AIMS Mathematics

http://www.aimspress.com/journal/Math

AIMS Mathematics, 7(9): 16112–16146.
DOI: 10.3934/math.2022882
Received: 17 March 2022
Revised: 26 April 2022
Accepted: 14 May 2022
Published: 01 July 2022

Research article

An active-set with barrier method and trust-region mechanism to solve a nonlinear Bilevel programming problem

B. El-Sobky[1,*], **G. Ashry**[1] **and Y. Abo-Elnaga**[2]

[1] Department of Mathematics and Computer Science, Alexandria University, Faculty of Science, Egypt

[2] Department of basic science, Tenth of Ramadan City, Higher Technological Institute, Egypt

* **Correspondence:** Email: bothinaelsobky@yahoo.com.

Abstract: Nonlinear Bilevel programming (NBLP) problem is a hard problem and very difficult to be resolved by using the classical method. In this paper, Karush-Kuhn-Tucker (KKT) condition is used with Fischer-Burmeister function to convert NBLP problem to an equivalent smooth single objective nonlinear programming (SONP) problem. An active-set strategy is used with Barrier method and trust-region technique to solve the smooth SONP problem effectively and guarantee a convergence to optimal solution from any starting point. A global convergence theory for the active-set barrier trust-region (ACBTR) algorithm is studied under five standard assumptions. An applications to mathematical programs are introduced to clarify the effectiveness of ACBTR algorithm. The results show that ACBTR algorithm is stable and capable of generating approximal optimal solution to the NBLP problem.

Keywords: nonlinear Bilevel programming problem; active-set; barrier method; trust-region mechanism; projected Hessian mechanism; global convergence
Mathematics Subject Classification: 49N10, 49N35, 65K05, 93D22, 93D52

1. Introduction

The NBLP problem is a nonlinear optimization problem that is constrained by another nonlinear optimization problem. This mathematical programming model arises when two independent decision makers, ordered within a hierarchical structure, have conflicting objectives. The decision maker at the lower level has to optimize her objective under the given parameters from the upper level decision maker, who, in return, with complete information on the possible reactions of the lower, selects the parameters so as to optimize her own objective. The decision maker with the upper level objective, $f_u(t, v)$ takes the lead, and chooses her decision vector t. The decision maker with lower level objective,

$f_l(t,v)$, reacts accordingly by choosing her decision vector v to optimize her objective, parameterized in t. Note that the upper level decision maker is limited to influencing, rather than controlling, the lower level's outcome. In fact, the problem has been proved to be NP-hard [5]. However, the NBLP problem is used so extensively in transaction network, finance budget, resource allocation, price control etc. Various approaches have been devoted to study this field, which leads to a a speedy development in theories and algorithms, see [1,3,30,32,41]. For detailed exposition, the reader can review [23,25,35].

A mathematical formulation for the NBLP problem is

$$\begin{array}{ll} \min & f_u(t,v) \\ s.t. & g_u(t,v) \leq 0, \\ \quad \min & f_l(t,v), \\ \quad s.t. & g_l(t,v) \leq 0, \\ & t \geq 0, \quad v \geq 0, \end{array} \tag{1.1}$$

where $t \in \Re^{n_1}$ and $v \in \Re^{n_2}$.

Let $n = n_1 + n_2$, and assume that the functions $f_u : \Re^n \to \Re$, $f_l : \Re^n \to \Re$, $g_u : \Re^n \to \Re^{m_1}$, and $g_l : \Re^n \to \Re^{m_2}$ are at least twice continuously differentiable function in our method.

Several approaches have been proposed to solve the NBLP problem 1.1, see [2, 13, 14, 25, 27, 37, 40, 42]. KKT conditions one of these approaches and used in this paper to convert the original NBLP problem 1.1 to the following one-level programming problem:

$$\begin{array}{ll} \min_{t,v} & f_u(t,v) \\ s.t. & g_u(t,v) \leq 0, \\ & \nabla_v f_l(t,v) + \nabla_v g_l(t,v)\lambda_l = 0, \\ & g_l(t,v) \leq 0, \\ & (\lambda_l)_j g_{l_j}(t,v) = 0, \quad j = 1,...,m_2, \\ & (\lambda_l)_j \geq 0, \quad j = 1,...,m_2, \\ & t \geq 0 \;\; and \quad v \geq 0, \end{array} \tag{1.2}$$

where $\lambda_l \in \Re^{m_2}$ is a multiplier vector associated with inequality constraint $g_l(t,v)$. problem 1.2 is non-convex and non-differentiable, moreover the regularity assumptions which are needed to successfully handle smooth optimization problems are never satisfied and it is not good to use our approach to solve problem 1.2. Dempe [13] presents smoothing method for the NBLP problem and the same method is also presented in [28] for programming with complementary constraints. Following this smoothing method we can propose our approach for the NBLP problem. Before presenting our approach for the NBLP problem, we give some definitions firstly.

Definition 1.1. *The Fischer-Burmeister function is* $\Psi(\tilde{a},\tilde{b}) : \Re^2 \to \Re$ *defined by* $\Psi(\tilde{a},\tilde{b}) = \tilde{a} + \tilde{b} - \sqrt{\tilde{a}^2 + \tilde{b}^2}$ *and the perturbed Fischer-Burmeister function is* $\Psi(\tilde{a},\tilde{b},\hat{\varepsilon}) : \Re^3 \to \Re$ *defined by* $\Psi(\tilde{a},\tilde{b},\epsilon) = \tilde{a} + \tilde{b} - \sqrt{\tilde{a}^2 + \tilde{b}^2 + \epsilon}$.

The Fischer-Burmeister function has the property that $\Psi(\tilde{a},\tilde{b}) = 0$ if and only if $\tilde{a} \geq 0$, $\tilde{b} \geq 0$, and $\tilde{a}\tilde{b} = 0$. It is non-differentiable at $\tilde{a} = \tilde{b} = 0$. Its perturbed variant satisfies $\Psi(\tilde{a},\tilde{b},\epsilon) = 0$ if and only if $\tilde{a} > 0$, $\tilde{b} > 0$, and $\tilde{a}\tilde{b} = \frac{\epsilon}{2}$ for $\epsilon > 0$. This function is smooth with respect to $\tilde{a}$,$\tilde{b}$, for $\epsilon > 0$. for more details see [8–10, 28].

In this paper, to allow the proposed algorithm ACBTR solve the NBLP problem 1.1 and satisfy the asymptotic stability conditions, we use the following changed perturbed Fischer-Burmeister function:

$$\tilde{\Psi}(\tilde{a}, \tilde{b}, \epsilon) = \sqrt{\tilde{a}^2 + \tilde{b}^2 + \epsilon} - \tilde{a} - \tilde{b}. \tag{1.3}$$

It is obvious that the changed perturbed Fischer-Burmeister function $\tilde{\Psi}(\tilde{a}, \tilde{b}, \epsilon)$ has the same property with the function $\Psi(\tilde{a}, \tilde{b}, \epsilon)$. Using the Fischer-Burmeister function 1.3, problem 1.2 equivalents to the following single objective constrained optimization problem

$$\begin{array}{ll} \min_{t,v} & f_u(t, v) \\ s.t. & g_u(t, v) \leq 0, \\ & \nabla_v f_l(t, v) + \nabla_v g_l(t, v)\mu = 0, \\ & \sqrt{g_{l_j}^2 + (\lambda_l)_j^2 + \epsilon} - (\lambda_l)_j + g_{l_j} = 0, \quad j = 1, ..., m_2, \\ & t \geq 0 \;\; and \;\; v \geq 0. \end{array} \tag{1.4}$$

Let $x = (t, v)^T$, $m = n_2 + m_2$ then the above problem can be written as SONP problem as follows

$$\begin{array}{ll} minimize & f_u(x) \\ subject\ to & h_l(x) = 0, \\ & g_u(x) \leq 0, \\ & x \geq 0, \end{array} \tag{1.5}$$

where $f_u : \Re^n \to \Re$, $h_l : \Re^n \to \Re^m$, and $g_u : \Re^n \to \Re^{m_1}$ are at least twice continuously differentiable functions.

