^{-1}(\mathbf{p}^*(t+1))\| \\ &\leq \frac{1}{\sigma}\|\mathbf{p}(t+1) - \mathbf{p}^*(t)\| + \frac{1}{\sigma}\|\mathbf{p}^*(t) - \mathbf{p}^*(t+1)\| \\ &\leq \frac{1}{\sigma}c\|\mathbf{p}(t) - \mathbf{p}^*(t)\| + \frac{1}{\sigma}\frac{L^2}{\sigma}\left(\frac{\gamma}{N} + \frac{\alpha}{\sigma}\right) \\ &\leq \frac{c^t}{\sigma}\|\mathbf{p}(0) - \mathbf{p}^*(0)\| + \frac{L^2}{\sigma^2}\left(\frac{\gamma}{N} + \frac{\alpha}{\sigma}\right),\end{aligned}$$

where the second inequality comes from the fact that $\nabla U_i^t(\cdot)$ are $1/\sigma$ Lipschitz continuous, see Lemma 4-b), and the third inequality from Theorem 2. □

APPENDIX E
ADDITIONAL LEMMAS AND ASSOCIATED PROOFS

Lemma 4. *Suppose that U_i is σ-concave has L-Lipschitz gradient, for all $i = 1, \cdots, N$. Then the following holds:*

a) ∇U_i is bijective on $\mathbb{R}^R$ for all $i=1, \cdots, N$, i.e., ∇U_i^{-1} exists.

b) ∇U_i^{-1} is $1/\sigma$-Lipschitz continuous and monotone decreasing with parameter σ/L^2, i.e., for all $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^R$

$$-\langle \nabla U_i^{-1}(\mathbf{x}_1) - \nabla U_i^{-1}(\mathbf{x}_2), \mathbf{x}_1 - \mathbf{x}_2\rangle \geq (\sigma/L^2)||\mathbf{x}_1 - \mathbf{x}_2||^2.$$

c) $\Gamma := \sum_{i=1}^N (\nabla U_i)^{-1}$ is bijective, N/σ-Lipschitz continuous and $N\sigma/L^2$ monotone decreasing.

d) Γ^{-1} is $L^2/(N\sigma)$-Lipschitz continuous.

Proof. For space constraints, we omit the proofs and refer readers to classic texts in analysis such as [26] and [27].

a) We refer readers to [26] for proof. We first show that ∇U_i is injective. For any $\mathbf{q}_1, \mathbf{q}_2 \in \mathbb{R}^R$, since U_i is σ-strongly-concave we have from (2) and the Cauchy-Schwarz inequality that

$$||\nabla U_i(\mathbf{q}_1) - \nabla U_i(\mathbf{q}_2)|| \geq \sigma\|\mathbf{q}_1 - \mathbf{q}_2\|. \tag{20}$$

Hence, $\nabla U_i(\mathbf{q}_1) = \nabla U_i(\mathbf{q}_2)$ implies that $\mathbf{q}_1 = \mathbf{q}_2$ and ∇U_i is injective. Next, we show that ∇U_i is surjective or equivalently, that $-\nabla U_i$ is surjective. Since $-U_i$ is a convex function and its gradient is continuous we have by [27, Corollary 20.25], ∇U_i is maximally monotone. From [27, Proposition 22.8], we conclude that each ∇U_i is surjective on $\mathbb{R}^N$.

b) From a) ∇U_i is bijective and for any $\mathbf{y}_1, \mathbf{y}_2 \in \mathbb{R}^R$ there exists $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^R$ such that $\nabla U_i(\mathbf{x}_1) = \mathbf{y}_1$ and $\nabla U_i(\mathbf{x}_2) = \mathbf{y}_2$. Using that ∇U_i is Lipschitz continuous and σ-strongly concave, one obtains

$$||\nabla U_i(\mathbf{x}_1) - \nabla U_i(\mathbf{x}_2)|| \leq L||\mathbf{x}_1 - \mathbf{x}_2||, \tag{21}$$

$$\sigma||\mathbf{x}_1 - \mathbf{x}_2||^2 \leq -\langle \nabla U_i(\mathbf{x}_1) - \nabla U_i(\mathbf{x}_2), \mathbf{x}_1 - \mathbf{x}_2\rangle. \tag{22}$$

