Distributed Resource Allocation Using One-Way Communication with Applications to Power Networks

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

Abstract—**Typical coordination schemes for future power grids require two-way communications. Since the number of end power-consuming devices is large, the bandwidth requirements for such two-way communication schemes may be prohibitive. Motivated by this observation, we study distributed coordination schemes that require only one-way limited communications. In particular, we investigate how dual descent distributed optimization algorithm can be employed in power networks using one-way communication. In this iterative algorithm, system coordinators broadcast coordinating (or pricing) signals to the users/devices who update power consumption based on the received signal. Then system coordinators update the coordinating signals based on the physical measurement of the aggregate power usage. We provide conditions to guarantee the feasibility of the aggregated power usage at each iteration so as to avoid blackout. Furthermore, we prove the convergence of algorithms under these conditions, and establish its rate of convergence. We illustrate the performance of our algorithms using numerical simulations. These results show that one-way limited communication may be viable for coordinating/operating the future smart grids.**

I. INTRODUCTION

Network infrastructures have a central role in communication, finance, technology and other areas of the economy, for instance, electricity transmission and distribution. Due to the massive scale of real-world networks, distributed optimization algorithms are expected to increasingly play a leading role in their operation. Distributed optimization problems are commonly solved by dual-based (sub)gradient methods where primal and dual decision variables are optimized at each iteration with some sort of message passing protocol between agents in the network [1]–[3]. In power networks, for example, consumers and suppliers communicate power consumption request and pricing signals back and forth. This requires ubiquitous, two-way communication among consumers and suppliers. *Since the number of end power-consuming devices is large, the bandwidth requirements for two-way schemes may be prohibitive*. However, the communication infrastructure for the grid, especially the power distribution networks, is still under-developed and the non-availability of communication bandwidth remains a challenge [4]–[9].

Given communication bandwidth constraints, it is critical that algorithms for distributed coordination in power networks

use significantly less communication overhead – for instance, one-way message passing, as opposed to the conventional two-way message passing scheme typically used in distributed coordination algorithms. That said, a one-way message passing protocol for power distribution networks in which suppliers are only able to send out a price signal can be challenging to implement especially since a widespread system breakdown (blackouts) can occur if the aggregate power consumption exceeds the supply capacity. Given *1)* limits on power capacity at the supply end, and *2)* a one-way message passing protocol, with many power demanding devices on the network, the question of interest is how to ensure that the aggregate power consumption from users does not exceed the available supply capacity. In developing distributed algorithms to solve this problem, it is important that while the algorithm runs and users locally compute their optimal power allocation, the aggregate power consumption does not exceed the supply capacity to avoid a system failure and catastrophic blackout.

Our work relates closely to [10] where a distributed resource allocation algorithm was presented based on a fast dual gradient method. It showed that dual gradient iteration has a convergence rate of $\mathcal{O}(1/t^2)$ and yet the primal decision variables converge with a rate of $\mathcal{O}(1/t)$ (where t denotes a time-step of the algorithm). The framework presented there, however, gives no guarantee of primal feasibility at each iteration of the algorithm. The focus of this paper is on developing a distributed power allocation algorithm in which one-way communication, from the power suppliers to users, is used to coordinate the allocation. Due to the critical nature of power distribution systems, the power allocation algorithm presented in this paper

- maintains primal feasibility at each iteration and satisfies the capacity constraint, to avoid a blackout,
- assumes a one-way communication or information sharing model, in which the supplier (or network operator) sends out a *price* signal to users, users update the consumption based on the signal but do not report the consumption back to the supplier, and
- has fast convergence rates and properties.

A. Contributions of This Work

We present a distributed, dual gradient power allocation algorithm for power distribution networks. The algorithm uses only one way communication, where the supplier iteratively broadcasts a price/dual variable to the users and then measures the aggregate power usage to compute the dual gradient. Since the users actually consume power during operation (*en route* convergence) of the algorithm, it essential that the aggregate

power usage does not exceed the suppliers capacity to avoid a blackout. We show how such blackouts can be avoided by providing step-sizes that ensure that the primal iterates remain feasible at ever iteration. To ensure the primal feasibility we must sacrifice the optimal convergence rate $O(1/t^2)$ [10] for gradient methods for convex problems with Lipschitz continuous gradients and instead only get the convergence rate $O(1/t)$. Nevertheless, we prove that under mild conditions on the problem structure our algorithm attains a convergence rate $O(c^t)$, where $0 < c < 1$ and t is the time-step of the algorithm in the dual variable, and still guarantee primal feasibility. Moreover, we provide conditions which ensure that the linear convergence rate, i.e., the constant c, does not depend on the number of users, which demonstrate excellent scaling properties of the algorithm.

