Distributed Lagrangian Methods for Network Resource Allocation

Thinh T. Doan\textsuperscript{1} and Carolyn L. Beck\textsuperscript{2}

Abstract—Motivated by broad applications in various fields of engineering, we study network resource allocation problems where the goal is to optimally allocate a fixed portion of resources over a network of nodes. In these problems, due to the large scale of the network and complicated interconnections between nodes, any solution must be implemented in parallel and based only on local data resulting in a need for distributed algorithms. In this paper, we propose a distributed Lagrangian method, which requires only local computation and communication. Our focus is to understand the performance of this algorithm on the underlying network topology. Specifically, we obtain an upper bound on the rate of convergence of the algorithm as a function of the size and the topology of the underlying network. The effectiveness and applicability of the proposed method is demonstrated by its use in solving the important economic dispatch problem in power systems, specifically on the benchmark IEEE-14 and IEEE-118 bus test systems.

I. INTRODUCTION

We study network resource allocation problems, where the goal is to optimally allocate a fixed portion of resources over a network of nodes. Each node suffers a cost as a function of the portion of resources allocated to it. The goal of this problem is to seek an optimal allocation such that the total cost incurred over the network is minimized, while satisfying each node’s local constraints. This problem is sometimes described in terms of utility functions, with the goal being to maximize the total utility. In this problem, due to the large scale and complex interconnection structure of the network, central coordination is undesirable or even impossible, resulting in the necessity for distributed algorithms. Our focus, therefore, is to study distributed algorithms which can be implemented in parallel and require only local data.

Resource allocation is a fundamental and important problem that arises in a variety of application domains within engineering. One standard example is the problem of congestion control where the global objective is to route and schedule information in an internet network such that a fair resource allocation between users is achieved [20]. Another example is coverage control in wireless sensor networks, where the goal is to optimally allocate a large number of sensors in an unknown environment such that the coverage area is maximized [7], [19]. A simplified version of the important economic dispatch problem in power systems can also be viewed as a resource allocation problem, wherein geographically distributed generators of electricity must coordinate to meet fixed demand while maintaining the stability of the systems [12], [25].

Given the large application domains of network resource allocation problems, distributed algorithms for such problems have received a surge of interest in recent years. The first algorithm which could be implemented in a distributed way was the “center-free” protocol introduced in [13] for a relaxation of this problem, in which the authors consider the case where local constraints are excluded. The term “center-free” was originally meant to refer to the absence of any central coordination. In this relaxed problem, the necessary and sufficient optimality conditions are the consensus of the nodes’ incremental costs and the feasibility constraint on the total amount of resource available [21]. The work in [13] has been followed by work in [9], [16], [17], [21], with the main focus on analyzing the performance of the algorithm and its variants for the relaxed problem.

Although the results of the relaxed problem are well-developed, they are impractical since the local constraints of the nodes are critical in real applications. For example, in economic dispatch problems these local constraints, which represent the limited capacity of generators, are inevitable. Recently, there has been a number of studies on resource allocation problems which utilize the necessary and sufficient conditions of the relaxed problem. In particular, the authors in [5], [10], [15], [23], [24] focus on studying economic dispatch problems where the objective functions are quadratic. Their methods are mainly based on utilizing distributed consensus algorithms to achieve an agreement on the incremental costs between generators while using a proper projection for the generators’ local constraints. The authors in [6], [16] relax the assumption on quadratic cost functions to allow for any convex cost functions with Lipschitz continuous gradients. These methods are based on considering a relaxed problem that uses appropriate penalty functions for the nodes’ local constraints. While the authors in [16] only show a convergence to a near-optimal solution, the authors in [6] show that the solutions of the relaxed problem coincide with resource allocation problems with an appropriate choice of the penalty level.

In this paper, we propose a fully distributed method for resource allocation problems with general convex objective functions. We focus on analyzing the performance of the method to understand how fast the optimal value may be obtained. In particular, we provide an upper bound for the convergence rate of the method as a function of the size and the topology of underlying networks. This fundamental question, which is necessary to analyze distributed systems,
has not been considered explicitly in the literature. Moreover, unlike the work previously cited where the authors consider either a relaxation of the problem or specific quadratic objective functions, our minimal assumption on the convexity of the objective functions is general enough to cover a larger class of problems.