By substituting $\mathbf{y_1}$ and $\mathbf{y}_2$ into (21) and (22) we have

$$||\mathbf{y}_1 - \mathbf{y}_2|| \leq L||\nabla U_i^{-1}(\mathbf{y}_1) - \nabla U_i^{-1}(\mathbf{y}_2)||,$$

$$\sigma||\nabla U_i^{-1}(\mathbf{y}_1) - \nabla U_i^{-1}(\mathbf{y}_2)||^2 \leq -\langle \mathbf{y}_1 - \mathbf{y}_1, \nabla U_i^{-1}(\mathbf{y}_1) - \nabla U_i^{-1}(\mathbf{y}_1)\rangle.$$

By combining the inequalities above, we get

$$-\langle \nabla U_i^{-1}(\mathbf{y}_1) - \nabla U_i^{-1}(\mathbf{y}_2), \mathbf{y}_1 - \mathbf{y}_2\rangle \geq (\sigma/L^2)||\mathbf{y}_1 - \mathbf{y}_2||^2.$$

c) We start by showing that Γ is $N\sigma/L^2$-strongly monotone decreasing. i.e., (2) holds where $\nabla U = \Gamma$, with the concavity parameter σ/L^2. For all $\mathbf{x}_1, \mathbf{x} \in \mathbb{R}^R$ we have

$$\begin{aligned}-\langle \Gamma(\mathbf{x}_1) - \Gamma(\mathbf{x}_2), \mathbf{x}_1 - \mathbf{x}_2\rangle &= \sum_{i=1}^N -\langle \nabla U_i^{-1}(\mathbf{x}_1) \\ &\quad - \nabla U_i^{-1}(\mathbf{x}_2), \mathbf{x}_1 - \mathbf{x}_2\rangle \\ &\geq \frac{N\sigma}{L^2}\|\mathbf{x}_1 - \mathbf{x}_2\|^2.\end{aligned}$$

Since Γ is $N\sigma/L^2$-strongly monotone decreasing it follows by the same arguments as a) that Γ is bijective. The fact that Γ is N/σ-Lipschitz continuous follows from the fact that the sum of Lipschitz continuous functions is also Lipschitz continuous.

d) Using the same argument as in b) $\Gamma(\cdot)$ is $N\sigma/L^2$ monotone decreasing. Hence, for any $\mathbf{y}_1, \mathbf{y}_2, \mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}$ such that $\mathbf{y}_1 = \Gamma(\mathbf{x}_1)$ and $\mathbf{y}_2 = \Gamma(\mathbf{x}_2)$, we get by using Cauchy-Schwarz that

$$\frac{N\sigma}{L^2}||\mathbf{x}_1 - \mathbf{x}_2|| \leq ||\Gamma(\mathbf{x}_1) - \Gamma(\mathbf{x}_2)||. \tag{23}$$

By rearranging we get

$$||\Gamma^{-1}(\mathbf{y}_1) - \Gamma^{-1}(\mathbf{y}_2)|| \leq \frac{L^2}{N\sigma}||\mathbf{y}_1 - \mathbf{y}_2||. \tag{24}$$

□

Lemma 5. *Suppose U_i^t are σ-strongly concave and the gradients are L-Lipschitz continuous for all i and t and Assumption 4 holds. Then for all $\mathbf{p}, Q \in \mathbb{R}^N$ the following inequalities hold:*

$$\left\|[\nabla U_i^t]^{-1}(\mathbf{p}) - [\nabla U_i^{t+1}]^{-1}(\mathbf{p})\right\| \leq \frac{\alpha}{\sigma}, \tag{25}$$

$$\left\|\Gamma_t^{-1}(Q) - \Gamma_{t+1}^{-1}(Q)\right\| \leq \frac{\alpha L^2}{\sigma^2}, \tag{26}$$

where $\Gamma_t(\mathbf{p}) = \sum_{i=1}^N [\nabla U_i^t]^{-1}(\mathbf{p})$.