The rest of the paper is organized as follows: Following notation and definitions, we introduce the system model, underlying assumptions and distributed solution in Section II. Sections III-A and III-B respectively present a convergence analysis and convergent rate analysis of our proposed algorithm. In Section IV, illustration of our algorithm is presented and discussed. We make our conclusions in Section V.

B. Notation and Definitions

Definition 1. *We say that a function $f : \mathbb{R}^n \to \mathbb{R}$ is L-smooth on $\mathcal{X} \subseteq \mathbb{R}^n$ if its gradient is L-Lipschitz continuous on $\mathcal{X}$, i.e., for all $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{X}$ we have*

$$||\nabla f(\mathbf{x}_1) - \nabla f(\mathbf{x}_2)|| \leq L||\mathbf{x}_1 - \mathbf{x}_2||. \tag{1}$$

Definition 2. *We say that a function $f : \mathbb{R}^n \to \mathbb{R}$ is μ-convex on $\mathcal{X} \subseteq \mathbb{R}^n$ if for all $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{X}$,*

$$f(\mathbf{x}_2) \geq f(\mathbf{x}_1) + \langle \nabla f(\mathbf{x}_1), \mathbf{x}_2 - \mathbf{x}_1 \rangle + \frac{\mu}{2}||x_2 - x_1||^2. \tag{2}$$

Similarly, we say a function is μ-concave if it is μ-strongly concave.

II. System Model and Algorithm

In this paper we focus on an abstract power distribution model, and for simple exposition we ignore some specific power network constraints. We consider an electric network with N users whose set is denoted by $\mathcal{N}$, and a single supplier. The power allocation of user $i \in \mathcal{N}$ is denoted by $q_i \in \mathbb{R}_+$ and the total power capacity is denoted by $Q \in \mathbb{R}_+$. The value of the power q_i of each user $i \in \mathcal{N}$ is decided by a *private* utility function $U_i(q_i)$, which the user would like to maximize. By private, we mean the function $U_i(q_i)$ of each user is unknown to the power supplier. The *Resource*

Allocation problem for the power distribution system is given by the following optimization program [3] [11]:

$$
\boxed{
\begin{array}{ll}
\text{Resource Allocation (RA):} \\
\hline
\underset{q_1, \cdots, q_N}{\text{maximize}} & \sum_{i=1}^{N} U_i(q_i), \\
\text{subject to} & \sum_{i=1}^{N} q_i \leq Q, \\
& m_i \leq q_i \leq M_i.
\end{array}
} \tag{RA}
$$

We make the following assumptions on (RA):

Assumption 1. *(Convexity) The function $U_i(q_i)$ is μ-concave and increasing for all $i \in \mathcal{N}$.*

Assumption 2. *(Well Posed) We have that*

$$\sum_{i=1}^{N} m_i \leq Q \leq \sum_{i=1}^{N} M_i. \tag{3}$$

In other words, Problem (RA) *is feasible and the constraint* $\sum_{i=1}^{N} q_i \leq Q$ *is not redundant.*

As motivated in the introduction, in this work we focus on one-way communication protocols for solving (RA). For each user $i \in \mathcal{N}$ the problem data U_i, q_i, m_i, and M_i are private to user i. In particular, the supplier can not access this information. We consider protocols based on the following two operations.

Operation 1 (One-Way Communication)**.** *At each time-step (iteration), the supplier can broadcast to all users one scalar message referred to as the price (of power).*

Operation 2 (Feedback Information)**.** *After broadcasting the price, the supplier can measure the deviation between total power load and supply, i.e. $\sum_{i=1}^{N} q_i(t) - Q$, where t indicates the time-step.*

Algorithms for solving (RA) that only use Operations 1 and 2 can be achieved using duality theory [1][3]. In particular, we achieve the solution by solving the dual problem of (RA) that is given as follows:

$$
\boxed{
\begin{array}{ll}
\text{The Dual of (RA):} \\
\hline
\underset{p}{\text{minimize}} & D(p), \\
\text{subject to} & p \geq 0,
\end{array}
} \tag{Dual-RA}
$$

where D and p are the dual function and dual variables [12, Chapter 5], respectively, and D is given by

$$
\begin{aligned}
D(p) &= \underset{\mathbf{q} \in \mathcal{M}}{\text{maximize}} \sum_{i=1}^{N} U_i(q_i) - p\left(\sum_{i=1}^{N} q_i - Q\right) \\
&= \sum_{i=1}^{N} U_i(q_i(p)) - p\left(\sum_{i=1}^{N} q_i(p) - Q\right)
\end{aligned} \tag{4}
$$

where m_i and M_i are respectively the lower and upper bounds on the power loads of user i; $\mathcal{M} = \prod_{i=1}^{N}[m_i, M_i]$, and $\mathbf{q} = (q_1, \ldots, q_N)^T$. The local problem for user i is to solve

$$q_i(p) = \underset{q_i \in [m_i, M_i]}{\operatorname{argmax}} U_i(q_i) - p \; q_i = \left[(U_i')^{-1}(p)\right]_m^M. \quad (5)$$

