Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2012, Article ID 471854, 20 pages
doi:10.1155/2012/471854

Research Article

Modified Lagrangian Methods for Separable Optimization Problems

Abdelouahed Hamdi and Aiman A. Mukheimer

Department of Mathematics and Physical Sciences, Prince Sultan University, P.O. Box 66833, Riyadh 11586, Saudi Arabia

Correspondence should be addressed to Abdelouahed Hamdi, whamdi99@gmail.com

Received 29 November 2011; Revised 17 January 2012; Accepted 18 January 2012

Academic Editor: Muhammad Aslam Noor

We propose a convergence analysis of a new decomposition method to solve structured optimization problems. The proposed scheme is based on a class of modified Lagrangians combined with the allocation of resources decomposition algorithm. Under mild assumptions, we show that the method generates convergent primal-dual sequences.

1. Introduction

The interest for large-scale optimization problems [1] has grown over the past twenty years and it is likely to continue to increase during the upcoming decades; while on the one hand, the systemic approach to modeling real systems results in more and more complicated models, the increasing computing power of the microprocessors, along with the recent advancements in parallel architectures on the other hand, seems to push back the limits of practical treatment of these models. Decomposition methods (including splitting, partitioning, and parallel methods) are practical candidates to solve large problems where an internal structure allows to identify the "weakly coupled subsystems." It should be clear that the reduction of the dimension of the original problem is not the only motivation for decomposing it into subproblems. Other important motivations are as follows:

(i) partitioning heterogeneous models when it is the juxtaposition of various parts of the model which turns its numerical treatment difficult (as in mixed models with continuous and discrete variables) [2];

(ii) decentralizing the decision-making such that the decomposition procedure could lead to autonomous subsystems, which are capable of computing their own part of the global solution independently without the need of a centralized decision level [2, 3];

(iii) parallelizing the computation among different processors [1, 2].

Our objective in this paper is to present a family of decomposition algorithms based on proximal-like techniques, which are suitable for decentralized and parallelized computations. The algorithm can be seen as a separable version of the nonlinear rescaling principle method [4] and is closely related to three known techniques: the Partial Inverse method, proposed by Spingarn in 1985 [5, 6], the Alternate Direction Multiplier method [7, 8], proposed originally by Gabay and Mercier [9] for the splitting of variational inequalities, and the Separable Augmented Lagrangian Algorithm (SALA) developed by Hamdi [10–12].

We will use a simplified framework to present the algorithm called the φ-SALA. It is tailored towards minimizing a convex separable function with separable constraints. A nice feature of the φ-SALA is that it preserves separability, as each iteration splits into a proximal step on the dual function and a projection step on the subspace. On the other hand, it yields the proximal decomposition on the graph of a maximal operator that has been introduced by Mahey et al. [13] for quadratic choice of the auxiliary functions φ.

The first drawback associated with the classical quadratic multiplier method (augmented lagrangian [14, 15]) and/or the nonlinear rescaling principle algorithm is that the augmented Lagrangian function and also the rescaled Lagrangian function are no long separable even when the original problem is separable. In other words, when applied to separable constraints like $g(x) = \sum_{i=1}^{p} g_i(x_i)$, the terms $\| \sum_{i=1}^{p} g_i(x_i) \|$ or $(1/k)\varphi(k \sum_{i=1}^{p} g_i(x_i))$ are no long separable. However, some careful reformulation of the problem (e.g., by introducing additional variables) may preserve some of the given separable structure, thus giving a chance to the augmented Lagrangian framework, to play again an important role in the development of efficient decomposition schemes.

The second drawback associated with the multiplier methods is the fact that it is only differentiable once even when the problem's data allow for higher differentiability, disabling the application of efficient Newton type methods. In fact such a lack of continuity in the second derivative can significantly slow down the rate of convergence of these methods and thus causing algorithmic failure. One way of coping with this difficulty is to use the recently developed nonquadratic multiplier methods based on entropy-like proximal methods (see [3, 16–20]), and thereby leading to multiplier methods of which are twice continuously differentiable as opposed to the classical quadratic multiplier (if the original problem is also C^2). This is an important advantage since Newton type methods can then be applied.

With respect to the first drawback, in [10, 11], Hamdi has proposed a separable augmented Lagrangian algorithm (SALA) which can be derived from the resource directive subproblems associated with the coupling constraints ((SALA) was developed for non-convex and convex separable problems.). And in this paper, we alleviate the second drawback by using the Nonlinear Re-scaling NR method, so as to get a separable augmented Lagrangian function, which is at least twice continuously differentiable.

It is worth citing here some recent works where decomposition methods related to our subject were developed (to the best of our knowledge). The most recent one is a modification of the original algorithm (SALA) [10–12] where Guèye et al. [21] replaced the scalar penalty parameter in [10] by a diagonal positive definite matrix of scaling factors. Their convergence analysis was done for only affine constraints. Auslender and Teboulle [22] proposed the entropic proximal decomposition method inducing C^∞-lagrangians for solving the structured convex minimization problems and variational inequalities based on the combination of their recent logarithmic-quadratic proximal point theory [23, 24] with the Chen-Teboulle decomposition scheme [25]. Kyono and Fukushima [26] proposed an extension of the Chen-Teboulle decomposition scheme [25] combined with the Bregman-based proximal point algorithm. Kyona and Fukushima their method was developed for solving large-scale Variational Inequalities (VIPs) (see [26] and for more developments on VIPs see [27–29] and references therein.). Hamdi and Mahey [11] proposed a stabilized version of the original (SALA) by using a primal proximal regularization that yields to better numerical stability, specially for nonconvex minimization problems. For other references on decomposition methods, one may refer to [5, 7, 11, 12, 25, 26, 30–34] and to Hamdi's survey on decomposition methods based on augmented lagrangian functions [35] and references therein.

The remainder of this paper is organized as follows. In Section 2, we present the nonlinear rescaling principle of Polyak. In Section 3, we describe the application of the nonlinear rescaling method in conjunction with (SALA) to a general structured convex program in order to yield to the decomposition method presented in [30], which can be seen as separable augmented lagrangian method. Section 4 is dedicated to the extended convergence analysis of the proposed algorithm.

2. Nonlinear Rescaling Principle

Let f be a convex real-valued function and let $(g_1(x), \dots, g_p(x))^\top$ be finite concave real-valued functions on R^n, and consider the convex programming problem:

$$\min_{x \in \mathbb{R}^n} \left\{ f(x) : g_i(x) \geq 0, i = \overline{1, m} \right\}. \tag{C}$$

Recently nonquadratic augmented lagrangian have received much attention in the literature; see, for example, ([3, 4, 16–20, 36] and references therein.). These methods basically rely on applying new class of proximal-like maps on the dual of (C), see [20], and in turns are in fact equivalent to using the nonlinear re-scaling method, as shown in [37]. In this paper we use the nonlinear re-scaling principle to construct smooth lagrangian decomposition methods. Thus, we begin by summarizing this approach; for details see [4, 18, 37] and references therein.

The main idea of the nonlinear re-scaling principle (called here the NR method) is considering a class of strictly concave and smooth enough scalar function with particular properties, and using it to transform the constraints terms of the classical lagrangian. The NR method alternates at each step the unconstrained minimization of the classical lagrangian for the equivalent problem with the lagrange multiplier update. It allows for generating a wide class of augmented lagrangian methods (for instance, the exponential multiplier method, the modified log-barrier method, and the Log-sigmoid multiplier method [18], etc.).

Let us consider the following class of C^2 functions φ defined of $\mathbb{R}$ and satisfying the following properties:

(P1) $\varphi'(t) > 0$, for all $t \in \mathbb{R}$;

(P2) $\varphi''(t) < 0$, for all $t \in \mathbb{R}$;

(P3) $\varphi(0) = 0$;

(P4) $\varphi'(0) = 1$;

(P5) $m^{-1} \leq \varphi''(t) < 0$, for all $t \in \mathbb{R}$ and $\varphi''(t) \leq -M^{-1}$, for all $t < 0$, where $M, m > 0$;

(P6) $(-\varphi)_{\infty}(-1) \geq 0$;

(P7) $(-\varphi)_{\infty}(1) = 0$.