Proof. We start by showing that for $\mathbf{p} \in \mathbb{R}^N$ it holds that

$$\left\|[\nabla U_i^t]^{-1}(\mathbf{p}) - [\nabla U_i^{t+1}]^{-1}(\mathbf{p})\right\| \leq \frac{\alpha}{\sigma}. \tag{27}$$

By Lemma 4-a) there exist $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^N$ such that $\nabla U_i^t(\mathbf{x}_1) = \nabla U_i^{t+1}(\mathbf{x}_2) = \mathbf{p}$. Hence using the triangle inequality we get that

$$\begin{aligned}0 = \|\mathbf{p} - \mathbf{p}\| &= \ \|\nabla U_i^{t+1}(\mathbf{x}_2) - \nabla U_i^t(\mathbf{x}_1)\| \\ &\geq \|\nabla U_i^{t+1}(\mathbf{x}_2) - \nabla U_i^{t+1}(\mathbf{x}_1)\| - \|\nabla U_i^{t+1}(\mathbf{x}_1) - \nabla U_i^t(\mathbf{x}_1)\| \\ &\geq \ \sigma\|\mathbf{x}_2 - \mathbf{x}_1\| - \alpha,\end{aligned} \tag{28}$$

where the second inequality comes by that $U_i(\cdot)$ is σ-concave, (see Equation 2.1.17 in [24, Theorem 2.1.9]. By rearranging (28) we get that

$$\frac{\alpha}{\sigma} \geq ||\mathbf{x}_2 - \mathbf{x}_1|| = \left\| [\nabla U_i^t]^{-1}(\mathbf{p}) - [\nabla U_i^{t+1}]^{-1}(\mathbf{p}) \right\|. \quad (29)$$

By summing over (27) and using triangle inequality we get also that

$$\|\Gamma_t(\mathbf{p}) - \Gamma_{t+1}(\mathbf{p})\| \leq \frac{\alpha N}{\sigma}. \quad (30)$$

Now take any $Q \in \mathbb{R}^R$. Using that Γ_t and Γ_{t+1} are bijective, (from Lemma 4-c)), there exists $\mathbf{p}_1, \mathbf{p}_2 \in \mathbb{R}^N$ such that $Q = \Gamma_t(\mathbf{p}_1) = \Gamma_{t+1}(\mathbf{p}_2)$. Similarly, we obtain

$$\begin{aligned} 0 = \|Q - Q\| = & \|\Gamma_{t+1}(\mathbf{p}_2) - \Gamma_t(\mathbf{p}_1)\| \\ & \geq \|\Gamma_{t+1}(\mathbf{p}_2) - \Gamma_{t+1}(\mathbf{p}_1)\| - \|\Gamma_{t+1}(\mathbf{p}_1) - \Gamma_t(\mathbf{p}_1)\| \\ & \geq \frac{N\sigma}{L^2}\|\mathbf{p}_2 - \mathbf{p}_1\| - \frac{\alpha N}{\sigma}, \end{aligned} \quad (31)$$

since from Lemma 4-c) $\Gamma_{t+1}(\cdot)$ is $N\sigma/L^2$-convex. Finally, by rearranging (31) and using that Γ_t and Γ_{t+1} are bijective, (from Lemma 4-c)), we get that

$$\frac{\alpha L^2}{\sigma^2} \geq \|\mathbf{p}_2 - \mathbf{p}_1\| = \|\Gamma_t^{-1}(Q) - \Gamma_{t+1}^{-1}(Q)\|. \quad (32)$$

□

REFERENCES

[1] Corinna Klessmann, Christian Nabe, and Karsten Burges. Pros and cons of exposing renewables to electricity market risks - a comparison of the market integration approaches in germany, spain, and the uk. *Energy Policy*, 36(10):3646–3661, 2008.

[2] Ruggero Schleicher-Tappeser. How renewables will change electricity markets in the next five years. *Energy policy*, 48:64–75, 2012.

[3] C Lindsay Anderson and Judith B Cardell. Reducing the variability of wind power generation for participation in day ahead electricity markets. In *Hawaii International Conference on System Sciences, Proceedings of the 41st Annual*, pages 178–178. IEEE, 2008.

[4] Lion Hirth. The market value of variable renewables: The effect of solar wind power variability on their relative price. *Energy economics*, 38:218–236, 2013.

[5] J Cox. Impact of intermittency: how wind variability could change the shape of the british and irish electricity markets. *Poyry Energy (Oxford) Ltd, Oxford*, 2009.

[6] Tim Mount, Alberto Lamadrid, Surin Maneevitjit, Bob Thomas, and Ray Zimmerman. The hidden system costs of wind generation in a deregulated electricity market. In *System Sciences (HICSS), 2010 43rd Hawaii International Conference on*, pages 1–10. IEEE, 2010.

