fractal and fractional

Article

An Active-Set Fischer–Burmeister Trust-Region Algorithm to Solve a Nonlinear Bilevel Optimization Problem

Bothina Elsobky * and Gehan Ashry

Department of Mathematics, Faculty of Science, Alexandria University, Alexandria 5424041, Egypt; gehan.ashry@alexu.edu.eg
* Correspondence: bothina_elsobky@alexu.edu.eg

Abstract: In this paper, the Fischer–Burmeister active-set trust-region (FBACTR) algorithm is introduced to solve the nonlinear bilevel programming problems. In FBACTR algorithm, a Karush–Kuhn–Tucker (KKT) condition is used with the Fischer–Burmeister function to transform a nonlinear bilevel programming (NBLP) problem into an equivalent smooth single objective nonlinear programming problem. To ensure global convergence for the FBACTR algorithm, an active-set strategy is used with a trust-region globalization strategy. The theory of global convergence for the FBACTR algorithm is presented. To clarify the effectiveness of the proposed FBACTR algorithm, applications of mathematical programs with equilibrium constraints are tested.

Keywords: a bilevel optimization problem; Fischer–Burmeister function; a Karush–Kuhn–Tucker conditions; active-set strategy; trust-region strategy; global convergence

MSC: 65Dxx; 65Kxx; 65Zxx

Citation: Elsobky, B.; Ashry, G. An Active-Set Fischer–Burmeister Trust-Region Algorithm to Solve a Nonlinear Bilevel Optimization Problem. *Fractal Fract.* **2022**, *6*, 412. https://doi.org/10.3390/fractalfract6080412

Academic Editor: Yongguang Yu

Received: 18 May 2022
Accepted: 18 July 2022
Published: 27 July 2022

1. Introduction

The mathematical formulation for NBLP problem which we will consider it is

$$\begin{array}{ll} \min_v & f_u(v,w) \\ s.t. & g_u(v,w) \leq 0, \\ \quad \min_w & f_l(v,w), \\ \quad s.t. & g_l(v,w) \leq 0, \end{array} \tag{1}$$

where $v \in \Re^{n_1}$ and $w \in \Re^{n_2}$. In our approach, the functions $f_u : \Re^{n_1+n_2} \to \Re$, $f_l : \Re^{n_1+n_2} \to \Re$, $g_u : \Re^{n_1+n_2} \to \Re^{m_1}$, and $g_l : \Re^{n_1+n_2} \to \Re^{m_2}$ must have a twice continuously differentiable function at least.

The NBLP problem (1) is utilized so extensively in transaction network, resource allocation, finance budget, price control, etc., see [1–4]. The NBLP problem (1) has two levels of optimization problems, upper and lower levels. A decision maker with the upper level objective function $f_u(v,w)$ takes the lead, and so he chooses the decision vector v. According to this, the decision maker with lower level objective function $f_l(v,w)$, chooses the decision vector w to optimize her objective, parameterized in v.

To obtain the solution of problem (1), number of different approaches have been offered, see (1), see [5–9]. In our method, we utilize one of these approaches to transforme NBLP problem (1) to a single level one by replacing the lower level optimization problem with its KKT conditions, see [10,11].

Utilizing KKT optimality conditions for the lower level problem, the NBLP problem (1) is reduced to the following single-objective optimization problem:

$$\begin{array}{ll} \min_{v,w} & f_u(v,w) \\ s.t. & g_u(v,w) \leq 0, \\ & \nabla_w f_l(v,w) + \nabla_w g_l(v,w)\lambda = 0, \\ & g_l(v,w) \leq 0, \\ & \lambda_j g_{l_j}(v,w) = 0, \;\; j = 1,\ldots,m_2, \\ & \lambda_j \geq 0, \;\; j = 1,\ldots,m_2, \end{array} \tag{2}$$

where $\lambda \in \Re^{m_2}$ a multiplier vector which is associated with inequality constraint $g_l(v,w)$.

Problem (2) is non-differentiable and non-convex. Furthermore, the regularity assumption prerequisites to successfully handle smooth optimization problems are never satisfied. Following the smoothing method which is proposed by [2], we can introduce the FBACTR algorithm to solve the problem (2). Before introducing FBACTR algorithm, we need the following definition.

Definition 1. *A Fischer–Burmeister function is the function $\Psi(e,d) : \Re^2 \to \Re$ and it is defined by $\Psi(e,d) = e + d - \sqrt{e^2 + d^2}$. A perturbed Fischer–Burmeister function is the function $\psi(e,d,\hat{\varepsilon}) : \Re^3 \to \Re$ and it is defined by $\psi(e,d,\hat{\varepsilon}) = e + d - \sqrt{e^2 + d^2 + \hat{\varepsilon}}$.*

The Fischer–Burmeister function has the property that $\Psi(e,d) = 0$ if and only if $e \geq 0$, $d \geq 0$, and $ed = 0$. It is non-differentiable at $e = d = 0$. Its perturbed variant satisfies $\psi(e,d,\hat{\varepsilon}) = 0$ if and only if $e > 0$, $d > 0$, and $ed = \frac{\hat{\varepsilon}}{2}$ for $\hat{\varepsilon} > 0$. This function is smooth with respect to e,d, for $\hat{\varepsilon} > 0$, and for more details see [12–15].

The next perturbed Fischer–Burmeister function is used to satisfy the asymptotic stability conditions, and allow the FBACTR algorithm to solve problem (2).

$$\psi(e,d,\hat{\varepsilon}) = \sqrt{e^2 + d^2 + \hat{\varepsilon}} - e - d. \tag{3}$$

Using the perturbed Fischer–Burmeister function (3), problem (2) can be approximated by:

$$\begin{array}{ll} \min_{v,w} & f_u(v,w) \\ s.t. & g_u(v,w) \leq 0, \\ & \nabla_w f_l(v,w) + \nabla_w g_l(v,w)\lambda = 0, \\ & \sqrt{g_{l_j}^2 + \lambda_j^2 + \hat{\varepsilon}} - \lambda_j + g_{l_j} = 0, \;\; j = 1,\ldots,m_2. \end{array} \tag{4}$$

The following notations are introduced to simplify our discussion. These notations are $n = n_1 + n_2 + m_2$, $x = (v,w,\lambda)^T \in \Re^n$ and $c(x) = (\nabla_w f_l(v,w) + \nabla_w g_l(v,w)\lambda, \sqrt{g_{l_j}^2 + \lambda_j^2 + \hat{\varepsilon}} - \lambda_j + g_{l_j})^T, j = 1,\ldots,m_2$. Hence problem (4) can be reduced as follows:

$$\begin{array}{ll} \text{minimize} & f_u(x) \\ \text{subject to} & g_u(x) \leq 0, \\ & c(x) = 0, \end{array} \tag{5}$$

where $f_u : \Re^n \to \Re$, $g_u : \Re^n \to \Re^{m_1}$, and $c : \Re^n \to \Re^{n_2+m_2}$.

A set of indices of binding or violated inequality constraints at x is defined by $I(x) = \{i : g_{u_i}(x) \geq 0\}$. A regular point is the point x_* at which the vectors of the set $\{\nabla c_i(x_*), i = 1,2,\ldots,n_2+m_2\} \bigcup \{\nabla g_{u_i}(x_*), i \in I(x_*)\}$ are linearly independent.

A regular point x_* is KKT point of problem (5) if there exist Lagrange multiplier vectors $\mu_* \in \Re^{n_2+m_2}$ and $\lambda_* \in \Re^{m_1}$ such that the following KKT conditions hold:

$$\begin{aligned} \nabla f_u(x_*) + \nabla c(x_*)\mu_* + \nabla g_u(x_*)\lambda_* &= 0, &(6)\\ c(x_*) &= 0, &(7)\\ g_u(x_*) &\leq 0, &(8)\\ (\lambda_*)_i g_{u_i}(x_*) &= 0, \quad i = 1, \ldots, m_1, &(9)\\ (\lambda_*)_i &\geq 0, \quad i = 1, \ldots, m_1. &(10) \end{aligned}$$

To solve the nonlinear single-objective constrained optimization problem (5), various approaches have been proposed; for more details, see [16–22].

An active-set strategy is utilized to reduce problem (5) to equality constrained optimization problem. The idea beyond the active-set method is to identify at every iteration, the active inequality constraints and treat them as equalities and this allows to utilize the improved methods which are used to solve the equality constrained problems, see [21,23,24]. Most of the methods that are used to solve the equality constrained problems, may not converge if the starting point is far away from the stationary point, so it is called a local method.

To ensure a convergence to the solution from any starting point, a trust-region strategy which is strongly global convergence can be induced. It is very important strategy to solve a smooth optimization. It is more robust when it deals with rounding errors. It does not require the objective function of the model be convex. For more details see [11,21–32].

To treat the difficult of having infeasible trust-region subproblem in FBACTR algorithm, a reduced Hessian technique which is suggested by [33,34] and used by [22,24,35] is utilized.

Under five assumptions, a theory of global convergence for FBACTR algorithm is proved. Moreover, numerical experiments display that FBACTR algorithm performers effectively and efficiently in pursuance.

We shall use the following notation and terminology. We use $\|.\|$ to denote the Euclidean norm $\|.\|_2$. Subscript k refers to iteration indices. For example, $f_{u_k} \equiv f_u(x_k)$, $g_{u_k} \equiv g_u(x_k)$, $c_k \equiv c(x_k)$, $Y_k \equiv Y(x_k)$, $P_k \equiv P(x_k)$, $\nabla_x \ell_k \equiv \nabla_x \ell(x_k, \mu_k)$, and so on to denote the function value at a particular point.

The rest of the paper is organized as follows. Section 2 is devoted to the description of an active-set trust-region algorithm to solve problem (5) and summarized to FBACTR algorithm to solve NBLP problem (1) is introduced. In Section 3 the analysis of the theory of global convergence of the FBACTR algorithm is presented. Section 4 contains an implementation of the FBACTR algorithm and the results of test problems. Finally, some further remarks are given in Section 5.

2. Active-Set with Trust-Region Technique

A detailed description for active-set with the trust-region strategy to solve problem (5) and summarized to FBACTR algorithm to solve problem (1) are introduced in this section.

Based on the active-set method which is suggested by [36] and used with [21–24], we define a 0–1 diagonal matrix $P(x) \in \Re^{m_1 \times m_1}$, whose diagonal entries are:

$$p_i(x) = \begin{cases} 1 & \text{if } g_{u_i}(x) \geq 0, \\ 0 & \text{if } g_{u_i}(x) < 0. \end{cases} \tag{11}$$

Using the previous definition of the matrix $P(x)$, a smooth and simple function is utilized to replace problem (5) with the following simple problem

$$\begin{array}{ll} \text{minimize} & f_u(x) + \frac{r}{2}\|P(x)g_u(x)\|^2 \\ \text{subject to} & c(x) = 0, \end{array} \tag{12}$$

where $r > 0$ is a parameter, see [21–23]. The Lagrangian function associated with problem (12) is given by:

$$L(x, \mu; r) = \ell(x, \mu) + \frac{r}{2}\|P(x)g_u(x)\|^2, \tag{13}$$

where

$$\ell(x, \mu) = f_u(x) + \mu^T c(x), \tag{14}$$

and $\mu \in \Re^{n_2+m_2}$ represents a Lagrange multiplier vector which is associated with the constraint $c(x)$. A KKT point (x_*, μ_*) for problem (12) is the point at which the following conditions are satisfied

$$\nabla \ell(x_*, \mu_*) + r\nabla g_u(x_*)P(x_*)g_u(x_*) = 0, \tag{15}$$

$$h(x_*) = 0, \tag{16}$$

where $\nabla \ell(x_*, \mu_*) = \nabla f_u(x_*) + \nabla c(x_*)\mu_*$.

If the KKT point (x_*, μ_*) satisfies conditions (6)–(10), we notice that it is also satisfies conditions (15) and (16), but the converse is not necessarily true. So, we design FBACTR algorithm in a way that, if (x_*, μ_*) satisfies conditions (15) and (16), then it is also satisfies KKT conditions (6)–(10).

Various approaches which were proposed to solve the equality constrained are local methods. By local method, we mean a method such that if the starting point is sufficiently close to a solution, then under some reasonable assumptions the method is guaranteed by theory to converge to the solution. There is no guarantee that the local method converges starting from the remote. Globalizing a local method means modifying the method in such a way that is guaranteed to converge from any starting point without sacrificing its fast local rate of convergence. To ensure convergence from the remote, the trust-region technique is utilized.

2.1. A Trust-Region Technique

To solve problem (12) and to convergence from remote with any starting point, the trust-region strategy is used. A naive trust-region quadratic subproblem associated with problem (12) is:

$$\begin{array}{ll} \text{minimize} & q_k(s) = \ell_k + \nabla_x \ell_k^T s + \frac{1}{2} s^T H_k s + \frac{r}{2}\|P_k(g_{u_k} + \nabla g_{u_k})^T s)\|^2 \\ \text{subject to} & c_k + \nabla c_k^T s = 0, \\ & \|s\| \le \delta_k, \end{array} \tag{17}$$

where $0 < \delta_k$ represents the trust-region radius and H_k is the Hessian matrix of the Lagrangian function (14) or an approximation to it.

Subproblem (17) may be infeasible because there may be no intersecting points between hyperplane of the linearized constraints $c(x) + \nabla c(x)^T s = 0$ and the constraint $\|s\| \le \delta_k$. Even if they intersect, there is no guarantee that this will keep true if δ_k is reduced, see [37]. To overcome this problem, a reduced Hessian technique which was suggested by [33,34] and used by [22,23,35] is used. In this technique, to obtain the trial step s_k, it is decomposed into two orthogonal components: the tangential component s_k^t to improve optimality and the normal component s_k^n to improve feasibility. To evaluate each of s_k^n and s_k^t, two unconstrained trust-region subproblems are solved.

- **To obtain the normal component s^n**

To evaluate the normal component s_k^n, the following trust-region subproblem must be solved:

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\|c_k + \nabla c_k^T s^n\|^2 \\ \text{subject to} & \|s^n\| \le \zeta\delta_k, \end{array} \tag{18}$$

for some $\zeta \in (0, 1)$.

Any method can be used to solve subproblem (18), as long as a fraction of the normal predicted decrease obtained by the Cauchy step s_k^{ncp} is less than or equal to the normal predicted decrease obtained by s_k^n. That is, the following condition must be held:

$$\|c_k\|^2 - \|c_k + \nabla c_k^T s_k^n\|^2 \geq \vartheta_1 \{\|c_k\|^2 - \|c_k + \nabla c_k^T s_k^{ncp}\|^2\}, \tag{19}$$

for some $\vartheta_1 \in (0, 1]$. The normal Cauchy step s_k^{ncp} is given by:

$$s_k^{ncp} = -\tau_k^{ncp} \nabla c_k c_k, \tag{20}$$

where the parameter τ_k^{ncp} is given by:

$$\tau_k^{ncp} = \begin{cases} \frac{\|\nabla c_k c_k\|^2}{\|(\nabla c_k)^T \nabla c_k c_k\|^2} & \text{if } \frac{\|\nabla c_k c_k\|^3}{\|\nabla c_k^T \nabla c_k c_k)\|^2} \leq \delta_k \\ & \text{and } \|\nabla c_k^T \nabla c_k c_k)\| > 0, \\ \frac{\delta_k}{\|\nabla c_k c_k\|} & \text{otherwise.} \end{cases} \tag{21}$$

A dogleg method is used to solve subproblem (18). It is very cheap if the Hessian is indefinite. The dogleg algorithm approximates the solution curve to subproblem (18) by piecewise linear function connecting the Newton point to the Cauchy point. For more details, see [35].