Various approaches have been proposed to solve the SONP problem 1.5, see [4, 7, 16–19, 24]. In this paper, we use an active-set with barrier method to reduce SONP problem 1.5 to equivalent equality constrained optimization problem. So, we can use one of methods which are used for solving equality constrained optimization problem.

In this paper, we use a trust-region technique which is successful approach for solving SONP problem and is very important to ensure global convergence from any starting point. The trust-region strategy can induce strongly global convergence. It is more robust when it deals with rounding errors. It does not require the Hessian of the objective function must be positive definite or the objective function of the model must be convex. Also, some criteria are used to test the trial step is acceptable or no. If it is not acceptable, then the subproblem must be resolved with a reduced the trust-region radius. For the detailed expositions, the reader review [17, 20–24, 33, 36, 45–48].

A projected Hessian method which is suggested by [6, 38] and used by [19, 20, 22], utilizes in this paper to treat the difficulty of having an infeasible trust-region subproblem. In this method, the trial step is decomposed into two components and each component is computed by solving a trust-region unconstrained subproblem.

Under standard five assumptions, a global convergence theory for ACBTR algorithm is introduced. Moreover, numerical experiments display that ACBTR algorithm performers effectively and efficiently in pursuance.

The balance of this paper is organized as follows. A detailed description for the proposed method to solve SONP problem 1.5 is introduced in the next section. Section 3 is devoted to analysis of the global convergence of ACBTR algorithm. In Section 4, we report preliminary numerical results. Finally, some further remark is given in Section 5.

Notations: We use $\|.\|$ to denote the Euclidean norm $\|.\|_2$. The $i-th$ component of any vector such as x is written as $x^{(i)}$. The j^{th} trial iterate of iteration k is denoted by k^j. Subscript k refers to iteration indices. For example, $f_{u_k} \equiv f_u(x_k)$, $h_{l_k} \equiv h_l(x_k)$, $g_{u_k} \equiv g_u(x_k)$, $W_k \equiv W(x_k)$, $\nabla_x L^s_k \equiv \nabla_x L^s(x_k, \lambda_k; \sigma_k)$, and so on to denote the function value at a particular point.

2. An active-set with barrier method and trust-region strategy

In this section, firstly, we will introduce the detailed description for the active-set strategy with barrier method to reduce SONP problem 1.5 to equality constrained optimization problem. Secondly, to solve the equality constrained optimization problem and guarantee convergence from any starting point, we will introduce the detailed description for the trust-region algorithm. Finally, we will introduce the main steps for the main algorithm ACBTR to solve NBLP problem 1.1.

2.1. An active-set strategy and barrier method

Motivated by the active-set strategy which is introduced by [12] and used by [17–21], we define a 0-1 diagonal matrix $W(x) \in \Re^{m_2 \times m_2}$ whose diagonal entries are

$$w_i(x) = \begin{cases} 1, & \text{if } g_{u_i}(x) \geq 0, \\ 0, & \text{if } g_{u_i}(x) < 0, \end{cases} \tag{2.1}$$

where $i = 1, ..., m_2$. By Using the diagonal matrix $W(x) \in \Re^{m_2 \times m_2}$, we can transform problem 1.5 to the following equality constrained optimization problem with positive variables

$$\begin{array}{ll} minimize_x & f_u(x) \\ subject\ to & h_l(x) = 0, \\ & g_u(x)^T W(x) g_u(x) = 0, \\ & x \geq 0. \end{array}$$

Penalty methods usually more suitable on problems with equality constraints. These methods are usually generate a sequence of points that converges to a solution of the problem from the exterior of the feasible region. An advantage of penalty methods is that they do not request the iterates to be strictly feasible. In this paper we use the penalty method to reduce the above problem to the following equality constrained optimization problem with positive variables

$$\begin{array}{ll} minimize_x & f_u(x) + \frac{\sigma}{2}\|W(x)g_u(x)\|^2 \\ subject\ to & h_l(x) = 0, \\ & x \geq 0, \end{array} \tag{2.2}$$

where σ is a positive parameter. Let $F^+(x) = \{x | x > 0\}$.

Motivated by the barrier method which is discussed in [[7, 26, 43], problem 2.2, for any $x \in F^+$ can be written as follows

$$\begin{array}{ll} minimize_x & f_u(x) - s\sum_{i=1}^{n} \ln(x^{(i)}) + \frac{\sigma}{2}\|W(x)g_u(x)\|^2 \\ subject\ to & h_l(x) = 0, \end{array} \tag{2.3}$$

for decreasing sequence of barrier parameters s converging to zero, see [26].

The Lagrangian function associated with problem 2.3 is

$$L^s(x,\lambda;\sigma) = f_u(x) - s\sum_{i=1}^{n}\ln(x^{(i)}) + \lambda^T h_l(x) + \frac{\sigma}{2}\|W(x)g_u(x)\|^2, \tag{2.4}$$

where $\lambda \in \Re^m$ is a multiplier vector associated with the equality constraint $h_l(x) = 0$.

The first-order necessary condition for the strictly positive point x_* to be a local minimizer of problem 2.3 is that there exists a Lagrange multiplier vector $\lambda_* \in \Re^m$, such that (x_*, λ_*) satisfies the following nonlinear system

$$\begin{aligned} \nabla f_u(x_*) - sX_*^{-1}e + \nabla h_l(x_*)\lambda_* + \sigma\nabla g_u(x_*)W(x_*)g_u(x_*) &= 0 \\ h_l(x_*) &= 0, \end{aligned}$$

where X is diagonal matrix whose diagonal entries are $(x_1,...,x_n) \in F^+$. Let $sX_*^{-1}e = y \in \Re^n$ be an auxiliary variable, then the above system can be written as follows

$$\nabla f_u(x_*) - y_* + \nabla h_l(x_*)\lambda_* + \sigma\nabla g_u(x_*)W(x_*)g_u(x_*) = 0, \tag{2.5}$$

$$X_* y_* - se = 0, \tag{2.6}$$

$$h_l(x_*) = 0, \tag{2.7}$$

where $x_* \in F^+$. The conditions [2.5–2.7] are called the barrier KKT conditions. For more details see [26].

Applying Newton's method to the nonlinear system (2.5)–(2.7), we have

$$\begin{pmatrix} H + \sigma\nabla g_u(x)W(x)\nabla g_u(x)^T & \nabla h_l(x) & -I \\ Y & 0 & X \\ \nabla h_l(x)^T & 0 & 0 \end{pmatrix}\begin{pmatrix} d_x \\ d_\lambda \\ d_y \end{pmatrix} = -\begin{pmatrix} \nabla_x L^s(x,\lambda;\sigma) \\ Xy - se \\ h_l(x) \end{pmatrix}, \tag{2.8}$$

where H is the Hessian matrix of the following function or an approximation to it

$$\ell^s(x,\lambda) = f_u(x) + \lambda^T h_l(x) - s\sum_{i=1}^{n}\ln(x^{(i)}).$$

The matrix Y is a diagonal matrix whose diagonal entries are $(y_1,...,y_n)$ and $\nabla_x L^s(x,\lambda;\sigma) = \nabla f_u(x) - y + \nabla h_l(x)\lambda + \sigma\nabla g_u(x)W(x)g_u(x)$.

From the second equation of the system (2.8) we have

$$d_y = -y + sX^{-1}e - X^{-1}Yd_x. \tag{2.9}$$

To decrease the dimension of system 2.8, we eliminate d_y from the first equation of the system 2.8 by using Eq 2.9 as follows

$$(H + \sigma\nabla g_u(x)W(x)\nabla g_u(x)^T)dx + \nabla h_l(x)d_\lambda - I(-y + sX^{-1}e - X^{-1}Yd_x) = -\nabla_x L^s(x,\lambda;\sigma)$$

Using Eq 2.6, we have the following system

$$\begin{pmatrix} B & \nabla h_l(x) \\ \nabla h_l(x)^T & 0 \end{pmatrix}\begin{pmatrix} d_x \\ d_\lambda \end{pmatrix} = -\begin{pmatrix} \nabla_x L^s(x,\lambda;\sigma) \\ h_l(x) \end{pmatrix}. \tag{2.10}$$

where, $B = H + X^{-1}Y + \sigma\nabla g_u(x)W(x)\nabla g_u(x)^T$.

