Decentralized Primal-Dual Proximal Operator Algorithm for Constrained Nonsmooth Composite Optimization Problems over Networks

Liping Feng [1], Liang Ran [2], Guoyang Meng [3], Jialong Tang [2], Wentao Ding [2] and Huaqing Li [1,2,*]

[1] Department of Computer Science, Xinzhou Teachers University, Xinzhou 034000, China
[2] Chongqing Key Laboratory of Nonlinear Circuits and Intelligent Information Processing,
College of Electronic and Information Engineering, Southwest University, Chongqing 400715, China
[3] Department of Mathematics, Xinzhou Teachers University, Xinzhou 034000, China
[*] Correspondence: huaqingli@swu.edu.cn

Abstract: In this paper, we focus on the nonsmooth composite optimization problems over networks, which consist of a smooth term and a nonsmooth term. Both equality constraints and box constraints for the decision variables are also considered. Based on the multi-agent networks, the objective problems are split into a series of agents on which the problems can be solved in a decentralized manner. By establishing the Lagrange function of the problems, the first-order optimal condition is obtained in the primal-dual domain. Then, we propose a decentralized algorithm with the proximal operators. The proposed algorithm has uncoordinated stepsizes with respect to agents or edges, where no global parameters are involved. By constructing the compact form of the algorithm with operators, we complete the convergence analysis with the fixed-point theory. With the constrained quadratic programming problem, simulations verify the effectiveness of the proposed algorithm.

1. Introduction

Recently, the distributed data processing methods based on multi-agent networks have received much attention. The traditional methods put all the data into one machine and perform the computation centrally. However, as the size of data continues to grow, this kind of centralized strategy is limited by the computing power of the hardware. In contrast to this, the distributed methods distribute computing tasks to agents over decentralized networks [1,2]. Each agent keeps an arithmetic unit and a memory unit. The agents interact with each other through communication links, and this communication occurs only among the neighboring agents. Under these conditions, the distributed methods can effectively solve the optimization problems common to sensor networks [3], economic dispatch [4–6], machine learning [7,8] and dynamic control [9].

The existing decentralized algorithms have included some successful results [10–15]. Previous works considered the problem models composed of a single function. With a fixed stepsize, Shi et al. designed EXTRA [10], which can exactly converge to the optimal solution. Lei et al. studied problems with bound constraints and proposed the primal-dual algorithm [11]. In addition, recent works [16–18] investigated a general distributed optimization with an objective function by designing decentralized subgradient-based algorithms, but diminishing or non-summable step-sizes are utilized, which may cause slow convergence rates [19].

In order to make full use of these special properties, some scholars have studied the nonsmooth composite optimization problems, which possess smooth and nonsmooth structures. By extending EXTRA to the nonsmooth combinational optimization, Shi et al.

proposed PG-EXTRA [20]. Li et al. introduced the network-independent stepsize to PG-EXTRA and then developed NIDS [21]. In addition, Aybat et al. proposed DPDA-D [22] for time-varying networks. Considering the situation that the nonsmooth term cannot be split, Xu et al. proposed the [23]. PG-ADMM [24] was designed based on the distributed alternating direction multiplier method. Particularly, the nonsmooth combinational optimization problems also include a class of problems consisting of three functions and a linear operator. This structure is mainly discussed in the centralized optimization [25–28], and recently some distributed works also appear in [29,30]. In this paper, inspired by the constrained optimization problem [31], we study the constrained nonsmooth composite optimization problems over networks.

The contributions of this paper can be summarized as follows:

1. This paper focuses on an optimization problem with partially smooth and nonsmooth objective functions, where the decision variable satisfies local equality and feasible constraints, unlike these works [10,16,18–21] without considering any constraints. Then, to solve this problem, we propose a novel decentralized algorithm by combining primal-dual frame with the proximal operators, which avoids the estimation of subgradients for nonsmooth terms.
2. Different from existing node-based methods [16–21], the proposed algorithm adopts an edge-based communication pattern that explicitly highlights the process of information exchange among neighboring agents and further gets rid of the dependence on Laplacians [13]. Such a consideration also makes it possible to use uncoordinated stepsizes instead of commonly global or dynamic ones [10,12,16,18,19,21].
3. By employing the first-order optimal conditions and fixed-point theory of operators, the convergence is proved, and its sublinear rate $O(1/k)$ (k is the number of iteration); i.e., at most, $O(1/\epsilon)$ iterations in order to reach an accuracy of ϵ is established.

Organization: The rest of this paper is organized as follows. In Section 2, the necessary notations and basic knowledge are first provided, and then we describe the optimization problem over the networks and necessary assumptions. Section 3 supplies the development of the proposed decentralized algorithm. In Section 4, the convergence analysis for the proposed algorithm is provided. In Section 5, we use the simulation experiments to verify the theoretical analysis. Finally, conclusions are given in Section 6.

2. Preliminaries

In this section, we introduce the notations involved in this paper. Meanwhile, the objective problem and its explanation are also supplied.

2.1. Graph Theory and Notations

The knowledge of graph theory is used to construct the mathematical model of the communication network. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ describe the network as a graph, where $\mathcal{V}$ is the set of vertices and $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ is the set of edges. For an agent $i \in \mathcal{V}$, $\mathcal{N}_i$ denotes the set of its neighbors. Let the unordered pair $(i, j) \in \mathcal{E}$ represent the edge between agent i and agent j. However, (i, j) or (j, i) is still order, i.e., the variables with respect to them are different.

Next, we explain the notations that appear in this paper. Let $\mathbb{R}$ represent the set of real numbers. Therefore, $\mathbb{R}^n$ denotes the n-dimensional vector space, and $\mathbb{R}^{n \times m}$ denotes the set of all n-row and m-column real matrices. We define I_n as the n-dimensional identity operator, $\mathbf{0}_n$ as the n-dimensional null vector, and $\mathbf{0}_{n \times n}$ as the null matrix. If their dimensions are clear from the context, we omit their subscript. Then, blkdiag$\{P, Q\}$ is the block diagonal matrix grouped by matrices P and Q. For a matrix P, let M be its transpose. We denote $\|x\|_P = \sqrt{x^\top P x}$ as the induced norm with matrix P. The subdifferential of function f is ∂f, where $\partial f(x) = \{p | \forall q \in \mathbb{R}^n, f(x) + p^\top(q - x) \le f(q)\}$. The conjugate function f^* is defined by $f^*(p) = \sup_{q \in \mathbb{R}^n}\{\langle p, q \rangle - f(q)\}$. For a positive constant μ, the resolvent of the proximal operator is $\text{prox}_{\mu f}(y) = (I + \mu \partial f)^{-1}(y) = \arg\min_x f(x) + \frac{1}{2\mu}\|x - y\|^2$, while the resolvent with respect to the matrix M is $\text{prox}_f^M(y) = (I + M^{-1}\partial f)^{-1}(y) =$

$\arg\min_x f(x) + \frac{1}{2}\|x - y\|_M^2$. Moreover, let $\mathcal{S}$ represent the optimal solution set of a solvable optimization problem over networks.