Once s_k^n is estimated, we will compute $s_k^t = Y_k \bar{s}_k^t$. A matrix Y_k is the matrix whose columns form a basis for the null space of $(\nabla c_k)^T$.

- **To obtain the tangential component s_k^t.**

To evaluate the tangential component s_k^t, the following subproblem is solved by using the dogleg method

$$\begin{array}{ll} \text{minimize} & (Y_k^T \nabla q_k(s_k^n))^T \bar{s}^t + \frac{1}{2} \bar{s}^{t^T} Y_k^T B_k Y_k \bar{s}^t \\ \text{subject to} & \|Y_k \bar{s}^t\| \leq \Delta_k, \end{array} \tag{22}$$

where $\nabla q_k(s_k^n) = \nabla_x \ell_k + B_k s_k^n + r_k \nabla g_{u_k} P_k g_{u_k}$, $B_k = H_k + r_k \nabla g_{u_k} P_k \nabla g_{u_k}^T$, and $\Delta_k = \sqrt{\delta_k^2 - \|s_k^n\|^2}$.

Since the dogleg method is used to solve the above subproblem, then a fraction of the tangential predicted decrease obtained by a tangential Cauchy step $\bar{s}_k^{tcp}$ is less than or equal to the tangential predicted decrease which is obtained by tangential step $\bar{s}_k^t$. That is, the following conditions hold

$$q_k(s_k^n) - q_k(s_k^n + Y_k \bar{s}_k^t) \geq \vartheta_2 [q_k(s_k^n) - q_k(s_k^n + Y_k \bar{s}_k^{tcp})], \tag{23}$$

for some $\vartheta_2 \in (0, 1]$. The tangential Cauchy step s_k^{tcp} is defined as follows

$$\bar{s}_k^{tcp} = -\tau_k^{tcp} Y_k^T \nabla q_k(s_k^n), \tag{24}$$

where the parameter τ_k^{tcp} is given by

$$\tau_k^{tcp} = \begin{cases} \frac{\|Y_k^T \nabla q_k(s_k^n)\|^2}{(Y_k^T \nabla q_k(s_k^n))^T \bar{B}_k Y_k^T \nabla q_k(s_k^n)} & \text{if } \frac{\|Y_k^T \nabla q_k(s_k^n)\|^3}{(Y_k^T \nabla q_k(s_k^n))^T \bar{B}_k Y_k^T \nabla q_k(s_k^n)} \leq \Delta_k \\ & \text{and } (Y_k^T \nabla q_k(s_k^n))^T \bar{B}_k Y_k^T \nabla q_k(s_k^n) > 0, \\ \frac{\Delta_k}{\|Y_k^T \nabla q_k(s_k^n)\|} & \text{otherwise,} \end{cases} \tag{25}$$

such that $\bar{B}_k = Y_k^T B_k Y_k$.

To be decided whether the step $s_k = s_k^n + s_k^t$ will be accepted or not, a merit function is needed to tie the objective function with the constraints in such a way that progress in the merit function means progress in solving the problem. The following augmented Lagrange function is used in FBACTR algorithm as a merit function,

$$\Phi(x,\mu;r;\sigma) = \ell(x,\mu) + \frac{r}{2}\|P(x)g_u(x)\|^2 + \sigma\|c(x)\|^2, \tag{26}$$

where $\sigma > 0$ is a penalty parameter.

To test whether the point $(x_k + s_k, \mu_{k+1})$ will be taken in the next iterate, an actual reduction and a predicted reduction are defined.

The actual reduction $Ared_k$ in the merit function in moving from (x_k,μ_k) to $(x_k + s_k, \mu_{k+1})$ is defined as follows:

$$Ared_k = \Phi(x_k,\mu_k;r_k;\sigma_k) - \Phi(x_k + d_k, \mu_{k+1}; r_k; \sigma_k).$$

$Ared_k$ can also be written as follows:

$$Ared_k = \ell(x_k,\mu_k) - \ell(x_{k+1},\mu_k) - \Delta\mu_k^T c_{k+1} + \frac{r_k}{2}[\|P_k g_u(x_k)\|^2 - \|P_{k+1} g_{u_{k+1}}\|^2] + \sigma_k[\|c_k\|^2 - \|c_{k+1}\|^2], \tag{27}$$

where $\Delta\mu_k = (\mu_{k+1} - \mu_k)$.

The predicted reduction in the merit function is defined to be:

$$\begin{aligned} Pred_k \quad = \quad & -(\nabla_x \ell(x_k,\mu_k))^T s_k - \frac{1}{2} s_k^T H_k s_k - \Delta\mu_k^T (c_k + \nabla c_k^T s_k) + \frac{r_k}{2}[\|P_k g_{u_k}\|^2 - \|P_k(g_{u_k} + \nabla g_{u_k}^T s_k)\|^2] \\ & + \sigma_k[\|c_k\|^2 - \|c_k + \nabla c_k^T s_k\|^2]. \end{aligned} \tag{28}$$

$Pred_k$ can be written as:

$$Pred_k \quad = \quad q_k(0) - q_k(s_k) - \Delta\mu_k^T(c_k + \nabla c_k^T s_k) + \sigma_k[\|c_k\|^2 - \|c_k + \nabla c_k^T s_k\|^2]. \tag{29}$$

- **To update the penalty parameter σ_k**

To update the penalty parameter σ_k to ensure that $Pred_k \geq 0$, the following schemeis used (see Algorithm 1):

Algorithm 1 To update the penalty parameter σ_k

If

$$Pred_k \leq \frac{\sigma_k}{2}[\|c_k\|^2 - \|c_k + \nabla c_k^T \mu_k s_k\|^2], \tag{30}$$

then, set

$$\sigma_k = \frac{2[q_k(s_k) - q_k(0) + \Delta\mu_k^T(c_k + \nabla c_k^T s_k)]}{\|c_k\|^2 - \|c_k + \nabla c_k^T s_k\|^2} + \beta_0, \tag{31}$$

where $\beta_0 > 0$ is a small fixed constant. Else, set

$$\sigma_{k+1} = \max(\sigma_k, r_k^2). \tag{32}$$

End if.

For more details, see [22].

- **To test the step s_k and update δ_k**

The framework to test the step s_k and update δ_k is clarified in the Algorithm 2.

Algorithm 2 (To test the step s_k and update δ_k)

Choose $0 < \alpha_1 < \alpha_2 < 1$, $0 < \tau_1 < 1 < \tau_2$, and $\delta_{min} \le \delta_0 \le \delta_{max}$.
While $\frac{Ared_k}{Pred_k} \in (0, \alpha_1)$ or $Pred_k \le 0$.
Set $\delta_k = \tau_1 \|s_k\|$.
Evaluate a new trial step s_k.
End while.
If $\frac{Ared_k}{Pred_k} \in [\alpha_1, \alpha_2)$.
Set $x_{k+1} = x_k + s_k$ and $\delta_{k+1} = \max(\delta_k, \delta_{min})$.
End if. If $\frac{Ared_k}{Pred_k} \in [\alpha_2, 1]$.
Set $x_{k+1} = x_k + s_k$ and $\delta_{k+1} = \min\{\delta_{max}, \max\{\delta_{min}, \tau_2 \delta_k\}\}$.
End if.

- **To update the positive parameter r_k**

  To update the positive parameter r_k, we use the following scheme (see Algorithm 3)

Algorithm 3 To update the positive parameter r_k

If

$$\frac{1}{2}[q_k(s_k^n) - q_k(s_k)] \le \|\nabla g_u(x_k) P(x_k) g_u(x_k)\| \min\{\|\nabla g_u(x_k) P(x_k) g_u(x_k)\|, \delta_k\}, \tag{33}$$

Set $r_{k+1} = r_k$.
Else, set $r_{k+1} = 2r_k$.
End if.

For more details see, [25].

Finally, the algorithm stopped if the termination criteria $\|Y_k^T \nabla_x \ell_k\| + \|\nabla g_{u_k} P_k g_{u_k}\| + \|c_k\| \le \varepsilon_1$ or $\|s_k\| \le \varepsilon_2$, for some $\varepsilon_1, \varepsilon_2 > 0$ is satisfied.

- **A trust-region algorithm**

  The framework of the trust-region algorithm to solve subproblem (17) are summarized as follows (see Algorithm 4).

Algorithm 4 Trust-region algorithm

Step 0. (*Initialization*)

Starting with x_0. *Evaluate* μ_0 and P_0. Set $r_0 = 1$, $\sigma_0 = 1$, and $\beta_0 = 0.1$.
Choose ε_1, ε_2, τ_1, τ_2, α_1, *and* α_2 *such that* $0 < \varepsilon_1$, $0 < \varepsilon_2$, $0 < \tau_1 < 1 < \tau_2$, *and* $0 < \alpha_1 < \alpha_2 < 1$.
Choose δ_{min}, δ_{max}, *and* δ_0 *such that* $\delta_{min} \le \delta_0 \le \delta_{max}$. Set $k = 0$.

Step 1. *If* $\|Y_k^T \nabla_x \ell_k\| + \|\nabla g_{u_k} P_k g_{u_k}\| + \|c_k\| \le \varepsilon_1$, *then stop.*
Step 2. (*How to compute* s_k)

(a) Evaluate the normal component s_k^n *by solving subproblem* (18).
(b) Evaluate the tangential component $\bar{s}_k^t$ *by solving subproblem* (22).
(c) Set $s_k = s_k^n + Y_k \bar{s}_k^t$.

Step 3. *If* $\|s_k\| \le \varepsilon_2$, *then stop.*
Step 4. *Set* $x_{k+1} = x_k + s_k$.
Step 5. *Compute* P_{k+1} *given by* (11).
Step 6. *Evaluate* μ_{k+1} *by solving the following subproblem*

$$minimize \ \ \|\nabla f_{u_{k+1}} + \nabla c_{k+1} \mu + r_k \nabla g_{u_{k+1}} P_{k+1} g_{u_{k+1}}\|^2. \tag{34}$$

Step 7. *To update the penalty parameter* σ_k, *using Algorithm* 1.
Step 8. *To test the step* s_k *and update the radius* δ_k, *using Algorithm* 2.
Step 9. *To update the positive parameter* r_k, *using Algorithm* 3.
Step 10. *Set* $k = k + 1$ *and go to Step 1.*

The main steps for solving the NBLP problem (1) are clarified in the following algorithm.

2.2. Fischer–Burmeister Active-Set Trust-Region Algorithm

The framework to solve NBLP problem (1) is summarized in the Algorithm 5 .

Algorithm 5 FBACTR algorithm

Step 1 . Use KKT optimality conditions for the lower level of problem (1) and convert it to a single objective constrained optimization problem (2).
Step 2. Using Fischer–Burmeister function (3) with $\epsilon = 0.001$ to obtain the smooth problem (4).
Step 3. Summarize problem (4) to the form of nonlinear optimization problem (5).
Step 4. Use the active set strategy to reduce problem (5) to problem (12).
Step 5. Use trust-region Algorithm 4 to solve problem (12) and obtained approximate solution for problem (5) which is approximate solution for problem (1).

The next section is dedicated to the global convergence analysis for the active-set with the trust-region algorithm.

3. Global Convergence Analysis

Let $\{(x_k, \mu_k)\}$ be the sequence of points generated by FBACTR Algorithm 5. Let $\Omega \subseteq \Re^n$ be a convex set which is contained all iterates $x_k \in \Re^n$ and $x_k + s_k \in \Re^n$.

Standard assumptions which are needed on the set Ω to demonstrate global convergence theory for FBACTR Algorithm 5 are stated in the following section.

3.1. A Standard Assumptions

The next standard assumptions are required to demonstrate the global convergence theory for the FBACTR Algorithm 5.

[SA_1.] Functions $f_u : \Re^n \to \Re$, $g_u : \Re^n \to \Re^m_1$, $f_l : \Re^n \to \Re^{n_2}$, and $g_l : \Re^n \to \Re^{m_2}$ are twice continuously differentiable functions for all $x \in \Omega$.
[SA_2.] The sequence of the Lagrange multiplier vectors $\{\mu_k\}$ is bounded.
[SA_3.] All of $c(x)$, $\nabla c(x)$, $\nabla^2 c_i(x)$ for $i = 1, 2, \ldots, n_2 + m_2$, $g_u(x)$, $\nabla g_u(x)$, $\nabla^2 g_{u_i}(x)$ for $i = 1, 2, \ldots, m_1$, and $(\nabla c(x)^T \nabla c(x))^{-1}$ are uniformly bounded on Ω.
[SA_4.] The matrix $\nabla c(x)$ has full column rank.
[SA_5.] The sequence of Hessian matrices $\{H_k\}$ is bounded.

Some fundamental lemmas which are needed in the proof of the main theorem introduced in the following section.

3.2. Main Lemmas

Some basic lemmas which are required to demonstrate the main theorems are presented in this section.

Lemma 1. *Under standard assumption SA_1–SA_5 and at any iteration k, there exists a positive constant K_1 such that:*

$$\|s_k^n\| \leq K_1 \|c_k\|. \tag{35}$$

Proof. Since the normal component s_k^n is normal to the tangent space, then we have:

$$\begin{aligned}
\|s_k^n\| &= \|\nabla c_k (\nabla c_k^T \nabla c_k)^{-1} \nabla c_k^T s_k\| \\
&= \|\nabla c_k (\nabla c_k^T \nabla c_k)^{-1} [c_k + \nabla c_k^T s_k - c_k]\| \\
&\leq \|\nabla c_k (\nabla c_k^T \nabla c_k)^{-1}\| [\|c_k + \nabla c_k^T s_k\| + \|c_k\|] \\
&\leq \|\nabla c_k (\nabla c_k^T \nabla c_k)^{-1}\| \|c_k\|,
\end{aligned}$$

where $\|c_k + \nabla c_k^T s_k\| \leq \|c_k\|$. Using standard assumptions SA_1–SA_5, we have the desired result. $\square$

Lemma 2. *Under standard assumptions SA_1 and SA_3, the functions $P(x)g_u(x)$ are Lipschitz continuous in Ω.*

Proof. See Lemma (4.1) in [36]. $\square$

From Lemma 2, we conclude that $g_u(x)^T P(x) g_u(x)$ is differentiable and $\nabla g_u(x) P(x) g_u(x)$ is Lipschitz continuous in Ω.