In words, $q_i(p)$ in (5) denotes the power demand of user i when the price of power is p. We have the following relationship between (RA) and (Dual-RA):

Lemma 1. *(Strong Duality) Suppose Assumptions 1 and 2 hold; if $p^\star$ is an optimal solution of* (Dual-RA)*, then $\mathbf{q}(p^\star) = (q_i(p^\star))_{i \in \mathcal{N}}$ (cf. (5)) is the optimal solution to* (RA)*.*

Proof. The proof follows by simply noting that $\sum_{i=1}^{N} m_i \leq Q$, and $m_i \leq M_i$ ensures that (RA) satisfies Slater's condition, which is sufficient for the zero duality gap for convex problems [12, Section 5.2.3]. $\square$

Let us now demonstrate benefits of studying the dual problem (Dual-RA), which takes advantage of the following property:

Proposition 1. *Suppose Assumption 1 holds, then $D(\cdot)$ is differentiable and N/μ-smooth.*

Proof. See Lemma II.2 in [10] $\square$

Due to Proposition 1, we can solve (Dual-RA) using a dual descent method, given by the iterates:

$$p(t+1) = \lceil p(t) - \gamma D'(p(t)) \rceil^+, \quad (6)$$

where $\gamma > 0$ is step-size and $D'(\cdot)$ is the gradient of $D(\cdot)$, which is given by

$$D'(p) = Q - \sum_{i=1}^{N} q_i(p). \quad (7)$$

Notice that (7) is exactly the feedback provide in Operation (2) above. In Algorithm 1 below we demonstrate how the the dual descent algorithm can be implemented between the supplier and the users using only Operations 1 and 2.

Algorithm 1: Dual descent, using only Operations 1 and 2 and local computation at the users and the supplier.

Supplier: *Decides the initial prize $p(0) \geq 0$;*
for $t = 0, 1 \cdots$ **do**
    Supplier: *Broadcast $p(t)$ to all users;*
    for $i = 1, \cdots, N$ **do**
        User i**:** *Receives $p(t)$;*
        User i**:** $q_i(t) = \left[(U_i')^{-1}(p(t))\right]_m^M$;
    Supplier: *Measures $D'(p(t)) = Q - \sum_{i=1}^{N} q_i(t)$;*
    Supplier: *Updates price:$p(t+1) = \lceil p(t) + \gamma D'(p(t)) \rceil^+$;*

Remark 1. *Due to the one-way communication, the primal iterates $q_i(t)$ are not communicated to the supplier. Instead the users take the action $q_i(t)$ and the aggregate of their actions is measured by the supplier. Therefore, it is essential*

that the primal problem (RA) *is feasible during every iteration of Algorithm 1, i.e. $\sum_{i=1}^{N} q_i(t) \leq Q$ for all $t \in \mathbb{N}$. Otherwise, a heightened demand of power in the network can result in a system overload and eventual blackout.*

In what follows, we study the convergence of Algorithm 1 and how to ensure that the primal problem is feasible during every iteration.

III. CONVERGENCE ANALYSIS OF ALGORITHM 1

In this section, we provide rules on choosing the initial price and the step-size to ensure that the primal variables are feasible throughout the algorithm (Section III-A). We also further prove that the algorithm has a convergence rate c^t, where $0 < c < 1$ and t is the time-step of the algorithm, under certain conditions (Section III-B).

A. General Convergence Result

The following result on the convergence of Algorithm 1 is standard in the literature [3].

Proposition 2. *If $\gamma \in]0, 2\mu/N[$ in Algorithm 1, then every limit point of $p(t)$ is a solution to* (Dual-RA) *and $\lim_{t \to \infty} \mathbf{q}(t) = \mathbf{q}^\star$ where $\mathbf{q}^\star$ is the solution to* (RA)*.*

Proof. Follows directly from Theorem 1 in [3]. $\square$

Proposition 2 ensures that the primal sequence $\mathbf{q}(t)$ converges to $\mathbf{q}^\star$. However, there is no guarantee that that the iterates $\mathbf{q}(t)$ are feasible to (RA) at each iteration except in the limit. Such infeasibilities can cause blackouts in the power network (cf. Remark 1) and therefore it is essential to find conditions that ensure the feasibility of $\mathbf{q}(t)$ for all $t \in \mathbb{N}$. Such conditions are now established.