We notice that, the system 2.10 is equivalent to the first order necessary condition for the following sequential quadratic programming problem

$$\begin{array}{ll} minimize & L^s(x,\lambda;\sigma) + \nabla_x L^s(x,\lambda;\sigma)^T d + \frac{1}{2} d^T B d \\ subject\ to & h_I(x) + \nabla h_I(x)^T d = 0. \end{array} \tag{2.11}$$

That is, the point (x_*, λ_*) that satisfies the KKT conditions for subproblem 2.11 will satisfy the KKT conditions for problem 1.5. A methods which are used to solve subproblem 2.11 is a local methods. That is, it may not converge to a stationary point if the starting point is far away from the solution. To guarantee convergence from any starting point, we use the trust-region technique.

2.2. Trust-region strategy

By using trust-region technique to ensure convergence of subproblem 2.11 and estimate the step d_k, we solve the following subproblem

$$\begin{array}{ll} minimize & \nabla_x L_k^{sT} d + \frac{1}{2} d^T B_k d \\ subject\ to & h_I(x_k) + \nabla h_I(x_k)^T d = 0, \\ & \|d\| \le \delta_k, \end{array} \tag{2.12}$$

where $0 < \delta_k$ represents the radius of the trust-region. The subproblem 2.12 may be infeasible because there may be no intersecting points between hyperplane of the linearized constraints $h_I(x) + \nabla h_I(x)^T d$ and the constraint $\|d\| \le \delta_k$. Even if they intersect, there is no guarantee that this will keep true if δ_k is reduced, see [11]. So, a projected Hessian technique is used in our approach to overcome this problem. This technique was suggested by [6, 38] and used by [19, 20, 22]. In this technique, the trial step d_k is decomposed into two orthogonal components: the normal component d_k^n to improve feasibility and the tangential component d_k^t to improve optimality. Each of d_k^n and d_k^t is evaluated by solving unconstrained trust-region subproblem.

- **To compute the normal component d^n**

$$\begin{array}{ll} minimize & \|h_{I_k} + \nabla h_{I_k}^T d^n\|^2 \\ subject\ to & \|d^n\| \le \zeta\delta_k, \end{array} \tag{2.13}$$

for some $\zeta \in (0,1)$. To solve the subproblem 2.13, we use a conjugate gradient method which is introduced by [39] and used by [23], see Algorithm 2.1 in [23]. It is very cheap if the problem is large-scale and the Hessian is indefinite. By using the conjugate gradient method, the following condition is hold

$$\|h_{I_k}\|^2 - \|h_{I_k} + \nabla h_{I_k}^T d_k^n\|^2 \ge \vartheta_1\{\|h_{I_k}\|^2 - \|h_{I_k} + \nabla h_{I_k}^T d_k^{ncp}\|^2\}, \tag{2.14}$$

for some $\vartheta_1 \in (0,1]$. That is, the normal predicted decrease obtained by the normal component d_k^n is greater than or equal to a fraction of the normal predicted decrease obtained by the normal Cauchy step d_k^{ncp}. The normal Cauchy step d_k^{ncp} is defined as

$$d_k^{ncp} = -\alpha_k^{ncp} \nabla h_{I_k} h_{I_k}, \tag{2.15}$$

where the parameter α_k^{ncp} is given by

$$\alpha_k^{ncp} = \begin{cases} \frac{\|\nabla h_{I_k} h_{I_k}\|^2}{\|(\nabla h_{I_k})^T \nabla h_{I_k} h_{I_k}\|^2} & \text{if } \frac{\|\nabla h_{I_k} h_{I_k}\|^3}{\|\nabla h_{I_k}^T \nabla h_{I_k} h_{I_k})\|^2} \leq \delta_k \\ & \text{and } \|\nabla h_{I_k}^T \nabla h_{I_k} h_{I_k})\| > 0, \\ \frac{\delta_k}{\|\nabla h_{I_k} h_{I_k}\|} & \text{otherwise.} \end{cases} \tag{2.16}$$

Once d_k^n is estimated, we will compute $d_k^t = Z_k \bar{d}_k^t$. A matrix Z_k is the matrix whose columns form a basis for the null space of $(\nabla h_{I_k})^T$.

- **To compute the tangential component** d_k^t

To estimate the tangential component d_k^t, let

$$q(d) = L^s(x, \lambda; \sigma) + \nabla_x L^s(x, \lambda; \sigma)^T d + \frac{1}{2} d^T B d. \tag{2.17}$$

and using the conjugate gradient method [23] to solve the following trust-region subproblem

$$\begin{array}{ll} minimize & [Z_k^T \nabla q_k(d_k^n)]^T \bar{d}^t + \frac{1}{2} \bar{d}^{t^T} Z_k^T B_k Z_k \bar{d}^t \\ subject\ to & \|Z_k \bar{d}^t\| \leq \Delta_k, \end{array} \tag{2.18}$$

where $\nabla q_k(d_k^n) = \nabla_x L_k^s + B_k d_k^n$ and $\Delta_k = \sqrt{\delta_k^2 - \|d_k^n\|^2}$.

Let a tangential predicted decrease which is obtained by the tangential component d_k^t be

$$Tpred_k(\bar{d}_k^t) = q_k(d_k^n) - q_k(d_k^n + Z_k \bar{d}_k^t). \tag{2.19}$$

Since the conjugate gradient method is used to solve subproblem (2.18) and estimate d_k^t, then the following condition holds

$$Tpred_k(\bar{d}_k^t) \geq \vartheta_2 \ \ Tpred_k(\bar{d}_k^{tcp}), \tag{2.20}$$

for some $\vartheta_2 \in (0, 1]$. This condition clarified that the tangential predicted decrease which is obtained by tangential step $\bar{d}_k^t$ is greater than or equal to a fraction of the tangential predicted decrease obtained by a tangential Cauchy step $\bar{d}_k^{tcp}$. The tangential Cauchy step $\bar{d}_k^{tcp}$ is defined as follows

$$\bar{d}_k^{tcp} = -\alpha_k^{tcp} Z_k^T \nabla q_k(d_k^n), \tag{2.21}$$

where the parameter α_k^{tcp} is given by

$$\alpha_k^{tcp} = \begin{cases} \frac{\|Z_k^T \nabla q_k(d_k^n)\|^2}{(Z_k^T \nabla q_k(d_k^n))^T \bar{B}_k Z_k^T \nabla q_k(d_k^n)} & \text{if } \frac{\|Z_k^T \nabla q_k(d_k^n)\|^3}{(Z_k^T \nabla q_k(d_k^n))^T \bar{B}_k Z_k^T \nabla q_k(d_k^n)} \leq \Delta_k \\ & \text{and } (Z_k^T \nabla q_k(d_k^n))^T \bar{B}_k Z_k^T \nabla q_k(d_k^n) > 0, \\ \frac{\Delta_k}{\|Z_k^T \nabla q_k(d_k^n)\|} & \text{otherwise,} \end{cases} \tag{2.22}$$

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

Once estimating d_k^t, we set $d_k = d_k^n + d_k^t$ and $x_{k+1} = x_k + d_k$. To guarantee that $x_{k+1} \in F^+$ at every iteration k, we need to evaluate the damping parameter μ_k.

- **To estimate the damping parameter μ_k**

The damping parameter μ_k is defined as follows:

$$\mu_k = \min\{\min_i\{\theta_k^{(i)}\}, 1\}, \tag{2.23}$$

where

$$\theta_k^{(i)} = \begin{cases} \frac{-x_k^{(i)}}{d_k^{(i)}}, & \text{if } d_k^{(i)} < 0 \\ 1 & \text{otherwise.} \end{cases}$$

To be decided whether the scale step $\mu_k d_k$ will be accepted or no, we need to a merit function. The merit function is the function which is tie the objective function $f_u(x)$ with the constraints $h_l(x)$ and $g_u(x)$ in such a way that progress in the merit function means progress in solving problem. In the proposed algorithm, we use the following an augmented Lagrange function as a merit function, see [31].

$$\Phi^s(x, \lambda; \sigma; \rho) = \ell^s(x, \lambda) + \frac{\sigma}{2}\|W(x)g_u(x)\|^2 + \rho\|h_l(x)\|^2, \tag{2.24}$$

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

To be decided whether the point $(x_k + \mu_k d_k, \lambda_{k+1})$ will be taken as a next iterate or no, we need to define the actual reduction $Ared_k$ and the predicted reduction $Pred_k$ in the merit function $\Phi^s(x, \lambda; \sigma; \rho)$.