Note that from (P2) and (P3) the functions φ and φ' are one-to-one and φ is strictly concave. Let $\varphi \in \Phi$, then we have for any $k > 0$,

$$t \geq 0 \quad \frac{1}{k}\varphi(tk) \geq 0. \tag{2.1}$$

The nonlinear re-scaling principle is based on the idea of transforming the original problem (C) to an equivalent problem, namely, to one which has the same set of optimal solutions as (C). To this end, let us consider here a parameterized transformed problems written as follows:

$$\min_{x \in R^n}\left\{ f(x) : k^{-1}\varphi\left(kg_i(x)\right) \geq 0, i = \overline{1,m} \right\}. \tag{$[C_k,\varphi]$}$$

Clearly, problem $([C_k,\varphi])$ is also convex and has the same feasible set as (C). The nonlinear re-scaling iterations are based on the classical lagrangian function denoted here by $P_k(x,u)$ associated to the problem $([C_k,\varphi])$; that is,

$$P_k(x,u) = f(x) - \sum_{i=1}^{m} k^{-1}u_i\varphi\left(kg_i(x)\right) \tag{2.2}$$

and can be resumed as follows.

Algorithm 2.1. Given $\varphi \in \Phi$, $u^0 > 0$, $k > 0$, generate the sequence $\{x^k, u^k\}$:
Find $x^{k+1} \in \arg\min_{x \in R^n} P_k(x,u^k) = f(x) - \sum_{i=1}^{m} k^{-1}u_i^k\varphi(kg_i(x))$.
Update $u_i^{k+1} = u_i^k\varphi'(kg_i(x^{k+1}))$, $i = 1, \ldots, m$.

Remark 2.2. Note that the multipliers are nonnegative for all k by (P1). Also, it is worth to mention the possibility to change the penalty parameter k at each iteration s. In [37], the authors propose a dynamic scaling parameters update as follows: $k_i^{k+1} = k/u_i^{k+1}$. This update will be used in our decomposition scheme in the next section.

The NR algorithm allows a generation of a wide class of augmented lagrangian. Typical examples include the choices $\varphi(t) = e^t - 1$ to get the exponential multiplier method, or $\varphi(t) = -\log(1-t)$ to get the modified log-barrier function method. Convergence analysis of Algorithm 2.1 is given in [4, 18, 37, 38].

3. Separable Modified Lagrangian Algorithm

In this section, we recall the generalization of the separable augmented lagrangian algorithm (SALA) algorithm (see [10, 11]) called (φ SALA) proposed in [30], to solve large-scale convex inequality constrained programs with separable structure. We are concerned here with block separable nonlinear constrained optimization problems:

$$\min_{x\in\mathbb{R}^n}\left\{F(x)=\sum_{i=1}^{p} f_i(x_i): x\in\Omega\right\}, \tag{SP}$$

where $f_i:\mathbb{R}^{n_i}\to\mathbb{R}$ are all convex functions, and

$$\Omega=\left\{x\in\mathbb{R}^n:\sum_{i=1}^{p} g_{ij}(x_i)\geq 0, j=\overline{1,m}\right\} \tag{3.1}$$

is the convex set where g_{ij} are defined from $\mathbb{R}^{n_i}\to\mathbb{R}$ for $j=\overline{1,m}$, $i=\overline{1,p}$, $\sum_{i=1}^{p} n_i=n$. Along this work, all the functions f_i, g_{ij} are C^2 and we assume the following.

(A1) The optimal set X^* of (SP) is nonempty and bounded.

(A2) The Slater's condition holds, that is,

$$\exists\overline{x}\in\mathbb{R}^n:\sum_{i=1}^{p} g_{ij}(\overline{x}_i)>0,\quad j=\overline{1,m}. \tag{3.2}$$

Now, to construct our decomposition algorithm, we use the m allocation vectors $y_j\in A=\{z\in\mathbb{R}^p\mid\sum_{i=1}^{p} z_i=0\}$ to get the equivalent problem (If (x^*,y^*) is an optimal solution to (SP_y) then x^* is an optimal solution to (SP).),

$$\begin{aligned}
&\min\ \sum_{i=1}^{p} f_i(x_i)\\
&\text{such that}\ \ g_{ij}(x_i)+y_{ij}\geq 0,\quad j=\overline{1,m},\ i=\overline{1,p},\\
&\qquad\sum_{i=1}^{p} y_{ij}=0\quad j=\overline{1,m},\\
&\qquad x_i\in\mathbb{R}^{n_i},\quad i=\overline{1,p},
\end{aligned} \tag{SP_y}$$

to which we propose to apply the nonlinear re-scaling principle with partial elimination of the constraints. We mean that only the constraints $g_{ij}(x_i)+y_{ij}\geq 0$ are replaced by the equivalent ones $(1/\lambda)\varphi(\lambda(g_{ij}(x_i)+y_{ij}))\geq 0$. Then for any $\lambda>0$, the following minimization problem:

$$\begin{aligned}
&\min_x\quad \sum_{i=1}^{p} f_i(x_i)\\
&\text{s.t}\quad \frac{1}{\lambda}\varphi\left[\lambda\left(g_{ij}(x_i)+y_{ij}\right)\right]\geq 0\quad j=\overline{1,m},\ i=\overline{1,p}\\
&\sum_{i=1}^{p} y_{ij}=0\quad j=\overline{1,m}
\end{aligned} \tag{$[SP(y,\varphi)]$}$$

is equivalent to the (SP_y) and according to Algorithm 2.1, for all $u^0 \in \mathbb{R}^{pm}_{++}$, $\lambda^0 = (\lambda^0_{11}, \dots, \lambda^0_{pm})$ where $\lambda^0_{ij} = \lambda(u^0_{ij})^{-1}$, we have the following iterative scheme:

$$\left(x^{k+1}, y^{k+1}\right) \longrightarrow \min_{\sum_{i=1}^{p} y_{ij}=0,\, j=\overline{1,m}} \Theta_{\lambda^k}\left(x, y, u^k\right), \tag{3.3}$$

$$u^{k+1}_{ij} = u^k_{ij}\varphi'\left(\lambda^k_{ij}\left(g_{ij}\left(x^{k+1}_i\right) + y^{k+1}_{ij}\right)\right), \tag{3.4}$$

$$\lambda^{k+1}_{ij} = \lambda\left(u^{k+1}_{ij}\right)^{-1}, \quad \forall i, j, \tag{3.5}$$

where $\Theta_\lambda(x, y, u)$ denotes the classical lagrangian for $([SP(y, \varphi)])$ given by

$$\Theta_\lambda(x, y, u) = \sum_{i=1}^{p} f_i(x_i) - \frac{1}{\lambda}\sum_{i=1}^{p}\sum_{j=1}^{m} u_{ij}\varphi[\lambda(g_{ij}(x_i) + y_{ij})], \tag{3.6}$$

where $u = (u_{11}, u_{21}, \dots, u_{p1}, u_{12}, \dots, u_{p2}, \dots, u_{1m}, \dots, u_{pm}) \in \mathbb{R}^{pm}_{++}$.