Proposition 3. *Let $\mathcal{P}^\star = [\underline{p}^\star, \overline{p}^\star]$ be the set of optimal solutions to* (Dual-RA)*: If Assumptions 1 and 2 hold and $\gamma \in]0, \mu/N]$ in Algorithm 1, then the following hold.*

a) *If $p(0) \geq \overline{p}^\star$, then $\lim_{t \to \infty} p(t) = \overline{p}^\star$ and the primal variables $\mathbf{q}(t)$ are feasible to* (RA) *at every iteration $t \in \mathcal{N}$.*

b) *The convergence rate of the objective function values $D(p(t))$ is $\mathcal{O}(1/t)$. Moreover, the optimal convergence rate*

$$D(p(t)) - D(p^\star) \leq \frac{2||p(0) - p^\star||^2}{t + 4}, \quad (8)$$

is achieved when $\gamma = \mu/N$.

Proof. a) From Proposition 2 we know that every limit point of $p(t)$ is in $\mathcal{P}^\star$. Moreover, since $\mathbf{q}(p)$ is decreasing function (cf. Equation (5)) and $\mathbf{q}(\overline{p}^\star)$ is feasible point of (RA), then $\mathbf{q}(p)$ is also a feasible point of (RA) for all $p \geq \overline{p}^\star$. Therefore, if we can show that $p(t) \geq \overline{p}^\star$ for all $t \in \mathbb{N}$, then it holds that $\lim_{t \to \infty} p(t) = \overline{p}^\star$ and $\mathbf{q}(p(t))$ is feasible to (RA) for all $t \in \mathbb{N}$.

Let us now show that $\overline{p}^\star < p(t)$ for all $t \in \mathbb{N}$ by induction. By assumption $\overline{p}^\star < p(0)$. Let us now suppose $\overline{p}^\star < p(t)$. Then $D'(\overline{p}^\star) = 0$ implies that $D'(p(t)) > 0$, and the convexity of $D(\cdot)$ implies that $D'(\cdot)$ is increasing. Hence, we have $p(t+$

$1) = p(t) - \gamma D'(p(t)) < p(t)$. Moreover, using that $D(\cdot)$ is N/μ-smooth and $D'(\cdot)$ is increasing we get

$$D'(p(t)) - D'(p(t+1)) < \frac{N}{\mu}(p(t) - p(t+1)) = \frac{N}{\mu}\gamma D(p(t)),$$

and rearranging one obtain s

$$0 < \left(1 - \frac{N}{\mu}\gamma\right) D'(p(t)) \le D'(p(t+1)). \tag{9}$$

Hence, the step-size $\gamma \in]0, \mu/N[$ implies that $D'(p(t+1)) > 0$. Hence, since $D'(\cdot)$ is increasing and $D'(p^\star) = 0$ we must have $p^\star \le p(t+1)$.

b) This result is proved in [13, Theorem 2.1.14]. $\qquad \square$

Remark 2. *For general objective functions with Lipschitz continuous gradients, the optimal convergence rate is $\mathcal{O}(1/t^2)$ by using Nesterov fast gradient methods, see [13, chapter 2]. However, these optimal dual gradient methods cannot ensure the feasibility of the primal problems during the converging process, bringing blackout risk.*

B. Linear Convergence Rate

We now identify structures on (RA) which ensure convergence rate c^t, where $0 < c < 1$, on Algorithm 1. We start by providing a linear rate under the following assumptions on utility function and local constraints.

Assumption 3. *The function $U_i(q_i)$ is L-Lipschitz continuous.*

Assumption 4. *Let $\underline{p}_i = U_i'(M_i)$ and $\overline{p}_i = U_i'(m_i)$. Then the set $\mathcal{P} = \bigcup_{i \in \mathcal{N}} [\underline{p}_i, \overline{p}_i]$ is connected, i.e., an interval on $\mathbb{R}_+$. We write $\mathcal{P}$ in terms of its end points as $\mathcal{P} = [\underline{p}, \overline{p}]$.*

The linear convergence rate can now formally be stated as follows.

Proposition 4. *Suppose Assumptions 1 to 4 hold; then for any $p(0) \in [\underline{p}, \overline{p}]$ and $\gamma \in]0, \mu/N[$, Algorithm 1 converges at rate c^t to the set of optimal solutions $\mathcal{P}^\star$; and*

$$\mathrm{dist}(p(t), \mathcal{P}^\star) \le c^t \mathrm{dist}(p(0), \mathcal{P}^\star), \tag{10}$$

where $c = 1 - \gamma/L$. The optimal convergence rate $c = 1 - \mu/(NL)$ is obtained when $\gamma = \mu/N$.