2.2. Decentralized Optimization Problem

The constrained composite optimization problem over networks studied in this paper is based on the network $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ with m agents. Specifically, the formulation of the problem is established as follows:

$$\min_{\tilde{x} \in \mathbb{R}^n} \sum_{i=1}^m f_i(\tilde{x}) + g_i(\tilde{x}), \tag{1a}$$

$$\text{s.t.} \quad A_i \tilde{x} = b_i, \ i = 1, \cdots, m, \tag{1b}$$

$$\tilde{x} \in \bigcap_{i=1}^m \Omega_i. \tag{1c}$$

In problem (1), $\tilde{x} \in \mathbb{R}^n$ is the decision variable; $f_i : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ and $g_i : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ are two private cost functions to agent i, where the former has the Lipschitz continuous gradient, but the latter may be nonsmooth; $b_i \in \mathbb{R}^r$ is a vector and $A_i : \mathbb{R}^n \to \mathbb{R}^r$ is a linear operator. Convex set Ω_i gives the box constraints to the decision variable of agent i.

To clarify the properties of problem (1), the following necessary assumption is given.

Assumption 1. *For any agent $i \in \mathcal{V}$:*

(i)   The cost function f_i is Lipschitz continuous and convex; i.e., if we consider the positive Lipschitz constant β_i, then it holds the inequality for the gradient ∇f_i:

$$\|\nabla f_i(\tilde{x}) - \nabla f_i(\tilde{y})\|^2 \le \beta_i (\tilde{x} - \tilde{y})^\top (\nabla f_i(\tilde{x}) - \nabla f(\tilde{y})). \tag{2}$$

(ii)   The local cost function g_i is a nonsmooth and convex function.
(iii) The optimal solution $\tilde{x}^*$ to objective problem (1) exists, which satisfies both the equality constraints and the box constraints.
(iv)   The graph $\mathcal{G}$ is undirected and connected.

Note that the cost functions f_i and g_i are separable. Hence, we introduce the consensus constraint to transform problem (1) into the structure that can be computed in a decentralized manner:

$$\min_{x_1, \cdots, x_m} \sum_{i=1}^m f_i(x_i) + g_i(x_i), \tag{3a}$$

$$\text{s.t.} \quad A_i x_i = b_i, \ i = 1, \cdots, m, \tag{3b}$$

$$x_i = x_j, \ i = 1, \cdots, m, \ j \in \mathcal{N}_i, \tag{3c}$$

$$x_i \in \Omega_i. \tag{3d}$$

Define the set $\mathcal{A}_i = \{z \in \mathbb{R}^n | A_i z = b_i\}$ and consider the indicator function

$$\delta_C(e) = \begin{cases} 0, & e \in C, \\ +\infty, & e \notin C, \end{cases}$$

such that Problem (3) can be processed by the penalty function method. For $i \in \mathcal{V}$ and $j \in \mathcal{N}_i$, let $C_{ij} = I$ if $i < j$ and $C_{ij} = -I$ otherwise. Thus, Problem (3) is equivalent to the following problem:

$$\min_{x_1, \cdots, x_m} \sum_{i=1}^m f_i(x_i) + g_i(x_i) + \delta_{\mathcal{A}_i}(x_i) + \delta_{\Omega_i}(x_i), \tag{4a}$$

$$\text{s.t.} \quad C_{ij} x_i + C_{ji} x_j = 0, \ i = 1, \cdots, m, \ j \in \mathcal{N}_i. \tag{4b}$$

Then, let $x = col(x_1, \ldots, x_m)$ be the global variable. For $i \in \mathcal{V}$ and $j \in \mathcal{N}_i$, we introduce a linear operator $N_{(i,j)} : x \mapsto ((C_{ij}x_i)^\top, (C_{ji}x_j)^\top)^\top$, which generates the edge-based variable from x. With the set $\mathcal{C}_{(i,j)} = \left\{ (z_1^\top, z_2^\top)^\top | z_1 + z_2 = 0 \right\}$, the constraint in the problem (4) can be transformed into another penalty function. Therefore, the problem (1) is finally equivalent to the following problem:

$$\min_{x_1, \cdots, x_m} \sum_{i=1}^m f_i(x_i) + g_i(x_i) + \delta_{\mathcal{A}_i}(x_i) + \delta_{\Omega_i}(x_i) + \sum_{i=1}^m \sum_{(i,j) \in \mathcal{E}} \delta_{\mathcal{C}_{(i,j)}} \left(N_{(i,j)} x \right). \tag{5}$$

Based on the problem (5), we design a novel decentralized algorithm to solve the constrained composite optimization problem over networks in the next section.

3. Algorithm Development

The introduction with respect to the design process of the proposed algorithm is provided in this section.

Notice that Problem (5) is an unconstrained problem. According to [32] (Proposition 19.20), we obtain the following Lagrangian function:

$$\begin{aligned}
\mathcal{L} = &\sum_{i=1}^m \left(f_i(x_i) + g_i(x_i) + v_i^\mathrm{T} x_i - \delta_{\mathcal{A}_i}^*(v_i) + u_i^\mathrm{T} x_i - \delta_{\Omega_i}^*(u_i) \right) \\
&+ \sum_{i=1}^m \sum_{(i,j) \in \mathcal{E}} \left(w_{(i,j)}^\mathrm{T} N_{(i,j)} x - \delta_{\mathcal{C}_{(i,j)}}^* \left(w_{(i,j)} \right) \right),
\end{aligned} \tag{6}$$

where $v_i \in \mathbb{R}^n$, $u_i \in \mathbb{R}^n$ and $w_{(i,j)} \in \mathbb{R}^{2n}$ are dual variables, and $\delta_{\mathcal{A}_i}^*$, $\delta_{\Omega_i}^*$ and $\delta_{\mathcal{C}_{(i,j)}}^*$ are the conjugate functions of $\delta_{\mathcal{A}_i}$, δ_{Ω_i}, $\delta_{\mathcal{C}_{(i,j)}}$, respectively. Notice that $w_{(i,j)} = (w_{ij}^\top, w_{ji}^\top)^\top \in \mathbb{R}^{2n}$ is an edge-based variable, where $w_{ij} \in \mathbb{R}^n$ is the local variable of agent i and $w_{ji} \in \mathbb{R}^n$ is for agent j. Then, the last term of the Lagrangian function (6) satisfies:

$$\sum_{i=1}^m \sum_{(i,j) \in \mathcal{E}} \left(w_{(i,j)}^\mathrm{T} N_{(i,j)} x - \delta_{\mathcal{C}_{(i,j)}}^* \left(w_{(i,j)} \right) \right) = \sum_{i=1}^m \sum_{j \in \mathcal{N}_i} \left(w_{(i,j)}^\mathrm{T} N_{(i,j)} x - \delta_{\mathcal{C}_{(i,j)}}^* \left(w_{(i,j)} \right) \right),$$