Lemma 3. *At any iteration k, let $A(x_k) \in \Re^{(m_1)\times(m_1)}$ be a diagonal matrix whose diagonal entries are:*

$$(a_k)_i = \begin{cases} 1 & \text{if } (g_{u_k})_i < 0 \text{ and } (g_{u_{k+1}})_i \geq 0, \\ -1 & \text{if } (g_{u_k})_i \geq 0 \text{ and } (g_{u_{k+1}})_i < 0, \\ 0 & \text{otherwise}, \end{cases} \tag{36}$$

where $i = 1, 2, \ldots, m_1$. Then

$$P_{k+1} = P_k + A_k. \tag{37}$$

Proof. See Lemma (6.2) in [21]. $\square$

Lemma 4. *Under standard assumptions SA_1 and SA_3, there exists a positive constant K_2 such that*

$$\|A_k g_{u_k}\| \leq K_2 \|s_k\|. \tag{38}$$

Proof. See Lemma (6.3) in [21]. $\square$

Lemma 5. *Under standard assumptions SA_1–SA_5, there exists a positive constant K_3 such that:*

$$\mid Ared_k - Pred_k \mid \leq K_3 \sigma_k \|s_k\|^2. \tag{39}$$

Proof. From (37) and (27) we have:

$$Ared_k = \ell(x_k, \mu_k) - \ell(x_{k+1}, \mu_k) - \Delta\mu_k^T c_{k+1} + \frac{r_k}{2}[g_{u_k}^T P_k g_{u_k} - g_{u_{k+1}}^T (P_k + A_k) g_{u_{k+1}}] + \sigma_k[\|c_k\|^2 - \|c_{k+1}\|^2]. \tag{40}$$

From (40), (28), and using Cauchy–Schwarz inequality, we have:

$$\begin{aligned} \mid Ared_k - Pred_k \mid \leq\; & \mid \ell(x_k, \mu_k) + \nabla_x \ell(x_k, \mu_k)^T s_k - \ell(x_{k+1}, \mu_k) \mid + \mid \Delta\mu_k^T [\, c_k + \nabla c_k^T s_k - c_{k+1} \,] \mid \\ & + \frac{r_k}{2} \mid \|P_k(g_{u_k} + \nabla g_{u_k}^T s_k)\|^2 - g_{u_{k+1}}^T (P_k + A_k) g_{u_{k+1}} \mid + \sigma_k \mid \|c_k + \nabla c_k^T s_k\|^2 - \|c_{k+1}\|^2 \mid . \end{aligned}$$

Hence,

$$\begin{aligned} |Ared_k - Pred_k| \leq\; & \frac{1}{2} \mid s_k^T (H_k - \nabla^2 \ell(x_k + \xi_1 s_k, \mu_k)) s_k \mid + \frac{1}{2} \mid s_k^T [\nabla^2 c(x_k + \xi_2 s_k) \Delta\mu_k] s_k \mid \\ & + \frac{r_k}{2} \mid s_k^T [\nabla g_{u_k} P_k \nabla g_{u_k}^T - \nabla g_u(x_k + \xi_4 s_k) P_k \nabla g_u(x_k + \xi_4 s_k)^T] s_k \mid \\ & + \frac{r_k}{2} \mid s_k^T \nabla^2 g_u(x_k + \xi_4 s_k) P_k g_u(x_k + \xi_4 s_k) s_k \mid + \frac{r_k}{2} \|A_k [g_{u_k} + \nabla g_u(x_k + \xi_5 s_k)^T s_k]\|^2 \\ & + \sigma_k \mid s_k^T [\nabla c_k \nabla c_k^T - \nabla c(x_k + \xi_6 s_k) \nabla c(x_k + \xi_6 s_k)^T] s_k \mid \\ & + \sigma_k \mid s_k^T \nabla^2 c(x_k + \xi_6 s_k) c(x_k + \xi_6 s_k) s_k \mid, \end{aligned}$$

for some ξ_1, ξ_2, ξ_3, ξ_4, ξ_5, and $\xi_6 \in (0, 1)$. Using standard assumptions SA_1–SA_5, $\sigma_k \geq r_k$, $\sigma_k \geq 1$, and inequality (38), we have:

$$\mid Ared_k - Pred_k \mid \leq \kappa_1 \|s_k\|^2 + \kappa_2 \sigma_k \|s_k\|^2 \|c_k\| + \kappa_3 \sigma_k \|s_k\|^3, \tag{41}$$

where $\kappa_1 > 0$, $\kappa_2 > 0$, and $\kappa_3 > 0$ are constants and independent of the iteration k. From inequality (41), $\sigma_k \geq 1$, $\|s_k\|$, and $\|c_k\|$ are uniformly bounded, we obtain the desired result. □

Lemma 6. *Under standard assumptions SA_1–SA_5, there exists a positive constant K_4 such that:*

$$\|c_k\|^2 - \|c_k + \nabla c_k^T s_k^n\|^2 \geq K_4 \|c_k\| \min\{\delta_k, \|c_k\|\}. \tag{42}$$

Proof. We consider two cases:

Firstly, from (20), if $s_k^{ncp} = -\frac{\delta_k}{\|\nabla c_k c_k\|}(\nabla c_k c_k)$ and $\delta_k \|\nabla c_k^T \nabla c_k c_k\|^2 \leq \|\nabla c_k c_k\|^3$, then we have:

$$\begin{aligned} \|c_k\|^2 - \|c_k + \nabla c_k^T s_k^{ncp}\|^2 &= -2(\nabla c_k c_k)^T s_k^{ncp} - s_k^{ncp^T} \nabla c_k \nabla c_k^T s_k^{ncp} \\ &= 2\delta_k \|\nabla c_k c_k\| - \frac{\delta_k^2 \|\nabla c_k^T \nabla c_k c_k\|^2}{\|\nabla c_k c_k\|^2} \\ &\geq 2\delta_k \|\nabla c_k c_k\| - \delta_k \|\nabla c_k c_k\| \\ &\geq \delta_k \|\nabla c_k c_k\|. \end{aligned} \tag{43}$$

Secondly, from (20), if $s_k^{ncp} = -\frac{\|\nabla c_k c_k\|^2}{\|\nabla c_k^T \nabla c_k c_k\|^2}(\nabla c_k c_k)$ and $\delta_k \|\nabla c_k^T \nabla c_k c_k\|^2 \geq \|\nabla c_k c_k\|^3$, then we have:

$$\begin{aligned} \|c_k\|^2 - \|c_k + \nabla c_k^T s_k^{ncp}\|^2 &= -2(\nabla c_k c_k)^T s_k^{ncp} - s_k^{ncp^T} \nabla c_k \nabla c_k^T s_k^{ncp} \\ &= \frac{2\|\nabla c_k c_k\|^4}{\|\nabla c_k^T \nabla c_k c_k\|^2} - \frac{\|\nabla c_k c_k\|^4}{\|\nabla c_k^T \nabla c_k c_k\|^2} \\ &= \frac{\|\nabla c_k c_k\|^4}{\|\nabla c_k^T \nabla c_k c_k\|^2} \\ &\geq \frac{\|\nabla c_k c_k\|^2}{\|\nabla c_k^T \nabla c_k c_k\|^2}. \end{aligned} \tag{44}$$

Using standard assumption SA_3, we have $\|\nabla c_k c_k\| \geq \frac{\|c_k\|}{\|(\nabla c_k^T \nabla c_k)^{-1} \nabla c_k\|}$. From inequalities (19), (43), (44), and using standard assumption SA_2, we obtain the desired result.

From Algorithm 1 and Lemma 6, we have, for all k:

$$Pred_k \geq \frac{\sigma_k}{2} K_4 \|c_k\| \min\{\delta_k, \|c_k\|\}. \tag{45}$$

□

Lemma 7. *Under standard assumptions SA_1–SA_5, there exists a constant $K_5 > 0$, such that:*

$$q_k(s_k^n) - q_k(s_k^n + Y_k \bar{s}_k^t) \geq \frac{1}{2} K_5 \|Y_k^T \nabla q_k(s_k^n)\| \min\{\Delta_k, \frac{\|Y_k^T \nabla q_k(s_k^n)\|}{\|\bar{B}_k\|}\}, \tag{46}$$

where $\bar{B}_k = Y_k^T B_k Y_k$.

Proof. We consider two cases:

Firstly, from (24), if $\bar{s}_k^{tcp} = -\frac{\Delta_k}{\|Y_k^T \nabla q_k(s_k^n)\|} Y_k^T \nabla q_k(s_k^n)$ and $\Delta_k (Y_k^T \nabla q_k(s_k^n))^T \bar{B}_k Y_k^T \nabla q_k(s_k^n) \leq \|Y_k^T \nabla q_k(s_k^n)\|^3$, then we have:

$$
\begin{aligned}
q_k(s_k^n) - q_k(s_k^n + Y_k \bar{s}_k^{tcp}) &= -(Y_k^T \nabla q_k(s_k^n))^T \bar{s}_k^{tcp} - \frac{1}{2} \bar{s}_k^{tcp^T} \bar{B}_k \bar{s}_k^{tcp} \\
&= \Delta_k \|Y_k^T \nabla q_k(s_k^n)\| \\
&\quad - \frac{\Delta_k^2}{2\|Y_k^T \nabla q_k(s_k^n)\|^2} [(Y_k^T \nabla q_k(s_k^n))^T \bar{B}_k Y_k^T \nabla q_k(s_k^n)] \\
&\geq \Delta_k \|Y_k^T \nabla q_k(s_k^n)\| - \frac{1}{2} \Delta_k \|Y_k^T \nabla q_k(s_k^n)\| \\
&\geq \frac{1}{2} \Delta_k \|Y_k^T \nabla q_k(s_k^n)\|.
\end{aligned}
\tag{47}
$$

Secondly, from (24), if $\bar{s}_k^{tcp} = -\frac{\|Y_k^T \nabla q_k(s_k^n)\|^2}{Y_k^T \nabla q_k(s_k^n))^T \bar{B}_k Y_k^T \nabla q_k(s_k^n)} Y_k^T \nabla q_k(s_k^n)$ and $\Delta_k (Y_k^T \nabla q_k(s_k^n))^T \bar{B}_k$ $Y_k^T \nabla q_k(s_k^n) \geq \|Y_k^T \nabla q_k(s_k^n)\|^3$, then we have:

$$
\begin{aligned}
q_k(s_k^n) - q_k(s_k^n + Y_k \bar{s}_k^{tcp}) &= -(Y_k^T \nabla q_k(s_k^n))^T \bar{s}_k^{tcp} - \frac{1}{2} \bar{s}_k^{tcp^T} \bar{B}_k \bar{s}_k^{tcp} \\
&= \frac{\|Y_k^T \nabla q_k(s_k^n)\|^4}{(Y_k^T \nabla q_k(s_k^n))^T \bar{B}_k Y_k^T \nabla q_k(s_k^n)} \\
&\quad - \frac{\|Y_k^T \nabla q_k(s_k^n)\|^4}{2(Y_k^T \nabla q_k(s_k^n))^T \bar{B}_k Y_k^T \nabla q_k(s_k^n)} \\
&= \frac{\|Y_k^T \nabla q_k(s_k^n)\|^4}{2(Y_k^T \nabla q_k(s_k^n))^T \bar{B}_k Y_k^T \nabla q_k(s_k^n)} \\
&\geq \frac{\|Y_k^T \nabla q_k(s_k^n)\|^2}{2\|\bar{B}_k\|}.
\end{aligned}
\tag{48}
$$

□

From inequalities (23), (47), (48), and using standard assumptions SA_1–SA_5, we obtain the desired result.

The next lemma shows that FBACTR algorithm cannot be looped infinitely without finding an acceptable step.

Lemma 8. *Under standard assumptions SA_1–SA_5, if there exists $\varepsilon > 0$ such that $\|c_k\| \geq \varepsilon$, then $\frac{Ared_{kj}}{Pred_{kj}} \geq \alpha_1$ for some finite j.*

Proof. Using (39), (45), and from $\|c_k\| \geq \varepsilon$, we have:

$$
\left| \frac{Ared_k}{Pred_k} - 1 \right| = \frac{|\ Ared_k - Pred_k\ |}{Pred_k} \leq \frac{2K_3 \delta_k^2}{K_4 \varepsilon \min\{\varepsilon, \delta_k\}}.
$$

δ_{kj} becomes small as s_{kj} gets rejected and eventually we will have:

$$
\left| \frac{Ared_{kj}}{Pred_{kj}} - 1 \right| \leq \frac{2K_3 \delta_{kj}}{K_4 \varepsilon}.
$$

Then, the acceptance rule will be met for finite j. This completes the proof. □

Lemma 9. *Under standard assumptions SA_1–SA_5 and if jth trial step of iteration k satisfies*

$$
\|s_{kj}\| \leq \min\{\frac{(1-\alpha_1)K_4}{4K_3}, 1\} \|c_k\|, \tag{49}
$$

then it must be accepted.

Proof. This lemma is proved by contradiction. In case of assuming that inequality (49) holds, the trial step s_{k^j} is rejected, by using inequalities (39), (45), and (49), then:

$$(1-\alpha_1) < \frac{\mid Ared_{k^j} - Pred_{k^j} \mid}{Pred_{k^j}} < \frac{2K_3\|s_{k^j}\|^2}{K_4\|c_k\|\|s_{k^j}\|} \leq \frac{(1-\alpha_1)}{2}.$$

This contradiction and therefore the lemma is proved.
□

Lemma 10. *Under standard assumptions SA_1–SA_5, there exists δ_{k^j} satisfies:*

$$\delta_{k^j} \geq \min\{\frac{\delta_{min}}{\beta_1}, \frac{\tau_1(1-\alpha_1)K_4}{4K_3}, \tau_1\}\|c_k\|, \tag{50}$$

for all trial steps j of any iteration k where β_1 is a positive constant independent of k or j.

Proof. At any trial iterate k^j of any iteration k, we consider two cases:

Firstly, if $j = 1$, then the step is accepted. That is $\delta_k^1 \geq \delta_{min}$ and take $\beta_1 = \sup_{x\in\Omega} \|c_k\|$, we have

$$\delta_k \geq \delta_{min} \geq \frac{\delta_{min}}{\beta_1}\|c_k\|. \tag{51}$$

Secondly, if $j > 1$, then there is at least one trial step which is rejected . From Lemma 9, we have

$$\|s_{k^i}\| > \min\{\frac{(1-\alpha_1)K_4}{4K_3}, 1\}\|c_k\|,$$

for all trial steps $i = 1, 2, \ldots j-1$ which are rejected . Since s_{k^i} is a trial step which is rejected, then from the previous inequality and Algorithm 2, we have:

$$\delta_{k^j} = \tau_1\|s_{k^{j-1}}\| > \tau_1 \min\{\frac{(1-\alpha_1)K_4}{4K_3}, 1\}\|c_k\|.$$

From inequality (51) and the above inequality, we obtain the desired result.

The next lemma obviously shows that as long as $\|c_k\|$ is bounded away from zero, the radius of the trust-region is bounded away from zero. □

Lemma 11. *Under standard assumptions SA_1–SA_5, if there exists $\varepsilon > 0$ such that $\|c_k\| \geq \varepsilon$. Then there exists $K_6 > 0$ such that:*

$$\delta_{k^j} \geq K_6.$$

Proof. Let

$$K_6 = \varepsilon \min\{\frac{\delta_{min}}{\beta_1}, \frac{\tau_1(1-\alpha_1)K_4}{4K_3}, \tau_1\}, \tag{52}$$

and using (50), the proof follows directly. □

In the next section, the iteration sequence convergence is studied when $r_k \to \infty$.

3.3. Convergence When the Positive Parameter $r_k \to \infty$

This section is devoted to the convergence of the iteration sequence when the positive parameter r_k goes to infinity.

Notice that, we do not require $[\nabla g_{u_i}(x), i \in I(x)]$ has full column rank in standard assumption SA_4, so, we may have other kinds of stationary points, which are defined in the following definitions.

Definition 2. *A feasible Fritz John (FFJ) point is a point x_* that satisfies the following FFJ conditions:*

$$\begin{aligned}
\eta_* \nabla f_u(x_*) + \nabla c(x_*)\mu_* + \nabla g_u(x_*)\lambda_* &= 0, \\
c(x_*) &= 0, \\
P(x_*)g(x_*) &= 0, \\
(\lambda_*)_i g_{u_i}(x_*) &= 0, \quad i = 1, \ldots, m_1, \\
\eta_*,\ (\lambda_*)_i &\geq 0, \quad i = 1, \ldots, m_1.
\end{aligned}$$

where η_*, μ_*, and λ_* are not all zeros. For more details see [18].