Proof. Let us start by showing that $D(\cdot)$ is $1/L$-convex on $[\underline{p}, \overline{p}]$. From Lemma 2, we can write $D(p) = \sum_{i=1}^N D_i(p)$ where $D_i(\cdot)$ are convex on $\mathbb{R}_+$ for all $i \in \mathcal{N}$ and $D_i(\cdot)$ is $1/L$-convex on $[\underline{p}_i, \overline{p}_i]$. Therefore, from Assumption 4 and the fact that sums of a convex and $1/L$-convex function are $1/L$-convex [13, Lemma 2.1.4] we get that $D(p)$ is $1/L$ convex on $[\underline{p}, \overline{p}]$.

Since $D(\cdot)$ is strongly convex on $\mathcal{P}$, $\mathcal{P}^\star \cap [\underline{p}, \overline{p}]$ can have at most one element. In fact, $\mathcal{P}^\star \cap [\underline{p}, \overline{p}]$ is non-empty, since from (7) and Assumption 2 we have

$$D'(\underline{p}) = Q - \sum_{i=1}^N M_i \le 0 \quad \text{and} \tag{11}$$

$$D'(\overline{p}) = Q - \sum_{i=1}^N m_i \ge 0, \tag{12}$$

and therefore by the continuity of $D'(\cdot)$ and the intermediate value theorem there exists $p^\star \in [\underline{p}, \overline{p}]$ such that $D'(p^\star) = 0$. We now show the c^t convergence rate of Algorithm 1 to $p^\star$. Without loss of generality suppose that $p(0) \ge p^\star$. Then from Proposition 3 and the step-size choice of $\gamma \in]0, \mu/N]$ we know $p(t) \ge p^\star$ for all $t \in \mathbb{N}$. Hence, suppose $p(0) \ge p^\star$ then we get that

$$p(t+1) - p^\star = p(t) - p^\star - \gamma D'(p(t)),$$

$$\le p(t) - p^\star - \frac{\gamma}{L}(p(t) - p^\star) \tag{13}$$

$$= \left(1 - \frac{\gamma}{L}\right)(p(t) - p^\star) \tag{14}$$

where the inequality follows from the fact that $p(t+1), p(t), p^\star \in [\underline{p}, \overline{p}]$ and that D is $1/L$-strongly convex on $[\underline{p}, \overline{p}]$, which implies that [13, Theorem 2.1.9.]

$$\frac{1}{L}(p - p^\star) \le D'(p), \text{ for all } p \in [\underline{p}, \overline{p}]. \tag{15}$$

Applying Inequality (14) t times gives (10), which concludes the proof. $\qquad \square$

Note that the optimal convergence rate $c = 1 - \mu/(NL)$ in (10) depends on the number of users N. However, the following mild assumption yields a convergence rate in the dual variable that is independent of the number of users, which illustrate excellent scaling properties Algorithm 1.

Assumption 5. *Let $\underline{p} = U_i'(M_i)$ and $\overline{p} = U_i'(m_i)$ for all $i \in \mathcal{N}$.*

Assumption 5 essentially gives each user the freedom to choose their utility functions U_i as well as upper and lower bounds m_i and M_i with the constraint that $U_i'(M_i) = \underline{p}$ and $U_i'(m_i) = \overline{p}$. These utility functions can be customized to reflect the preferences or priorities of the user, for example the times of day that each user will require the most power.

Proposition 5. *Suppose Assumptions 1–3 and 5 hold. Then for any $p(0) \in [\underline{p}, \overline{p}]$ and $\gamma \in]0, \mu/N]$, Algorithm 1 has a c^t convergence rate which does not depend on the number of users. In particular, we have*

$$\mathrm{dist}(p(t), \mathcal{P}^\star) \le c^t \mathrm{dist}(p(0), \mathcal{P}^\star), \tag{16}$$

where $c = 1 - N\gamma/L$. The optimal convergence rate $c = 1 - \mu/L$ is obtained when $\gamma = \mu/N$.

Proof. The proof follows almost the same steps as the proof of Proposition 4. The only difference is that here the dual function D is N/L-convex on $[\underline{p}, \overline{p}]$ instead of $1/L$-convex as in Proposition 4. To see why D is N/L-convex on $[\underline{p}, \overline{p}]$, we have from Lemma 2 (in the Appendix) that $D(p) = \sum_{i=1}^N D_i(p)$ where $D_i(p)$ is $1/L$-convex on $[\underline{p}, \overline{p}]$ for each $i \in \mathcal{N}$. $\qquad \square$