**Main Contributions** The main contributions of the paper are the following. We first propose a novel distributed Lagrangian method for resource allocation problems on undirected and connected networks. Second, we demonstrate that under fairly standard assumptions on the convexity of objective functions and network connectivity, the distributed Lagrangian method achieves a convergence rate \( O(n \ln(k)/(1 - \sigma_2)\sqrt{k}) \) to the optimal value, where \( \sigma_2 \) is a parameter representing the spectral properties of network connectivity of the nodes, \( n \) is the number of nodes, and \( k \) is the number of iterations. Unlike the work in [6] and [16], we assume neither any relaxation nor Lipschitz continuity of the gradient of objective functions in our problem, meaning that our method is general enough to cover a larger class of problems. As one illustration of the effectiveness of the proposed method for control and optimization applications, we apply it to solve the important economic dispatch problem in power systems, and specifically consider the benchmark IEEE-14 and IEEE-118 bus test systems. Simulations show that our method outperforms the method proposed in [15].

The remainder of this paper is organized as follows. We give a formal statement of the problem in Section II. Our proposed method is described in Section III and its convergence analysis is given in Section IV. In Section V, we discuss the results of applying our method to specific economic dispatch problems via simulations. We conclude the paper with some potential future directions in Section VI.

**Notation 1:** We use boldface to distinguish between vectors \( \mathbf{x} \) in \( \mathbb{R}^n \) and scalars \( x \) in \( \mathbb{R} \). Given a vector \( \mathbf{x} \), we write \( \mathbf{x} = (x_1, x_2, \ldots, x_n) \) and denote by \( \|\mathbf{x}\|_2 \) its Euclidean norm. Finally, let \( 1 \in \mathbb{R}^n \) be the vector whose entries are 1.

**II. Problem Formulation**

In this paper we consider resource allocation problems over a network of \( n \) nodes. The goal of this problem is to seek an optimal allocation for a prespecified portion or quantity of a given resource such that the total cost incurred over the network is minimized while satisfying the nodes’ local constraints. In general, this problem can be formulated as the following optimization problem:

\[
P : \begin{cases}
\min_{\mathbf{x} \in \mathbb{R}^n} & \sum_{i=1}^{n} f_i(x_i) \\
\text{s.t.} & \sum_{i=1}^{n} x_i = b, \\
& x_i \in \mathcal{X}_i, \quad \forall i = 1, \ldots, n,
\end{cases}
\]

where \( f_i : \mathbb{R} \to \mathbb{R} \) is a cost function known only by node \( i \) and the constant \( b \) is the total portion of resources shared by the nodes. Moreover, the sets \( \mathcal{X}_i \) are assumed to be convex and compact. We note that the formulation of \( P \) is general enough to cover the models studied in [6], [9], [10], [15], [16].

In the sequel, will use \( \mathcal{X} \) to denote the Cartesian product of \( \mathcal{X}_i \), i.e., \( \mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2 \times \ldots \times \mathcal{X}_n \). The feasible set \( S \) of \( P \) is given as \( S = \{ x \in \mathcal{X} | \sum_{i=1}^{n} x_i = b \} \) and it is assumed to be nonempty. For short, we denote \( f : \mathbb{R}^n \to \mathbb{R} \) as the sum of the functions \( f_i \), i.e.,

\[
f(\mathbf{x}) = \sum_{i=1}^{n} f_i(x_i).
\]

We will make the following assumptions on the objective functions \( f_i \) throughout the paper:

**Assumption 1:** For each \( i = 1, \ldots, n \) the function \( f_i : \mathbb{R} \to \mathbb{R} \) is proper, closed, and convex.

**Assumption 2 (Slater’s condition [4]):** There exists a point \( \mathbf{x} \) that belongs to the relative interior of \( \mathcal{X} \) and satisfies \( \sum_{i=1}^{n} x_i = b \).

We note that since the constraint set in \( P \) is compact and the objective function \( f \) is continuous, there exists a vector \( \mathbf{x}^* = (x_1^*, x_2^*, \ldots, x_n^*) \in \mathcal{S} \) which achieves the minimum of this problem. However, this solution may not be unique. We denote the set of solutions of \( P \) as \( \mathcal{S}^* \). Moreover, under the Slater’s condition strong duality holds and the set of dual optimal solutions is nonempty.