Thus, the Lagrangian function (6) can also be written as

$$\begin{aligned}
\mathcal{L} = &\sum_{i=1}^m \left(f_i(x_i) + g_i(x_i) + v_i^\mathrm{T} x_i - \delta_{\mathcal{A}_i}^*(v_i) + u_i^\mathrm{T} x_i - \delta_{\Omega_i}^*(u_i) \right) \\
&+ \sum_{i=1}^m \sum_{j \in \mathcal{N}_i} \left(w_{(i,j)}^\mathrm{T} N_{(i,j)} x - \delta_{\mathcal{C}_{(i,j)}}^* \left(w_{(i,j)} \right) \right),
\end{aligned} \tag{7}$$

Taking the partial derivatives of the Lagrangian function (7) and combining the operator splitting method [29], we propose a new update flow as follows:

$$
\begin{cases}
\bar{w}_{(i,j)}^k = \text{prox}_{\omega_{(i,j)}\delta^*_{\mathcal{C}_{(i,j)}}} \left(w_{(i,j)}^k + \omega_{(i,j)} \left(N_{(i,j)} x^k \right) \right), \\
\bar{u}_i^k = \text{prox}_{\mu_i \delta^*_{\Omega_i}} \left(u_i^k + \mu_i x_i^k \right), \\
\bar{v}_i^k = \text{prox}_{\sigma_i \delta^*_{\mathcal{A}_i}} \left(v_i^k + \sigma_i x_i^k \right), \\
x_i^{k+1} = \text{prox}_{\gamma_i g_i} \left(x_i^k - \gamma_i \nabla f_i \left(x_i^k \right) - \gamma_i \bar{v}_i^k - \gamma_i \bar{u}_i^k - \gamma_i \sum_{j \in \mathcal{N}_i} C_{ij}^\top \bar{w}_{ij}^k \right), \\
w_{(i,j)}^{k+1} = \bar{w}_{(i,j)}^k + \omega_{(i,j)} \left(N_{(i,j)} x^{k+1} - N_{(i,j)} x^k \right), \\
u_i^{k+1} = \bar{u}_i^k + \mu_i \left(x_i^{k+1} - x_i^k \right), \\
v_i^{k+1} = \bar{v}_i^k + \sigma_i \left(x_i^{k+1} - x_i^k \right),
\end{cases}
\tag{8}
$$

where $\bar{w}_{(i,j)} = (\bar{w}_{(i,j)}^\top, \bar{w}_{(j,i)}^\top)^\top \in \mathbb{R}^{2n}$, $\bar{u}_i \in \mathbb{R}^n$, $\bar{v}_i \in \mathbb{R}^n$ are the auxiliary variables, and γ_i, σ_i, and μ_i are positive stepsizes. Notice that the stepsizes are uncoordinated, which can be selected independently related to different agents and enjoy their own acceptable ranges. Additionally, the edge-based parameters $\omega_{(i,j)}$ can be seen as inherent parameters of the communication network, revealing the quality of the communication.

The steps related to the edge-based variables in update flow (8) cannot be conducted directly, so we next replace them with the agent-based variables. We apply the Moreau decomposition to the first step in update flow (8) such that for the second term on the right side, we have

$$
\text{prox}_{\omega_{(i,j)}^{-1}\delta_{\mathcal{C}_{(i,j)}}} \left(\omega_{(i,j)}^{-1} w_{(i,j)}^k + N_{(i,j)} x^k \right) = \underset{y \in \mathcal{C}_{(i,j)}}{\arg\min} \left\{ \left\| y - \left(\omega_{(i,j)}^{-1} w_{(i,j)}^k + N_{(i,j)} x^k \right) \right\| \right\}. \tag{9}
$$

Define (9) as the projection $\mathcal{P}_{\mathcal{C}_{(i,j)}} \left(\omega_{(i,j)}^{-1} w_{(i,j)}^k + N_{(i,j)} x^k \right)$. Then, according to the definition of the set $\mathcal{C}_{(i,j)}$, the projection has the following explicit expression:

$$
\mathcal{P}_{\mathcal{C}_{(i,j)}} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} a_1 - a_2 \\ a_2 - a_1 \end{bmatrix}.
$$

Thus, for $i \in \mathcal{V}$, $j \in \mathcal{N}_i$, the update step for $\bar{w}_{(i,j)}$ can be decomposed into

$$
\bar{w}_{ij}^k = \frac{1}{2} \left(w_{ij}^k + w_{ji}^k \right) + \frac{\omega_{(i,j)}}{2} \left(C_{ij} x_i^k + C_{ji} x_j^k \right). \tag{10}
$$

Moreover, the update step for $w_{(i,j)}$ can be replaced by

$$
w_{ij}^{k+1} = \bar{w}_{ij}^k + \omega_{(i,j)} C_{ij} \left(x_i^{k+1} - x_i^k \right). \tag{11}
$$

Combining the update flow (8), (10) and (11), we finally propose the decentralized algorithm for Problem (1) in Algorithm 1.

Here, we directly give the stepsize condition of Algorithm 1 in the following assumption. The specific theoretical origin of this condition can be found in the convergence analysis section.

Assumption 2. *(Stepsize conditions)*
For any agent $i \in \mathcal{V}$ and $j \in \mathcal{N}_i$, the stepsizes γ_i, μ_i, σ_i and $\omega_{(i,j)}$ are positive. Let the following condition hold:

$$
\gamma_i < \frac{1}{\frac{\beta_i}{2} + \mu_i + \sigma_i + \sum_{j \in \mathcal{N}_i} \omega_{(i,j)}},
$$

where β_i is the Lipschitz constant for the gradient ∇f_i.

Algorithm 1 The Decentralized Algorithm

Initialization : For each agent $i \in \mathcal{V}$ and all $j \in \mathcal{N}_i$, let $\bar{w}_{ij}^0 \in \mathbb{R}^n$, $\bar{u}_i^0 \in \mathbb{R}^n$, $\bar{v}_i^0 \in \mathbb{R}^n$, $x_i^0 \in \mathbb{R}^n$, $w_{ij}^0 \in \mathbb{R}^n$, $u_i^0 \in \mathbb{R}^n$ and $v_i^0 \in \mathbb{R}^n$.