In the next section we will consider a specific form of utility function that is commonly used in the resource allocation literature [14]. In particular, we assume a utility function of the form

$$U_i(q_i) = a_i \log(b_i + q_i), \tag{17}$$

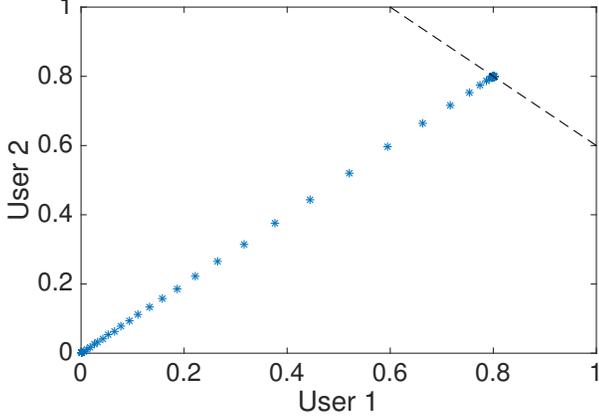**Demand estimates over time for Two-User Case**

Fig. 1: Using a sufficiently small step size, the feasibility of the primal problem is maintained. The upper boundary of this feasible set is denoted by the dotted line.

where the parameters a_i, b_i may be unique to the different users $i \in \mathcal{N}$. Applying Proposition 4 to this logarithmic utility function, we can assert specific conditions under which the set $\mathcal{P} = \bigcup_{i \in \mathcal{N}} [\underline{p}_i, \overline{p}_i]$ is connected.

Proposition 6. *Let a_i in Equation (17) be ordered as $a_1 \leq \cdots \leq a_N$. Let $M_i = M$ and $m_i = 0$ for all i. Then if $(b_i + M)a_i \geq a_{i+1}$ for $i = 1, \cdots, N-1$ and $p(0) \in [a_1/(b_i + M), a_N]$ then if we choose the step-size $\gamma = \mu/N$ we have*

$$\text{dist}(\mathbf{p}(t), \mathcal{P}^\star) \leq \left(1 - \frac{\mu}{NL}\right)^t \text{dist}(\mathbf{p}(0), \mathcal{P}^\star) \quad (18)$$

Proof. The result is a corollary of Proposition 4 and can be obtained by showing that Assumption 4 holds.

To see that Assumption 4 holds, note that $\underline{p}_i = U_i'(M) = a_i/(b_i + M)$ and $\overline{p}_i = U_i'(m) = a_i$. Therefore, $(b_i + M)a_i \geq a_{i+1}$ for $i = 1, \cdots, N-1$ implies that $\mathcal{P} = \bigcup_{i \in \mathcal{N}} [\underline{p}_i, \overline{p}_i]$ is connected. $\square$

IV. NUMERICAL RESULTS

In this section we illustrate the theoretical findings using two numerical examples. In both examples, we set $m_i = 0$ and M_i and use the same utility function of the form in Equation (17) with $a_i = 20$, $b_i = 1$.

Figure 1 depicts the first example which is a system consisting of two users. The power suppliers capacity is $Q = 1.6$ and the step size is $\gamma = \mu/N$. As observed, both users are assigned a socially optimal rate of $q_i = 0.8$ power units each and primal feasibility is satisfied at each iteration of the algorithm, i.e., $\sum_{i=1}^N q_i(t) \leq Q$ for all t.

Figure 2 illustrate a scenario with multiple users $N = 5, 10, 20, 30, 40, 150, 1000$. The step-size is $\gamma = \mu/N$, the initial price is $p(0) = 30$ in all cases, and the suppliers capacity is $Q = 4N/5$. The assumptions of Proposition 5 hold in this case. Figure 2a depicts the aggregate rates $\sum_{i=1}^N q_i$ allocated to all users scaled by the power capacity Q. As observed, the primal variables remain feasible at each iteration

of the algorithm, as proved in Proposition 3. Figures 2b and 2c respectively illustrate the convergence of the algorithm to the optimal primal and dual variables. As observed, the dual variables converge at a linear rate, which is to be expected due to Proposition 5. Moreover, the convergence is not sensitive to the number of users N which also agrees with the results of Proposition 5. Interestingly, the primal variables also converge at a linear rate which suggest that similar convergence results can be obtained for the primal variables.

As can be observed, the results in figures 2a and 2c are identical for $N = 5, 10, 20, 30, 40, 150, 1000$. This can be explained by that the utilities of of the users are identical and the power supply increases with number of users. However, Proposition 5 also holds under much more general assumptions. In future work we will consider more complicated and realistic problems for which Proposition 5 holds.