The minimization in (3.3) is done by alternating the minimization with respect to x, then followed by the one w.r.t. the allocation variable y; that is, we fix $y = y^k$ and find

$$x^{k+1} \in \operatorname{Arg}\min_x \Theta_{\lambda^k}\left(x, y^k, u^k\right) = \operatorname{Arg}\min_x \sum_{i=1}^{p}\left\{ f_i(x_i) - \sum_{j=1}^{m} \frac{1}{\lambda^k_{ij}} u^k_{ij}\varphi\left[\lambda\left(g_{ij}(x_i) + y^k_{ij}\right)\right]\right\}. \tag{3.7}$$

Then we can split the above minimization into p independent sub-problems with low dimension. That is,

$$x^{k+1}_i \in \operatorname{Arg}\min_{x_i}\left\{ f_i(x_i) - \sum_{j=1}^{m} \frac{1}{\lambda^k_{ij}} u^k_{ij}\varphi\left[\lambda^k_{ij}\left(g_{ij}(x_i) + y^k_{ij}\right)\right]\right\}. \tag{3.8}$$

And now we fix $x = x^{k+1}$ to solve for y^{k+1}

$$y^{k+1} \in \operatorname{Arg}\min\left\{ \sum_{i=1}^{p}\sum_{j=1}^{m} -\frac{1}{\lambda^k_{ij}} u^k_{ij}\varphi\left[\lambda^k_{ij}\left(g_{ij}\left(x^{k+1}_i\right) + y_{ij}\right)\right] : \sum_{i=1}^{p} y_{ij} = 0\right\}. \tag{3.9}$$

The following lemma gives an important link between the allocation variable and the lagrange dual variable.

Lemma 3.1. *According to* (3.9), *u^{k+1} and y^{k+1} are orthogonal and satisfy*

$$y^{k+1}_{ij} = -g_{ij}\left(x^{k+1}_i\right) + \delta^{k+1}_j, \quad i = \overline{1,p},\ j = \overline{1,m}, \tag{3.10}$$

$$u^{k+1}_j = u^k_j \varphi^{'(\lambda^k_j \delta^{k+1}_j)}, \quad j = \overline{1,m}, \tag{3.11}$$

where $\delta^{k+1}_j = p^{-1}\sum_{i=1}^{p} g_{ij}(x^{k+1}_i)$.

Proof. For any $j = \overline{1,m}$, by writing the classical lagrangian to (3.9)

$$L^k(z_j, t_j) = \sum_{i=1}^{p} -\frac{1}{\lambda_{ij}^k} u_{ij}^k \varphi\left[\lambda_{ij}^k\left(g_{ij}\left(x_i^{k+1}\right) + y_{ij}\right)\right] + \sum_{i=1}^{p} y_{ij} t_j, \tag{3.12}$$

where $t_j \in \mathbb{R}$, $j = \overline{1,m}$, and using the optimality, with (3.4), we show that

$$t_j = u_{ij}^{k+1} = u_{ij}^k \varphi'\left[\lambda_{ij}^k\left(g_{ij}\left(x_i^{k+1}\right) + y_{ij}^{k+1}\right)\right], \tag{3.13}$$

which means that u^{k+1} does not depend on i; that is, $u_{ij}^{k+1} = u_{lj}^{k+1}$, for all $i, l = \overline{1,p}$, and u_{ij} can be replaced by u_j for all $i = \overline{1,p}$.

Now, according to (P1) and after straightforward calculations, we reach (3.10). Equation (3.11) is obtained directly using (3.4) and (3.10). The orthogonality of the vectors u^{k+1} and y^{k+1} is direct.

$$\langle u, y\rangle = \sum_{j=1}^{m}\sum_{i=1}^{p} u_{ij} y_{ij} = \sum_{j=1}^{m} u_j \sum_{i=1}^{p} y_{ij} = 0. \tag{3.14}$$

□

One can observe that also the penalty parameters $(\lambda_{1j}, \lambda_{2j}, \ldots, \lambda_{pj})$ belong to the set $V = \{(a_1, \ldots, a_p) : a_1 = \cdots = a_p\}$, and finally our algorithm (φSALA) can be stated as follows.

Algorithm 3.2. We have the following steps.

Step 1. Select $\varphi \in \Phi$, $u^0 \in \mathbb{R}_{++}^{mp}$ where $u_j \in V$, $j = \overline{1,m}$, $\lambda > 0$, $y^0 = (y_1, \ldots, y_m)$, where $y_j \in A$, $j = \overline{1,m}$ and $\lambda^0 = (\lambda_{11}, \ldots, \lambda_{pm})$ where $\lambda_j^0 = \lambda(u_j^0)^{-1}$, $\epsilon_1 > 0$, $\epsilon_2 > 0$, $\epsilon_3 > 0$.

Step 2. Determine: for any $i = \overline{1,p}$

$$x_i^{k+1} := \arg\min_{x_i \in \mathbb{R}^{n_i}} \left\{ f_i(x_i) - \sum_{j=1}^{m} \frac{1}{\lambda_j^k} u_j^k \varphi\left(\lambda_j^k\left(g_{ij}(x_i) + y_{ij}^k\right)\right)\right\}. \tag{3.15}$$

Step 3. If $v_1 \leq \epsilon_1$, $v_2 \leq \epsilon_2$, $v_3 \leq \epsilon_3$ where

$$\begin{gathered} v_1 = \sum_{j=1}^{m}\sum_{i=1}^{p}\left|u_j^k\left(g_{ij}\left(x_i^{k+1}\right) + y_{ij}^k\right)\right|, \qquad v_2 = \max_{1\leq j\leq m}\left\{\left(-\delta_j^{k+1}\right)\right\}, \\ v_3 = \left|F\left(x^k\right) - F\left(x^{k+1}\right)\right|. \end{gathered} \tag{3.16}$$

Stop.

Else: go to Step 4.

Step 4. Update and go back to Step 2:

$$y_{ij}^{k+1} = -g_{ij}\left(x_i^{k+1}\right) + \delta_j^{k+1} \quad i = \overline{1,p},\ j = \overline{1,m}$$
$$u_j^{k+1} = u_j^k \varphi'\left(\lambda_j^k \delta_j^{k+1}\right), \quad \lambda_j^{k+1} = \lambda\left(u_j^{k+1}\right)^{-1}. \tag{3.17}$$

The following proposition gives us some properties of Θ_λ.

Proposition 3.3. (1) $\Theta_\lambda(x,y,u)$ *is strictly convex in* $x \in \mathbb{R}^n$ *for any* $u \in \mathbb{R}^{mp}_{++}$, $y \in \mathbb{R}^{mp}$, $\lambda > 0$.
(2) *For any K.K.T point* (x^*, y^*, u^*) *of* (SP_y) *One has*

(i) $\Theta_\lambda(x^*, y^*, u^*) = L(x^*, y^*, u^*) = \sum_{i=1}^p f_i(x_i^*)$,

(ii) $\nabla_x \Theta_\lambda(x^*, y^*, u^*) = \nabla_x L(x^*, y^*, u^*) = 0$,

(iii) $\nabla_x^2 \Theta_\lambda(x^*, y^*, u^*) = \nabla_x^2 L(x^*, y^*, u^*) - \varphi''(0)\nabla G(x^*)\Lambda\nabla G(x^*)'$, *where* $\Lambda = \operatorname{diag}(u_{ij}^*)_{(i,j)=(1,1)}^{(p,m)}$ *and* $G(x) = (\lambda_{11}g_{11}(x_1), \ldots, \lambda_{mp}g_{mp}(x_p))$.