For $k = 0, 1, 2, \cdots$ **do**

Each agent i repeats, for all $j \in \mathcal{N}_i$,

$$\bar{w}_{ij}^k = \frac{1}{2}\left(w_{ij}^k + w_{ji}^k\right) + \frac{\omega_{(i,j)}}{2}\left(C_{ij}x_i^k + C_{ji}x_j^k\right),$$

$$\bar{u}_i^k = \mathrm{prox}_{\mu_i \delta_{\Omega_i}^*}\left(u_i^k + \mu_i x_i^k\right),$$

$$\bar{v}_i^k = \mathrm{prox}_{\sigma_i \delta_{\mathcal{A}_i}^*}\left(v_i^k + \sigma_i x_i^k\right),$$

$$x_i^{k+1} = \mathrm{prox}_{\tau_i g_i}\left(x_i^k - \gamma_i \nabla f_i\left(x_i^k\right) - \gamma_i \bar{u}_i^k - \gamma_i \bar{v}_i^k - \gamma_i \sum_{j \in \mathcal{N}_i} C_{ij}^\top \bar{w}_{ij}^k\right),$$

$$w_{ij}^{k+1} = \bar{w}_{ij}^k + \omega_{(i,j)} C_{ij}\left(x_i^{k+1} - x_i^k\right),$$

$$u_i^{k+1} = \bar{u}_i^k + \mu_i\left(x_i^{k+1} - x_i^k\right),$$

$$v_i^{k+1} = \bar{v}_i^k + \sigma_i\left(x_i^{k+1} - x_i^k\right).$$

Agent i sends w_{ij}^{k+1}, $\bar{w}_{ij}^k$, $C_{ij}x_i^{k+1}$ to all of its neighbors.

End

Output : The sequence $\left(x_i^k\right)_{k=1}^\infty$ to estimate the optimal solution.

---

4. Convergence Analysis

In this section, we first establish the compact form with operators of the proposed algorithm. Then, the results of the theoretical analysis are provided.

For $i \in \mathcal{V}$, $j \in \mathcal{N}_i$, we make the following definitions. Let w and $\bar{w}$ represent the variables stacked by $w_{(i,j)}$ and $\bar{w}_{(i,j)}$, respectively. Define vectors $u = col\left(u_1, u_2, \ldots, u_m\right)$, $v = col\left(v_1, v_2, \ldots, v_m\right)$, $\bar{u} = col\left(\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_m\right)$ and $\bar{v} = col\left(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_m\right)$. Then, we let $W = \mathrm{blkdiag}\left\{\omega_{(i,j)} I_{2n}\right\}_{(i,j)\in\mathcal{E}}$, $\Gamma = \mathrm{blkdiag}\{\gamma_i I_n\}_{i\in\mathcal{V}}$, $\Sigma = \mathrm{diag}\{\sigma_i I_n\}_{i\in\mathcal{V}}$ and $M = \mathrm{blkdiag}\{\mu_i I_n\}_{i\in\mathcal{V}}$ be the stepsize matrices. Then $\mathcal{C} = \prod_{(i,j)\in\mathcal{E}} \mathcal{C}_{(i,j)}$, $\Omega = \prod_{i\in\mathcal{V}} \Omega_i$ and $\mathcal{A} = \prod_{i\in\mathcal{V}} \mathcal{A}_i$ hold such that there exist $\delta_{\mathcal{C}} = \sum_{(i,j)\in\mathcal{E}} \delta_{\mathcal{C}_{(i,j)}}$, $\delta_\Omega = \sum_{i\in\mathcal{V}} \delta_{\Omega_i}$ and $\delta_{\mathcal{A}} = \sum_{i\in\mathcal{V}} \delta_{\mathcal{A}_i}$. The linear operator $N : x \mapsto \left(N_{(i,j)}x\right)_{(i,j)\in\mathcal{E}}$ is stacked by $N_{(i,j)}$. Considering the resolvent of the proximal operator, the update flow (8) leads to the following equalities:

$$
\begin{cases}
\partial\delta_{\mathcal{C}}^*\left(\bar{w}^k\right) + W^{-1}\bar{w}^k = W^{-1}w^k + Nx^k, \\
\partial\delta_\Omega^*\left(\bar{u}^k\right) + M^{-1}\bar{u}^k = M^{-1}u^k + x^k, \\
\partial\delta_{\mathcal{A}}^*\left(\bar{v}^k\right) + \Sigma^{-1}\bar{v}^k = \Sigma^{-1}v^k + x^k, \\
\partial g\left(\bar{x}^k\right) + \Gamma^{-1}\bar{x}^k + \bar{u}^k + \bar{v}^k + N^\top \bar{w}^k = \Gamma^{-1}x^k - \nabla f\left(x^k\right), \\
w^{k+1} = \bar{w}^k + WN\bar{x}^k - WNx^k, \\
u^{k+1} = \bar{u}^k + M\bar{x}^k - Mx^k, \\
v^{k+1} = \bar{v}^k + \Sigma\bar{x}^k - \Sigma x^k,
\end{cases}
\tag{12}
$$

where $\bar{x}^k = x^{k+1}$ is the auxiliary variable.

Define two variables $U = col(w, u, v, x)$ and $\bar{U} = col(\bar{w}, \bar{u}, \bar{v}, \bar{x})$. Based on the equalities in (12), Algorithm 1 is equivalent to the following compact form described by the operators:

$$
\begin{cases}
\bar{U}^k = (T_H + T_A)^{-1}(T_H - T_Q - T_D)U^k, \\
U^{k+1} = U^k + T_S^{-1}(T_H - T_Q)\left(\bar{U}^k - U^k\right),
\end{cases}
\tag{13}
$$

where the operators are given as follows:

$$
T_H : U \mapsto \left(\left(W^{-1}w\right)^\top, \left(M^{-1}u\right)^\top, \left(\Sigma^{-1}v\right)^\top, \left(N^\top w + u + v + \Gamma^{-1}x\right)^\top\right)^\top,
$$

$$
T_A : U \mapsto \left(\partial \delta_{\mathcal{C}}^*(w)^\top, \partial \delta_{\Omega}^*(u)^\top, \partial \delta_{\mathcal{A}}^*(v)^\top, \partial g(x)^\top\right)^\top,
$$

$$
T_Q : U \mapsto \left((-Nw)^\top, -u^\top, -v^\top, \left(N^\top w + u + v\right)^\top\right)^\top,
$$

$$
T_D : U \mapsto \left(\mathbf{0}_n, \mathbf{0}_n, \mathbf{0}_n, \nabla f(x)^\top\right)^\top,
$$

$$
T_S : U \mapsto \left((Ww)^\top, (Mu)^\top, (\Sigma v)^\top, (\Gamma x)^\top\right)^\top.
$$

Consider one iteration of the proposed algorithm as an operator T. Then we let $U^* = col(x^*, w^*, u^*, v^*)$ be the fixed point of the operator T such that $U^* = TU^*$. Next, we conduct the convergence analysis.