V. SUMMARY AND CONCLUSION

We considered the problem of allocating power to users in a network by one supplier and presented a distributed algorithm that uses a one-way communication between the supplier and users to coordinate the optimal solution. We showed how to choose step sizes and derive appealing convergence properties in the dual domain. Furthermore, we identified mild modifications to the power allocation problem that *1)* results in a c^t convergence rate of our algorithm, and *2)* yields a convergence rate independent of the number of users in the network. All these indicate excellent scaling properties of the algorithm. The results presented here propose a paradigm for solving distributed resource allocation problems that achieve fast convergence rates, even with a one-way information passing mechanism. Based on this work, of interest to us, is to derive similar convergence rates in the primal domain. In addition, within the context of limited communication, we would like to study the tradeoffs between the need for primal feasibility, convergence rates and actual bits transmitted (bandwidth) using the one-way coordination protocol.

APPENDIX

The following Lemmas are used in our derivations.

Lemma 2. *Suppose Assumption 1 holds and set $\underline{p}_i = U_i'(M_i)$ and $\overline{p}_i = U_i'(m_i)$ for $i \in \mathcal{N}$. Then the dual function $D(\cdot)$ can be decomposed as $D(p) = \sum_{i=1}^N D_i(p)$, where $D_i(\cdot)$ is $1/L$-convex on $[\underline{p}_i, \overline{p}_i]$ and the derivative $D'(\cdot)$ is constant on $[0, \underline{p}_i]$ and on $[\overline{p}_i, \infty]$. In particular, $D_i'(p) = Q/N - M_i$ if $p \in [0, \underline{p}_i]$ and $D_i'(p) = Q/N - m_i$ if $p \in [\overline{p}_i, \infty]$.*

Proof. a) We set

$$D_i(p) = U_i(q_i(p)) - pq_i(p) - p\frac{Q}{N}, \quad (19)$$

where $q_i(p)$ is defined as in (5). By summing (19) over $i \in \mathcal{N}$ and recalling the definition of the dual function D in (4), we get

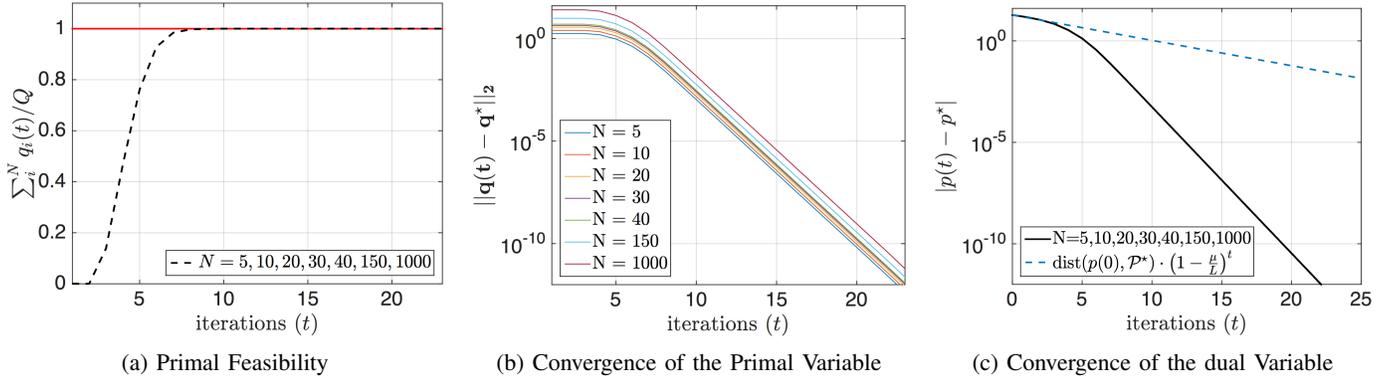$$D(p) = \sum_{i=1}^N D_i(p). \quad (20)$$

Fig. 2: Convergence of the algorithm for different number of users $N = 5, 10, 20, 30, 40, 150, 1000$. In Figures 2a and 2c the behaviour of the algorithm is identical for all N. The blue dotted line in 2c is the theoretical bound from Proposition 5.

Let us now show that $D_i(\cdot)$ is $\frac{1}{L}$-convex on $[\underline{p}_i, \overline{p}_i]$. From [13, Theorem 2.1.9.] it is $\frac{1}{L}$-convex on $[\underline{p}_i, \overline{p}_i]$ if and only if for all $y_1, y_2 \in [\underline{p}_i(t), \overline{p}_i(t)]$ it holds that

$$\frac{1}{L}|y_2 - y_1|^2 \le \langle y_2 - y_1, D_i'(y_2) - D_i'(y_1)\rangle$$
$$= \langle y_2 - y_1, \left(-(U_i')^{-1}(y_2)\right) - \left(-(U_i')^{-1}(y_1)\right)\rangle, \quad (21)$$

where the equality comes by using that (cf. (7) and (5))