Proof. (1) $\Theta_\lambda(x,y,u) = \sum_{i=1}^p f_i(x_i) - \sum_{j=1}^m (1/\lambda_{ij})u_{ij}\varphi(\lambda_{ij}(g_{ij}(x_i) + y_{ij}))$. Since g_{ij}, φ, $i = \overline{1,p}$, $j = \overline{1,m}$ are concave, strictly concave (resp.) and φ is increasing. Then Θ is strictly convex in $x \in \mathbb{R}^n$ for any $u \in \mathbb{R}^{mp}_{++}$, $y \in \mathbb{R}^{mp}$, $\lambda > 0$.
(2) Let (x^*, y^*, u^*) be any K.K.T point of (SP_y) then

(i) we have

$$\Theta_\lambda(x^*, y^*, u^*) = \sum_{i=1}^p \left[f_i(x_i^*) - \sum_{j=1}^m \frac{1}{\lambda_{ij}} u_{ij}^* \varphi\left(\lambda_{ij}\left(g_{ij}(x_i^*) + y_{ij}^*\right)\right)\right],$$
$$L(x^*, y^*, u^*) = \sum_{i=1}^p \left[f_i(x_i^*) - \sum_{j=1}^m u_{ij}^*\left(g_{ij}(x_i^*) + y_{ij}^*\right)\right] = \sum_{i=1}^p f_i(x_i^*) \tag{3.18}$$

from the complementary condition.

$$\Longrightarrow \varphi\left(\lambda_{ij}\left(g_{ij}(x_i^*) + y_{ij}^*\right)\right) = 0$$
$$\Longrightarrow u_{ij}^*\varphi\left(\lambda_{ij}\left(g_{ij}(x_i^*) + y_{ij}^*\right)\right) = 0, \tag{3.19}$$

then we get

$$\Theta_\lambda(x^*, y^*, u^*) = \sum_{i=1}^p f_i(x_i^*). \tag{3.20}$$

(ii) We have

$$\nabla_x \Theta_\lambda(x^*, y^*, u^*) = \sum_{i=1}^{p} \left\{ \nabla_x f_i(x_i^*) - \sum_{j=1}^{m} u_{ij}^* \varphi'\left(\lambda_{ij}\left(g_{ij}(x_i^*) + y_{ij}^*\right)\right) \nabla_x g_{ij}(x_i^*) \right\}. \tag{3.21}$$

Similarly, if $u_{ij}^* \neq 0$ then $g_{ij}(x^*) + y_{ij}^* = 0$ therefore $\varphi'(\lambda_{ij}(g_{ij}(x_i^*) + y_{ij}^*)) = 1$. This means that

$$\nabla_x \Theta_\lambda(x^*, y^*, u^*) = \sum_{i=1}^{p} \left(\nabla_x f_i(x_i^*) - \sum_{j=1}^{m} u_{ij}^* \nabla_x g_{ij}(x_i^*) \right) = \nabla_x L(x^*, y^*, u^*) = 0. \tag{3.22}$$

(iii) From (ii) we can calculate $\nabla_x^2 \Theta_\lambda(x^*, y^*, u^*)$,

$$\begin{aligned}\nabla_x^2 \Theta_\lambda(x, y, u) = \sum_{i=1}^{p} \Bigg[& \nabla_x^2 f_i(x_i) - \sum_{j=1}^{m} \lambda_{ij} u_{ij} \varphi''(\lambda_{ij}(g_{ij}(x_i) + y_{ij}))(\nabla_x g_{ij}(x_i))(\nabla_x g_{ij}(x_i))' \\ & - \sum_{j=1}^{m} u_{ij} \varphi'(\lambda_{ij}(g_{ij}(x_i) + y_{ij})) \nabla_x^2 g_{ij}(x_i) \Bigg].\end{aligned} \tag{3.23}$$

At (x^*, y^*, u^*), we have

$$\begin{aligned}\nabla_x^2 \Theta_\lambda(x^*, y^*, u^*) &= \sum_{i=1}^{p} \Bigg(\nabla_x^2 f_i(x_i^*) - \sum_{j=1}^{m} \Big[\lambda_{ij} u_{ij}^* \varphi''\left(\lambda_{ij}\left(g_{ij}(x_i^*) + y_{ij}^*\right)\right)(\nabla_x g_{ij}(x_i^*))(\nabla_x g_{ij}(x_i^*))' \\ &\qquad + u_{ij}^* \varphi'\left(\lambda_{ij}\left(g_{ij}(x_i^*) + y_{ij}^*\right)\right) \nabla_x^2 g_{ij}(x_i^*) \Big] \Bigg) \\ &= \sum_{i=1}^{p} \left\{ \nabla_x^2 f_i(x_i^*) - \sum_{j=1}^{m} u_{ij}^* \nabla_x^2 g_{ij}(x_i^*) \right\} \\ &\quad - \sum_{i=1}^{p} \sum_{j=1}^{m} \lambda_{ij} u_{ij}^* \varphi''\left(\lambda_{ij}\left(g_{ij}(x_i^*) + y_{ij}^*\right)\right)(\nabla_x g_{ij}(x_i^*))(\nabla_x g_{ij}(x_i^*))' \\ &= \nabla_x^2 L(x^*, y^*, u^*) - \sum_{i=1}^{p} \sum_{j=1}^{m} \lambda_{ij} u_{ij}^* \varphi''\left(\lambda_{ij}\left(g_{ij}(x_i^*) + y_{ij}^*\right)\right)(\nabla_x g_{ij}(x_i^*)) \\ &\quad \times (\nabla_x g_{ij}(x_i^*))' \\ &= \nabla_x^2 L(x^*, y^*, u^*) - \varphi''(0) \sum_{i=1}^{p} \sum_{j=1}^{m} \lambda_{ij} u_{ij}^* (\nabla_x g_{ij}(x_i^*))(\nabla_x g_{ij}(x_i^*))'.\end{aligned} \tag{3.24}$$

Let

$$G(x) = (\lambda_{11} g_{11}(x_1), \dots, \lambda_{p1} g_{p1}(x_p), \lambda_{12} g_{12}(x_1), \dots, \lambda_{mp} g_{mp}(x_p)) \tag{3.25}$$

then

$$\nabla G(x) = \left(\lambda_{11} \nabla g_{11}(x_1) \;\; \lambda_{21} \nabla g_{21}(x_2) \;\; \cdots \;\; \nabla \lambda_{mp} g_{mp}(x_p)\right)_{n \times pm}. \tag{3.26}$$

That is, the first column is the gradient $\nabla g_{11}(x_1)$ and so on. Let $\Lambda = \operatorname{diag}\,(u_{ij}^*)_{(i,j)=(1,1)}^{(p,m)}$ that is, the diagonal is

$$\left(u_{11}^*, u_{21}^*, \dots, u_{p1}^*, u_{12}^*, \dots, u_{mp}^*\right). \tag{3.27}$$

Then,

$$\begin{aligned}
\nabla G(x^*)\Lambda &= \left(\nabla g_{11}(x_1^*) u_{11}^* \lambda_{11} \;\; \nabla g_{21}(x_2^*) u_{21}^* \lambda_{21} \;\; \cdots \;\; \nabla g_{mp}\left(x_p^*\right) u_{mp}^* \lambda_{mp}\right)_{n \times pm}, \\
\nabla G(x^*)\Lambda(\nabla G(x^*))' &= \left(\lambda_{11} \nabla g_{11}(x_1^*) u_{11}^* \;\; \cdots \;\; \lambda_{pm} \nabla g_{mp}\left(x_p^*\right) u_{mp}^*\right) \begin{pmatrix} (\nabla g_{11}(x_1^*))' \\ \cdots \\ \left(\nabla g_{mp}\left(x_p^*\right)\right)' \end{pmatrix} \\
&= \sum_{i=1}^{p} \sum_{j=1}^{m} \lambda_{ij} u_{ij}^* (\nabla g_{ij}(x_i^*))(\nabla g_{ij}(x_i^*))'.
\end{aligned} \tag{3.28}$$

Then

$$\nabla_x^2 \Theta_\lambda(x^*, y^*, u^*) = \nabla_x^2 L(x^*, y^*, u^*) - \varphi''(0) \nabla G(x^*)\Lambda(\nabla G(x^*))'. \tag{3.29}$$

□

Remark 3.4. The analysis presented in this paper differs from the short one presented in [30]. In this paper, our analysis is made possible by the strong tool of recession functions.