Our focus is on designing a distributed method to solve \( P \) in a network, which means that each node is only allowed to interact with nodes, referred to as their local neighbors, that are connected to it in a graph. Specifically, we assume we are given a graph \( G = (\mathcal{V}, \mathcal{E}) \) with \( \mathcal{V} = \{1, \ldots, n\} \); nodes \( i \) and \( j \) can exchange messages if and only if \( (i, j) \in \mathcal{E} \). We denote by \( \mathcal{N}_i \) the set of neighbors of node \( i \). We make the following fairly standard assumption regarding connectivity of graph \( G \).

**Assumption 3:** The graph \( G \) is undirected and connected.

**III. MAIN ALGORITHM**

In this section, we propose a novel distributed Lagrangian method, namely, a distributed Lagrangian method, for \( P \). Our method can be presented as a distributed version of well-known classical Lagrangian methods [3]. We start this section by giving a brief motivation for our method.

Without loss of generality, we assume that each node \( i \) knows the constant \( b_i \) such that \( \sum_{i=1}^{n} b_i = b \). Here \( b_i \) can be interpreted as the initial resource allocation at node \( i \). One specific choice is to initially distribute \( b \) equally to all nodes, i.e., \( b_i = b/n \) for all \( i = 1, \ldots, n \). We note that the design of our algorithm as well as our analysis given later does not depend on the choice of these \( b_i \) since they are only used for notational convenience.

We now explain the mechanics of our approach. Consider the following Lagrangian function \( \mathcal{L} : \mathcal{X} \times \mathbb{R} \to \mathbb{R} \) of \( P \)

\[
\mathcal{L}(\mathbf{x}, \lambda) := \sum_{i=1}^{n} f_i(x_i) + \lambda \left( \sum_{i=1}^{n} x_i - b \right),
\]

where \( \lambda \in \mathbb{R} \) is the Lagrangian multiplier associated with the coupling equality constraint in (1). The dual function
\(d : \mathbb{R} \to \mathbb{R}\) of problem \(P\) for some \(\lambda\) value is then defined as
\[
d(\lambda) := \min_{x \in \mathbb{R}^n} \left( \sum_{i=1}^{n} f_i(x_i) + \lambda \left( \sum_{i=1}^{n} (x_i - b_i) \right) \right)
\]
where \(f_i^*\) is the Fenchel conjugate of \(f_i\), defined by
\[
f_i^*(u) = \sup_{x \in X_i} \{ux - f_i(x)\}.
\]
The dual problem of \(P\), denoted by \(\text{DP}\), is given by
\[
\text{DP} : \max_{\lambda \in \mathbb{R}} \sum_{i=1}^{n} \{-f_i^*(-\lambda) - \lambda b_i\},
\]
which is then equivalent to solving
\[
\min_{\lambda \in \mathbb{R}} \sum_{i=1}^{n} f_i^*(-\lambda) + \lambda b_i, \quad \text{where } q_i : \mathbb{R} \to \mathbb{R}
\]
is a convex function since \(f_i^*\) is a convex function [3].

A common approach to solve problem (1) is Lagrangian methods [3], which require a central coordinator to update and distribute the multiplier \(\lambda\) to the nodes. The key idea of our approach is to eliminate this requirement by utilizing the distributed subgradient method presented in [18] to compute the solution of (4). In particular, we have that each node \(i\) stores a local copy \(\lambda_i\) of the Lagrange multiplier \(\lambda\), and then iteratively updates \(\lambda_i\) upon communicating with its neighbors. The update of \(\lambda_i\) is comprised of two steps, namely, a consensus step and a so-called “local gradient descent” step. These two steps, coupled with the primal update in the Lagrangian approach, results in our distributed Lagrangian algorithm, presented in Algorithm 1.

The updates in Algorithm 1 have a simple implementation: first, at time \(k \geq 0\), each node \(i\) in step (6) broadcasts the current value of \(\lambda_i(k)\) to its neighbors. Node \(i\) computes \(v_i\), as given by the weighted average of the current local values of its neighbors and its own value. The updates of \(x_i(k+1)\) and \(\lambda_i(k+1)\) then do not require any additional communications among the nodes at step \(k\). In particular, the update of the primal variables \(x_i\) is executed in step (7). Finally, in step (8) \(\lambda_i\) is updated by moving a distance of \(\alpha(k)\) along a subgradient of \(q_i\) at \(v_i(k+1)\), given by
\[
b_i - x_i(k+1) \in \partial q_i(v_i(k+1)),
\]
where \(\partial q_i(v_i(k+1))\) is the sub-differential set of \(q_i\) at \(v_i(k+1)\). We note that our method maintains the feasibility of the nodes’ local constraints at every iteration, i.e., \(x_i(k) \in X_i\) for all \(k \geq 0\). Furthermore, the nodes do not need to store the variables \(v_i\) in step (6) since these are used only for notational convenience.