In the proposed algorithm, $Ared_k$ is defined as follows

$$Ared_k = \Phi^s(x_k, \lambda_k; \sigma_k; \rho_k) - \Phi^s(x_k + \mu_k d_k, \lambda_{k+1}; \sigma_k; \rho_k).$$

Also $Ared_k$ can be written as follows,

$$Ared_k = \ell^s(x_k, \lambda_k) - \ell^s(x_{k+1}, \lambda_k) - \Delta\lambda_k^T h_{l_{k+1}} + \frac{\sigma_k}{2}[\|W_k g_u(x_k)\|^2 - \|W_{k+1} g_{u_{k+1}}\|^2] + \rho_k[\|h_{l_k}\|^2 - \|h_{l_{k+1}}\|^2], \tag{2.25}$$

where $\Delta\lambda_k = (\lambda_{k+1} - \lambda_k)$.

In the proposed algorithm, $Pred_k$ is defined as follows

$$\begin{aligned} Pred_k = {} & -\nabla_x \ell^s(x_k, \lambda_k)^T \mu_k d_k - \frac{1}{2}\mu_k^2 d_k^T \tilde{H}_k d_k - \Delta\lambda_k^T (h_{l_k} + \nabla h_{l_k}^T \mu_k d_k) \\ & + \frac{\sigma_k}{2}[\|W_k g_u(x_k)\|^2 - \|W_k(g_u(x_k) + \nabla g_u(x_k)^T \mu_k d_k)\|^2] \\ & + \rho_k[\|h_{l_k}\|^2 - \|h_{l_k} + \nabla h_{l_k}^T \mu_k d_k\|^2]. \end{aligned} \tag{2.26}$$

where $\nabla_x l^s(x, \lambda) = \nabla f_u(x) - y + \nabla h_l(x)\lambda$ and $\tilde{H} = H + X^{-1}Y$.

Also, $Pred_k$ can be written as follows

$$Pred_k = q_k(0) - q_k(\mu_k d_k) - \Delta\lambda_k^T (h_{l_k} + \nabla h_{l_k}^T \mu_k d_k) + \rho_k[\|h_{l_k}\|^2 - \|h_{l_k} + \nabla h_{l_k}^T \mu_k d_k\|^2]. \tag{2.27}$$

where the quadratic form $q(d)$ in 2.17 can be written as follows

$$q(d) = \ell^s(x, \lambda) + \nabla_x \ell^s(x, \lambda)^T d + \frac{1}{2} d^T \tilde{H} d + \frac{\sigma}{2}[\|W(x) g_u(x)\|^2 - \|W(x)(g_u(x) + \nabla g_u(x)^T d)\|^2]. \tag{2.28}$$

- **To update** ρ_k

To ensure that $Pred_k \geq 0$, we update the penalty parameter ρ_k utilizing the following scheme.

Algorithm 2.1. *If*

$$Pred_k \leq \frac{\rho_k}{2}[\|h_{l_k}\|^2 - \|h_{l_k} + \nabla h_{l_k}^T \mu_k d_k\|^2], \tag{2.29}$$

then, set

$$\rho_k = \frac{2[q_k(\mu_k d_k) - q_k(0) + \Delta\lambda_k^T(h_{l_k} + \nabla h_{l_k}^T \mu_k d_k)]}{\|h_{l_k}\|^2 - \|h_{l_k} + \nabla h_{l_k}^T \mu_k d_k\|^2} + \beta_0, \tag{2.30}$$

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

$$\rho_{k+1} = \max\{\rho_k, \sigma_k^2\}. \tag{2.31}$$

End if.

For more details, see [15–19].

- **To test the scaling step** $\mu_k d_k$ **and update** δ_k

The framework to test the scaling step $\mu_k d_k$ and update δ_k is presented in the following algorithm.

Algorithm 2.2. *Choose* $0 < \gamma_1 < \gamma_2 < 1$, $0 < \alpha_1 < 1 < \alpha_2$, *and* $\delta_{min} \leq \delta_0 \leq \delta_{max}$.
While $\frac{Ared_k}{Pred_k} \in (0, \gamma_1)$ *or* $Pred_k \leq 0$.
Set $\delta_k = \alpha_1 \|d_k\|$ *and return to evaluate a new trial step and end while.*
If $\frac{Ared_k}{Pred_k} \in [\gamma_1, \gamma_2)$. *Set* $x_{k+1} = x_k + \mu_k d_k$ *and* $\delta_{k+1} = \max(\delta_k, \delta_{min})$.
End if.
If $\frac{Ared_k}{Pred_k} \in [\gamma_2, 1]$. *Set* $x_{k+1} = x_k + \mu_k d_k$ *and* $\delta_{k+1} = \min\{\delta_{max}, \max\{\delta_{min}, \alpha_2 \delta_k\}\}$.
End if.

- **To update the positive parameter** σ_k

To update the positive parameter σ_k, we use the following scheme.

Algorithm 2.3. *If*

$$\frac{1}{2} Tpred_k(\bar{d}_k^t) \geq \|\nabla g_u(x_k) W_k g_u(x_k)\| \min\{\|\nabla g_u(x_k) W_k g_u(x_k)\|, \Delta_k\}. \tag{2.32}$$

Set $\sigma_{k+1} = \sigma_k$.
Else, set $\sigma_{k+1} = 2\sigma_k$. *End if.*

For more details see [18, 23].
Finally, the algorithm is stopped when $\|Z_k^T \nabla_x \ell_k^s\| + \|\nabla g_u(x_k) W_k g_u(x_k)\| + \|h_{l_k}\| \leq \varepsilon_1$ or $\|d_k\| \leq \varepsilon_2$, for some $\varepsilon_1, \varepsilon_2 > 0$.

- **A trust-region algorithm**

The framework of the trust-region algorithm to solve subproblem 2.12 are summarized as follows.

Algorithm 2.4. *(Trust-region algorithm)*
Step 0. *Starting with* $x_0 \in F^+$. *Evaluate* y_0 *and* λ_0. *Set* $s_0 = 0.1$, $\rho_0 = 1$, $\sigma_0 = 1$, *and* $\beta_0 = 0.1$. *Choose* ε_1, ε_2, α_1, α_2, γ_1, *and* γ_2 *such that* $0 < \varepsilon_1$, $0 < \varepsilon_2$, $0 < \alpha_1 < 1 < \alpha_2$, *and* $0 < \gamma_1 < \gamma_2 < 1$. *Choose* δ_{min}, δ_{max}, *and* δ_0 *such that* $\delta_{min} \le \delta_0 \le \delta_{max}$. *Set* $k = 0$.
Step 1. *If* $\|Z_k^T \nabla_x \ell_k^s\| + \|\nabla g_u(x_k) W_k g_u(x_k)\| + \|h_{l_k}\| \le \varepsilon_1$, *then stop.*
Step 2. *(How to compute* d_k*)*

a). Evaluate the normal component d_k^n *by solving subproblem (2.13).*
b). Evaluate the tangential component $\bar{d}_k^t$ *by solving subproblem (2.18).*
c). Set $d_k = d_k^n + Z_k \bar{d}_k^t$.

Step 3. *If* $\|d_k\| \le \varepsilon_2$, *then stop.*
Step 4. *(How to compute* μ_k*)*

a). Compute the damping parameter μ_k *using (2.23).*
b). Set $x_{k+1} = x_k + \mu_k d_k$.