Lemma 1. *(Optimal analysis) Let Assumption 1 be satisfied. The fixed point U^* related to the operator T meets the first-order optimal conditions of the objective problem, and $x^* \in \mathcal{S}$ is an optimal solution.*

Proof. Substituting the fixed point into (12), we have the following set of equalities:

$$
\begin{cases}
\partial \delta_{\mathcal{C}}^*(w^*) - Nx^* = 0, \\
\partial \delta_{\Omega}^*(u^*) - x^* = 0, \\
\partial \delta_{\mathcal{A}}^*(v^*) - x^* = 0, \\
\partial g(x^*) + \nabla f(x^*) + u^* + v^* + N^\top w^* = 0,
\end{cases}
$$

which is also the KKT condition of the Lagrangian function (6). Therefore, x^* is an optimal solution to problem (1). $\square$

The relationship between the fixed point and the optimal solution is ensured by Lemma 1. Split the operator T_H as $T_H = T_P + T_K$, where we let

$$
T_P = \begin{bmatrix}
W^{-1} & 0 & 0 & \frac{1}{2}N \\
0 & M^{-1} & 0 & \frac{1}{2}I \\
0 & 0 & \Sigma^{-1} & \frac{1}{2}I \\
\frac{1}{2}N^\top & \frac{1}{2}I & \frac{1}{2}I & \Gamma^{-1}
\end{bmatrix},
$$

$$
T_K = \begin{bmatrix}
0 & 0 & 0 & -\frac{1}{2}N \\
0 & 0 & 0 & -\frac{1}{2}I \\
0 & 0 & 0 & -\frac{1}{2}I \\
\frac{1}{2}N^\top & \frac{1}{2}I & \frac{1}{2}I & 0
\end{bmatrix},
$$

and further define another linear operator

$$T_{\tilde{P}} = \begin{bmatrix} W^{-1} & 0 & 0 & -\frac{1}{2}N \\ 0 & M^{-1} & 0 & -\frac{1}{2}I \\ 0 & 0 & \Sigma^{-1} & -\frac{1}{2}I \\ -\frac{1}{2}N^{\top} & -\frac{1}{2}I & -\frac{1}{2}I & \Gamma^{-1} \end{bmatrix}.$$

With these definitions above, the following lemma provides the property of the operator T for convergence analysis.

Lemma 2. *Under Assumption 1, there exists the following inequality for U^*:*

$$\left(\bar{U}^k - U^*\right)^{\top} T_D \left(U^k - U^*\right) \geq -\frac{1}{4}\left\|x^k - \bar{x}^k\right\|_B^2,$$

where $B = \mathrm{blkdiag}\{\beta_i I_n\}$ for $i \in \mathcal{V}$ is the Lipschitz parameter matrix.

Proof. With the definition of operator T_D, we have the equality

$$\begin{aligned}
&\left(\bar{U}^k - U^*\right)^{\top} T_D \left(U^k - U^*\right) \\
&= -\left(x^k - \bar{x}^k\right)^{\top}\left(\nabla f\left(x^k\right) - \nabla f(x^*)\right) + \left(x^k - x^*\right)^{\top}\left(\nabla f\left(x^k\right) - \nabla f(x^*)\right).
\end{aligned} \tag{14}$$

According to [32] (Theorem 18.16), for $i \in \mathcal{V}$, ∇f_i is cocoercive, i.e., it holds

$$\left\|\nabla f\left(x^k\right) - \nabla f(x^*)\right\|_{B^{-1}}^2 \leq \left(x^k - x^*\right)^{\top}\left(\nabla f\left(x^k\right) - \nabla f(x^*)\right). \tag{15}$$

Note that for any vector a and b in the same dimension and a diagonal positive definite matrix V, then there exists the inequality $x^{\top}y \leq \|x\|_V^2 + \frac{1}{4}\|y\|_{V^{-1}}^2$. Hence, we have

$$\left(x^k - \bar{x}^k\right)^{\top}\left(\nabla f\left(x^k\right) - \nabla f(x^*)\right) \leq \left\|\nabla f\left(x^k\right) - \nabla f(x^*)\right\|_{B^{-1}}^2 + \frac{1}{4}\|x_k - \bar{x}_k\|_B^2. \tag{16}$$

Combining (14)–(16), we can obtain the objective inequality and end the proof. □

Lemma 3. *Under Assumption 1, there exists the following inequality for U^*:*

$$\left\|U^{k+1} - U^*\right\|_{T_S}^2 - \left\|U^k - U^*\right\|_{T_S}^2 \leq \left\|U^{k+1} - U^k\right\|_{T_S - 2T_{\tilde{P}}}^2 + \frac{1}{2}\left\|x^k - \bar{x}^k\right\|_B^2, \tag{17}$$

where $T_{\tilde{P}}$ is defined before Lemma 2.

Proof. Considering the change of the optimal residual before and after one iteration, we have

$$\begin{aligned}
&\left\|U^{k+1} - U^*\right\|_{T_S}^2 - \left\|U^k - U^*\right\|_{T_S}^2 \\
&= -\left\|U^{k+1} - U^k\right\|_{T_S}^2 + 2\left(U^{k+1} - U^*\right)^{\top} T_S \left(U^{k+1} - U^k\right) \\
&= \left\|U^{k+1} - U^k\right\|_{T_S}^2 + 2\left(U^k - U^*\right)^{\top} T_S \left(U^{k+1} - U^k\right).
\end{aligned} \tag{18}$$

From the second step of the update flow (13), there exists

$$T_S \left(U^{k+1} - U^k\right) = (T_H - T_Q)\left(\bar{U}^k - U^k\right),$$

such that the equality (18) leads to

$$\left\| U^{k+1} - U^* \right\|_{T_S}^2 - \left\| U^k - U^* \right\|_{T_S}^2$$

$$= \left\| U^{k+1} - U^k \right\|_{T_S}^2$$

$$+ 2 \left(\bar{U}^k - U^* \right)^\top (T_H - T_Q) \left(\bar{U}^k - U^k \right)$$

$$- 2 \left(\bar{U}^k - U^k \right)^\top (T_H - T_Q) \left(\bar{U}^k - U^k \right)$$

$$\leq \left\| U^{k+1} - U^k \right\|_{T_S}^2$$

$$+ 2 \left(\bar{U}^k - U^* \right)^\top (T_H - T_Q) \left(\bar{U}^k - U^k \right)$$

$$- 2 \left(\bar{U}^k - U^k \right)^\top T_P \left(\bar{U}^k - U^k \right).$$

From the first step of the update flow (13), it holds that

$$(T_H - T_Q) \left(\bar{U}^k - U^k \right) = - \left(T_D U^k + T_Q \bar{U}^k + T_A \bar{U}^k \right).$$

Thus, we further have

$$\left\| U^{k+1} - U^* \right\|_{T_S}^2 - \left\| U^k - U^* \right\|_{T_S}^2$$

$$\leq \left\| U^{k+1} - U^k \right\|_{T_S}^2 - 2 \left(\bar{U}^k - U^k \right)^\top T_P \left(\bar{U}^k - U^k \right)$$

$$- 2 \left(\bar{U}^k - U^* \right)^\top \left(T_D U^k + T_Q \bar{U}^k + T_A U^* \right)$$

$$- 2 \left(\bar{U}^k - U^* \right)^\top \left(T_A \bar{U}^k - T_A U^* \right). \tag{19}$$