Assumption (A1) can be written in terms of recession functions ($f_\infty(d) = \lim_{t \to \infty}(f(x + td) - f(x))/t$, for all $x \in \operatorname{dom} f$ for convex proper and lower semicontinuous functions. In our case the functions are convex proper and continuous.)

$$F_\infty(d) \leq 0, \quad \left(-\sum_{i=1}^{p} g_{ij}\right)_\infty (d) \leq 0, \quad \forall j = \overline{1, m} \Longrightarrow d = 0. \tag{3.30}$$

Further, since $\sum_{i=1}^{p} g_{ij}$ is not identically $+\infty$, [39, Theorem 9.3, page 77], allows us to reformulate (A1) as follows:

$$F_\infty(d) > 0, \quad \sum_{i=1}^{p} (-g_{ij})_\infty (d_i) > 0, \quad \forall j = \overline{1, m}, \; \forall d \neq 0. \tag{3.31}$$

The next proposition shows that the minimization subproblems are solvable.

Proposition 3.5. *If X^* is nonempty and bounded, then for any $u \in \mathbb{R}^{mp}_{++}$, $y \in \mathbb{R}^{pm}$ and $\lambda > 0$*

$$\operatorname{Arg\,min}\{\Theta_\lambda(x,y,u) : x \in \mathbb{R}^n\} \neq \emptyset \text{ and bounded.} \tag{3.32}$$

Proof. To this goal, we need to show that $(\Theta_\lambda)_\infty(d) > 0$ for any $d \in \mathbb{R}^n$, $d \neq 0$. According to Proposition 2.1 in [40] and Proposition 2.6.1 in [41], we can express the recession function of Θ_λ as follows

$$(\Theta_\lambda)_\infty(d) = F_\infty(d) + \sum_{j=1}^{m}\sum_{i=1}^{p}\left(\frac{u_j^k}{\lambda_j}\psi\left(\lambda_j\left(g_{ij}(x_i) + y_{ij}\right)\right)\right)_\infty (d_i), \tag{3.33}$$

where $\psi(t) = -\varphi(-t)$.

If, we denote $I_+ = \{(i,j) : (-g_{ij})_\infty(d_i) > 0\}$, $I_- = \{(i,j) : (-g_{ij})_\infty(d_i) \leq 0\}$, then

$$(\Theta_\lambda)_\infty(d) = F_\infty(d) + \sum_{(i,j)\in I_+}^{m} \frac{u_j^k}{\lambda_j}\psi_\infty\left(\lambda_j(-g_{ij})_\infty(d_i)\right) + \sum_{(i,j)\in I_-}^{m} \frac{u_j^k}{\lambda_j}\psi_\infty\left(\lambda_j(-g_{ij})_\infty(d_i)\right) \tag{3.34}$$

and by using

$$\psi_\infty(s) = \begin{cases} s\psi_\infty(1) & \text{if } s > 0, \\ -s\psi_\infty(-1) & \text{if } s \leq 0, \end{cases} \tag{3.35}$$

we have

$$(\Theta_\lambda)_\infty(d) = F_\infty(d) + \psi_\infty(1)\sum_{(i,j)\in I_+} u_j^k(-g_{ij})_\infty(d_i) - \psi_\infty(-1)\sum_{(i,j)\in I_-} u_j^k(-g_{ij})_\infty(d_i). \tag{3.36}$$

Now, since $\psi_\infty(-1) = (-\varphi)_\infty(1) = 0$, the above relation becomes

$$(\Theta_\lambda)_\infty(d) = F_\infty(d) + \psi_\infty(1)\sum_{(i,j)\in I_+} u_j^k(-g_{ij})_\infty(d_i), \tag{3.37}$$

and finally, using (3.31) and (P6) the proof is complete. □

4. Convergence Analysis

In this section, we present the convergence analysis of the sequence (x^k, y^k, u^k) for a wide class of constraint transformations $\varphi \in \Phi$ under some assumptions on the input data. To this goal, we give the following main two propositions.

Proposition 4.1. *Under A_1 and A_2, the dual sequence $\{u^k\}$ is bounded.*

Proof. Let $D_\lambda(u') = \min_{x,y}\{\Theta_\lambda(x,y,u') : \sum_i y_{ij} = 0.\}$ and let the vector π such that $\pi_{ij} \leq g_{ij}(x_i) + y_{ij}$, for all $i = \overline{1,p}$, $j = \overline{1,m}$. Then, it is easy to see that

$$D_\lambda(u') = \min_\pi \left\{ \widetilde{P}(\pi) - \sum_{j=1}^{m} \frac{1}{\lambda_j} \sum_{i=1}^{p} u'_{ij}\varphi(\lambda\pi_{ij}) \right\}, \tag{4.1}$$

where $\widetilde{P}$ denotes the perturbation function associated to $([SP(y,\varphi)])$ and defined by

$$\widetilde{P}(\pi) = \min_{x,y}\{F(x) : g_{ij}(x_i) + y_{ij} \geq \pi_{ij},\ y \in A\}. \tag{4.2}$$

So, by adding and subtracting the term $\sum_{j=1}^{m}(1/\lambda_j)w'_{ij}\varphi[\lambda\pi_{ij}]$, for w' with the same structure as u', and if we set $B = \{(x,y) : g_{ij}(x_i) + y_{ij} \geq \pi_{ij}, y \in A\}$, we obtain

$$D_\lambda(u') = \min_\pi \left\{ \min_{(x,y)\in B} F(x) - \sum_{i,j} \frac{1}{\lambda_j} w'_{ij}\varphi(\lambda\pi_{ij}) + \sum_{i,j} \frac{1}{\lambda_j} w'_{ij}\varphi(\lambda\pi_{ij}) - \sum_{i,j} \frac{1}{\lambda_j} u'_{ij}\varphi(\lambda\pi_{ij}) \right\}. \tag{4.3}$$

That is,

$$D_\lambda(u') = \min_{\pi,(x,y)\in B} \left\{ F(x) - \sum_{j=1}^{m} \frac{1}{\lambda_j} \sum_{i=1}^{p} w'_{ij}\varphi(\lambda\pi_{ij}) + \sum_{j=1}^{m} \frac{1}{\lambda_j} \sum_{i=1}^{p} \left[w'_{ij} - u'_{ij}\right]\varphi(\lambda\pi_{ij}) \right\} \tag{4.4}$$

which can be rewritten also as follows:

$$D_\lambda(u') = \min_\pi \left\{ \widetilde{P}(\pi) - \sum_{j=1}^{m} \frac{1}{\lambda_j} \sum_{i=1}^{p} w'_{ij}\varphi(\lambda\pi_{ij}) + \sum_{j=1}^{m} \frac{1}{\lambda_j} \sum_{i=1}^{p} \left(w'_{ij} - u'_{ij}\right)\varphi(\lambda\pi_{ij}) \right\} \tag{4.5}$$

and then we have

$$D_\lambda(u') \geq D_\lambda(w') + \min_\pi \sum_{j=1}^{m} \frac{1}{\lambda_j} \sum_{i=1}^{p} \left(w'_{ij} - u'_{ij}\right)\varphi(\lambda\pi_{ij}). \tag{4.6}$$

Using (P1), (P3), and (P4) and if we take $w' = u^k$, $u' = u^{k+1}$, we can show that when x^{k+1} is not feasible, $u^{k+1} \geq u^k$ and it is easy to see that the minimum in this case equals zero and we have $D_\lambda(u^{k+1}) \geq D_\lambda(u^k) \geq \cdots \geq D_\lambda(u^0)$, for all $k \geq 0$. Thus, using the fact that D_λ is concave, L^* is bounded, and u^{k+1} is in the dual level set, the sequence $\{u^k\}$ is bounded. □

Proposition 4.2. *Under A_1, A_2, the primal sequence $\{x^k\}$ is bounded.*

Proof. Let $(x^{k+1}, y^{k+1}) \in \operatorname{Arg\,min}_x\{\Theta_{\lambda_k}(x, y, u^k) : \sum_i y_{ij} = 0\}$ and fix $\overline{x} \in X^*$ and $\overline{y}_{ij} = -g_{ij}(\overline{x_i}) + p^{-1}\sum g_{ij}(\overline{x_i})$. It is clear that

$$F\left(x^{k+1}\right) - \sum_j \sum_i \frac{u_j^k}{\lambda_j^k}\varphi\left[\lambda_j^k\left(g_{ij}\left(x_i^{k+1}\right) + y_{ij}^{k+1}\right)\right] \leq \Theta_{\lambda_k(\overline{x},\overline{y},u^k)}, \tag{4.7}$$

and using the feasibility of $\overline{x}$, we obtain directly

$$F\left(x^{k+1}\right) - \sum_j \sum_i \frac{u_j^k}{\lambda_j^k}\varphi\left[\lambda_j^k\left(g_{ij}\left(x_i^{k+1}\right) + y_{ij}^{k+1}\right)\right] \leq F(\overline{x}). \tag{4.8}$$

Now, let us assume that the primal sequence is unbounded, then the sequence $w_k = x^k/\|x^k\|$ is bounded and $\lim_{k\to\infty} w_k = d \neq 0$.