Step 5. *Compute the vector* y_{k+1}, *by using the following equation*

$$y_{k+1} = s_k X_k^{-1} e - X_k^{-1} Y_k \mu_k d_k. \tag{2.33}$$

The above equation is obtained from (2.9).
Step 6. *Compute* W_{k+1} *given by (2.1).*
Step 7. *Evaluate* λ_{k+1} *by solving the following subproblem*

$$minimize \ \ \|\nabla f_{k+1} - y_{k+1} + \nabla h_{l_{k+1}} \lambda + \rho_k \nabla g_{u_{k+1}} W_{k+1} g_{u_{k+1}}\|^2. \tag{2.34}$$

Step 8. *Using scheme 2.1 to update the penalty parameter* ρ_k.
Step 9. *Using Algorithm (2.2) to test the scaled step* $\mu_k d_k$ *and update the radius* δ_k.
Step 10. *Update the positive parameter* σ_k *using scheme 2.3.*
Step 11. *To Update the barrier parameter* s_k, *set* $s_{k+1} = \frac{s_k}{10}$.
Step 12. *Set* $k = k + 1$ *and go to Step 1.*

In the following subsection we will clarify the main steps for solving NBLP problem 1.1.

2.3. An Active-set-barrier-trust-region algorithm

The framework to solve NBLP problem 1.1 are summarized in the following algorithm.

Algorithm 2.5. *(An active-set-barrier-trust-region (ACBTR) algorithm)*
Step 1. *Using KKT optimality conditions for the lower level problem 1.1 to reduce problem 1.1 to one-level problem 1.2.*
Step 2. *Using Fischer-Burmeister function 1.3 with* $\epsilon = 0.001$ *to obtain the smooth problem 1.4 and which is equivalent problem 1.5.*
Step 3. *Using An active set strategy with Barrier method to obtain subproblem 2.11.*
Step 4. *Using trust-region Algorithm 2.4 to solve subproblem 2.11 and obtained approximate solution for problem 1.5.*

In the following section we will introduce a global convergence analysis for ACBTR algorithm.

3. Global convergence analysis for ACBTR algorithm

let Ω be a convex subset of $\Re^n$ that contains all iterates $x_k \in F^+$ and $(x_k + \mu_k d_k) \in F^+$. To prove the global convergence theory of ACBTR algorithm on Ω, we assume that the following assumptions are hold.

- **Assumptions**

$[A_1]$. The functions $f_u(x)$, $h_l(x)$, and $g_u(x)$ are twice continuously differentiable function for all $0 < x \in S$.
$[A_2]$. All of $f_u(x)$, $\nabla f_u(x)$, $\nabla^2 f_u(x)$, $g_u(x)$, $\nabla g_u(x)$, $h_l(x)$, $\nabla h_l(x)$, $\nabla^2 h_{l_i}(x)$ for $i = 1, ..., m$, and $(\nabla h_{l_k})[(\nabla h_{l_k})^T(\nabla h_{l_k})]^{-1}$ are uniformly bounded in S.
$[A_3]$. The columns of the matrix $\nabla h_l(x)$ are linearly independent.
$[A_4]$. The sequence $\{\lambda_k\}$ is bounded.
$[A_5]$. The sequence of matrices $\{\tilde{H}_k\}$ is bounded.

In the above assumptions, even though we assume that $\nabla h_l(x)$ has full column rank for all $x_k \in F^+$, we do not require $\nabla g_u(x)$ has full column rank for all $x_k \in F^+$. So, we may have other kinds of stationary points which are presented in the following definitions.

Definition 3.1. *A point $x_* \in F^+$ is called a Fritz John (FJ) point if there exist γ_*, λ_*, and ν_*, not all zeros, such that*

$$\begin{aligned}
\tau_* \nabla f(x_*) + \nabla h_l(x_*)\lambda_* + \nabla g_u(x_*)\nu_* &= 0, && (3.1)\\
h_l(x_*) &= 0, && (3.2)\\
W_* g_u(x_*) &= 0, && (3.3)\\
(\nu_*)_i (g_u(x_*))_i &= 0, \quad i = 1, ..., m_2, && (3.4)\\
\tau_*,\ (\nu_*)_i &\geq 0, \quad i = 1, ..., m_2. && (3.5)
\end{aligned}$$

Equations (3.1)–(3.5) are called FJ conditions. More details see [4].

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

Definition 3.2. *A point $x_* \in F^+$ is called an infeasible Fritz John (IFJ) point if there exist τ_*, λ_*, and ν_* such that*

$$\begin{aligned}
\tau_* \nabla f_u(x_*) + \nabla h_l(x_*)\lambda_* + \nabla g_u(x_*)\nu_* &= 0, && (3.6)\\
h_l(x_*) &= 0, && (3.7)\\
\nabla g_u(x_*) W_* g_u(x_*) &= 0 \quad \textit{but} \quad \|W_* g_u(x_*)\| > 0, && (3.8)\\
(\nu_*)_i (g_u(x_*))_i &\geq 0, \quad i = 1, ..., m_2, && (3.9)\\
\tau_*,\ (\nu_*)_i &\geq 0, \quad i = 1, ..., m_2. && (3.10)
\end{aligned}$$

Equations (3.6)–(3.10) are called IFJ conditions.

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

Lemma 3.1. *Under assumptions A_1–A_5, a subsequence $\{x_{k_i}\}$ of the iteration sequence asymptotically satisfies IFJ conditions if it satisfies:*
1). $\lim_{k_i\to\infty} h_l(x_{k_i}) = 0$.
2). $\lim_{k_i\to\infty} \|W_{k_i} g_u(x_{k_i})\| > 0$.
3). $\lim_{k_i\to\infty} \left\{ min_{d\in\Re^{n-m_2}} \|W_{k_i}(g_{u_{k_i}} + \nabla g_{u_{k_i}}^T Z_{k_i}\mu_{k_i}\bar{d}^t)\|^2 \right\} = \lim_{k_i\to\infty} \|W_{k_i} g_{u_{k_i}}\|^2$.

Proof. To simplify the notations, let the subsequence $\{k_i\}$ be renamed to $\{k\}$. Let $\hat{d}_k$ be a minimizer of $minimize_{\bar{d}^t} \|W_k(g_u(x_k) + \nabla g_u(x_k)^T Z_k \mu_k \bar{d}^t)\|^2$, then it satisfies

$$Z_k^T \nabla g_u(x_k) W_k g_u(x_k)\mu_k + Z_k^T \nabla g_u(x_k) W_k \nabla g_u(x_k)^T Z_k \mu_k^2 \hat{d}_k = 0. \tag{3.11}$$

From condition 3, we have

$$\lim_{k\to\infty} \{2\mu_k \hat{d}_k{}^T Z_k^T \nabla g_u(x_k) W_k g_u(x_k) + \mu_k^2 \hat{d}_k{}^T Z_k^T \nabla g_u(x_k) W_k \nabla g_u(x_k)^T Z_k \hat{d}_k\} = 0. \tag{3.12}$$

Now, we will consider two cases:
Firstly, if $\lim_{k\to\infty} \hat{d}_k = 0$, then from (3.11) we have $\lim_{k\to\infty} \mu_k Z_k^T \nabla g_u(x_k) W_k g_u(x_k) = 0$.
Secondly, if $\lim_{k\to\infty} \hat{d}_k \neq 0$, then multiplying (3.11) from the left by $2\hat{d}_k^T$ and subtract it from the limit (3.12), we have $\lim_{k\to\infty} \|W_k \nabla g_u(x_k)^T Z_k \mu_k \hat{d}_k\|^2 = 0$. This implies $\lim_{k\to\infty} \mu_k Z_k^T \nabla g_u(x_k) W_k g_u(x_k) = 0$.

That is, in either case, we have

$$\lim_{k\to\infty} Z_k^T \nabla g_u(x_k) W_k g_u(x_k) = 0. \tag{3.13}$$

Take $(\nu_k)_i = (W_k g_u(x_k))_i$, $i = 1, ..., p$. Since $\lim_{k\to\infty} \|W_k g_u(x_k)\| > 0$, then $\lim_{k\to\infty} (\nu_k)_i \geq 0$, for $i = 1, ..., p$ and $\lim_{k\to\infty} (\nu_k)_i > 0$, for some i. Therefore $\lim_{k\to\infty} Z_k^T \nabla g_u(x_k)\nu_k = 0$. But this implies the existence of a sequence $\{\lambda_k\}$ such that $\lim_{k\to\infty}\{\nabla h_{l_k}\lambda_k + \nabla g_u(x_k)\nu_k\} = 0$. Thus IFJ conditions are hold in the limit with $\tau_* = 0$.