Then, we discuss the right side of (19). Note that Lemma 1 proves the equivalence between the fixed point and the optimal solution. Substituting the property of fixed points into the update flow (13), we obtain $U^* = \bar{U}^*$ and

$$T_A U^* = -T_Q U^* - T_D U^*.$$

Hence, the third term on the right side of (19) satisfies

$$\left(\bar{U}^k - U^* \right)^\top \left(T_D U^k + T_Q \bar{U}^k + T_A U^* \right)$$

$$= \left(\bar{U}^k - U^* \right)^\top T_D \left(U^k - U^* \right) + \left(\bar{U}^k - U^* \right)^\top T_Q \left(\bar{U}^k - U^* \right)$$

$$\geq -\frac{1}{4} \left\| x^k - \bar{x}^k \right\|_B^2 + \left(\bar{U}^k - U^* \right)^\top T_Q \left(\bar{U}^k - U^* \right), \tag{20}$$

where the inequality is based on Lemma 2. Notice that the operator T_A is monotone [32] (Theorem 21.2 and Proposition 20.23), i.e., it holds

$$\left(\bar{U}^k - U^* \right)^\top \left(T_A \bar{U}^k - T_A U^* \right) \geq 0. \tag{21}$$

Since the linear operator T_Q is a skew-symmetric matrix, it is monotone [29]. Combining (19)–(21), we obtain

$$\left\|U^{k+1}-U^*\right\|_{T_S}^2 - \left\|U^k-U^*\right\|_{T_S}^2$$
$$\leq \left\|U^{k+1}-U^k\right\|_{T_S}^2 - 2\left(\bar{U}^k-U^k\right)^\top T_P\left(\bar{U}^k-U^k\right) + \frac{1}{2}\left\|x^k-\bar{x}^k\right\|_B^2. \tag{22}$$

From the second step of the update flow (13), it holds

$$\bar{U}^k - U^k = (T_H - T_Q)^{-1}T_S\left(U^{k+1}-U^k\right),$$

where T_H, T_Q and T_S are the linear operators. Considering that T_P is also a linear operator, the second term on the right side of (22) has an equivalent form:

$$\left(\bar{U}^k-U^k\right)^\top T_P\left(\bar{U}^k-U^k\right) = \left(U^{k+1}-U^k\right)^\top T_{\tilde{P}}\left(U^{k+1}-U^k\right). \tag{23}$$

Substituting (23) into (22), we complete the proof. $\square$

Summarizing the above lemmas, the following theorem supplies the convergence results.

Theorem 1. *When Assumption 1 and 2 are satisfied, for the sequence $\left(U^k\right)_{k\geq0}$ generated by the operator T, we have*

$$\left\|U^{k+1}-U^*\right\|_{T_s}^2 - \left\|U^k-U^*\right\|_{T_s}^2 \leq -\left\|U^{k+1}-U^k\right\|_{2\left(T_{\tilde{P}}-\tilde{B}\right)-T_S}^2, \tag{24}$$

where $\tilde{B} = \mathrm{blkdiag}\left\{\mathbf{0}_{n\times n}, \mathbf{0}_{n\times n}, \mathbf{0}_{n\times n}, \frac{1}{4}B\right\}$. Then, the sequence $\left(U^k\right)_{k\geq0}$ has sublinear rate $O(1/k)$, and the sequence $\left(x^k\right)_{k\geq0}$ converges to an optimal solution $x^* \in \mathcal{S}$.

Proof. With the definition of $\tilde{B}$, we have the following equality:

$$\frac{1}{4}\left\|x^k-\bar{x}^k\right\|_B^2 = \left\|U^{k+1}-U^k\right\|_{\tilde{B}}^2. \tag{25}$$

Substituting (25) into (17), we obtain the inequality (24). Note that under Assumption 2, the matrix $2\left(T_{\tilde{P}}-\tilde{B}\right)-T_S$ is positive definite. Hence, the sequence $\left(U^k\right)_{k\geq0}$ converges to the fixed point U^*. Meanwhile, utilizing [10] (Proposition 1) results in the $O(1/k)$ rate, and based on Lemma 1, the convergence of $\left(x^k\right)_{k\geq0}$ holds. $\square$

In Theorem 1, the positive definite property is needed for the induced matrices, which leads to the stepsize conditions in Assumption 2.

5. Numerical Simulation

The correctness of the theoretical analysis is verified through numerical simulation on a constrained optimization problem over networks in this section.

The constrained quadratic programming problem [33] is considered in the experiments, which has the formulation as follows:

$$\min_{\tilde{x}} \quad \sum_{i=1}^{m} \|\tilde{x}\|_{E_i}^2 + e_i^\top\tilde{x} + \rho_i\|\tilde{x}\|_1, \tag{26a}$$

$$\text{s.t.} \quad A_i\tilde{x} = b_i, \quad i \in \mathcal{V}, \tag{26b}$$

$$\tilde{x}^{\min} \leq \tilde{x} \leq \tilde{x}^{\max}, i \in \mathcal{V}, \tag{26c}$$

where matrix $E_i \in \mathbb{R}^{n \times n}$ is diagonal and positive definite, $e_i \in \mathbb{R}^n$ is a vector, and ρ_i is the penalty factor. Both $x_i^{\min}$ and $x_i^{\max}$ are vectors with constants, which give the bounds of the decision variable $\tilde{x}$. In the light of (1), we can set $f_i(\tilde{x}) = \|\tilde{x}\|_{E_i}^2 + e_i^\mathrm{T} \tilde{x}$ and $g_i(\tilde{x}) = \rho_1 \|\tilde{x}\|_1$.

In this case, the dimension of the decision variable is set as $n = 4$, and we let $r = 1$. For $i \in \mathcal{V}$, the paramount data of Problem (26) are selected randomly. The elements of matrix E_i are in $[1, 2]$, and the elements of the linear operator A_i are in $[1, 15]$. Both vectors e_i and b_i take values in $[-5, 5]$. The box constraints are considered as $[-2.5, 2.5]$. Then, we set the uncoordinated setpsizes randomly as $\gamma_i \in [0.005, 0.006]$, while σ_i, μ_i and $\omega_{(i,j)}$ are in $[5, 6]$. The numerical experiments are performed over the generated network with eight agents, which is displayed in Figure 1. The simulations are carried by running the distributed algorithms on a laptop with Intel(R) Core i5-5500U CPU @ 2.40 GHz, 8.0 GB of RAM, and Matlab R2016a on Windows 10 operating system.
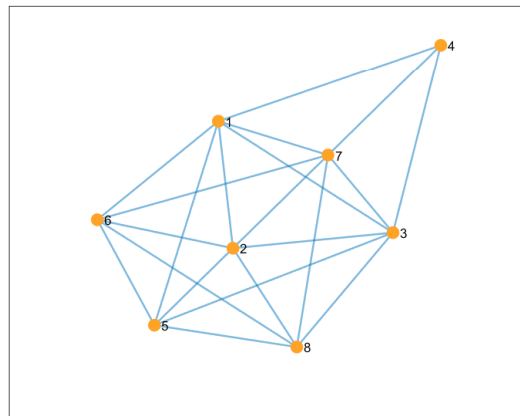

Figure 1. The 8-agents communication network.