$$D_i'(p) = Q/N - q_i(p) = Q/N - [(U_i')^{-1}(p)]_{m_i}^{M_i}. \quad (22)$$

Due to the strong concavity of $U_i(\cdot)$ on $[m_i, M_i]$ (cf. Assumption 1), U_i' is bijective from $[m_i, M_i]$ to $[\underline{p}_i, \overline{p}_i]$ and we can choose $x_1, x_2 \in [m_i, M_i]$ such that $U_i'(x_1) = y_1$ and $U_i'(x_2) = y_2$. Then since U_i is concave and L-smooth we have from (23) in Lemma 3 that

$$- \langle U_i'(x_2) - U_i'(x_1), x_2 - x_1\rangle \ge \frac{1}{L}|U_i'(x_2) - U_i'(x_1)|^2,$$

or by using $U'(x_1) = y_1$ and $U'(x_1) = y_1$, we get

$$\langle y_2 - y_1, \left(-(U_i')^{-1}(y_2)\right) - \left(-(U_i')^{-1}(y_1)\right)\rangle \ge \frac{1}{L}|y_2 - y_1|^2.$$

Lemma 3. *A function f is concave and L-smooth on $\mathcal{X}$ if for all $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{X} \subseteq \mathbb{R}^n$,*

$$\langle f'(\mathbf{x}_1) - f'(\mathbf{x}_2), \mathbf{x}_2 - \mathbf{x}_1\rangle \ge \frac{1}{L}||f'(\mathbf{x}_2) - f'(\mathbf{x}_1)||^2, \quad (23)$$

Proof. The proof follows directly from Theorem 2.1.5. in [13] and using that $-f$ is convex. $\square$

Hence, (21) holds and we can conclude that $D_i(\cdot)$ is $1/L$-convex on $[\underline{p}_i, \overline{p}_i]$. Let us now show that the gradient $D_i'(\cdot)$ is constant on $[0, \underline{p}_i]$ and on $[\overline{p}_i, \infty]$. The result follows by combining (22) with U_i' is decreasing and that $U'(M_i) = \underline{p}_i$ and $U'(m_i) = \underline{p}_i$. $\square$

References

[1] F. P. Kelly, A. K. Maulloo, and D. K. Tan, "Rate control for communication networks: shadow prices, proportional fairness and stability," *Journal of the Operational Research society*, pp. 237–252, 1998.

[2] A. Nedić and A. Ozdaglar, "Subgradient methods in network resource allocation: Rate analysis," in *Information Sciences and Systems, 2008. CISS 2008. 42nd Annual Conference on*. IEEE, 2008, pp. 1189–1194.

[3] S. Low, "Optimization flow control with on-line measurement or multiple paths," in *Proceedings of the 16th International Teletraffic Congress*. Citeseer, 1999, pp. 237–249.

[4] S. Galli, A. Scaglione, and Z. Wang, "For the grid and through the grid: The role of power line communications in the smart grid," *Proceedings of the IEEE*, vol. 99, no. 6, pp. 998–1027, 2011.

[5] E. Biglieri, "Coding and modulation for a horrible channel," *Communications Magazine, IEEE*, vol. 41, no. 5, pp. 92–98, 2003.

[6] R. Clarke, "Internet privacy concerns confirm the case for intervention," *Communications of the ACM*, vol. 42, no. 2, pp. 60–67, 1999.

[7] A. P. Snow, "Network reliability: the concurrent challenges of innovation, competition, and complexity," *Reliability, IEEE Transactions on*, vol. 50, no. 1, pp. 38–40, 2001.

[8] G. N. Ericsson, "Cyber security and power system communicationessential parts of a smart grid infrastructure," *Power Delivery, IEEE Transactions on*, vol. 25, no. 3, pp. 1501–1507, 2010.

[9] V. C. Gungor, D. Sahin, T. Kocak, S. Ergut, C. Buccella, C. Cecati, and G. P. Hancke, "A survey on smart grid potential applications and communication requirements," *Industrial Informatics, IEEE Transactions on*, vol. 9, no. 1, pp. 28–42, 2013.

[10] A. Beck, A. Nedic, A. Ozdaglar, and M. Teboulle, "An gradient method for network resource allocation problems," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 1, pp. 64–73, 2014.

[11] X. Liu, K. Ren, Y. Yuan, Z. Li, and Q. Wang, "Optimal budget deployment strategy against power grid interdiction," in *INFOCOM, 2013 Proceedings IEEE*, April 2013, pp. 1160–1168.

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

[13] Y. Nesterov, *Introductory Lectures on Convex Optimization*. Springer, 2004.

[14] R. Yim, O.-S. Shin, and V. Tarokh, "Reverse-link rate control algorithms with fairness guarantees for cdma systems," *Wireless Communications, IEEE Transactions on*, vol. 6, no. 4, pp. 1386–1397, 2007.