Let ϵ_1 and ϵ_2 such that

$$F_\infty(d) > \epsilon_1, \quad \left(\sum_i -g_{ij}\right)_\infty (d) > \epsilon_2, \tag{4.9}$$

and take k_0 such that

$$F\left(x^{k+1}\right) \geq \epsilon_1 \left\|x^k\right\|, \quad \forall k \geq k_0, \tag{4.10}$$

$$\sum_i (-g_{ij})\left(x_i^{k+1}\right) > \epsilon_2 \left\|x^{k+1}\right\|, \quad \forall k \geq k_0. \tag{4.11}$$

By dividing both sides in (4.8) by $\|x^{k+1}\|$, we get

$$\frac{F(x^{k+1})}{\|x^{k+1}\|} - \sum_{i,j} \frac{u_j^k}{\lambda_j^k} \frac{\varphi\left[\lambda_j^k\left(g_{ij}\left(x_i^{k+1}\right) + y_{ij}^{k+1}\right)\right]}{\|x^{k+1}\|} \leq \frac{F(\overline{x})}{\|x^{k+1}\|}, \tag{4.12}$$

and according to (4.11) and the monotonicity of φ, and by denoting $\nu_k = \lambda_j^k \|x^{k+1}\|/p$, (4.12) becomes

$$\frac{F(x^{k+1})}{\|x^{k+1}\|} - \sum_{i,j} \frac{u_j^k}{\lambda_j^k \|x^{k+1}\|}\varphi[-\epsilon_2 \nu_k] \leq \frac{F(\overline{x})}{\|x^{k+1}\|}. \tag{4.13}$$

Since the dual sequence is bounded, at the limit we have

$$\epsilon_1 - \sum_{i,j} \frac{\overline{u_j}}{p} \lim_{k\to\infty} \frac{\varphi(-\epsilon_2 \nu_k)}{\nu_k} \leq 0. \tag{4.14}$$

Since $\lim_{k\to\infty} \nu_k = +\infty$ and using $\varphi(t) = -\psi(-t)$, we can rewrite (4.14) as follows:

$$\epsilon_1 + \sum_{i,j} \frac{\overline{u_j}}{p} \lim_{k\to\infty} \frac{\psi(\epsilon_2 \nu_k)}{\nu_k} \leq 0. \tag{4.15}$$

Since

$$\lim_{k\to\infty} \frac{\psi(\epsilon_2 \nu_k)}{\nu_k} = \lim_{t\to\infty} \frac{\psi(\epsilon_2 t)}{t} = \psi_\infty(\epsilon_2), \tag{4.16}$$

and (4.15) is equivalent to

$$\epsilon_1 + \sum_j \overline{u_j} \psi_\infty(\epsilon_2) \leq 0. \tag{4.17}$$

If $\epsilon_2 \leq 0$ then $\psi_\infty(\epsilon_2) = -\epsilon_2 \psi_\infty(-1) = 0$ and then $\epsilon_1 \leq 0$.

If $\epsilon_2 > 0$ then $\psi_\infty(\epsilon_2) = \epsilon_2 \psi_\infty(1)$.

Now by letting $\epsilon_1 \to F_\infty(d)$ and $\epsilon_2 \to (\sum_i (-g_{ij})_\infty(d))$, we deduct that

$$F_\infty(d) \leq 0, \quad \sum_i (-g_{ij})_\infty(d) \leq 0, \tag{4.18}$$

and since X^* is not empty and is bounded, then $d = 0$ which contradicts the fact that $d \neq 0$. Thus, what we assumed is false, and the primal sequence $\{x^k\}_k$ is bounded. □

Proposition 4.3. *Let the sequences $\{x^k, y^k, u^k\}$ generated by ΨSALA and assume that there exist x^* a primal solution to the original problem (SP). Then the following inequality holds:*

$$\sum_{i=1}^{p} f_i\left(x_i^{k+1}\right) \leq \sum_{i=1}^{p} f_i(x_i^*) + \sum_{i=1}^{p} u^{k+1} \theta_i^{k+1} \left(g_i(x_i^*) - g_i\left(x_i^{k+1}\right)\right), \tag{4.19}$$

where

$$\theta_i^{k+1} = \Psi'\left[\lambda\left(g_i\left(x_i^{k+1}\right) + y_i^k\right)\right] \mid \Psi'\left[\lambda\left(g_i\left(x_i^{k+1}\right) + y_i^{k+1}\right)\right]. \tag{4.20}$$

Proof. See [30]. □

Proposition 4.4. *Let the sequences $\{x^k, y^k, u^k\}$generated by ΨSALA and assume that there exists a saddle point (x^*, y^*, u^*) of the Lagrangian $l(x, y, u)$ associated to problem (SP). Then the following inequality holds:*

$$\sum_{i=1}^{p} f_i(x_i^*) \leq \sum_{i=1}^{p} f_i\left(x_i^{k+1}\right) + \left\langle u^*, \delta^{k+1} \right\rangle. \tag{4.21}$$

Proof. See [30]. □

Proposition 4.5. *All the three sequences $\{x^k\}$, $\{u^k\}$, and $\{y^k\}$ generated in Algorithm 2.1 are bounded.*

Theorem 4.6. *Let $\overline{y}$ and $\overline{u}$, the respective limit point of the bounded sequences $\{y^k\}$ and $\{u^k\}$ generated by ΨSALA. Then one has the following properties:*

(i) $\limsup_{k\to\infty} \sum_{i=1}^{p} g_i(x_i^{k+1}) \geq 0$;

(ii) $\limsup_{k\to\infty} (g_i(x_i^{k+1}) + y_i^{k+1}) \geq 0$;

(iii) $\lim_{k\to\infty} \langle u^{k+1}, g_i(x_i^{k+1}) + y_i^{k+1}\rangle = 0$.

Proof. See [30]. □

Theorem 4.7. *If the assumptions A_1, A_2, and A_3 are satisfied, then one has the following.*

(1) *Any limit point (x, u) of the sequence $\{(x^k, u^k)\}$ is in the set $X^* \times L^*$.*

(2) *The sequences $f(x^k)$ and $d(u^k)$ are convergent and*

$$\lim_{s\to\infty} f\left(x^k\right) = f(x^*), \qquad \lim_{s\to\infty} d\left(u^k\right) = d(u^*). \tag{4.22}$$

Proof. Let $\overline{u}$ be any limit point of the sequence $\{u^k\}$, then there exists a subsequence converging to $\overline{u}$. Without restricting the generality, we can assume $u^k \to \overline{u}$. For j such that $\overline{u}_{ij} > 0$, then

$$\lim_{s\to\infty} \frac{u_{ij}^{k+1}}{u_{ij}^{k}} = 1, \tag{4.23}$$

and since $u_{ij}^{k+1} = u_{ij}^{k}\varphi'((\lambda_{ij}^{k}/p)\sum_{i=1}^{p} g_{ij}(x_i^k))$, $i = \overline{1,p}$, $j = \overline{1,m}$, then

$$\lim_{s\to\infty} \varphi'\left(\frac{\lambda_{ij}^{k}}{p}\sum_{i=1}^{p} g_{ij}\left(x_i^k\right)\right) = 1. \tag{4.24}$$