The simulation results are shown in Figures 2 and 3. The transient behaviors of each component of x_i^k is displayed in Figure 2, in which a node-based consensus algorithm [34] is introduced as a comparative profile. Note that the obtained optimal solution from the proposed algorithm is in line with that of the node-based consensus one, i.e., $x^* = [0.6900, 0.6270, 0.8046, 0.4400]^\mathrm{T}$, but the latter achieves a stable consensus after 15,000 iterations. Figure 3 shows that our proposed algorithm outperforms the node-based and subgradient algorithms [35] in terms of convergence performance by evaluating the relative errors $\sum_{i=1}^m \|x_i^k - \tilde{x}^*\| / m \|\tilde{x}^*\|$.
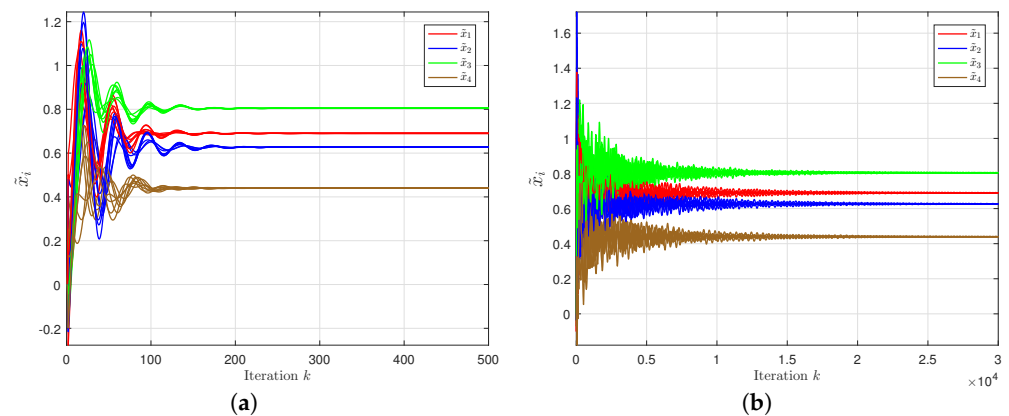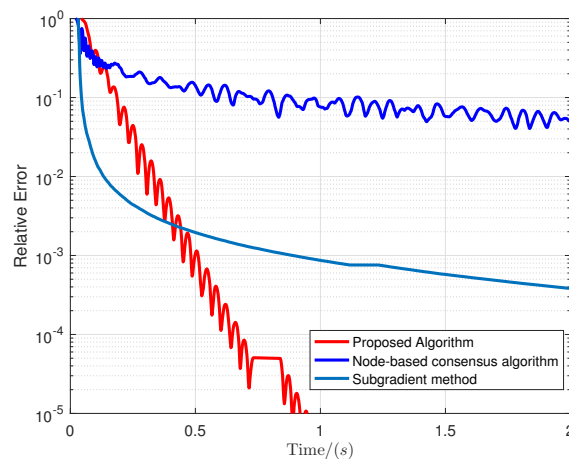

Figure 2. Trajectory of variable x_i^k. (**a**) Proposed algorithm. (**b**) Node-based consensus algorithm.

Figure 3. Convergence performance comparison.

6. Conclusions

In this paper, a distributed algorithm based on proximal operators has been designed to deal with discussed a class of distributed composite optimization problems, in which the local function has a smooth and nonsmooth structure and the decision variable abides by both affine and feasible constraints. Distinguishing attributes of the proposed algorithm include the use of uncoordinated stepsizes and the edge-based communication that avoids the dependency on Laplacian weight matrices. Meanwhile, the algorithm has been verified in theory and simulation. However, there are still some aspects worthy of improvement in this paper. For example, it is worth adopting efficient accelerated protocols (such as the Nesterov-based method and heavy ball method) to improve the convergence rate and developing asynchronous distributed algorithms to deal with the issue of communication latency. In addition, more general optimization models and more efficient algorithms should be investigated in order to address potential applications, e.g., [36–38] with nonconvex objectives, coupled and nonlinear constraints.

Author Contributions: L.F.: Conceptualization, Investigation, Project administration, Software, Writing—original draft. L.R.: Data curation, Software. G.M.: Software. J.T.: Project administration, Software. W.D.: Software. H.L.: Funding acquisition, Investigation, Methodology, Project administration, Resources. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Li, H.; Su, E.; Wang, C.; Liu, J.; Xia, D. A primal-dual forward-backward splitting algorithm for distributed convex optimization. *IEEE Trans. Emerg. Top. Comput. Intell.* **2021**, 1–7. [CrossRef]
2. Li, H.; Hu, J.; Ran, L.; Wang, Z.; Lü, Q.; Du, Z.; Huang, T. Decentralized Dual Proximal Gradient Algorithms for Non-Smooth Constrained Composite Optimization Problems. *IEEE Trans. Parallel Distrib. Syst.* **2021**, *32*, 2594–2605. [CrossRef]
3. Zhang, Y.; Lou, Y.; Hong, Y.; Xie, L. Distributed projection-based algorithms for source localization in wireless sensor networks. *IEEE Trans. Wirel. Commun.* **2015**, *14*, 3131–3142. [CrossRef]
4. Li, B.; Wang, Y.; Li, J.; Cao, S. A fully distributed approach for economic dispatch problem of smart grid. *Energies* **2018**, *11*, 1993. [CrossRef]
5. Cao, C.; Xie, J.; Yue, D.; Huang, C.; Wang, J.; Xu, S.; Chen, X. Distributed economic dispatch of virtual power plant under a non-ideal communication network. *Energies* **2017**, *10*, 235. [CrossRef]