Let \(A\) be the communication matrix whose the \((i,j)\)-th entries are \(a_{ij}\). Moreover, for short we denote the function \(q : \mathbb{R}^n \to \mathbb{R}\) as
\[
q(\lambda) = \sum_{i=1}^{n} q_i(\lambda_i).
\]
The updates (6)–(8) then can be written compactly in vector form as,
\[
v(k+1) = A\lambda(k)
\]
\[
x_i(k+1) \in \arg \min_{x \in \mathbb{R}^n} \sum_{i \in V} f_i(x_i) + v_i(k+1)(x_i - p_i)
\]
\[
\lambda(k+1) = v(k+1) - \alpha(k)\nabla q(v(k+1)),
\]
where \(\nabla q(\cdot)\), with some abuse of notation, is the subgradient of the dual function \(q\), i.e.,
\[
\nabla q(\lambda) = [\partial q_1(\lambda_1), \ldots, \partial q_n(\lambda_n)]^T.
\]
Throughout the remainder of this paper, we assume that \(A\) satisfies the following assumption.

\textbf{Assumption 4:} \(A\) is a doubly stochastic matrix, i.e., \(\sum_{i=1}^{n} a_{ij} = \sum_{j=1}^{n} a_{ij} = 1\) with \(a_{ii} > 0\). Moreover, the weights \(a_{ij} > 0\) if and only if \((i,j) \in E\) otherwise \(a_{ij} = 0\).

In the sequel, let \(\sigma_2\) be the second largest singular value of \(A\). By the Courant-Fisher theorem [14], since \(A\) is doubly stochastic and the graph is connected, we have \(\sigma_2 \in (0, 1)\).

\textbf{Algorithm 1} A Distributed Lagrangian Method (DLM)

1. \textbf{Initialize:} Each node \(i\) initializes its variables as \(x_i(0) \in \mathbb{R}, \lambda_i(0) \in \mathbb{R}\).
2. \textbf{Iteration:} For \(k \geq 0\) each node \(i\) implements the following steps
\[
v_i(k+1) = \sum_{j \in N_i} a_{ij} \lambda_j(k)
\]
\[
x_i(k+1) \in \arg \min_{x_i \in X_i} f_i(x_i) + v_i(k+1)(x_i - b_i)
\]
\[
\lambda_i(k+1) = v_i(k+1) - \alpha(k)(x_i(k+1) - b_i).
\]

\textbf{IV. CONVERGENCE ANALYSIS}

The goal of this section is to provide an analysis for the convergence of the distributed Lagrangian method proposed in the previous section. Specifically, we will show that the method achieves an asymptotic convergence to the optimal value of \(P\). In addition, to understand how fast the algorithm obtains the optimal objective value we utilize the standard technique from centralized subgradient methods. Specifically, the method achieves an asymptotic convergence to the optimal value at the rate \(O(n \ln(k)/(1-\sigma_2)\sqrt{k})\). Here the parameter \(\sigma_2\) represents the spectral properties of network connectivity. We note that such an explicit upper bound on the convergence rate for distributed resource allocation is not available in the literature. More detail will be given shortly.

We start our analysis by introducing more notation. Given an optimal solution \(x^* \in S^*\) of \(P\), let \(\lambda^*\) be a dual optimal of \(\text{DP}\), i.e., \((x^*, \lambda^*)\) is a saddle point of \(\mathcal{L}(x, \lambda)\)
\[
\mathcal{L}(x^*, \lambda) \leq \mathcal{L}(x, \lambda^*) \leq \mathcal{L}(x, \lambda^*) \quad \forall x \in \mathbb{R}^n, \lambda \in \mathbb{R}.
\]
Moreover, let \( \lambda^* \) be a vector such that \( \lambda^* = (\lambda^*, \ldots, \lambda^*)^T \in \mathbb{R}^n \). Denote by \( \mathcal{L}_i : X_i \times \mathbb{R} \to \mathbb{R} \) the local Lagrangian function at node \( i \)

\[
\mathcal{L}_i(x_i, v_i) = f_i(x_i) + v_i(x_i - b_i),
\]

and with some abuse of notation, the global Lagrangian function \( \mathcal{L} : \mathcal{X} \times \mathbb{R}^n \to \mathbb{R} \) is defined as

\[
\mathcal{L}(x, v) = \sum_{i \in V} \mathcal{L}_i(x_i, v_i).
\]

(9)

Our first result is the Lipschitz continuity of the functions \( q_i \) given in the following proposition.