Since $\lim_{s\to\infty} u_{ij}^k = \overline{u}_{ij} > 0$, we have $\lambda_{ij}^k u_{ij}^k = \lambda > 0$. Therefore, $\lambda_{ij}^k = (\lambda/u_{ij}^k) > 0$. Then,

$$\lim_{s\to\infty} \lambda_{ij} \geq \frac{\lambda}{\overline{u}_{ij}} > 0 \tag{4.25}$$

and then,

$$\lim_{s\to\infty} \sum_{i=1}^{p} g_{ij}\left(x_i^k\right) = 0; \tag{4.26}$$

therefore,

$$\lim_{s\to\infty}\left(u_{ij}^{k+1}\sum_{i=1}^{p} g_{ij}\left(x_i^{k+1}\right)\right)=0. \tag{4.27}$$

Now, we will prove that the set $\{\sum_{i=1}^{p} g_{ij}(x_i^k) \mid j=\overline{1,m}, s\in\mathbb{N}\}$ is bounded. Since $\{x^k\}$ is bounded, there exists $\delta\in\mathbb{R}$ such that $\|x^k\|\le\delta$. Let $\overline{B}(0,\delta)\subseteq\mathbb{R}^n$ be the closed ball of center zero and radius δ then

$$x^k\in\overline{B}(0,\delta),\quad \forall k\in\mathbb{N}. \tag{4.28}$$

Since $C_j(x)=\sum_{i=1}^{p} g_{ij}(x_i)$ is continuous for all $j=\overline{1,m}$, then, $C_j(\overline{B}(0,\delta))$ is closed and bounded for all $j=\overline{1,m}$. Then there exists $M_j\ge 0$ such that

$$\left|\sum_{i=1}^{p} g_{ij}(x_i)\right|\le M_j,\quad \forall x\in\overline{B}(0,\delta),\ \forall j=\overline{1,m}. \tag{4.29}$$

Let $\rho=\max\{M_j : j=\overline{1,m}\}$, then

$$\left|\sum_{i=1}^{p} g_{ij}(x_i)\right|\le\rho,\quad \forall x\in\overline{B}(0,\delta),\ \forall j=\overline{1,m}. \tag{4.30}$$

Therefore, $|\sum_{i=1}^{p} g_{ij}(x_i^k)|\le\rho$, for all $j=\overline{1,m}$, for all $k\in\mathbb{N}$, which means $\{\sum_{i=1}^{p} g_{ij}(x_i^k) \mid j=\overline{1,m}, s\in\mathbb{N}\}$ is a bounded set.

Thus, for j such that $\overline{u}_{ij}=0$, and the above result, we get

$$\lim_{s\to\infty}\left(u_{ij}^{k+1}\sum_{i=1}^{p} g_{ij}\left(x_i^{k+1}\right)\right)=0. \tag{4.31}$$

That is,

$$\lim_{s\to\infty}\left(u_{ij}^{k+1}\sum_{i=1}^{p} g_{ij}\left(x_i^{k+1}\right)\right)=0\quad \forall j=\overline{1,m}. \tag{4.32}$$

Also, if $\overline{x}$ is a limit point of $\{x^k\}$, keeping in mind that (see the proof given by Polyak in [38]) $\max\{-g_{ij}(x_i^k\}\to 0$, then

$$\lim_{s\to\infty}\sum_{i=1}^{p} g_{ij}\left(x_i^{k+1}\right)\ge 0\quad \forall j=\overline{1,m}. \tag{4.33}$$

On another hand, from (4.32) we get

$$\begin{aligned}
&\nabla_x \overline{L}(\overline{x},\overline{u}) = 0 \\
&\Longrightarrow \nabla F(\overline{x}) - \sum_{j=1}^{m}\sum_{i=1}^{p} \overline{u}_{ij} \nabla g_{ij}(\overline{x}_i) = 0,
\end{aligned} \tag{4.34}$$

$$\begin{aligned}
&\overline{u}_{ij} \geq 0, \quad j = \overline{1,m}, \\
&\overline{u}_{ij} \sum_{i=1}^{p} g_{ij}(\overline{x}_i) = 0, \quad \text{from (4.32)}, \\
&\sum_{i=1}^{p} g_{ij}(\overline{x}_i) \geq 0, \quad \text{from (4.33)}.
\end{aligned} \tag{4.35}$$

Since the Lagrange of problem (SP) is

$$\overline{L}(x,u) = F(x) - \sum_{j=1}^{m}\sum_{i=1}^{p} u_{ij} g_{ij}(x_i) \tag{4.36}$$

and it is convex, then from (4.34)

$$\overline{x} \in \operatorname{Arg}\min_{x\in\mathbb{R}^n} \overline{L}(x,\overline{u}) \tag{4.37}$$

and then

$$F(\overline{x}) = F(\overline{x}) - \sum_{j=1}^{m} \overline{u}_{ij} \left(\sum_{i=1}^{p} g_{ij}(\overline{x}_i) \right) = \overline{L}(\overline{x},\overline{u}) \leq \overline{L}(x,\overline{u}), \quad \forall x \in \mathbb{R}^n. \tag{4.38}$$

Also, we know that $\sum_{i=1}^{p} g_{ij}(\overline{x}_i) \geq 0$, for all $j = \overline{1,m}$, and for any $u_{ij} \geq 0$: $-\sum_{j=1}^{m}\sum_{i=1}^{p} u_{ij} g_{ij}(\overline{x}_i) \leq 0$ which implies

$$\Longrightarrow \overline{L}(\overline{x},u) = F(\overline{x}) - \sum_{j=1}^{m}\sum_{i=1}^{p} u_{ij} g_{ij}(\overline{x}_i) \leq F(\overline{x}) = \overline{L}(\overline{x},\overline{u}) \tag{4.39}$$

and then

$$\overline{L}(\overline{x},u) \leq \overline{L}(\overline{x},\overline{u}) \leq \overline{L}(x,\overline{u}) \tag{4.40}$$

for all $x \in \mathbb{R}^n$, for all $u \geq 0$, and by the saddle point theorem we have

$$(\overline{x},\overline{u}) \in X^* \times L^*. \tag{4.41}$$

Since $\{d(u^k)\}$ is an increasing sequence bounded above by $d(u^*)$ then it is convergent. Let $\{(x^{k_l}, u^{k_l})\}$ be any convergent subsequence of $\{(x^k, u^k)\}$. Then from (3.4)

$$\lim_{l\to\infty} L\left(x^{k_l}, u^{k_l}\right) = F(x^*) = d(u^*). \tag{4.42}$$

Since $\nabla_x \overline{L}(x^k, u^k) = 0$

$$\begin{aligned}
&\Longrightarrow \min_{x\in\mathbb{R}^n} \overline{L}\left(x, u^k\right) = \overline{L}\left(x^k, u^k\right) \\
&\Longrightarrow d\left(u^k\right) = \overline{L}\left(x^k, u^k\right) \\
&\Longrightarrow \lim_{l\to\infty} \overline{L}\left(x^{k_l}, u^{k_l}\right) = \lim_{l\to\infty} d\left(u^{k_l}\right) = d(u^*).
\end{aligned} \tag{4.43}$$

Then $\{d(u^k)\}$ converges to $d(u^*)$. The remaining proof is similar to the one given in [30]. □

In the previous theorem, we proved the boundedness of the primal and dual sequences, but for the sequence of allocation vectors $\{y^k\}$ we can get the boundedness directly by proving that the set

$$\left\{g_{ij}\left(x_i^k\right) \mid j = \overline{1,m}, i = \overline{1,p},\ k \in \mathbb{N}\right\} \tag{4.44}$$

is bounded. (The same way as we proved that the set $\{\sum_{i=1}^{p} g_{ij}(x_i^k) \mid j = \overline{1,m}, k \in \mathbb{N}\}$ is bounded.).