If $\eta_* \neq 0$, then the point $(x_*, 1, \frac{\mu_*}{\eta_*}, \frac{\lambda_*}{\eta_*})$ is called a KKT point and FFJ conditions are called KKT conditions.

Definition 3. *An infeasible Fritz John (IFJ) point is a point x_* that satisfies the following IFJ conditions:*

$$\begin{aligned}
\eta_* \nabla f_u(x_*) + \nabla c(x_*)\mu_* + \nabla g_u(x_*)\lambda_* &= 0, \\
c(x_*) &= 0, \\
\nabla g_u(x_*)P(x_*)g_u(x_*) &= 0 \quad \textit{but} \ \|P(x_*)g_u(x_*)\| > 0, \\
(\lambda_*)_i g_{u_i}(x_*) &\geq 0, \quad i = 1, \ldots, m_1, \\
\eta_*,\ (\lambda_*)_i &\geq 0, \quad i = 1, \ldots, m_1,
\end{aligned}$$

where η_*, μ_*, and λ_* are not all zeros. For more details see [18].

If $\eta_* \neq 0$, then the point $(x_*, 1, \frac{\mu_*}{\eta_*}, \frac{\lambda_*}{\eta_*})$ is called an infeasible KKT point and IFJ conditions are called infeasible KKT conditions.

Lemma 12. *Under standard assumptions SA_1–SA_5, a subsequence $\{k_i\}$ of the sequence of the iteration satisfies IFJ conditions if the following conditions satisfied :*

(i) $\lim_{k_i \to \infty} c(x_{k_i}) = 0$.

(ii) $\lim_{k_i \to \infty} \|P_{k_i} g_u(x_{k_i})\| > 0$.

(iii) $\lim_{k_i \to \infty} \left\{ min_{s \in \Re^{n-m_1+1}} \|P_{k_i}(g_{u_{k_i}} + \nabla g_{u_{k_i}}^T Y_{k_i} \bar{s}^t)\|^2 \right\} = \lim_{k_i \to \infty} \|P_{k_i} g_{u_{k_i}}\|^2$.

Proof. For simplification and without loss of generality, let $\{k_i\}$ represents the whole sequence $\{k\}$. Assume that $\tilde{s}_k$ is the solution of the subproblem $minimize_{\bar{s}^t} \|P_k(g_{u_k} + \nabla g_{u_k}^T Y_k \bar{s}^t)\|^2$, then it satisfies the following equation:

$$Y_k^T \nabla g_{u_k} P_k g_{u_k} + Y_k^T \nabla g_{u_k} P_k \nabla g_{u_k}^T Y_k \tilde{s}_k = 0. \tag{53}$$

It also satisfies the right hand side of Condition (iii). That is,

$$\lim_{k \to \infty} \{2\tilde{s}_k{}^T Y_k^T \nabla g_{u_k} P_k g_{u_k} + \tilde{s}_k{}^T Y_k^T \nabla g_{u_k} P_k \nabla g_{u_k}^T Y_k \tilde{s}_k\} = 0. \tag{54}$$

We will consider two cases:

Firstly, if $\lim_{k \to \infty} \tilde{s}_k = 0$, then from Equation (53) we have $\lim_{k \to \infty} Y_k^T \nabla g_{u_k} P_k g_{u_k} = 0$.

Secondly, if $\lim_{k \to \infty} \tilde{s}_k \neq 0$, then by multiplying Equation (53) from the left by $2\tilde{s}_k^T$ and subtract it from Equation (54), we have $\lim_{k \to \infty} \|P_k \nabla g_{u_k}^T Y_k \tilde{s}_k\|^2 = 0$. Hence $\lim_{k \to \infty} Y_k^T \nabla g_{u_k} P_k g_{u_k} = 0$. That is, in two cases, we have

$$\lim_{k \to \infty} Y_k^T \nabla g_{u_k} P_k g_{u_k} = 0. \tag{55}$$

Since $\lim_{k\to\infty}\|P_k g_{u_k}\| > 0$, then $\lim_{k\to\infty}(P_k g_{u_k})_i \geq 0$, for $i = 1,\ldots,m_1$ and $\lim_{k\to\infty}(P_k g_{u_k})_i > 0$, for some i. Let $(\lambda_k)_i = (P_k g_{u_k})_i$, $i = 1,\ldots,m_1$, then $\lim_{k\to\infty} Y_k^T \nabla g_{u_k} \lambda_k = 0$. Hence, there exists a sequence of $\{\mu_k\}$ such that $\lim_{k\to\infty} Y_k^T\{\nabla c_k \mu_k + \nabla g_{u_k}\lambda_k\} = 0$. That is, IFJ conditions hold in the limit with $\eta_* = 0$, see Definition 3. □

Lemma 13. *Under standard assumptions, SA_1–SA_5, a subsequence $\{k_i\}$ of the sequence of the iteration satisfies FFJ conditions if the following conditions are satisfied:*

(i) $\lim_{k_i\to\infty} c(x_{k_i}) = 0$.

(ii) *For all k_i, $\|P_{k_i} g_{u_{k_i}}\| > 0$ and* $\lim_{k_i\to\infty} P_{k_i} g_{u_{k_i}} = 0$.

(iii) $\lim_{k_i\to\infty}\left\{\min_{s\in\Re^{n-m_1+1}} \frac{\|P_{k_i}(g_{u_{k_i}} + \nabla g_{u_{k_i}}^T Y_{k_i}\bar{s}^t)\|^2}{\|P_{k_i} g_{u_{k_i}}\|^2}\right\} = 1$.

Proof. For simplification and without loss of generality, let $\{k_i\}$ represents the whole sequence $\{k\}$. Notice that the following equation,

$$\lim_{k\to\infty}\left\{\min_{\nu\in\Re^{n-m_1+1}}\left\{\|U_k + P_k \nabla g_k^T Y_k \nu\|^2\right\}\right\} = 1, \tag{56}$$

is equivalent to Condition (iii), where U_k is a unit vector in the direction of $P_k g_{u_k}$ and $\nu = \frac{\bar{s}^t}{\|P_k g_{u_k}\|}$. Let $\tilde{\nu}_k$ be a solution of the following problem:

$$\min_{\nu\in\Re^{n-m_1+1}}\left\{\|U_k + P_k \nabla g_k^T Y_k \nu\|^2\right\}. \tag{57}$$

Hence,

$$Y_k^T \nabla g_{u_k} P_k \nabla g_{u_k}^T Y_k \tilde{\nu}_k + Y_k^T \nabla g_{u_k} P_k U_k = 0. \tag{58}$$

Now two cases are considered:

Firstly, if $\lim_{k\to\infty} Y_k \tilde{\nu}_k = 0$ and using (58), then $\lim_{k\to\infty} Y_k^T \nabla g_{u_k} P_k U_k = 0$.

Secondly, if $\lim_{k\to\infty} Y_k \tilde{\nu}_k \neq 0$, then from (56) and the fact that $\tilde{\nu}_k$ is a solution of problem (57) we have:

$$\lim_{k\to\infty}\{\bar{\nu_k}^T Y_k^T \nabla g_{u_k} P_k \nabla g_{u_k}^T Y_k \tilde{\nu}_k + 2U_k^T P_k \nabla g_{u_k}^T Y_k \tilde{\nu}_k\} = 0.$$

Multiplying Equation (58) from the left by $2\tilde{\nu}_k^T$ and subtracting it from the above limit, we have the following equation: $\lim_{k\to\infty} \tilde{\nu}_k^T Y_k^T \nabla g_{u_k} P_k \nabla g_{u_k}^T Y_k \tilde{\nu}_k = 0$. That is $\lim_{k\to\infty}\{Y_k \nabla g_{u_k} P_k U_k\} = 0$. Hence in both cases, we have $\lim_{k\to\infty}\{Y_k \nabla g_{u_k} P_k U_k\} = 0$. The remnant of the proof follows using cases similar to those in Lemma 12. □

Lemma 14. *Under standard assumptions SA_1–SA_5, if k represents the index of iteration at which σ_k is increased, then we have:*

$$r_k\|c_k\|^2 \leq K_7, \tag{59}$$

where K_7 is a positive constant.

Proof. Since σ_k is increased, then from Algorithm 1 we have:

$$\frac{\sigma_k}{2}[\|c_k\|^2 - \|c_k + \nabla c_k^T s_k\|^2] = [q_k(s_k) - q_k(0) + \Delta\mu_k^T(c_k + \nabla c_k^T s_k)] + \frac{\beta_0}{2}[\|c_k\|^2 - \|c_k + \nabla c_k^T s_k\|^2].$$

From (42), (50), and using the above equality, we have:

$$\begin{aligned}\frac{\sigma_k}{2}K_4\|c_k\|^2 \quad \min \quad & \{\frac{\delta_{min}}{\beta_1}, \frac{\tau_1(1-\alpha_1)K_4}{4K_3}, \tau_1, 1\} \leq \nabla_x \ell_k^T s_k + \frac{1}{2}s_k^T H_k s_k + \Delta\mu_k^T(c_k + \nabla c_k^T s_k) \\ &+ \frac{r_k}{2}[\|P_k(g_{u_k} + \nabla g_{u_k}^T s_k)\|^2 - \|P_k g_{u_k}\|^2] + \frac{\beta_0}{2}[\|c_k\|^2 - \|c_k + \nabla c_k^T s_k\|^2].\end{aligned}$$

However, $\sigma_k \geq r_k^2$, then:

$$\begin{aligned}\frac{r_k^2}{2}K_4\|c_k\|^2\min\{\frac{\delta_{min}}{\beta_1},\frac{\tau_1(1-\alpha_1)K_4}{4K_3},\tau_1,1\} &\leq \nabla_x\ell(x_k,\mu_k)^T s_k+\frac{1}{2}s_k^T H_k s_k+\Delta\mu_k^T(c_k+\nabla c_k^T s_k)\\ &+ \frac{r_k}{2}[\|P_k(g_{u_k}+\nabla g_{u_k}^T s_k)\|^2+\frac{\beta_0}{2}\|c_k\|^2.\end{aligned}$$

Hence,

$$\begin{aligned}\frac{r_k}{2}K_4\|c_k\|^2\min\{\frac{\delta_{min}}{\beta_1},\frac{\tau_1(1-\alpha_1)K_4}{4K_3},\tau_1,1\} &\leq \frac{1}{r_k}[\nabla_x\ell_k^T s_k+\frac{1}{2}s_k^T H_k s_k+\Delta\mu_k^T(c_k+\nabla c_k^T s_k)\\ &+ \frac{\beta_0}{2}\|c_k\|^2]+\frac{1}{2}\|P_k(g_{u_k}+\nabla g_{u_k}^T s_k)\|^2\\ &\leq \frac{1}{r_k}[|\nabla_x\ell_k^T s_k|+\frac{1}{2}|s_k^T H_k s_k|+|\Delta\mu_k^T(c_k+\nabla c_k^T s_k)|\\ &+ \frac{\beta_0}{2}\|c_k\|^2]+\frac{1}{2}\|P_k(g_{u_k}+\nabla g_{u_k}^T s_k)\|^2.\end{aligned}$$

From Cauchy–Schwarz inequality, standard assumptions SA_3–SA_5, and the fact that $\|s_k\| \leq \delta_{max}$, the proof is completed. □

Lemma 15. *Under standard assumptions SA_1–SA_5, if $r_k \to \infty$ and there is an infinite subsequence $\{k_i\}$ of the sequence of the iteration at which σ_k is increased, then:*

$$\lim_{k_i\to\infty}\|c_{k_i}\| = 0. \tag{60}$$

Proof. From lemma (59) and using r_k is unbounded, the proof is completed. □

Theorem 1. *Under standard assumptions SA_1–SA_5, if $r_k \to \infty$ as $k \to \infty$, then*

$$\lim_{k\to\infty}\|c_k\| = 0. \tag{61}$$

Proof. See Theorem 4.18 [22]. □

Lemma 16. *Under standard assumptions SA_1–SA_5, if there exists a subsequence $\{k_j\}$ of indices indexing iterates that satisfy $\|P_k g_{u_k}\| \geq \varepsilon > 0$ for all $k \in \{k_j\}$ and $r_k \to \infty$ as $k \to \infty$. Then a subsequence of the iteration sequence indexed $\{k_j\}$ satisfies IFJ conditions in the limit.*

Proof. For simplification and without loss of generality, the total sequence $\{k\}$ denotes to $\{k_j\}$. This lemma is proved by contradiction, therefor we suppose there is no subsequence of the sequence $\{k\}$ satisfies IFJ conditions in the limit. Using Lemma 12, we have for all k, $|\|P_k g_{u_k}\|^2 - \|P_k(g_{u_k}+\nabla g_{u_k}^T Y_k \bar{s}_k^t)\|^2| \geq \varepsilon_1$ for some $\varepsilon_1 > 0$. From (55), we have $\|Y_k\nabla g_{u_k}P_k g_{u_k}\| \geq \varepsilon_2$, for some $\varepsilon_2 > 0$, hence:

$$\begin{aligned}\|Y_k^T\nabla g_{u_k}P_k g_{u_k}+Y_k^T\nabla g_{u_k}P_k\nabla g_{u_k}^T s_k^n\| &\geq \|Y_k^T\nabla g_{u_k}P_k g_{u_k}\|-\|Y_k^T\nabla g_{u_k}P_k\nabla g_{u_k}^T\|\|s_k^n\|\\ &\geq \varepsilon_2 - K_1\|Y_k^T\nabla g_{u_k}P_k\nabla g_{u_k}^T\|\|c_k\|.\end{aligned}$$

Since $\{\|c_k\|\}$ converges to zero and $\|Y_k^T\nabla g_{u_k}P_k\nabla g_{u_k}^T\|$ is bounded, then we can write: $\|Y_k^T\nabla g_{u_k}P_k g_{u_k}+Y_k^T\nabla g_{u_k}P_k\nabla g_{u_k}^T s_k^n\| \geq \frac{\varepsilon_2}{2}$. Therefore,

$$\begin{aligned}\|Y_k^T\nabla q_k(s_k^n)\| &\geq r_k\|Y_k^T\nabla g_{u_k}P_k g_{u_k}+Y_k^T\nabla g_{u_k}P_k\nabla g_{u_k}^T s_k^n\|-\|Y_k^T\nabla_x\ell_k+Y_k^T H_k s_k^n\|\\ &\geq r_k\frac{\varepsilon_2}{2}-\|Y_k^T\nabla_x\ell_k+Y_k^T H_k s_k^n\|\\ &\geq r_k[\frac{\varepsilon_2}{2}-\frac{1}{r_k}\|Y_k^T\nabla_x\ell_k+Y_k^T H_k s_k^n\|].\end{aligned}$$

From (46), we have:

$$q_k(s_k^n) - q_k(s_k) \geq \frac{K_5}{2} r_k[\frac{\varepsilon_2}{2} - \frac{1}{r_k}\|Y_k^T[\nabla_x \ell_k + H_k s_k^n]\|] \min\{\Delta_k, \frac{\frac{\varepsilon_2}{2} - \frac{1}{r_k}\|Y_k^T[\nabla_x \ell_k + H_k s_k^n]\|}{\|Y_k^T \nabla g_{u_k} P_k \nabla g_{u_k}^T Y_k\| + \frac{1}{r_k}\|Y_k^T H_k Y_k\|}\}.$$

For a k sufficiently large, we have:

$$q_k(s_k^n) - q_k(s_k) \geq \frac{K_5 \varepsilon_2}{4} r_k \min\{\Delta_k, \frac{\varepsilon_2}{2\|Y_k^T \nabla g_{u_k} P_k \nabla g_{u_k}^T Y_k\|}\}.$$