**Proposition 1**: Let the sequences \( \{x_i(k)\} \) and \( \{\lambda_i(k)\} \) for all \( i \in V \) be generated by the DLM algorithm. Then there exists a positive constant \( C \) such that for all \( k \geq 0 \),

\[
|\partial q_i(v_i(k + 1))| \leq C, \quad \text{for all } i \in V.
\]

(10)

**Proof**: Since \( X_i \) is a compact set and \( x_i(k) \in X_i \) \( \forall k \geq 0 \), by (5) there exists a positive constant \( C_i \) such that

\[
|\partial q_i(v_i(k + 1))| \leq C_i \quad \forall k \in \mathbb{N}.
\]

Letting \( C = \max_i C_i \) gives (10). \( \blacksquare \)

Recall that the updates (6) and (8) are the distributed subgradient methods, studied in [18] for problem (4). Moreover, since the subgradient of \( q_i \) is bounded, we can now apply the result in [18, Proposition 4] to show the convergence of \( \lambda_i \) to \( \lambda^* \), a dual optimal. Due to the limited space, we skip the proof here and refer the readers to [18].

**Lemma 1 ([18]):** Let Assumptions 1–4 hold. Let the sequences \( \{x_i(k)\} \) and \( \{\lambda_i(k)\} \) for all \( i \in V \) be generated by the DLM algorithm. Assume that the step size \( \alpha(k) \) is positive, nonincreasing with \( \alpha(0) = 1 \), and satisfies the following conditions,

\[
\sum_{k=0}^{\infty} \alpha(k) = \infty, \quad \sum_{k=0}^{\infty} \alpha^2(k) < \infty.
\]

(11)

Then, given a dual optimal \( \lambda^* \) of (4), we have

\[
\lim_{k \to \infty} \lambda_i(k) = \lambda^*, \quad \text{for all } i \in V.
\]

(12)

A specific choice for the stepsize sequence is \( \alpha(k) = 1/k \), which obviously satisfies (11). We are now ready to state our first main result, which is an asymptotic convergence of DLM to the optimal value of \( P \). We present here a sketch of its proof for brevity while the full analysis can be found in the longer version of this paper [8].

**Theorem 1**: Let Assumptions 1–4 hold. Let \( \{x_i(k)\} \) and \( \{\lambda_i(k)\} \) for all \( i \in V \) be generated by the DLM algorithm. Then, for all integers \( k \geq 0 \) we have,

\[
\lim_{k \to \infty} f(x(k)) = f(x^*).
\]

(13)

**Proof**: We first have \((x^*, \lambda^*)\) is a saddle point of the Lagrangian (3) where \( \lambda^* \) satisfies (12). By (9) one can show the following relation

\[
0 \leq \sum_{i \in V} \mathcal{L}_i(x_i^*, v_i(k + 1)) - \mathcal{L}_i(x_i(k + 1), v_i(k + 1)) \leq C|v(k + 1) - \lambda^*|_1 + \sum_{i \in V} v_i(k + 1)(x_i^* - p_i).
\]

(14)

By Lemma 1, we have \( \lim_{k \to \infty} \lambda_i(k) = \lambda^* \) implying \( \lim_{k \to \infty} v_i(k + 1) = \lambda^* \) since \( v_i(k + 1) = \sum_{j \in X_i} a_{ij} \lambda_j(k) \) and \( A \) is doubly stochastic. In addition, since \( x_i^* \) satisfies \( \sum_{i \in V} x_i^* = b \) we have that the right hand side of (14) goes to 0 as \( k \to \infty \). Thus, we obtain

\[
0 \leq \limsup_{k \to \infty} \sum_{i \in V} \mathcal{L}_i(x_i^*, v_i(k + 1)) - \mathcal{L}_i(x_i(k + 1), v_i(k + 1)) = \limsup_{k \to \infty} \sum_{i \in V} \mathcal{L}_i(x_i^*, \lambda^*) - \mathcal{L}_i(x_i(k + 1), \lambda^*) = 0,
\]

which implies that

\[
0 \leq f(x^*) - \limsup_{k \to \infty} \mathcal{L}(x(k + 1), \lambda^*) = 0,
\]