The following lemma clarify that, for any subsequence $\{x_{k_i}\}$ of the iteration sequence that asymptotically satisfies the FJ conditions, the corresponding subsequence of smallest singular values of $\{Z_k^T \nabla g_u(x_k) W_k\}$ is not bounded away from zero. That is, asymptotically the gradient of the active constraints are linearly dependent.

Lemma 3.2. *Under assumptions A_1–A_5, a subsequence $\{x_{k_i}\}$ of the iteration sequence asymptotically satisfies FJ conditions if it satisfies:*
1). $\lim_{k_i\to\infty} h(x_{k_i}) = 0$.
2). For all k_i, $\|W_{k_i} g_{u_{k_i}}\| > 0$ *and* $\lim_{k_i\to\infty} W_{k_i} g_{u_{k_i}} = 0$.
3). $\lim_{k_i\to\infty} \left\{ \min_{d\in\Re^{n-p}} \frac{\|W_{k_i}(g_{u_{k_i}} + \nabla g_{u_{k_i}}^T Z_{k_i}\mu_{k_i}\bar{d}^t)\|^2}{\|W_{k_i} g_{u_{k_i}}\|^2} \right\} = 1$.

Proof. The proof of this lemma is similar to the proof of Lemma 4.4 in [19].

In the following section, we introduce some basic lemmas which are requisite to prove global convergence analysis for ACBTR algorithm.

3.1. Basic lemmas

In this section, we introduce some significant lemmas which are required to prove global convergence theory for ACBTR algorithm.

Lemma 3.3. *Under assumptions A_1 and A_3, $W(x)g_u(x)$ is Lipschitz continuous in Ω.*

Proof. The proof of this lemma is similar to the proof of Lemma 4.1 of [12].

From the above lemma, we conclude that $g_u(x)^T W(x)g_u(x)$ is differentiable and $\nabla g_u(x)W(x)g_u(x)$ is Lipschitz continuous in Ω.

Lemma 3.4. *At any iteration k, let $E(x_k) \in \Re^{m_2\times m_2}$ be a diagonal matrix whose diagonal entries are*

$$(e_k)_i = \begin{cases} 1 & \text{if } (g_u(x_k))_i < 0 \text{ and } (g_{u_{k+1}})_i \geq 0, \\ -1 & \text{if } (g_u(x_k))_i \geq 0 \text{ and } (g_{u_{k+1}})_i < 0, \\ 0 & \text{otherwise}, \end{cases} \tag{3.14}$$

where $i = 1, 2, ..., m_2$. Then

$$W_{k+1} = W_k + E_k. \tag{3.15}$$

Proof. See Lemma 6.2 of [17].

Lemma 3.5. *Under assumptions A_1–A_3, there exists at any iteration k, a constant $C_1 > 0$ independent of k such that*

$$\|E_k g_u(x_k)\| \leq C_1\|d_k\|, \tag{3.16}$$

where $E_k \in \Re^{m_2\times m_2}$ is the diagonal matrix whose diagonal entries are defined in (3.14).

Proof. See Lemma 6.3 of [17].

Lemma 3.6. *Under assumptions A_1–A_3, there exists at any iteration k, a constant $0 < C_2$ independent of k such that*

$$\|d_k^n\| \leq C_2\|h_{l_k}\|. \tag{3.17}$$

Proof. Since d_k^n is normal to the tangent space, then we have

$$\begin{aligned} \|d_k^n\| &= \|\nabla h_{l_k}(\nabla h_{l_k}^T \nabla h_{l_k})^{-1}\nabla h_{l_k}^T d_k\| \\ &= \|\nabla h_{l_k}(\nabla h_{l_k}^T \nabla h_{l_k})^{-1}[h_{l_k} + \nabla h_{l_k}^T d_k - h_{l_k}]\| \\ &\leq \|\nabla h_{l_k}(\nabla h_{l_k}^T \nabla h_{l_k})^{-1}\|[\|h_{l_k} + \nabla h_{l_k}^T d_k\| + \|h_{l_k}\|] \\ &\leq \|\nabla h_{l_k}(\nabla h_{l_k}^T \nabla h_{l_k})^{-1}\|\|h_{l_k}\|. \end{aligned}$$

where $\|h_{l_k} + \nabla h_{l_k}^T d_k\| \leq \|h_{l_k}\|$. Using the assumptions A_1–A_5, we have the desired result.

The next lemma clarifies how delicate the definition of $Ared_k$ is as an approximation to $Pred_k$.

Lemma 3.7. *Under assumptions A_1–A_5, there exists a constant $0 < C_3$, such that*

$$|Ared_k - Pred_k| \leq C_3\mu_k\rho_k\|d_k\|^2. \tag{3.18}$$

Proof. From the definition of $Ared_k$ (2.25) and using (3.15), we have

$$\begin{aligned} Ared_k &= \ell^s(x_k, \lambda_k) - \ell^s(x_{k+1}, \lambda_k) - \Delta\lambda_k^T h_{l_{k+1}} \\ &+ \frac{\sigma_k}{2}[g_u(x_k)^T W_k g_u(x_k) - g_{u_{k+1}}^T (W_k + E_k)g_{u_{k+1}}] + \rho_k[\|h_{l_k}\|^2 - \|h_{l_{k+1}}\|^2]. \end{aligned}$$

From the above equation, the definition of $Pred_k$ (2.26), and using the inequality of Cauchy-Schwarz, we have

$$\begin{aligned}
|Ared_k - Pred_k| \leq\ & \frac{1}{2}\mu_k^2|d_k^T[H_{l_k} - \nabla^2\ell^s(x_k+\xi_1 d_k)]d_k| + \frac{\mu_k^2}{2}\mid d_k^T X_k^{-1}Y_k d_k\mid \\
&+\sigma_k\mu_k \mid (\nabla g_u(x_k) - \nabla g(x_k+\xi_2\mu_k d_k))W_k g_u(x_k)d_k \mid \\
&+\frac{\sigma_k\mu_k^2}{2}\mid d_k^T[\nabla g_u(x_k)W_k\nabla g_u(x_k)^T - \nabla g(x_k+\xi_2 d_k)W_k\nabla g(x_k+\xi_2\mu_k d_k)^T]d_k \mid \\
&+\frac{\sigma_k\mu_k^2}{2}\|E_k g_u(x_k)\|^2 + \sigma_k\mu_k \mid \nabla g(x_k+\xi_2\mu_k d_k)E_k g_u(x_k)d_k \mid \\
&+\frac{\sigma_k\mu_k^2}{2}\mid d_k^T[\nabla g(x_k+\xi_2 d_k)E_k\nabla g(x_k+\xi_2\mu_k d_k)^T]d_k \mid \\
&+\mu_k|\Delta\lambda_k[\nabla h_{l_k} - \nabla h(x_k+\xi_2 d_k)]^T d_k| \\
&+2\rho_k\mu_k|[(\nabla h_{l_k} - \nabla h(x_k+\xi_2\mu_k d_k))h_{l_k}]^T d_k| \\
&+\rho_k\mu_k^2|d_k^T[\nabla h_{l_k}\nabla h_{l_k}^T - \nabla h(x_k+\xi_2\mu_k d_k)\nabla h(x_k+\xi_2\mu_k d_k)^T]d_k|,
\end{aligned}$$

for some ξ_1 and $\xi_2 \in (0,1)$. By using assumptions A_1–A_5, $\rho_k \geq \sigma_k$, $\rho_k \geq 1$, and inequality (3.16), we have

$$|Ared_k - Pred_k| \leq \mu_k[\kappa_1\|d_k\|^2 + \kappa_2\rho_k\|d_k\|^3 + \kappa_3\rho_k\|d_k\|^2\|h_{l_k}\|], \tag{3.19}$$

where κ_1, κ_2, and κ_3 are positive constants. Since $\rho_k \geq 1$, $\|d_k\| \leq \delta_{max}$, and $\|h_{l_k}\|$ is uniformly bounded, then inequality (3.18) hold.