From Algorithm 3, $\{r_k\}$ is boundless only if there exist an infinite subsequence of indices $\{k_i\}$, at which:

$$\frac{1}{2}[q_k(s_k^n) - q_k(s_k)] < \|\nabla g_{u_k} P_k g_{u_k}\| \min\{\|\nabla g_{u_k} P_k g_{u_k}\|, \delta_k\}. \tag{62}$$

Since $r_k \to \infty$, therefore an infinite number of acceptable iterates at which (62) holds and from the way of updating r_k, we have $r_k \to \infty$ as $k \to \infty$. This gives a contradiction unless $r_k \delta_k$ is bounded and hence $\delta_k \to 0$. Therefore $\|s_k\| \to 0$. We will consider two cases: Firstly, if $\|P_k g_{u_k}\|^2 - \|P_k(g_{u_k} + \nabla g_{u_k}^T Y_k \bar{s}_k^t)\|^2 > \varepsilon_1$, then we have

$$r_k\{\|P_k g_{u_k}\|^2 - \|P_k(g_{u_k} + \nabla g_{u_k}^T Y_k \bar{s}_k^t)\|^2\} > r_k \varepsilon_1 \to \infty. \tag{63}$$

Using (63) and standard assumptions $SA_3 - SA_5$, we have $[q_k(s_k^n) - q_k(s_k)] \to \infty$. That is, the left hand side of inequality (62) goes to infinity while the right hand side tends to zero and this is a contradiction in this case.
Secondly, if $\|P_k g_{u_k}\|^2 - \|P_k(g_{u_k} + \nabla g_{u_k}^T Y_k \bar{s}_k^t)\|^2 < -\varepsilon_1$, then

$$r_k\{\|P_k g_{u_k}\|^2 - \|P_k(g_{u_k} + \nabla g_{u_k}^T Y_k \bar{s}_k^t)\|^2|\} < -r_k \varepsilon_1 \to -\infty,$$

where $r_k \to \infty$ as $k \to \infty$ and similar to the first case, $[q_k(s_k^n) - q_k(s_k)] \to -\infty$. This is a contradiction with $[q_k(s_k^n) - q_k(s_k)] > 0$. The lemma is proved. $\square$

Lemma 17. *Under standard assumptions SA_1–SA_5, if $r_k \to \infty$ as $k \to \infty$, and there exists a subsequence indexed $\{k_j\}$ of iterates that satisfy $\|P_k g_{u_k}\| > 0$ for all $k \in \{k_j\}$ and $\lim_{k_j \to \infty} \|P_{k_j} g_{u_{k_j}}\| = 0$, then a subsequence of the sequence of iterates indexed $\{k_j\}$ satisfies FFJ conditions in the limit.*

Proof. Without loss of generality, let $\{k_j\}$ be the whole iteration sequence $\{k\}$ to simplify. This lemma is proved by contradiction and so suppose that there is no subsequence that satisfies FFJ conditions in the limit. From condition (iii) of Lemma 13, for all k sufficiently large, there exists a constant $\varepsilon_3 > 0$ such that:

$$\frac{|\ \|P_k g_{u_k}\|^2 - \|P_k(g_{u_k} + \nabla g_{u_k}^T Y_k \bar{s}_k^t)\|^2\ |}{\|P_k g_{u_k}\|^2} \geq \varepsilon_3. \tag{64}$$

The following two cases are considered:
Firstly, if $\lim inf_{k \to \infty} \frac{\bar{s}_k^t}{\|P_k g_{u_k}\|} = 0$, then there is a contradiction with inequality (64).
Secondly, if $\lim sup_{k \to \infty} \frac{\bar{s}_k^t}{\|P_k g_{u_k}\|} = \infty$, then from subproblem (22) we have:

$$Y_k^T \nabla q_k(s_k^n) = -Y_k^T (B_k + v_k I) Y_k \bar{s}_k^t, \tag{65}$$

where $v_k \geq 0$ represents the Lagrange multiplier vector, which is associated with the constraint $\|Y_k \bar{s}^t\| \leq \Delta_k$. From (65) and (46), we have:

$$q_k(s_k^n) - q_k(s_k) \geq \frac{K_5}{2} \|Y_k^T \nabla q_k(s_k^n)\| \min\{\Delta_k, \frac{\|Y_k^T[\frac{1}{r_k} H_k + (\frac{v_k}{r_k} I + \nabla g_{u_k} P_k \nabla g_{u_k}^T)] Y_k \bar{s}_k^t\|}{\|Y_k^T(\frac{1}{r_k} H_k + \nabla g_{u_k} P_k \nabla g_{u_k}^T) Y_k\|}\}. \quad (66)$$

Since $r_k \to \infty$ as $k \to \infty$, then there exists an infinite number of acceptable steps such that inequality (62) holds. However, inequality (62) can be written as follows:

$$\frac{1}{2}[q_k(s_k^n) - q_k(s_k)] < \beta_2^2 \|P_k g_{u_k}\|^2, \quad (67)$$

where $\beta_2 = sup_{x \in \Omega} \|Y_k \nabla g_{u_k}\|$.

From (66) and (67), we have:

$$\frac{K_5}{2} \|Y_k^T \nabla q_k(s_k^n)\| \min\{\frac{\Delta_k}{\|P_k g_{u_k}\|}, \frac{\|Y_k^T[\frac{1}{r_k} H_k + (\frac{v_k}{r_k} I + \nabla g_{u_k} P_k \nabla g_{u_k}^T)] Y_k \bar{s}_k^t\|}{\|Y_k^T(\frac{1}{r_k} H_k + \nabla g_{u_k} P_k \nabla g_{u_k}^T) Y_k\| \|P_k g_{u_k}\|}\} < 2\beta_2^2 \|P_k g_{u_k}\|.$$

However, in previous inequality, the right hand side tends to zero as $k \to \infty$ and also $\lim_{k_i \to \infty} \frac{\bar{s}_{k_i}^t}{\|P_{k_i} g_{k_i}\|} = \infty$ along the subsequence $\{k_i\}$. Therefore,

$$\|Y_{k_i}^T \nabla q_{k_i}(s_{k_i}^n)\| \frac{\|Y_{k_i}^T[\frac{1}{r_{k_i}} H_{k_i} + (\frac{v_{k_i}}{r_{k_i}} I + \nabla g_{k_i} P_{k_i} \nabla g_{k_i}^T)] Y_{k_i} \bar{s}_{k_i}^t\|}{\|Y_{k_i}^T(\frac{1}{r_{k_i}} H_{k_i} + \nabla g_{k_i} P_{k_i} \nabla g_{k_i}^T) Y_{k_i}\| \|P_{k_i} g_{k_i}\|},$$

is bounded. That is, either $\frac{\bar{s}_{k_i}^t}{\|P_{k_i} g_{k_i}\|}$ lies in the null space of $Y_{k_i}^T(\frac{v_{k_i}}{r_{k_i}} I + \nabla g_{k_i} P_{k_i} \nabla g_{k_i}^T) Y_{k_i}^T$ or $\|Y_{k_i} \nabla q_{k_i}(s_{k_i}^n)\| \to 0$.

The first possibility occurs only when $\frac{v_{k_i}}{r_{k_i}} \to 0$ as $k_i \to \infty$ and $\frac{\bar{s}_{k_i}^t}{\|P_{k_i} g_{k_i}\|}$ lie in the null space of the matrix $Y_{k_i}^T \nabla g_{k_i} P_{k_i} \nabla g_{k_i}^T Y_{k_i}$ which is contradicted with assumption (64). This means that, FFJ conditions are satisfied in the limit. As $k_i \to \infty$, the second possibility is, $\|Y_{k_i} \nabla q_{k_i}(s_{k_i}^n)\| \to 0$ and from (65), we have $\|\bar{s}_{k_i}^t\| \to 0$ which is contradicted with assumption (64). That is FFJ conditions are satisfied in the limit.

In the next section, the convergence of the sequence of the iteration sequence is studied when r_k bounded.

3.4. Global Convergence When r_k Is Bounded

Our analysis in this section is continued supposing that r_k is bounded. Therefore, let $\bar{k}$ be an integer at which $r_k = \bar{r} < \infty$ for all $k \geq \bar{k}$. That is,

$$\frac{1}{2}[q_k(s_k^n) - q_k(s_k)] \geq \|\nabla g_{u_k} P_k g_{u_k}\| \min\{\|\nabla g_{u_k} P_k g_{u_k}\|, \delta_k\}. \quad (68)$$

From assumptions SA_3 and SA_5, and using (68), then for all k, there is a constant $\beta_3 > 0$ such that:

$$\|B_k\| \leq \beta_3, \quad \|Y_k^T B_k\| \leq \beta_3, \quad and \quad \|Y_k^T B_k Y_k\| \leq \beta_3, \quad (69)$$

where $B_k = H_k + \bar{r} \nabla g_{u_k} P_k \nabla g_{u_k}^T$. □

Lemma 18. *Under standard assumptions SA_1–SA_5, there exists a constant $K_8 > 0$ such that:*

$$q_k(0) - q_k(s_k^n) - \Delta\mu_k^T(c_k + \nabla c_k^T s_k) \geq -K_8 \|c_k\|. \quad (70)$$

Proof. Since

$$
\begin{aligned}
q_k(0) - q_k(s_k^n) &= -\nabla_x \ell_k^T s_k^n - \frac{1}{2} s_k^{n^T} H_k s_k^n + \frac{\bar{r}}{2}[\|P_k g_{u_k}\|^2 - \|P_k(g_{u_k} + \nabla g_{u_k}^T s_k^n)\|^2] \\
&= -(\nabla_x \ell_k + \bar{r} \nabla g_{u_k} P_k g_{u_k})^T s_k^n - \frac{1}{2} s_k^{n^T}(H_k + \bar{r} \nabla g_{u_k} P_k \nabla g_{u_k}^T) s_k^n \\
&= -(\nabla_x \ell_k + \bar{r} \nabla g_{u_k} P_k g_{u_k})^T s_k^n - \frac{1}{2} s_k^{n^T} B_k s_k^n,
\end{aligned}
$$

then we have:

$$
\begin{aligned}
q_k(0) - q_k(s_k^n) &- \Delta\mu_k^T(c_k + \nabla c_k^T s_k) = -(\nabla_x \ell_k + \bar{r} \nabla g_{u_k} P_k g_{u_k})^T s_k^n - \frac{1}{2} s_k^{n^T} B_k s_k^n - \Delta\mu_k^T(c_k + \nabla c_k^T s_k) \\
&\geq -\|\nabla_x \ell_k\| \|s_k^n\| - \bar{r} \|\nabla g_{u_k} P_k g_{u_k}\| \|s_k^n\| - \|B_k\| \|s_k^n\|^2 - \|\Delta\mu_k\| \|c_k + \nabla c_k^T s_k\| \\
&\geq -[\|\nabla_x \ell_k\| + \bar{r} \|\nabla g_{u_k} P_k g_{u_k}\| + \|B_k\| \|s_k^n\|] \|s_k^n\| - \|\Delta\mu_k\| \|\nabla c_k\| \|s_k^n\|.
\end{aligned}
$$

From inequality (35) and the fact that $Y_k^T \nabla c(x_k) = 0$, then we have:

$$
q_k(0) - q_k(s_k^n) - \Delta\mu_k^T(c_k + \nabla c_k^T s_k) \geq [(\|\nabla_x \ell_k\| + \bar{r} \|\nabla g_{u_k} P_k g_{u_k}\| + \|B_k\| \|s_k^n\| + \|\Delta\mu_k\| \|\nabla c_k\|) K_1] \|c_k\|.
$$

From standard assumptions SA_2, SA_3, SA_5, the fact that $\|s_k^n\| \leq \delta_{max}$, and using (69), then there exists $K_8 > 0$, such that inequality (70) holds. $\square$

Lemma 19. *Under standard assumptions SA_1–SA_5, then for all k we have:*

$$
\begin{aligned}
Pred_k \geq\ & \frac{1}{2} K_5 \|Y_k^T \nabla q_k(s_k^n)\| \min\{\Delta_k, \frac{\|Y_k^T \nabla q_k(s_k^n)\|}{\|\bar{s}_k\|}\} + \|\nabla g_{u_k} P_k g_{u_k}\| \min\{\|\nabla g_{u_k} P_k g_{u_k}\|, \delta_k\} \\
& - K_8 \|c_k\| + \sigma_k[\|c_k\|^2 - \|c_k + \nabla c_k^T s_k\|^2].
\end{aligned} \tag{71}
$$

Proof. From (29), we have:

$$
\begin{aligned}
Pred_k &= [q_k(s_k^n) - q_k(s_k)] + [q_k(0) - q_k(s_k^n) - \Delta\mu_k^T(c_k + \nabla c_k^T s_k)] + \sigma_k[\|c_k\|^2 - \|c_k + \nabla c_k^T s_k\|^2] \\
&= \frac{1}{2}[q_k(s_k^n) - q_k(s_k)] + \frac{1}{2}[q_k(s_k^n) - q_k(s_k)] \\
&\quad + [q_k(0) - q_k(s_k^n) - \Delta\mu_k^T(c_k + \nabla c_k^T s_k)] + \sigma_k[\|c_k\|^2 - \|c_k + \nabla c_k^T s_k\|^2].
\end{aligned}
$$

Using inequalities (46), (68), and (70), we obtain the desired result. $\square$

Lemma 20. *Under standard assumptions SA_1–SA_5, if $\|Y_k^T(\nabla_x \ell_k + \bar{r} \nabla g_{u_k} P_k g_{u_k})\| + \|\nabla g_{u_k} P_k g_{u_k}\| \geq \varepsilon > 0$ and $\|c_k\| \leq \tau \delta_k$ where τ is a positive constant given by*

$$
\tau \leq \min\left\{ \frac{\varepsilon}{6\beta_3 K_1 \delta_{max}}, \frac{\sqrt{3}}{2K_1}, \frac{K_5 \varepsilon}{24 K_8} \min\{\frac{2\varepsilon}{3\delta_{max}}, 1\}, \frac{\varepsilon}{4K_8} \min\{\frac{\varepsilon}{2\delta_{max}}, 1\} \right\}, \tag{72}
$$

then there exists a constant $K_9 > 0$ such that:

$$
Pred_k \geq K_9 \delta_k + \sigma_k[\|c_k\|^2 - \|c_k + \nabla c_k^T s_k\|^2]. \tag{73}
$$

Proof. Since $\|Y_k^T(\nabla_x \ell_k + \bar{r} \nabla g_{u_k} P_k g_{u_k})\| + \|\nabla g_{u_k} P_k g_{u_k}\| \geq \varepsilon$, then we can say $\|Y_k^T(\nabla_x \ell_k + \bar{r} \nabla g_{u_k} P_k g_{u_k})\| \geq \frac{\varepsilon}{2}$ and $\|\nabla g_{u_k} P_k g_{u_k}\| \geq \frac{\varepsilon}{2}$. We will consider two cases:

Firstly, if $\|Y_k^T(\nabla_x \ell_k + \bar{r} \nabla g_{u_k} P_k g_{u_k})\| \geq \frac{\varepsilon}{2}$, then from inequalities (69), (35) and $\|c_k\| \leq \tau \delta_k$, we have:

$$
\begin{aligned}
\|Y_k^T(\nabla_x \ell_k + \bar{r} \nabla g_{u_k} P_k g_{u_k} + B_k s_k^n)\| &\geq \|Y_k^T(\nabla_x \ell_k + \bar{r} \nabla g_{u_k} P_k g_{u_k})\| - \|Y_k^T B_k s_k^n\| \\
&\geq \|Y_k^T(\nabla_x \ell_k + \bar{r} \nabla g_{u_k} P_k g_{u_k})\| - \beta_3 K_1 \|c_k\| \\
&\geq \frac{\varepsilon}{2} - \beta_3 K_1 \tau \delta_k.
\end{aligned}
$$