(15)

where we use the fact that \( \mathcal{L}(x^*, \lambda^*) = f(x^*) \). Using the standard analysis for the convergence of \( \limsup \) of the sequence \( x_i(k) \) and the definition of the subgradient of convex functions \( q_i \), we can show that

\[
\limsup_{k \to \infty} \mathcal{L}(x(k + 1), \lambda^*) = \lim_{k \to \infty} f(x(k)).
\]

Thus from (15) we obtain (13). \( \blacksquare \)

Finally, we present the convergence of the DLM algorithm to the optimal value at a rate \( O(n \ln(k)/(1 - \sigma_2)\sqrt{K}) \) on the dual function value estimated as the time-weighted average of any dual estimate \( \lambda_i \). The following corollary, studied in [18], is used to show this result.

**Corollary 1 ([18]):** Let Assumptions 1–4 hold. Let \( \{x_i(k)\} \) and \( \{\lambda_i(k)\} \) for all \( i \in V \) be generated by the DLM algorithm. Let \( \alpha(k) = 1/\sqrt{K} \) for \( k \geq 1 \) and \( \alpha(0) = 1 \). Then we have for all \( i = 1, \ldots, n \), and \( K \geq 1 \)

\[
q \left( \frac{\sum_{k=0}^{K} \alpha(k) \lambda_i(k)}{\sum_{k=0}^{K} \alpha(k)} - 1 \right) - q(\lambda^*)
\leq \frac{||\lambda(0) - \lambda^*||^2_2}{2\sqrt{K}} + \frac{4nC||\lambda(0)||_2 + 5nC^2(2+\ln(K))}{2(1-\sigma_2)\sqrt{K}}.
\]

(16)

Corollary 1 reveals that \( q \) evaluated at time-averaged weighted local copies of the dual variables (with each node) converges to the optimal value \( q^* \). Moreover, it shows that the difference of \( q \) at this ‘average’ \( \lambda \) from \( q^* \) scales as \( O(\ln(k)/\sqrt{K}) \). The \( 1/\sqrt{K} \) term mirrors the convergence results for centralized subgradient descent algorithms. In a sense, the distributed nature of the algorithm slows the convergence by a factor of \( \ln(k) \). Further, the convergence rate depends inversely on \( 1 - \sigma_2(A) \), the spectral gap of \( A \). Finding the matrix \( A \) with the maximal spectral gap that has a sparsity pattern of \( G \) amounts to solving a semidefinite program [22]. Such a choice for \( A \) will result in faster convergence.

**V. Power Application Studies**

Our main motivations for the studies in this paper are for use in power systems applications [12], [25]. In this section, we show the effectiveness of the Lagrangian distributed method by applying it to the economic dispatch problem, wherein geographically distributed generators of electricity must coordinate to meet fixed demand. We first apply the
method to the IEEE-14 bus test system [2] and compare the performance of our method with that in [15]. Then to illustrate its performance on a larger scale system, we study economic dispatch on the IEEE-118 bus test system [1]. Simulation results are given to demonstrate that our method works very well in both systems.

A. Case study 1: Economic dispatch on IEEE-14 bus systems

We first study the economic dispatch problem on IEEE-14 bus systems, where generators are located at buses 1, 2, 3, 6, and 8 as shown in Fig. 1. Each generator $i$ suffers a quadratic cost as a function of the amount of its generated power $P_i$, i.e., $f_i(P_i) = \gamma_i P_i^2 + \beta_i P_i$ where $\gamma_i, \beta_i$ are cost coefficients of generator $i$. We also assume that each generator $i$ can only generate a limited amount of power represented by $[0, P_{i\text{max}}]$. The coefficients of each generator are listed in Table I which are adopted from [15]. The total load demand required from the system is $P = 300\, \text{MW}$, which is a typical load value chosen for the purpose of this simulation. The goal now is to meet the load demand while minimizing the total cost incurred at the generators. Here we assume the connection between generators is represented by an undirected circle graph.

Since the local function at each generator is quadratic, the updates in the DLM algorithm can be simplified as

$$x_i(k+1) = P_{[0,P_{i\text{max}}]} \left[ -\sum_{j \in N_i} a_{ij} \lambda_j(k) - \frac{\beta_i}{2\gamma_i} \right] \quad \lambda_i(k+1) = \sum_{j \in N_i} a_{ij} \lambda_j(k) - \alpha(k)(x_i(k+1) - b_i),$$

where $P_{[0,P_{i\text{max}}]}[x]$ is the projection of $x$ on to the interval $[0, P_{i\text{max}}]$.