The proof of the following two lemmas depends on the fact that d_k^n and $\bar{d}_k^t$ satisfy the condition of the fraction of Cauchy decrease.

Lemma 3.8. *Under assumptions A_1–A_5, there exists a constant $0 < C_4$ such that*

$$\|h_{l_k}\|^2 - \|h_{l_k} + \nabla h_{l_k}^T d_k^n\|^2 \geq C_4\|h_{l_k}\|\min\{\|h_{l_k}\|, \delta_k\}. \tag{3.20}$$

Proof. From the definition of the normal Cauchy step (2.15), we will consider two cases:
Firstly, if $d_k^{ncp} = -\frac{\delta_k}{\|\nabla h_{l_k}h_{l_k}\|}(\nabla h_{l_k}h_{l_k})$ and $\delta_k\|\nabla h_{l_k}^T\nabla h_{l_k}h_{l_k}\|^2 \leq \|\nabla h_{l_k}h_{l_k}\|^3$, then we have

$$\begin{aligned}
\|h_{l_k}\|^2 - \|h_{l_k} + \nabla h_{l_k}^T d_k^{ncp}\|^2 &= -2(\nabla h_{l_k}h_{l_k})^T d_k^{ncp} - d_k^{ncp^T}\nabla h_{l_k}\nabla h_{l_k}^T d_k^{ncp} \\
&= 2\delta_k\|\nabla h_{l_k}h_{l_k}\| - \frac{\delta_k^2\|\nabla h_{l_k}^T\nabla h_{l_k}h_{l_k}\|^2}{\|\nabla h_{l_k}h_{l_k}\|^2} \\
&\geq 2\delta_k\|\nabla h_{l_k}h_{l_k}\| - \delta_k\|\nabla h_{l_k}h_{l_k}\| \\
&\geq \delta_k\|\nabla h_{l_k}h_{l_k}\|.
\end{aligned} \tag{3.21}$$

Secondly, if $d_k^{ncp} = -\frac{\|\nabla h_{l_k}h_{l_k}\|^2}{\|\nabla h_{l_k}^T\nabla h_{l_k}h_{l_k}\|^2}(\nabla h_{l_k}h_{l_k})$ and $\delta_k\|\nabla h_{l_k}^T\nabla h_{l_k}h_{l_k}\|^2 \geq \|\nabla h_{l_k}h_{l_k}\|^3$, then we have

$$\begin{aligned}
\|h_{l_k}\|^2 - \|h_{l_k} + \nabla h_{l_k}^T d_k^{ncp}\|^2 &= -2(\nabla h_{l_k}h_{l_k})^T d_k^{ncp} - d_k^{ncp^T}\nabla h_{l_k}\nabla h_{l_k}^T d_k^{ncp} \\
&= \frac{2\|\nabla h_{l_k}h_{l_k}\|^4}{\|\nabla h_{l_k}^T\nabla h_{l_k}h_{l_k}\|^2} - \frac{\|\nabla h_{l_k}h_{l_k}\|^4}{\|\nabla h_{l_k}^T\nabla h_{l_k}h_{l_k}\|^2}
\end{aligned}$$

$$= \frac{\|\nabla h_{l_k} h_{l_k}\|^4}{\|\nabla h_{l_k}^T \nabla h_{l_k} h_{l_k}\|^2}$$
$$\geq \frac{\|\nabla h_{l_k} h_{l_k}\|^2}{\|\nabla h_{l_k}^T \nabla h_{l_k} h_{l_k}\|^2}. \tag{3.22}$$

Using assumption A_3, we have $\|\nabla h_{l_k} h_{l_k}\| \geq \frac{\|h_{l_k}\|}{\|(\nabla h_{l_k}^T \nabla h_{l_k})^{-1} \nabla h_{l_k}\|}$. Hence, from inequalities (2.14), (3.21), (3.22), and using assumption A_2, we obtain the inequality (3.20).

From the above lemma and the fact that

$$\|h_{l_k}\|^2 - \|h_{l_k} + \nabla h_{l_k}^T \mu_k d_k^n\|^2 \geq \mu_k[\|h_{l_k}\|^2 - \|h_{l_k} + \nabla h_{l_k}^T d_k^n\|^2],$$

where $\mu_k \in (0, 1]$, then we have

$$\|h_{l_k}\|^2 - \|h_{l_k} + \nabla h_{l_k}^T \mu_k d_k^n\|^2 \geq C_4 \mu_k \|h_{l_k}\| \min\{\|h_{l_k}\|, \delta_k\}. \tag{3.23}$$

From the way of updating ρ_k shown in Step 8 in Algorithm (2.4) and above inequality, we have

$$Pred_k \geq \frac{1}{2} C_4 \mu_k \rho_k \|h_{l_k}\| \min\{\|h_{l_k}\|, \delta_k\}. \tag{3.24}$$

Lemma 3.9. *Under assumptions A_1–A_5, there exists a constant $0 < C_5$, such that*

$$Tpred_k(\bar{d}_k^t) \geq C_5 \|Z_k^T \nabla q_k(d_k^n)\| \min\{\frac{\|Z_k^T \nabla q_k(d_k^n)\|}{\|\bar{B}_k\|}, \Delta_k\}. \tag{3.25}$$

Proof. From the definition of the tangential Cauchy step (2.21), we will consider two cases:
Firstly, if $\bar{d}_k^{tcp} = -\frac{\Delta_k}{\|Z_k^T \nabla q_k(d_k^n)\|} Z_k^T \nabla q_k(d_k^n)$ and $\Delta_k (Z_k^T \nabla q_k(d_k^n))^T \bar{B}_k Z_k^T \nabla q_k(d_k^n) \leq \|Z_k^T \nabla q_k(d_k^n)\|^3$, then we have

$$\begin{aligned}
Tpred_k(\bar{d}_k^{tcp}) &= q_k(d_k^n) - q_k(d_k^n + Z_k \bar{d}_k^{tcp}) \\
&= -(Z_k^T \nabla q_k(d_k^n))^T \bar{d}_k^{tcp} - \frac{1}{2} \bar{d}_k^{tcp^T} \bar{B}_k \bar{d}_k^{tcp} \\
&= \Delta_k \|Z_k^T \nabla q_k(d_k^n)\| \\
&\quad - \frac{\Delta_k^2}{2\|Z_k^T \nabla q_k(d_k^n)\|^2} [(Z_k^T \nabla q_k(d_k^n))^T \bar{B}_k Z_k^T \nabla q_k(d_k^n)] \\
&\geq \Delta_k \|Z_k^T \nabla q_k(d_k^n)\| - \frac{1}{2} \Delta_k \|Z_k^T \nabla q_k(d_k^n)\| \\
&\geq \frac{1}{2} \Delta_k \|Z_k^T \nabla q_k(d_k^n)\|.
\end{aligned} \tag{3.26}$$

Secondly, if $\bar{d}_k^{tcp} = -\frac{\|Z_k^T \nabla q_k(d_k^n)\|^2}{Z_k^T \nabla q_k(d_k^n))^T \bar{B}_k Z_k^T \nabla q_k(d_k^n)} Z_k^T \nabla q_k(d_k^n)$ and $\Delta_k (Z_k^T \nabla q_k(d_k^n))^T \bar{B}_k Z_k^T \nabla q_k(d_k^n) \geq \|Z_k^T \nabla q_k(d_k^n)\|^3$, then we have

$$\begin{aligned}
Tpred_k(\bar{d}_k^{tcp}) &= q_k(d_k^n) - q_k(d_k^n + Z_k \bar{d}_k^{tcp}) \\
&= -(Z_k^T \nabla q_k(d_k^n))^T \bar{d}_k^{tcp} - \frac{1}{2} \bar{d}_k^{tcp^T} \bar{B}_k \bar{d}_k^{tcp}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{\|Z_k^T \nabla q_k(d_k^n)\|^4}{(Z_k^T \nabla q_k(d_k^n))^T \bar{B}_k Z_k^T \nabla q_k(d_k^n)} \\
&\quad - \frac{\|Z_k^T \nabla q_k(d_k^n)\|^4}{2(Z_k^T \nabla q_k(d_k^n))^T \bar{B}_k Z_k^T \nabla q_k(d_k^n)} \\
&= \frac{\|Z_k^T \nabla q_k(d_k^n)\|^4}{2(Z_k^T \nabla q_k(d_k^n))^T \bar{B}_k Z_k^T \nabla q_k(d_k^n)} \\
&\geq \frac{\|Z_k^T \nabla q_k(d_k^n)\|^2}{2\|\bar{B}_k\|}. \qquad (3.27)
\end{aligned}
$$

Hence, from inequalities (2.20), (3.26), (3.27), and using assumptions A_1–A_5, we obtain the desired result.