However, $\tau \leq \frac{\varepsilon}{6\beta_3 K_1 \delta_{max}}$, then

$$\|Y_k^T(\nabla_x \ell_k + \bar{r}\nabla g_{u_k} P_k g_{u_k} + A_k s_k^n)\| \geq \frac{\varepsilon}{3}. \tag{74}$$

From inequality (35), assumption $\|c_k\| \leq \tau\delta_k$, and using the value of τ in (72), we have $\|s_k^n\| \leq K_1\|c_k\| \leq K_1\tau\delta_k \leq K_1\frac{\sqrt{3}}{2K_1}\delta_k = \frac{\sqrt{3}}{2}\delta_k$. That is, $\Delta_k^2 = \delta_k^2 - \|s_k^n\|^2 \geq \delta_k^2 - \frac{3}{4}\delta_k^2 = \frac{1}{4}\delta_k^2$. This means that,

$$\Delta_k \geq \frac{1}{2}\delta_k. \tag{75}$$

From inequalities (71), (74), (75), and assumption $\|c_k\| \leq \tau\delta_k$, we have the following:

$$\begin{aligned} Pred_k &\geq \frac{1}{2}K_5\|Y_k^T(\nabla_x \ell_k + \bar{r}\nabla g_{u_k} P_k g_{u_k} + B_k s_k^n)\| \min\{\|Y_k^T(\nabla_x \ell_k + \bar{r}\nabla g_{u_k} P_k g_{u_k} + B_k s_k^n)\|, \frac{1}{2}\delta_k\} \\ &\quad -K_8\|c_k\| + \sigma_k[\|c_k\|^2 - \|c_k + \nabla c_k^T s_k\|^2] \\ &\geq \frac{K_5\varepsilon}{12}\delta_k \min\{\frac{2\varepsilon}{3\delta_{max}}, 1\} - K_8\tau\delta_k + \sigma_k[\|c_k\|^2 - \|c_k + \nabla c_k^T s_k\|^2]. \end{aligned}$$

However, $\tau \leq \frac{K_5\varepsilon}{24K_8}\min\{\frac{2\varepsilon}{3\delta_{max}}, 1\}$, then we have

$$Pred_k \geq \frac{K_5\varepsilon}{24}\min\{\frac{2\varepsilon}{3\delta_{max}}, 1\}\delta_k + \sigma_k[\|c_k\|^2 - \|c_k + \nabla c_k^T s_k\|^2].$$

Secondly, if $\|\nabla g_{u_k} P_k g_{u_k}\| \geq \frac{\varepsilon}{2}$ and using inequality (71), then

$$\begin{aligned} Pred_k &\geq \|\nabla g_{u_k} P_k g_{u_k}\| \min\{\|\nabla g_{u_k} P_k g_{u_k}\|, \delta_k\} - K_8\|c_k\| + \sigma_k[\|c_k\|^2 - \|c_k + \nabla c_k^T s_k\|^2] \\ &\geq \frac{\varepsilon}{2}\min\{\frac{\varepsilon}{2\delta_{\max}}, 1\}\delta_k - K_8\tau\delta_k + \sigma_k[\|c_k\|^2 - \|c_k + \nabla c_k^T s_k\|^2] \\ &\geq \frac{\varepsilon}{4}\min\{\frac{\varepsilon}{2\delta_{\max}}, 1\}\delta_k + \sigma_k[\|c_k\|^2 - \|c_k + \nabla c_k^T s_k\|^2], \end{aligned}$$

where $\tau \leq \frac{\varepsilon}{4K_8}\min\{\frac{\varepsilon}{2\delta_{\max}}, 1\}$. Let $K_9 = \min\left\{\frac{K_5\varepsilon}{24}\min\{\frac{2\varepsilon}{3\delta_{max}}, 1\}\ ,\ \frac{\varepsilon}{4}\min\{\frac{\varepsilon}{2\delta_{\max}}, 1\}\right\}$, then the result follows.

From the previous lemma, we notice that either $\|Y_k^T(\nabla_x \ell_k + \bar{r}\nabla g_{u_k} P_k g_{u_k})\| \geq \frac{\varepsilon}{2} > 0$ or $\|\nabla g_{u_k} P_k g_{u_k}\| \geq \frac{\varepsilon}{2} > 0$ and $\|c_k\| \leq \tau\delta_k$, where τ is given by (72) at any iteration k, the value of the penalty parameter σ_k is not needed to increase. That is the penalty parameter σ_k is increased only when $\|c_k\| \geq \tau\delta_k$.

□

Lemma 21. *Under standard assumptions SA_1–SA_5, if σ_k is increased at kth iteration , then there is a positive constant K_{10} such that:*

$$\sigma_k \min\{\|c_k\|, \delta_k\} \leq K_{10}. \tag{76}$$

Proof. From Algorithm 1, we have:

$$\begin{aligned} \frac{\sigma_k}{2}[\|c_k\|^2 - \|c_k + \nabla c_k^T s_k\|^2] &= [q_k(s_k) - q_k(s_k^n)] + [q_k(s_k^n) - q_k(0)] + \Delta\mu_k^T(c_k + \nabla c_k^T s_k) \\ &\quad + \frac{\beta_0}{2}[\|c_k\|^2 - \|c_k + \nabla c_k^T s_k\|^2] \\ &= -\frac{1}{2}[q_k(s_k^n) - q_k(s_k)] - \frac{1}{2}[q_k(s_k^n) - q_k(s_k)] \\ &\quad + [q_k(s_k^n) - q_k(0) + \Delta\mu_k^T(c_k + \nabla c_k^T s_k)] + \frac{\beta_0}{2}[\|c_k\|^2 - \|c_k + \nabla c_k^T s_k\|^2], \end{aligned}$$

where σ_k increased at any iteration and $r_k = \bar{r}$. From the previous equation, (42), (46), (68), and (70), we have

$$\begin{aligned}\frac{\sigma_k}{2}K_4\|c_k\|\min\{\delta_k,\|c_k\|\} &\leq -\frac{K_5}{2}\|Y_k^T\nabla q_k(s_k^n)\|\min\{\Delta_k,\frac{\|Y_k^T\nabla q_k(s_k^n)\|}{\|\bar{B}_k\|}\} \\ &\quad -\|\nabla g_{u_k}P_k g_{u_k}\|\min\{\|\nabla g_{u_k}P_k g_{u_k}\|,\delta_k\} + K_8\|c_k\| + \frac{\beta_0}{2}\|c_k\|^2 \\ &\leq K_8\|c_k\| + \frac{\beta_0}{2}\|c_k\|^2.\end{aligned}$$

Using assumption SA_3, we get the desired result. □

Lemma 22. *Under standard assumptions SA_1–SA_5 and at the jth trial iterate of any iteration k. If σ_{kj} is increased, then there is a constant $K_{11} > 0$, such that*

$$\sigma_{kj}\|c_k\| \leq K_{11}. \tag{77}$$

Proof. From (50) and (76), we get the desired result. □

Lemma 23. *Under standard assumptions SA_1–SA_5, if $\sigma_k \to \infty$, then*

$$\lim_{k_i\to\infty}\|c_{k_i}\| = 0, \tag{78}$$

where $\{k_i\}$ is a subsequence indexes the iterates at which σ_k is increased.

Proof. From Lemma 22 we obtain the desired result. □

3.5. *Main Results for Global Convergence*

In this section, main global convergence results for FBACTR algorithm are introduced.

Theorem 2. *Under standard assumptions SA_1–SA_5, the sequence of iterates which is generated by FBACTR algorithm satisfies*

$$\lim_{k\to\infty}\|c_k\| = 0. \tag{79}$$

Proof. This theorem is proved by contradiction and so we suppose that $\limsup_{k\to\infty}\|c_k\| \geq \varepsilon > 0$. This means that there exists an infinite subsequence of indices $\{k_j\}$ indexing iterates that satisfy $\|c_{k_j}\| \geq \frac{\varepsilon}{2}$. However, there exists an infinite sequence of acceptable steps from Lemma 8. Without loss of generality and to simplify, we suppose that all members of $\{k_j\}$ are acceptable iterates. Now, two cases are considered:

Firstly, if $\{\sigma_k\}$ is unbounded, then an infinite number of iterates $\{k_i\}$ exists and at which the penalty parameter σ_k is increased. So, for k that is sufficiently large and from Lemma 23, let $\{k_i\}$ and $\{k_j\}$ be the two sequences which are not have common elements. Let k_{ρ_1} and k_{ρ_2} be two consecutive iterates at which σ_k is increased and $k_{\rho_1} < k < k_{\rho_2}$, where $k \in \{k_j\}$. The penalty parameter σ_k is the same for all iterates that lie between k_{ρ_1} and k_{ρ_2}. Since all the iterates of $\{k_j\}$ are acceptable, then for all $k \in \{k_j\}$,

$$\Phi_k - \Phi_{k+1} = Ared_k \geq \alpha_1 Pred_k.$$

Using inequality (45), we have:

$$\frac{\Phi_k - \Phi_{k+1}}{\sigma_k} \geq \frac{\alpha_1 K_4}{2}\|c_k\|\min\{\|c_k\|,\delta_k\}.$$

Summing over all acceptable iterates that lie between k_{ρ_1} and k_{ρ_2}, we have:

$$\sum_{k=k_{\rho_1}}^{k_{\rho_2}-1} \frac{\Phi_k - \Phi_{k+1}}{\sigma_k} \geq \frac{\alpha_1 K_4 \varepsilon}{4} \min\{\hat{K_6}, \frac{\varepsilon}{2}\},$$

where $\hat{K_6}$ is as K_6 in (52)but ε is replaced by $\frac{\varepsilon}{2}$. Hence,

$$\frac{\ell(x_{k_{\rho_1}}, \mu_{k_{\rho_1}}; \bar{r}) - \ell(x_{k_{\rho_2}}, \mu_{k_{\rho_2}}; \bar{r})}{\sigma_{k_{\rho_1}}} + [\|c_{k_{\rho_1}}\|^2 - \|c_{k_{\rho_2}}\|^2] \geq \frac{\alpha_1 K_4 \varepsilon}{4} \min\{\hat{K_6}, \frac{\varepsilon}{2}\}.$$

Since $\sigma_k \to \infty$, then for k_{ρ_1} sufficiently large, we have:

$$\frac{\mid \ell(x_{k_{\rho_1}}, \mu_{k_{\rho_1}}; \bar{r}) - \ell(x_{k_{\rho_2}}, \mu_{k_{\rho_2}}; \bar{r}) \mid}{\sigma_{k_{\rho_1}}} < \frac{\alpha_1 K_4 \varepsilon}{8} \min\{\hat{K_6}, \frac{\varepsilon}{2}\}.$$

Therefore,

$$\|c_{k_{\rho_1}}\|^2 - \|c_{k_{\rho_2}}\|^2 \geq \frac{\alpha_1 K_4 \varepsilon}{8} \min\{\hat{K_6}, \frac{\varepsilon}{2}\}.$$

This leads to a contradiction with Lemma 23 unless $\varepsilon = 0$.

Secondly, If $\{\sigma_k\}$ is bounded, then for all an integer $\tilde{k}$ and $k \geq \tilde{k}$, we have $\sigma_k = \tilde{\sigma}$. Hence, for any $\hat{k} \in \{k_j\}$ where $\hat{k} \geq \tilde{k}$ and using (45), we have:

$$Pred_{\hat{k}} \quad \geq \quad \frac{\tilde{\sigma} K_4}{2} \|c_{\hat{k}}\| \min\{\delta_{\hat{k}}, \|c_{\hat{k}}\|\} \geq \frac{\varepsilon \tilde{\sigma} K_4}{4} \min\{\frac{\varepsilon}{2\delta_{max}}, 1\}\delta_{\hat{k}}. \tag{80}$$

Then for any $\hat{k} \in \{k_j\}$, we have:

$$\Phi_{\hat{k}} - \Phi_{\hat{k}+1} = Ared_{\hat{k}} \geq \alpha_1 Pred_{\hat{k}},$$

such that all the iterates of $\{k_j\}$ are acceptable. From above inequality, inequality (80) and using Lemma 11 we have:

$$\Phi_{\hat{k}} - \Phi_{\hat{k}+1} \geq \frac{\alpha_1 \varepsilon \tilde{\sigma} K_4}{4} \min\{\frac{\varepsilon}{2\delta_{max}}, 1\}\hat{K_6} > 0.$$

However, this is a contradiction of the fact that $\{\Phi_k\}$ is bounded when $\{\sigma_k\}$ is bounded. Therefore, we have a contradiction in both cases. Hence the supposition is not correct and this proves the theorem. □

Theorem 3. *Under standard assumptions SA_1–SA_5, the sequence of iterates generated by FBACTR algorithm satisfies:*

$$\liminf_{k\to\infty} \; [\ \|Y_k^T \nabla_x \ell_k\| + \|\nabla g_{u_k} P_k g_{u_k}\|\ ] \;=\; 0. \tag{81}$$

Proof. First, we prove that:

$$\liminf_{k\to\infty} [\|Y_k^T (\nabla_x \ell_k + \bar{r} \nabla g_{u_k} P_k g_{u_k})\| + \|\nabla g_{u_k} P_k g_{u_k}\ \|\ ] \;=\; 0. \tag{82}$$

The proof of (82) is by contradiction, so, for all k, assume that $\|\ Y_k^T (\nabla_x \ell_k + \bar{r} \nabla g_{u_k} P_k g_{u_k})\ \| + \|\nabla g_{u_k} P_k g_{u_k}\| > \varepsilon$. Let $\{k_i\}$ be an infinite subsequence at which $\|c_{k_i}\| > \tau \delta_{k_i}$, where τ is defined in (72). However, $\|c_k\| \to 0$, then

$$\lim_{k_i \to \infty} \ \delta_{k_i} = 0.$$

Let k^j be any trial iterate belonging to $\{k_i\}$ and we consider two cases:

Firstly, if $\{\sigma_k\}$ is unbounded, then for the rejected trial step $j-1$ of iteration $k \in \{k_i\}$, we have $\|c_k\| > \tau\delta_{kj} = \tau_1\tau\|s_{kj-1}\|$. Since the trial step s_{kj-1} is rejected and using inequalities (45) and (41), then

$$
\begin{aligned}
(1-\alpha_1) &\leq \frac{|Ared_{kj-1} - Pred_{kj-1}|}{Pred_{kj-1}} \\
&\leq \frac{[2\kappa_1\|s_{kj-1}\| + 2\kappa_2\sigma_{kj-1}\|s_{kj-1}\|\|c_k\| + 2\kappa_3\sigma_{kj-1}\|s_{kj-1}\|^2]}{\sigma_{kj-1}K_4\min(\tau_1\tau,1)\|c_k\|} \\
&\leq \frac{2\kappa_1}{\sigma_{kj-1}K_4\tau_1\tau\min(\tau_1\tau,1)} + \frac{2\kappa_2\tau_1\tau + 2\kappa_3}{K_4\tau_1\tau\min(\tau_1\tau,1)}\|s_{kj-1}\|.
\end{aligned}
$$

However, $\{\sigma_k\}$ is unbounded, hence for all $k \geq \hat{k}$, $\hat{k}$ is sufficiently large, we have:

$$
\sigma_{kj-1} > \frac{4\kappa_1}{K_4\tau_1\tau\min(\tau_1\tau,1)(1-\alpha_1)}.
$$

Therefore, for all $k \geq \hat{k}$, we have:

$$
\|s_{kj-1}\| \geq \frac{K_4\tau_1\tau\min(\tau_1\tau,1)(1-\alpha_1)}{4(\kappa_2\tau_1\tau + \kappa_3)}.
$$

From Algorithm 2, we have:

$$
\delta_{kj} = \tau_1\|s_{kj-1}\| \geq \frac{K_4\tau_1^2\tau\min(\tau_1\tau,1)(1-\alpha_1)}{4(\kappa_2\tau_1\tau + \kappa_3)}.
$$

This gives a contradiction and this leads to δ_{kj} not being able to go to zero in this case.