![Fig. 1: IEEE 14 bus systems.](image1)

![Fig. 2: Economic dispatch problem on IEEE-14 bus systems.](image2)

**TABLE I: Generator parameters (MU= Monetary units)**

| Gen. Bus | $\gamma_i[\text{MU/MW}^2]$ | $\beta_i[\text{MU/MW}]$ | $P_{i\text{max}}[\text{MW}]$ |
|----------|---------------------------|--------------------------|-----------------------------|
| 1        | 1                         | 0.04                     | 2.0                         |
| 2        | 2                         | 0.03                     | 3.0                         |
| 3        | 3                         | 0.035                    | 4.0                         |
| 4        | 6                         | 0.03                     | 4.0                         |
| 5        | 8                         | 0.04                     | 2.5                         |

In this study, we compare the performance of our DLM algorithm with the consensus + innovation method (CIM) studied in [15]. We select the same step size as that used in [15], i.e., $\alpha(k) = 0.08/k^{0.85}$, which also satisfies the conditions in Lemma 1. The simulations of these two methods when solving the economic dispatch problem on the IEEE-14 bus test system are shown in Figure 2. In each figure, our method is shown on the top plot while the CIM algorithm of [15] is shown in the bottom. The plots show that in this study DLM outperforms CIM. In particular, DLM schedules the power to the generators after 20 iterations while CIM requires 40 iterations. Moreover, the Lagrange multipliers $\lambda_i$ in DLM converge after 60 iterations, while those in CIM have not converged even after 100 iterations.

B. Case study 2: Economic dispatch on IEEE-118 bus systems

We now consider the economic dispatch problem on a larger system, the IEEE-118 bus test system [1]. This system has 54 generators connected by bus lines. Each generator $i$ suffers a quadratic cost as a function of the amount of its generated power $P_i$, i.e., $f_i(P_i) = \mu_i + \beta_i P_i + \gamma_i P_i^2$. The coefficients of functions $f_i$ belong to the ranges $\mu_i \in [6.78, 74.33]$, $\beta_i \in [8.3391, 37.6968]$,
Power allocation at generators

iteration

power allocation $P'$

... primary frequency control in power systems. IEEE Transactions on Automatic Control, 59(5):1177–1189, April 2014.

and $\gamma_i \in [0.0024, 0.0697]$. The units of $\mu_i, \beta_i, \gamma_i$ are MMBtu, MMBtu/MW and MMBtu/MW$^2$, respectively. Each $P_i$ is constrained on some interval $[P_{i\text{min}}, P_{i\text{max}}]$ where these values vary as $P_{i\text{min}} \in [5, 150]$ and $P_{i\text{max}} \in [150, 400]$. The unit of power in this system is MW. The total load demand required from the system is $P = 6000$ (MW), which again is a typical load demand chosen for the purpose of this simulation. In this study we select the step size $\alpha(k) = 1/k$ for $k \geq 1$ with $\alpha(0) = 1$. Moreover, the connections between generators are based on the physical connections given by bus data [1]. The simulation results are given in Figure 3 which demonstrate the convergence of our method for this system.

VI. CONCLUDING REMARKS

In this paper, we have proposed a distributed Lagrangian method for network resource allocation problems. Under a minimal assumption on the convexity of the nodes’ functions, we show that our method achieves asymptotic convergence to the optimal value with the rate, $O(n \; \ln(n/k)(1 - \sigma_2)\sqrt{k})$.

Finally, to show the effectiveness of our method, we applied our method to solve the economic dispatch problem on the benchmark IEEE-14 and IEEE-118 bus test systems.

Our main motivation comes from applications in power systems, where our method provides a useful tool for determining near optimal power allocation in a distributed power system. One direction following this work is to couple our method with other control tasks in power systems, for example, frequency control problems [11], [25]. A second interesting open direction is to consider the case of unknown or time-varying resources as is typically found in actual power systems where the demand is changing over time and the relevant data may be uncertain.

ACKNOWLEDGEMENT

The authors would like to thank Alex Olshevsky, Rafal Goebel, and Daniel Liberzon for their insightful comments which helped to improve the quality of this paper.