[7] Frank P Kelly, Aman K Maulloo, and David KH Tan. Rate control for communication networks: shadow prices, proportional fairness and stability. *Journal of the Operational Research society*, pages 237–252, 1998.

[8] Steven H Low and David E Lapsley. Optimization flow control – i: basic algorithm and convergence. *IEEE/ACM Transactions on Networking (TON)*, 7(6):861–874, 1999.

[9] Sindri Magnusson, Chinwendu Enyioha, Kathryn Heal, Carlo Fischione, Na Li, and Vahid Tarokh. Distributed resource allocation using one-way communication with applications to power networks. *In Proceedings of the 2016 IEEE Conference on Information Systems and Sciences (CISS)*, 2016.

[10] Minghong Lin, Zhenhua Liu, Adam Wierman, and Lachlan LH Andrew. Online algorithms for geographical load balancing. In *Green Computing Conference (IGCC), 2012 International*, pages 1–10. IEEE, 2012.

[11] Shengwei Mei, Yingying Wang, and Zhenquan Sun. Robust economic dispatch considering renewable generation. In *Innovative Smart Grid Technologies Asia (ISGT), 2011 IEEE PES*, pages 1–5. IEEE, 2011.

[12] Qing Ling and Alejandro Ribeiro. Decentralized dynamic optimization through the alternating direction method of multipliers. In *2013 IEEE 14th Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 170–174. IEEE, 2013.

[13] Andrea Simonetto, Aryan Mokhtari, Alec Koppel, Geert Leus, and Alejandro Ribeiro. A class of prediction-correction methods for time-varying convex optimization. 2015.

[14] Mardavij Roozbehani, Munther Dahleh, Sanjoy K Mitter, et al. Volatility of power grids under real-time pricing. *Power Systems, IEEE Transactions on*, 27(4):1926–1940, 2012.

[15] Mark G Lijesen. The real-time price elasticity of electricity. *Energy economics*, 29(2):249–258, 2007.

[16] Juan M Lujano-Rojas, Cláudio Monteiro, Rodolfo Dufo-López, and José L Bernal-Agustín. Optimum residential load management strategy for real time pricing (rtp) demand response programs. *Energy Policy*, 45:671–679, 2012.

[17] Zhi Chen, Lei Wu, and Yong Fu. Real-time price-based demand response management for residential appliances via stochastic optimization and robust optimization. *Smart grid, IEEE transactions on*, 3(4):1822–1831, 2012.

[18] Peng Yang, Gongguo Tang, and Arye Nehorai. A game-theoretic approach for optimal time-of-use electricity pricing. *Power Systems, IEEE Transactions on*, 28(2):884–892, 2013.

[19] Chinwendu Enyioha, Sindri Magnusson, Kathryn Heal, Na Li, Carlo Fischione, and Vahid Tarokh. Robustness analysis for an online decentralized descent power allocation algorithm. *In Proceedings of the 2016 IEEE Workshop on Information Theory and Applications (ITA)*, 2016.

[20] Georgios B Giannakis, Vassilis Kekatos, Nikolaos Gatsis, Seung-Jun Kim, Hao Zhu, and B Wollenberg. Monitoring and optimization for power grids: A signal processing perspective. *Signal Processing Magazine, IEEE*, 30(5):107–128, 2013.

[21] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, New York, NY, USA, 2004.

[22] Chunyan Wang and SM Shahidehpour. Effects of ramp-rate limits on unit commitment and economic dispatch. *Power Systems, IEEE Transactions on*, 8(3):1341–1350, 1993.

[23] C Wang and SM Shahidehpour. Ramp-rate limits in unit commitment and economic dispatch incorporating rotor fatigue effect. *Power Systems, IEEE Transactions on*, 9(3):1539–1545, 1994.

[24] Yurii Nesterov. *Introductory lectures on convex optimization*, volume 87. Springer Science & Business Media, 2004.

[25] IESO Canada Independent Electricy System Operator. IESO Power Data canada, 2015. http://www.ieso.ca/Pages/Power-Data/default.aspx#.

[26] Walter Rudin. *Real and complex analysis*. Tata McGraw-Hill Education, 1987.

[27] Patrick L. Combettes Heinz H. Bauschke. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer New York, 2011.