Secondly, if the sequence $\{\sigma_k\}$ is bounded, then there exists an integer $\bar{k}$ and $\bar{\sigma}$ such that for all $k \geq \bar{k}$, $\sigma_k = \bar{\sigma}$. Consider a trial step j of iteration $k \geq \bar{k}$ and $\|c_k\| > \tau\delta_{kj}$, we consider three cases:

(i) If $j = 1$, then $\delta_{kj} \geq \delta_{\min}$, see Algorithm 2. This means that, δ_{kj} is bounded in this case;

(ii) If $j > 1$, and $\|c_{k^l}\| > \tau\delta_{k^l}$ for $l = 1, \ldots, j$, then for all rejected trial steps $l = 1, \ldots, j-1$ of iteration $k \geq \bar{k}$, we have

$$
(1-\alpha_1) \leq \frac{|Ared_{k^l} - Pred_{k^l}|}{Pred_{k^l}} \leq \frac{2K_6\|s_{k^l}\|}{K_4\min(\tau,1)\|c_k\|}.
$$

Hence,

$$
\begin{aligned}
\delta_{kj} = \tau_1\|s_{kj-1}\| &\geq \frac{\tau_1K_4\min(\tau,1)(1-\alpha_1)\|c_k\|}{2K_3} \geq \frac{\tau_1K_4\min(\tau,1)(1-\alpha_1)\tau}{2K_3}\delta_{k^1} \\
&\geq \frac{\tau_1K_4\min(\tau,1)(1-\alpha_1)\tau}{2K_3}\delta_{\min}.
\end{aligned}
$$

That is, δ_{kj} is also bounded in this case.

(iii) If $j > 1$ and $\|c_{k^l}\| > \tau\delta_{k^l}$ does not hold for all l, then there exists an integer ϱ such that $\|c_{k^l}\| > \tau\delta_{k^l}$ holds for $l = \varrho+1, \ldots, j$ and $\|c_{k^l}\| \leq \tau\delta_{k^l}$ holds for all $l = 1, \ldots, \varrho$. As in case (ii), we can write:

$$
\delta_{kj} \geq \frac{\tau_1K_4\min(\tau,1)(1-\alpha_1)}{2K_3}\|c_k\| \geq \frac{\tau_1K_4\min(\tau,1)(1-\alpha_1)\tau}{2K_3}\delta_{k^{\varrho+1}}. \tag{83}
$$

From Algorithm 2, we have:

$$\delta_{k^{\varrho+1}} \geq \tau_1 \|s_{k^\varrho}\|. \tag{84}$$

From Lemma 20, if $\|c_{k^l}\| \leq \tau\delta_{k^l}$ and s_{k^ϱ} is rejected, then we have:

$$(1-\alpha_1) \leq \frac{|Ared_{k^\varrho} - Pred_{k^\varrho}|}{Pred_{k^\varrho}} \leq \frac{2K_6\bar{r}\|s_{k^\varrho}\|}{K_9}.$$

That is,

$$\|s_{k^\varrho}\| \geq \frac{K_9(1-\alpha_1)}{2K_6\bar{\sigma}}.$$

This implies that $\|s_{k^\varrho}\|$ is bounded and from (83) and (84) we have also δ_{k^j} is bounded in this case. That is in three cases, we have δ_{k^j} is bounded, but this leading to a contradiction. Hence, all the iterates satisfy $\|c_k\| \leq \tau\delta_{k^j}$ for k^j are sufficiently large. From Lemma 20, then the value of the penalty parameter is not needed to increase. Hence, $\{\sigma_k\}$ is bounded. Using Lemma 20 and for $k^j \geq \bar{k}$, we have:

$$\Phi_{k^j} - \Phi_{k^j+1} = Ared_{k^j} \geq \alpha_1 Pred_{k^j} \geq \alpha_1 K_9 \delta_{k^j}.$$

As $k \to \infty$, then:

$$\lim_{k\to\infty} \delta_{k^j} = 0. \tag{85}$$

That is the trust-region radius is not bounded below and this leading to a contradiction. Because at iteration $k^j > \bar{k}$, if the previous step was accepted; i.e., at $j = 1$, then $\delta_{k^1} \geq \delta_{\min}$. That is δ_{k^j} is bounded in this case.

If $j > 1$, then there exists at least one rejected trial step. From Lemmas 5 and 20, then for the rejected trial step $s_{k^{j-1}}$ we have:

$$(1-\alpha_1) < \frac{\bar{\sigma}K_3\|s_{k^{j-1}}\|^2}{K_9\delta_{k^{j-1}}}.$$

From Algorithm 2, we have:

$$\delta_{k^j} = \tau_1\|s_{k^{j-1}}\| > \frac{\tau_1 K_9(1-\alpha_1)}{\bar{\sigma}K_3}.$$

Hence δ_{k^j} is bounded and this contradicts (85). That is, the supposition is wrong and hence,

$$\liminf_{k\to\infty} \left[\, \|Y_k^T(\nabla_x \ell_k + \bar{r}\nabla g_{u_k} P_k g_{u_k})\| + \|\nabla g_{u_k} P_k g_{u_k}\| \,\right] = 0.$$

That is, (81) holds and the proof is completed.

From the above two theorems, we conclude that, given any $\varepsilon > 0$, the algorithm terminates because $\|Y_k^T \nabla_x \ell_k\| + \|\nabla g_{u_k} P_k g_{u_k}\| + \|c_k\| < \varepsilon$, for some finite k. □

4. Numerical Results and Comparisons

In this section, we introduce an extensive variety of possible numeric NBLP problems to illustrate the validity of the proposed Algorithm FBACTR Algorithm 5 to solve the NBLP problem. The proposed algorithm FBACTR experimented on 16 benchmark examples given in [4,7,38–40].

Ten independent runs with a distinct initial value starting points for every test example are performed to observe the matchmaking of the result. Statistical results of all examples are briefed in Table 1 which displays that the results found by the FBACTR Algorithm 5 are approximate or equal to those by the compered algorithms in method [11] and the literature.

Table 1. Comparisons of the results of FBACTR Algorithm 5 with the method [11] and methods in the reference.

Problem Name	(v_*, w_*) Method [11]	f_u^* f_l^* Method [11]	(v_*, w_*) FBACTR Algorithm 5	f_u^* f_l^* FBACTR Algorithm 5	(v_*, w_*) Ref.	f_u^* f_l^* Ref.
TP1	(0.8503, 0.0227, 0.03589)	-2.6764 0.0332	(0.8465, 0.7695, 0)	-2.0772 -0.5919	(0.8438, 0.7657, 0)	-2.0769 -0.5863
TP2	(0.609, 0.391, 0, 0, 1.828)	0.6086 1.6713	(0.6111, 0.3890, 0, 0, 1.8339)	0.64013 1.6816	(0.609, 0.391, 0, 0, 1.828)	0.6426 1.6708
TP3	(0.97, 3.14, 2.6, 1.8)	-8.92 -6.05	(0.97, 3.14 2.6, 1.8)	-8.92 -6.05	(0.97, 3.14, 2.6, 1.8)	-8.92 -6.05
TP4	(0.5, 0.5, 0.5, 0.5)	-1 0	(0.5, 0.5, 0.5, 0.5)	-1 0	(0.5, 0.5, 0.5, 0.5)	-1 0
TP5	(9.839, 10.059)	96.809 0.0019	(9.9953, 9.9955)	99.907 1.8628×10^{-4}	(10.03, 9.969)	100.58 0.001
TP6	(1.6879, 0.8805, 0)	-1.3519 7.4991	$(1.8889, 8.8889 \times 10^{-1}, 6.8157 \times 10^{-6})$	-1.4074 7.6172	NA	3.57 2.4
TP7	(1, 0)	17 1	(1, 0)	17 1	(1, 0)	17 1
TP8	(0.75, 0.75, 0.75, 0.75)	-2.25 0	(0.7513, 0.7513, 0.752, 0.752)	-2.2480 0	$(\sqrt{3}/2, \sqrt{3}/2, \sqrt{3}/2, \sqrt{3}/2)$	-2.1962 0
TP9	(11.138, 5)	2209.8 222.52	(11.25, 5)	2250 197.753	(11.25, 5)	2250 197.753
TP10	$(1, 0, 6.6387 \times 10^{-6})$	6.6387×10^{-6} -6.6387×10^{-6}	(1, 0, 1)	1 -1	(1, 0, 1)	1 -1
TP11	(24.972, 29.653, 5.0238, 9.7565)	4.9101 0.01332	(25, 30, 5, 10)	5 0	(25, 30, 5, 10)	5 0
TP12	(3, 5)	9 0	(3, 5)	9 0	(3, 5)	9 0
TP13	(0, 1.7405, 1.8497, 0.9692)	-15.548 -1.4247	(0, 2, 1.875, 0.9063)	-12.68 -1.016	(0, 2, 1.875, 0.9063)	-12.68 -1.016
TP14	(10.016, 0.81967)	81.328 -0.3359	(10, 0.011)	8.1978×10^{1} 0	(10.04, 0.1429)	82.44 0.271
TP15	(0, 0.9, 0, 0.6, 0.4)	-29.2 3.2	(0, 0.9, 0, 0.6, 0.4)	-29.2 3.2	(0, 0.9, 0, 0.6, 0.4)	-29.2 3.2
TP16	(0, 0.9, 0, 0.6, 0.4, 0, 0, 0)	-29.2 0.3148	(0, 0.9, 0, 0.6, 0.4, 0, 0, 0)	-29.2 0.3148	(0, 0.9, 0, 0.6, 0.4, 0, 0, 0)	-29.2 0.3148

For comparison, the corresponding results of the mean number of iterations (iter), the mean number of function evaluations (nfunc), and the mean value of CPU time (CPUs) in seconds obtained by Methods in [11,41,42] respectively are included and summarized in Table 2. These results show that results of the FBACTR Algorithm 5 are approximate or equal to those of the compared algorithms in the literature.

It is evident from the results that our approach is able to handle NBLP problems even if the upper and the lower levels are convex or not and the computed results converge to the optimal solution which is similar or approximate to the optimal reported in the literature. Finally, it is obvious from the comparison between the solutions obtained using the FBACTR Algorithm 5 with those in the literature, that the FBACTR Algorithm 5 is capable of finding the optimal solution to some problems by a small number of iterations, a small number of function evaluations, and less time.

We offered the numerical results of FBACTR Algorithm 5 using MATLAB (R2013a) (8.2.0.701)64-bit(win64) and a starting point $x_0 \in int(\tilde{\boldsymbol{F}})$. The following parameter setting is used: $\delta_{min} = 10^{-4}$, $\delta_0 = max(\|s_0^{cp}\|, \delta_{min})$, $\delta_{max} = 10^4\delta_0$, $\alpha_1 = 10^{-3}$, $\alpha_2 = 0.8$, $\tau_1 = 0.5$, $\tau_2 = 2$, $\varepsilon_1 = 10^{-10}$, and $\varepsilon_2 = 10^{-12}$.

Table 2. Comparisons of the results of FBACTR Algorithm 5 with method [11], method [41] and method [42] with respect to the number of iterations, the number of function evaluations, and time/s.

Problem Name	Iter Method [11]	nfunc Method [11]	CPUs Method [11]	Iter FBACTR Algorithm	nfunc FBACTR Algorithm	CPUs FBACTR Algorithm	CPUs Method [41]	CPUs Method [42]
TP1	11	12	1.43	10	13	1.62	1.734	-
TP2	10	14	1.987	9	12	1.87	2.375	-
TP3	6	8	2.9	7	8	2.52	3.315	11.854
TP4	10	14	1.68	12	13	1.92	1.576	-
TP5	6	9	1.635	6	7	1.523	1.825	5.888
TP6	6	11	4.1	8	10	3.95	4.689	25.332
TP7	12	13	1.9	11	12	1.652	1.769	-
TP8	10	11	1.002	11	12	0.953	1.124	-
TP9	10	13	1.95	8	10	1.87	-	-
TP10	5	7	2.987	5	6	3.31	-	-
TP11	9	12	3.742	10	13	3.632	-	37.308
TP12	8	9	1.23	7	9	1.33	-	-
TP13	5	7	2.1	5	8	1.998	-	14.42
TP14	6	8	2.12	5	6	1.97	-	4.218
TP15	5	6	20.512	6	7	20.125	-	45.39
TP16	5	7	40.319	4	5	35.21	-	107.55

5. Conclusions

In this paper, the FBACTR Algorithm 5 is presented to solve the NBLP problem (1). A KKT condition is used with the Fischer–Burmeister function and an active-set strategy to convert the NBLP problem to an equivalent smooth equality constrained optimization problem. To ensure global convergence for the FBACTR algorithm, a trust-region globalization strategy is used.

A global convergence theory for the FBACTR algorithm is introduced and applications to mathematical programs with equilibrium constraints are provided to clarify the effectiveness of the proposed approach. Numerical results reflect the good behavior of the FBACTR algorithm and the computed results converge to the optimal solutions. It is clear from the comparison between the solutions obtained using the FBACTR algorithm with algorithms [11,41,42] that the FBACTR can find the optimal solution to some problems with a small number of iterations, small number of function evaluations, and in less time.