REFERENCES

[1] IEEE 118 bus system [online]. http://motor.ece.iit.edu/data/JEAS_IEEE118.doc.
[2] IEEE 14 bus system [online]. http://icseg.iti.illinois.edu/ieee-14-bus-system/.
[3] D. Bertsekas. Nonlinear Programming: 2nd Edition. Cambridge, MA: Athena Scientific, 1999.
[4] D. Bertsekas, A. Nedić, and A. Ozdaglar. Convex Analysis and Optimization. Cambridge, MA: Athena Scientific, 2004.
[5] G. Binetti, A. Davoudi, F.L. Lewis, D. Naso, and B. Turchiano. Distributed consensus-based economic dispatch with transmission losses. IEEE Transactions on Power Systems, 29(4):1711–1720, June 2014.
[6] A. Cherukuri and J. Cortes. Distributed generator coordination for initialization and anytime optimization in economic dispatch. IEEE Transactions on Control of Network Systems, 2(3):226 – 237, Sept. 2015.
[7] J. Cortes, S. Martinez, T. Karatas, and F. Bullo. Coverage control for mobile sensing networks. IEEE Transactions on Automatic Control, 52(2):243 – 255, 2004.
[8] T. T. Doan and C. L. Beck. Distributed Lagrangian Methods for Network Resource Allocation. https://arxiv.org/abs/1609.06287.
[9] T. T. Doan and A. Olshevsky. Distributed Resource Allocation on Dynamic Networks in Quadratic Time. Systems & Control Letters, 99:5763, January 2017.
[10] A. Domínguez-García, S.T. Cady, and C.N. Hadjicostis. Decentralized optimal dispatch of distributed energy resources. In 51st IEEE Conference on Decision and Control, pages 3688–3693, 2012.
[11] F. Dörfler and S. Grammatico. Amidst centralized and distributed frequency control in power systems. In IEEE American Control Conference (ACC), Boston, USA, July 2016.
[12] F. Dörfler, J.W. Simpson-Porco, and F. Bullo. Breaking the hierarchy: Distributed control and economic optimality in microgrids. IEEE Transactions on Automatic Control, 3(3):241 – 253, Sept. 2016.
[13] Y. C. Ho, L. D. Servi, and R. Suri. A class of center-free resource allocation algorithms. Large Scale Systems, 1:51–62, 1980.
[14] R.A. Horn and C.R. Johnson. Matrix Analysis. Cambridge, U.K.: Cambridge Univ. Press, 1985.
[15] Soummya Kar and Gabriela Hug. Distributed robust economic dispatch in power systems: A consensus + innovations approach. In 2012 IEEE Power and Energy Society General Meeting, pages 1–8, 2012.
[16] H. Lakshmanan and D.P. de Farias. Decentralized resource allocation in dynamic networks of agents. SIAM Journal on Optimization, 19(2):911–940, 2008.
[17] I. Necsoiu. Random coordinate descent algorithms for multi-agent convex optimization over networks. IEEE Trans. on Automatic Control, 58(8):2001–2012, 2013.
[18] A. Nedić and A. Ozdaglar. Distributed subgradient methods for multi-agent optimization. IEEE Transactions on Automatic Control, 54(1):48–61, 2009.
[19] P. Sharma, S. M. Salapaka, and C. L. Beck. Entropy-based framework for dynamic coverage and clustering problems. IEEE Transactions on Automatic Control, 57(1):135–150, Jan 2012.
[20] Rayadurgam Srikanth. The Mathematics of Internet Congestion Control. Birkhauser, 2004.
[21] L. Xiao and S. Boyd. Optimal scaling of a gradient method for distributed resource allocation. J. of Optimization Theory and Appl., 129(3):469–488, 2006.
[22] L. Xiao and S. Boyd. Fast linear iterations for distributed averaging. Systems & Control Letters, 53:65–78, Sept. 2004.
[23] H. Xing, Y. Mou, M. Fu, and Z. Lin. Distributed bisection method for economic power dispatch in smart grid. IEEE Transactions on Power Systems, 30(6):3024–3035, Aug. 2015.
[24] S. Yang, S. Tan, and J.-X. Xu. Consensus based approach for economic dispatch problem in a smart grid. IEEE Transactions on Power Systems, 28(4):4416–4426, Oct. 2013.
[25] C. Zhao, U. Topcu, N. Li, and S.H. Low. Design and stability of load-side primary frequency control in power systems. IEEE Transactions on Automatic Control, 59(5):1177–1189, April 2014.