References

[1] L. S. Lasdon, *Optimization Theory for Large Systems*, The Macmillian, 1970.

[2] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*, Prentice-Hall, Upper Saddle River, NJ, USA, 1989.

[3] J. E. Eckstein, “Nonlinear proximal point algorithms using Bregman functions, with applications to convex programming,” Working Paper 91-004, Harvard Business School, 1990.

[4] R. Polyak and M. Teboulle, “Nonlinear rescaling and proximal-like methods in convex optimization,” *Mathematical Programming*, vol. 76, no. 2, pp. 265–284, 1997.

[5] J. E. Spingarn, “Applications of the method of partial inverses to convex programming: decomposition,” *Mathematical Programming*, vol. 32, no. 2, pp. 199–223, 1985.

[6] R. S. Burachik, C. Sagastizábal, and S. Scheimberg, “An inexact method of partial inverses and a parallel bundle method,” *Optimization Methods & Software*, vol. 21, no. 3, pp. 385–400, 2006.

[7] J. Eckstein, “Parallel alternating direction multiplier decomposition of convex programs,” *Journal of Optimization Theory and Applications*, vol. 80, no. 1, pp. 39–62, 1994.

[8] M. Fukushima, “Application of the alternating direction method of multipliers to separable convex programming problems,” *Computational Optimization and Applications*, vol. 1, no. 1, pp. 93–111, 1992.

[9] D. Gabay and B. Mercier, “A dual algorithm for the solution of nonlinear variational problems via finite element approximation,” *Computers and Mathematics with Applications*, vol. 2, no. 1, pp. 17–40, 1976.

[10] A. Hamdi, P. Mahey, and J. P. Dussault, “A new decomposition method in nonconvex programming via a separable augmented Lagrangian,” in *Recent Advances in Optimization*, P. Gritzmann, R. Horst, E. Sachs, and R. Tichatschke, Eds., vol. 452 of *Lecture Notes in Economics and Mathematical Systems*, pp. 90–104, Springer, Berlin, Germany, 1997.

[11] A. Hamdi and P. Mahey, "Separable diagonalized multiplier method for decomposing nonlinear programs," *Computational & Applied Mathematics*, vol. 19, no. 1, pp. 1–29, 2000.

[12] A. Hamdi, "Two-level primal-dual proximal decomposition technique to solve large scale optimization problems," *Applied Mathematics and Computation*, vol. 160, no. 3, pp. 921–938, 2005.

[13] P. Mahey, S. Oualibouch, and Pham Dinh Tao, "Proximal decomposition on the graph of a maximal monotone operator," *SIAM Journal on Optimization*, vol. 5, no. 2, pp. 454–466, 1995.

[14] D. P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press, 1982.

[15] R. T. Rockafellar, "Augmented Lagrangians and applications of the proximal point algorithm in convex programming," *Mathematics of Operations Research*, vol. 1, no. 2, pp. 97–116, 1976.

[16] A. N. Iusem and M. Teboulle, "Convergence rate analysis of nonquadratic proximal methods for convex and linear programming," *Mathematics of Operations Research*, vol. 20, no. 3, pp. 657–677, 1995.

[17] K. C. Kiwiel, "Proximal minimization methods with generalized Bregman functions," *SIAM Journal on Control and Optimization*, vol. 35, no. 4, pp. 1142–1168, 1997.

[18] R. A. Polyak, "Log-sigmoid multipliers method in constrained optimization," *Annals of Operations Research*, vol. 101, pp. 427–460, 2001.

[19] M. Teboulle, "Entropic proximal mappings with applications to nonlinear programming," *Mathematics of Operations Research*, vol. 17, no. 3, pp. 670–690, 1992.

[20] M. Teboulle, "Convergence of proximal-like algorithms," *SIAM Journal on Optimization*, vol. 7, no. 4, pp. 1069–1083, 1997.

[21] O. M. Guèye, J.-P. Dussault, and P. Mahey, "Separable augmented Lagrangian algorithm with multidimensional scaling for monotropic programming," *Journal of Optimization Theory and Applications*, vol. 127, no. 2, pp. 329–345, 2005.

[22] A. Auslender and M. Teboulle, "Entropic proximal decomposition methods for convex programs and variational inequalities," *Mathematical Programming*, vol. 91, no. 1, pp. 33–47, 2001.

[23] A. Auslender, M. Teboulle, and S. Ben-Tiba, "A logarithmic-quadratic proximal method for variational inequalities," *Computational Optimization and Applications*, vol. 12, no. 1–3, pp. 31–40, 1999.

[24] A. Auslender, M. Teboulle, and S. Ben-Tiba, "Interior proximal and multiplier methods based on second order homogeneous kernels," *Mathematics of Operations Research*, vol. 24, no. 3, pp. 645–668, 1999.

[25] G. Chen and M. Teboulle, "A proximal-based decomposition method for convex minimization problems," *Mathematical Programming*, vol. 64, no. 1, pp. 81–101, 1994.

[26] M. Kyono and M. Fukushima, "Nonlinear proximal decomposition method for convex programming," *Journal of Optimization Theory and Applications*, vol. 106, no. 2, pp. 357–372, 2000.

[27] M. Aslam Noor, "Some developments in general variational inequalities," *Applied Mathematics and Computation*, vol. 152, no. 1, pp. 199–277, 2004.

[28] M. A. Noor, "Extended general variational inequalities," *Applied Mathematics Letters*, vol. 22, no. 2, pp. 182–186, 2009.

[29] M. A. Noor, K. I. Noor, and E. Al-said, "Iterative methods for solving nonconvex equilibrium problems," *Applied Mathematics and Information Science*, vol. 6, no. 1, pp. 65–69, 2012.

[30] A. Hamdi, "Decomposition for structured convex programs with smooth multiplier methods," *Applied Mathematics and Computation*, vol. 169, no. 1, pp. 218–241, 2005.

[31] P. Mahey, A. Ouorou, L. LeBlanc, and J. Chifflet, "A new proximal decomposition algorithm for routing in telecommunication networks," *Networks*, vol. 31, no. 4, pp. 227–238, 1998.

[32] J. M. Mulvey and A. Ruszczyński, "A diagonal quadratic approximation method for large-scale linear programs," *Operations Research Letters*, vol. 12, no. 4, pp. 205–215, 1992.

[33] G. Stephanopoulos and A. W. Westerberg, "The use of Hestenes' method of multipliers to resolve dual gaps in engineering system optimization," *Journal of Optimization Theory and Applications*, vol. 15, pp. 285–309, 1975.

[34] A. Tanikawa and H. Mukai, "A new technique for nonconvex primal-dual decomposition of a large-scale separable optimization problem," *IEEE on Transactions on Automatic Control*, vol. 30, no. 2, pp. 133–143, 1985.

[35] A. Hamdi, "Decomposition methods based on augmented lagrangians: a survey," Technical Report, Kuwait University-Faculty of Science, 2006.

[36] A. Ben-Tal and M. Zibulevsky, "Penalty/barrier multiplier methods for convex programming problems," *SIAM Journal on Optimization*, vol. 7, no. 2, pp. 347–366, 1997.

[37] R. Polyak and I. Griva, "Primal-dual nonlinear rescaling method for convex optimization," *Journal of Optimization Theory and Applications*, vol. 122, no. 1, pp. 111–156, 2004.

[38] R. Polyak, "Nonlinear Rescaling as Interior Quadratic Prox Method in Convex Optimization," *Journal Computational Optimization and Applications* , vol. 35, no. 3, pp. 347–373, 2006.

[39] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, USA, 1970.

[40] A. Auslender, R. Cominetti, and M. Haddou, "Asymptotic analysis for penalty and barrier methods in convex and linear programming," *Mathematics of Operations Research*, vol. 22, no. 1, pp. 43–62, 1997.

[41] A. Auslender and M. Teboulle, *Asymptotic Cones and Functions in Optimization and Variational Inequalities*, Springer, 2003.