Test Problem 1 [41]:

$$\begin{array}{ll} \min_v & f_u = w_1^2 + w_2^2 + v^2 - 4v \\ s.t. & 0 \le v \le 2, \\ \min_w & f_l = w_1^2 + 0.5w_2^2 + w_1 w_2 + \\ & (1-3v)w_1 + (1+v)w_2, \\ s.t. & 2w_1 + w_2 - 2v \le 1, \\ & w_1 \ge 0, \quad w_2 \ge 0. \end{array}$$

Test Problem 2 [41]:

$$\begin{array}{ll} \min_v & f_u = w_1^2 + w_3^2 - w_1 w_3 - 4w_2 - 7v_1 + 4v_2 \\ s.t. & v_1 + v_2 \le 1, \\ & v_1 \ge 0, \quad v_2 \ge 0 \\ \min_w & f_l = w_1^2 + 0.5w_2^2 + 0.5w_3^2 + w_1 w_2 + \\ & (1-3v_1)w_1 + (1+v_2)w_2, \\ s.t. & 2w_1 + w_2 - w_3 + v_1 - 2v_2 + 2 \le 0, \\ & w_1 \ge 0; \quad w_2 \ge 0 \quad w_3 \ge 0. \end{array}$$

Test Problem 3 [41]:

$$\begin{array}{rl}
\min_v & f_u = 0.1(v_1^2 + v_2^2) - 3w_1 - 4w_2 + 0.5(w_1^2 + w_2^2) \\
s.t. & \\
\min_w & f_l = 0.5(w_1^2 + 5w_2^2) - 2w_1 w_2 - v_1 w_1 - v_2 w_2, \\
s.t. & -0.333w_1 + w_2 - 2 \leq 0, \\
& w_1 - 0.333w_2 - 2 \leq 0, \\
& w_1 \geq 0, \quad w_2 \geq 0,
\end{array}$$

Test Problem 4 [41]:

$$\begin{array}{rl}
\min_v & f_u = v_1^2 - 2v_1 + v_2^2 - 2v_2 + w_1^2 + w_2^2 \\
s.t. & v_1 \geq 0, \quad v_2 \geq 0 \\
\min_w & f_l = (w_1 - v_1)^2 + (w_2 - v_2)^2, \\
s.t. & 0.5 \leq w_1 \leq 1.5, \\
& 0.5 \leq w_2 \leq 1.5,
\end{array}$$

Test Problem 5 [41]:

$$\begin{array}{rl}
\min_v & f_u = v^2 + (w - 10)^2 \\
s.t. & -v + w \leq 0, \\
& 0 \leq v \leq 15, \\
\min_w & f_l = (v + 2w - 30)^2, \\
s.t. & v + w \leq 20, \\
& 0 \leq w \leq 20,
\end{array}$$

Test Problem 6 [41]:

$$\begin{array}{rl}
\min_v & f_u = (v_1 - 1)^2 + 2w_1^2 - 2v_1 \\
s.t. & v_1 \geq 0, \\
\min_w & f_l = (2w_1 - 4)^2 + (2w_2 - 1)^2 + v_1 w_1, \\
s.t. & 4v_1 + 5w_1 + 4w_2 \leq 12, \\
& -4v_1 - 5w_1 + 4w_2 \leq -4, \\
& 4v_1 - 4w_1 + 5w_2 \leq 4, \\
& -4v_1 + 4w_1 + 5w_2 \leq 4, \\
& w_1 \geq 0, \quad w_2 \geq 0,
\end{array}$$

Test Problem 7 [41]:

$$\begin{array}{rl}
\min_v & f_u = (v - 5)^2 + (2w + 1)^2 \\
s.t. & v \geq 0, \\
\min_w & f_l = (2w - 1)^2 - 1.5vw, \\
s.t. & -3v + w \leq -3, \\
& v - 0.5w \leq 4, \\
& v + w \leq 7, \\
& w \geq 0.
\end{array}$$

Test Problem 8 [41]:

$$\begin{array}{rl}
\min_v & f_u = v_1^2 - 3v_1 + v_2^2 - 3v_2 + w_1^2 + w_2^2 \\
s.t. & v_1 \geq 0, \quad v_2 \geq 0, \\
\min_w & f_l = (w_1 - v_1)^2 + (w_2 - v_2)^2, \\
s.t. & 0.5 \leq w_1 \leq 1.5, \\
& 0.5 \leq w_2 \leq 1.5,
\end{array}$$

Test Problem 9 [3]:

$$\begin{array}{ll} \min_v & f_u = 16v^2 + 9w^2 \\ s.t. & -4v + w \leq 0, \\ & v \geq 0, \\ \quad \min_w & f_l = (v + w - 20)^4, \\ \quad s.t. & 4v + w - 50 \leq 0, \\ & w \geq 0. \end{array}$$

Test Problem 10 [3]:

$$\begin{array}{ll} \min_v & f_u = v_1^3 w_1 + w_2 \\ s.t. & 0 \leq v_1 \leq 1, \\ \min_w & f_l = -w_2 \\ s.t. & v_1 w_1 \leq 10, \\ & w_1^2 + v_1 w_2 \leq 1, \\ & w_2 \geq 0. \end{array}$$

Test Problem 11 [42]:

$$\begin{array}{ll} \min_v & f_u = 2v_1 + 2v_2 - 3w_1 - 3w_2 - 60 \\ s.t. & v_1 + v_2 + w_1 - 2w_2 \leq 40, \\ & 0 \leq v_1 \leq 50, \\ & 0 \leq v_2 \leq 50, \\ \quad \min_w & f_l = (w_1 - v_1 + 20)^2 + (w_2 - v_2 + 20)^2, \\ \quad s.t. & v_1 - 2w_1 \geq 10, \\ & v_2 - 2w_2 \geq 10, \\ & -10 \leq w_1 \leq 20, \\ & -10 \leq w_2 \leq 20. \end{array}$$

Test Problem 12 [3]:

$$\begin{array}{ll} \min_v & f_u = (v - 3)^2 + (w - 2)^2 \\ s.t. & -2v + w - 1 \leq 0, \\ & v - 2w + 2 \leq 0, \\ & v + 2w - 14 \leq 0, \\ & 0 \leq v \leq 8, \\ \quad \min_w & f_l = (w - 5)^2 \\ s.t. & w \geq 0. \end{array}$$

Test Problem 13 [42]:

$$\begin{array}{ll} \min_v & f_u = -v_1^2 - 3v_2^2 - 4w_1 + w_2^2 \\ s.t. & v_1^2 + 2v_2 \leq 4, \\ & v_1 \geq 0, \quad v_2 \geq 0, \\ \quad \min_w & f_l = 2v_1^2 + w_1^2 - 5w_2, \\ \quad s.t. & v_1^2 - 2v_1 + 2v_2^2 - 2w_1 + w_2 \geq -3, \\ & v_2 + 3w_1 - 4w_2 \geq 4, \\ & w_1 \geq 0, \quad w_2 \geq 0. \end{array}$$

Test Problem 14 [42]:

$$\begin{array}{ll} \min_v & f_u = (v - 1)^2 + (w - 1)^2 \\ s.t. & v \geq 0, \\ \quad \min_w & f_l = 0.5w^2 + 500w - 50vw \\ s.t. & y \geq 0. \end{array}$$

Test Problem 15 [42]:

$$\begin{array}{ll}
\min_v & f_u = -8v_1 - 4v_2 + 4w_1 - 40w_2 - 4w_3 \\
s.t. & v_1 \geq 0, \quad v_2 \geq 0 \\
\quad \min_w & f_l = v_1 + 2v_2 + w_1 + w_2 + 2w_3, \\
\quad s.t. & w_2 + w_3 - w_1 \leq 1, \\
& 2v_1 - w_1 + 2w_2 - 0.5w_3 \leq 1, \\
& 2v_2 + 2w_1 - w_2 - 0.5w_3 \leq 1, \\
& w_i \geq 0, \quad i = 1, 2, 3.
\end{array}$$

Test Problem 16 [42]:

$$\begin{array}{ll}
\min_v & f_u = -8v_1 - 4v_2 + 4w_1 - 40w_2 - 4w_3 \\
s.t. & v_1 \geq 0, \quad v_2 \geq 0 \\
\quad \min_w & f_l = \frac{1+v_1+v_2+2w_1-w_2+w_3}{6+2v_1+w_1+w_2-3w_3}, \\
\quad s.t. & -w_1 + w_2 + w_3 + w_4 = 1, \\
& 2v_1 - w_1 + 2w_2 - 0.5w_3 + w_5 = 1, \\
& 2v_2 + 2w_1 - w_2 - 0.5w_3 + w_6 = 1, \\
& w_i \geq 0, \quad i = 1, \ldots, 6.
\end{array}$$

Author Contributions: B.E.: Conceptualization and software G.A.: formal analysis and writing. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: The author would like to thank the anonymous referees for their valuable comments and suggestions which have helped to greatly improve this paper.

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

References

1. Bialas, W.; Karwan, M. On two-level optimization. *IEEE Trans. Autom. Control.* **1982**, *27*, 211–214. [CrossRef]
2. Dempe, S. *Foundation of Bilevel Programming*; Kluwer Academic: London, UK, 2002.
3. Gumus, H.; Flouda, A. Global Optimization of Nonlinear Bilevel Programming Problems. *J. Glob. Optim.* **2001**, *20*, 1–31. [CrossRef]
4. Muu, D.; Quy, N. A Global Optimization Method for Solving Convex Quadratic Bilevel Programming Problems. *J. Glob. Optim.* **2003**, *26*, 199–219. [CrossRef]
5. Abo-Elnaga, Y.; El-Shorbagy, M. Multi-Sine Cosine Algorithm for Solving Nonlinear Bilevel Programming Problems. *Int. J. Comput. Intell. Syst.* **2020**, *13*, 421–432. [CrossRef]
6. Abo-Elnaga, Y.; Nasr, S. Modified Evolutionary Algorithm and Chaotic Search for Bilevel Programming Problems. *Symmetry* **2020**, *12*, 727. [CrossRef]
7. Falk, J.; Liu, J. On bilevel programming, Part I: General nonlinear cases. *Math. Program.* **1995**, *70*, 47–72. [CrossRef]
8. Ma, L.; Wang, G. A Solving Algorithm for Nonlinear Bilevel Programing Problems Based on Human Evolutionary Model. *Algorithms* **2020**, *13*, 260. [CrossRef]
9. Savard, G.; Gauvin, J. The steepest descent direction for the nonlinear bilevel programming problem. *Oper. Res. Lett.* **1994**, *15*, 265–272. [CrossRef]
10. Edmunds, T.; Bard, J. Algorithms for nonlinear bilevel mathematical programs. *IEEE Trans. Syst. Man Cybern.* **1991**, *21*, 83–89. [CrossRef]
11. El-Sobky, B.; Ashry, G. An interior-point trust-region algorithm to solve a nonlinear bilevel programming problem. *AIMS Math.* **2022**, *7*, 5534–5562. [CrossRef]
12. Chen, J. The semismooth-related properties of a merit function and a descent method for the nonlinear complementarity problem. *J. Glob. Optim.* **2006**, *36*, 565–580. [CrossRef]
13. Chen, J. On some NCP-functions based on the generalized Fischer–Burmeister function. *Asia-Pac. J. Oper. Res.* **2007**, *24*, 401–420. [CrossRef]
14. Chen, J.; Pan, S. A family of NCP-functions and a descent method for the nonlinear complementarity problem. *Comput. Optim. Appl.* **2008**, *40*, 389–404. [CrossRef]

15. Facchinei, F.; Jiang, H.; Qi, L. A smoothing method for mathematical programming with equilibrium constraints. *Math. Program.* **1999**, *85*, 107–134. [CrossRef]
16. Byrd, R.; Hribar, M.; Nocedal, J. An interior point algorithm for largescale nonlinear programming. *SIAM J. Optim.* **1999**, *9*, 877–900. [CrossRef]
17. Byrd, R.; Gilbert, J.; Nocedal, J. A trust region method based on interior point techniques for nonlinear programming. *Math. Program.* **2000**, *89*, 149–185. [CrossRef]
18. Bazaraa, M.; Sherali, H.; Shetty, C. *Nonlinear Programming Theory and Algorithms*; John Wiley and Sons: Hoboken, NJ, USA, 2006.
19. Curtis, F.E.; Schenk, O.; Wachter, A. An interior-point algorithm for large-scale nonlinear optimization with inexact step computations. *Siam J. Sci. Comput.* **2010**, *32*, 3447–3475. [CrossRef]
20. Esmaeili, H.; Kimiaei, M. An efficient implementation of a trust-region method for box constrained optimization. *J. Appl. Math. Comput.* **2015**, *48*, 495–517. [CrossRef]
21. El-Sobky, B. A Multiplier active-set trust-region algorithm for solving constrained optimization problem. *Appl. Math. Comput.* **2012**, *219*, 127–157. [CrossRef]
22. El-Sobky, B. An active-set interior-point trust-region algorithm. *Pac. J. Optim.* **2018**, *14*, 125–159. [CrossRef]
23. El-Sobky, B.; Abotahoun, A. An active-set algorithm and a trust-region approach in constrained minimax problem. *Comput. Appl. Math.* **2018**, *37*, 2605–2631. [CrossRef]
24. El-Sobky, B.; Abotahoun, A. A trust-region Algorithm for Solving Mini-Max Problem. *J. Comput. Math.* **2018**, *36*, 881–902.
25. El-Sobky, B.; Abouel-Naga, Y. A penalty method with trust-region mechanism for nonlinear bilevel optimization problem. *J. Comput. Appl. Math.* **2018**, *340*, 360–374. [CrossRef]
26. El-Sobky, B.; Abo-Elnaga, Y.; Mousa, A.; El-Shorbagy, A. trust-region based penalty barrier algorithm for constrained nonlinear programming problems: An application of design of minimum cost canal sections. *Mathematics* **2021**, *9*, 1551. [CrossRef]
27. Kouri, D.; Heinkenschloss, M.; Ridzal, D.; van Waanders, B. A trust-region Algorithm with Adaptive Stochastic Collocation for PDE Optimization under Uncertainty. *SIAM J. Sci. Comput.* **2020**, *35*, 1847–1879. [CrossRef]
28. Li, N.; Xue, D.; Sun, W.; Wang, J. A stochastic trust-region method for unconstrained optimization problems. *Math. Probl. Eng.* **2019**, *2019*, 8095054. [CrossRef]
29. Niu, L.; Yuan, Y. A new trust region algorithm for nonlinear constrained optimization. *J. Comput. Math.* **2020**, *28*, 72–86.
30. Wang, X.; Yuan, Y. A trust region method based on a new affine scaling technique for simple bounded optimization. *Optim. Methods Softw.* **2013**, *28*, 871–888. [CrossRef]
31. Wang, X.; Yuan, Y. An augmented Lagrangian trust region method for equality constrained optimization, *Optim. Methods Softw.* **2015**, *30*, 559–582. [CrossRef]
32. Zeng, M.; Ni, Q. A new trust region method for nonlinear equations involving fractional mode. *Pac. J. Optim.* **2019**, *15*, 317–329.
33. Byrd, R. Robust trust-region methods for nonlinearly constrained optimization. In Proceedings of the Second SIAM Conference on Optimization, Houston, TX, USA, 18–20 May 1987 .
34. Omojokun, E. Trust-Region Strategies for Optimization with Nonlinear Equality and Inequality Constraints. Ph.D. Thesis, Department of Computer Science, University of Colorado, Boulder, CO, USA, 1989.
35. El-Sobky, B.; Abouel-Naga, Y. Multi-objective optimal load flow problem with interior-point trust-region strategy. *Electr. Power Syst. Res.* **2017**, *148*, 127–135. [CrossRef]
36. Dennis, J.; El-Alem, M.; Williamson, K. A trust-region approach to nonlinear systems of equalities and inequalities. *SIAM J. Optim.* **1999**, *9*, 291–315. [CrossRef]
37. Dennis, J.; Heinkenschloss, M.; Vicente, L. trust-region interior-point SQP algorithms for a class of nonlinear programming problems. *SIAM J. Control. Optim.* **1998**, *36*, 1750–1794. [CrossRef]
38. Bard, J.F. Convex two-level optimization. *Math. Program.* **1988**, *40*, 15–27. [CrossRef]
39. Oduguwa, V.; Roy, R. Bi-level optimization using genetic algorithm. In Proceedings of the IEEE international Conference Artificial Intelligence Systems, Divnomorskoe, Russia, 5–10 September 2002; pp. 123–128.
40. Shimizu, K.; Aiyoshi, E. A new computational method for Stackelberg and min-max problems by use of a penalty method. *IEEE Trans. Autom. Control* **1981**, *26*, 460–466. [CrossRef]
41. Li, H.; Jiao, Y.; Zhang, L. Orthogonal genetic algorithm for solving quadratic bilevel programming problems. *J. Syst. Eng. Electron.* **2010**, *21*, 763–770. [CrossRef]
42. Wang, Y.; Jiao, Y.; Li, H. An evolutionary algorithm for solving nonlinear bilevel programming based on a new constraint-Handling scheme. *IEEE Trans. Syst. Man Cybern. Part C* **2005**, *35*, 221–232. [CrossRef]