6.   Li, H.; Zheng, Z.; Lü, Q.; Wang, Z.; Gao, L.; Wu, G.; Ji, L.; Wang, H. Primal-dual fixed point algorithms based on adapted metric for distributed optimization. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, 1–15. [CrossRef]
7.   Ababei, C.; Moghaddam, M.G. A survey of prediction and classification techniques in multicore processor systems. *IEEE Trans. Parallel Distrib. Syst.* **2019**, *30*, 1184–1200. [CrossRef]
8.   Liu, S.; Qiu, Z.; Xie, L. Convergence rate analysis of distributed optimization with projected subgradient algorithm. *Automatica* **2017**, *83*, 162–169. [CrossRef]
9.   Li, Y.; Liao, X.; Li, C.; Huang, T.; Yang, D. Impulsive synchronization and parameter mismatch of the three-variable autocatalator model. *Phys. Lett. A* **2007**, *366*, 52–60. [CrossRef]
10.   Shi, W.; Ling, Q.; Wu, G.; Yin, W. EXTRA: An exact first-order algorithm for decentralized consensus optimization. *SIAM J. Optim.* **2015**, *25*, 944–966. [CrossRef]
11.   Lei, J.; Chen, H.F.; Fang, H.T. Primal-dual algorithm for distributed constrained optimization. *Syst. Control Lett.* **2016**, *96*, 110–117. [CrossRef]
12.   Nedic, A.; Ozdaglar, A.; Parrilo, P.A. Constrained consensus and optimization in multi-agent networks. *IEEE Trans. Autom. Control* **2010**, *55*, 922–938. [CrossRef]
13.   Gharesifard, B.; Cortes, J. Distributed continuous-time convex optimization on weight-balanced digraphs. *IEEE Trans. Autom. Control* **2014**, *59*, 781–786. [CrossRef]
14.   Zhu, M.; Martinez, S. On distributed convex optimization under inequality and equality constraints. *IEEE Trans. Autom. Control* **2012**, *57*, 151–164.
15.   Wang, X.; Hong, Y.; Ji, H. Distributed optimization for a class of nonlinear multiagent systems with disturbance rejection. *IEEE Trans. Cybern.* **2016**, *46*, 1655–1666. [CrossRef]
16.   Nedic, A.; Ozdaglar, A. Distributed subgradient methods for multiagent optimization. *IEEE Trans. Autom. Control* **2009**, *54*, 48–61. [CrossRef]
17.   Zhang, L.; Liu, S. Projected subgradient based distributed convex optimization with transmission noises. *Appl. Math. Comput.* **2022**, *418*, 126794. [CrossRef]
18.   Ren, X.; Li, D.; Xi, Y.; Shao, H. Distributed subgradient algorithm for multi-agent optimization with dynamic stepsize. *IEEE/CAA J. Autom. Sin.* **2021**, *8*, 1451–1464. [CrossRef]
19.   Niu, Y.; Wang, H.; Wang, Z.; Xia, D.; Li, H. Primal-dual stochastic distributed algorithm for constrained convex optimization. *J. Frankl. Inst.* **2019**, *356*, 9763–9787. [CrossRef]
20.   Shi, W.; Ling, Q.; Wu, G.; Yin, W. A proximal gradient algorithm for decentralized composite optimization. *IEEE Trans. Singal Process.* **2015**, *63*, 6013–6023. [CrossRef]
21.   Li, Z.; Shi, W.; Yan, M. A decentralized proximal-gradient method with network independent step-sizes and separated convergence rates. *IEEE Trans. Singal Process.* **2019**, *67*, 4494–4506. [CrossRef]
22.   Aybat, N.S.; Hamedani, E.Y. A distributed ADMM-like method for resource sharing over time-varying networks. *SIAM J. Optim.* **2019**, *29*, 3036–3068. [CrossRef]
23.   Xu, J.; Tian, Y.; Sun, Y.; Scutari, G. Distributed algorithms for composite optimization: Unified framework and convergence analysis. *IEEE Trans. Signal Process.* **2021**, *69*, 3555–3570. [CrossRef]
24.   Aybat, N.S.; Wang, Z.; Lin, T.; Ma, S. Distributed linearized alternating direction method of multipliers for composite convex consensus optimization. *IEEE Trans. Autom. Control* **2018**, *63*, 5–20. [CrossRef]
25.   Latafat, P.; Patrinos, P. Asymmetric forward-backward-adjoint splitting for solving monotone inclusions involving three operators. *Comput. Optim. Appl.* **2017**, *68*, 57–93. [CrossRef]
26.   Combettes, P.L.; Pesquet, J.C. Primal-dual splitting algorithm for solving inclusions with mixtures of composite, lipschitzian, and parallel-sum type monotone operators. *Set-Valued Var. Anal.* **2012**, *20*, 307–330. [CrossRef]
27.   Condat, L. A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms. *J. Optim. Theory Appl.* **2013**, *158*, 460–479. [CrossRef]
28.   Vu, B.C. A splitting algorithm for dual monotone inclusions involving cocoercive operators. *Adv. Comput. Math.* **2013**, *38*, 667–681. [CrossRef]
29.   Latafat, P.; Freris, N.M.; Patrinos, P. A new randomized block-coordinate primal-dual proximal algorithm for distributed optimization. *IEEE Trans. Autom. Control* **2019**, *64*, 4050–4065. [CrossRef]
30.   Wei, Y.; Fang, H.; Zeng, X.; Chen, J.; Pardalos, P. A smooth double proximal primal-dual algorithm for a class of distributed nonsmooth optimization problems. *IEEE Trans. Autom. Control* **2020**, *65*, 1800–1806. [CrossRef]
31.   Liu, Q.; Yang, S.; Hong, Y. Constrained consensus algorithms with fixed step size for distributed convex optimization over multiagent networks. *IEEE Trans. Autom. Control* **2017**, *62*, 4259–4265. [CrossRef]
32.   Bauschke, H.H.; Combettes, P.L. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*; Springer: Cham, Switzerland, 2011.
33.   Boyd, S.; Vandenberghe, L. *Convex Optimization*; Cambridge University Press: Cambridge, UK, 2004.
34.   Zhao, Y.; Liu, Q. A consensus algorithm based on collective neurodynamic system for distributed optimization with linear and bound constraints. *Neural Netw.* **2020**, *122*, 144–151. [CrossRef] [PubMed]
35.   Yuan, D.; Xu, S.; Zhao, H. Distributed primal-dual subgradient method for multiagent optimization via consensus algorithms. *IEEE Trans. Syst. Man Cybern. Part B Cybern.* **2011**, *41*, 1715–1724. [CrossRef] [PubMed]

36. Ćalasan, M.; Micev, M.; Ali, Z.M.; Zobaa, A.F.; Aleem, S.H.E.A. Parameter estimation of induction machine single-cage and double-cage models using a hybrid simulated annealing–evaporation rate water cycle algorithm. *Mathematics* **2020**, *8*, 1024. [CrossRef]
37. Ali, Z.M.; Diaaeldin, I.M.; Aleem, S.H.E.A.; El-Rafei, A.; Abdelaziz, A.Y.; Jurado, F. Scenario-based network reconfiguration and renewable energy resources integration in large-scale distribution systems considering parameters uncertainty. *Mathematics* **2021**, *9*, 26. [CrossRef]
38. Rawa, M.; Abusorrah, A.; Bassi, H.; Mekhilef, S.; Ali, Z.M.; Aleem, S.; Hasanien, H.M.; Omar, A.I. Economical-technical-environmental operation of power networks with wind-solar-hydropower generation using analytic hierarchy process and improved grey wolf algorithm. *Ain Shams Eng. J.* **2021**, *12*, 2717–2734. [CrossRef]