From (2.19), (3.25), and the fact that

$$q_k(\mu_k d_k^n) - q_k(\mu_k(d_k^n + Z_k \bar{d}_k^t)) \geq \mu_k[q_k(d_k^n) - q_k(d_k^n + Z_k \bar{d}_k^t)],$$

where $\mu_k \in (0, 1]$, then we have

$$q_k(\mu_k d_k^n) - q_k(\mu_k d_k) \geq C_5 \mu_k \|Z_k^T \nabla q_k(d_k^n)\| \min\{\frac{\|Z_k^T \nabla q_k(d_k^n)\|}{\|\bar{B}_k\|}, \Delta_k\}. \qquad (3.28)$$

That is

$$T pred_k(\mu_k \bar{d}_k^t) \geq C_5 \mu_k \|Z_k^T \nabla q_k(d_k^n)\| \min\{\frac{\|Z_k^T \nabla q_k(d_k^n)\|}{\|\bar{B}_k\|}, \Delta_k\}. \qquad (3.29)$$

The following lemma clarifies that if at any iteration k, the point $x_k \in F^+$ is not feasible, then algorithm ACBTR can not loop infinitely without finding an acceptable step.

Lemma 3.10. *Under assumptions A_1–A_5. If $\|h_{l_k}\| \geq \varepsilon > 0$, then the condition $\frac{Ared_{k^j}}{Pred_{k^j}} \geq \gamma_1$ will be satisfied for some finite j.*

Proof. From inequalities (3.18), (3.24), and the condition $\|h_{l_k}\| \geq \varepsilon$, we have

$$| \frac{Ared_k}{Pred_k} - 1 | = \frac{| Ared_k - Pred_k |}{Pred_k} \leq \frac{2C_3 \delta_k^2}{C_4 \varepsilon \min\{\varepsilon, \delta_k\}}.$$

Now as the trial step d_{k^j} gets rejected, δ_{k^j} becomes small and eventually we will have

$$\left| \frac{Ared_{k^j}}{Pred_{k^j}} - 1 \right| \leq \frac{2C_3 \delta_{k^j}}{C_4 \varepsilon}.$$

For j finite, this inequality implies that, the acceptance rule will be met. This completes the proof.

Lemma 3.11. *Under assumptions A_1–A_5 and the j^{th} trial step of iteration k satisfies,*

$$\|d_{k^j}\| \leq \min\{\frac{(1-\gamma_1)C_4}{4C_3}, 1\}\|h_{l_k}\|, \qquad (3.30)$$

then the step accepted.

Proof. The proof of this lemma by contradiction. Assume that the inequality (3.30) holds and the step d_{k^j} is rejected. From inequalities (3.18), (3.24), and using inequality (3.30), we have

$$(1-\gamma_1) < \frac{|Ared_{k^j} - Pred_{k^j}|}{Pred_{k^j}} < \frac{2C_3\|d_{k^j}\|}{C_4\|h_{l_k}\|} \leq \frac{1}{2}(1-\gamma_1).$$

This is a contradiction and this completes the proof.

Lemma 3.12. *Under assumptions A_1–A_5 and for all j^{th} trial step of any iteration k, then δ_{k^j} satisfies*

$$\delta_{k^j} \geq \min\{\frac{\delta_{min}}{b_1}, \frac{\alpha_1(1-\gamma_1)C_4}{4C_3}, \alpha_1\}\|h_{l_k}\|, \tag{3.31}$$

where $b_1 > 0$ is a constant.

Proof. For all j^{th} trial step of any iteration k, we will consider tow cases:
Firstly, if $j = 1$ and the step accepted, then $\delta_k \geq \delta_{min}$. Hence,

$$\delta_k \geq \delta_{min} \geq \frac{\delta_{min}}{b_1}\|h_{l_k}\|, \tag{3.32}$$

where $b_1 = \sup_{x\in S} \|h_{l_k}\|$. Then (3.31) holds in this case.
Secondly, if $j > 1$, then there exists at least one rejected trial step and hence from Lemma (3.11) we have

$$\|d_{k^i}\| > \min\{\frac{(1-\gamma_1)C_4}{4C_3}, 1\}\|h_{l_k}\|,$$

for all $i = 1, 2, ...j-1$. From Algorithm 2.2 and d_{k^i} is a rejected trial step, then we have

$$\delta_{k^j} = \alpha_1\|d_{k^{j-1}}\| > \alpha_1 \min\{\frac{(1-\gamma_1)C_4}{4C_3}, 1\}\|h_{l_k}\|. \tag{3.33}$$

From inequalities (3.32) and (3.32) the desired result is obtained.

The next lemma prove that as long as $\|h_{l_k}\|$ is bound away from zero, the trust-region radius is also bound away from zero.

Lemma 3.13. *Under assumptions A_1–A_5 . If $\|h_{l_k}\| \geq \varepsilon > 0$, then*

$$\delta_{k^j} \geq C_6,$$

where $C_6 > 0$ is a constant.

Proof. The proof follows directly by taking

$$C_6 = \varepsilon \min\{\frac{\delta_{min}}{b_1}, \frac{\alpha_1(1-\gamma_1)C_4}{4C_3}, \alpha_1\}, \tag{3.34}$$

in inequality (3.31).

3.2. Global convergence theory when $\sigma_k \to \infty$

In this section, we clarify the convergence of the sequence of iteration when the positive parameter $\sigma_k \to \infty$.

Lemma 3.14. *Under assumptions A_1–A_5. If ρ_k is increased at any iteration k, then*

$$\sigma_k \mu_k \|h_{l_k}\|^2 \leq C_7, \tag{3.35}$$

where C_7 is a positive constant.

Proof. From the way of updating the positive penalty parameter ρ_k, we notice that ρ_k is increased at a given iteration k according to one of the two rules (2.31) or (2.30). Suppose that ρ_k is increased according to the rule (2.30), then

$$\begin{aligned}\frac{\rho_k}{2}[\|h_{l_k}\|^2 - \|h_{l_k} + \nabla h_{l_k}^T \mu_k d_k\|^2] &= [q_k(\mu_k d_k) - q_k(0) + \Delta\lambda_k^T(h_{l_k} + \nabla h_{l_k}^T \mu_k d_k)] \\ &+ \frac{b_0}{2}[\|h_{l_k}\|^2 - \|h_{l_k} + \nabla h_{l_k}^T \mu_k d_k\|^2].\end{aligned}$$

Using inequalities (3.23) and (3.31), then we have

$$\begin{aligned}\frac{\rho_k}{2} C_4 \mu_k \|h_{l_k}\|^2 \min\{\frac{\delta_{min}}{b_1}, \frac{\alpha_1(1-\gamma_1)C_4}{4C_3}, \alpha_1\} &\leq \nabla_x \ell^s(x_k,\lambda_k))^T \mu_k d_k + \frac{1}{2}\mu_k^2 d_k^T \tilde{H}_k d_k \\ &+ \Delta\lambda_k^T(h_{l_k} + \nabla h_{l_k}^T \mu_k d_k) \\ &+ \frac{\sigma_k}{2}[\|W_k(g_u(x_k) + \nabla g_u(x_k)^T \mu_k d_k)\|^2 - \|W_k g_u(x_k)\|^2] \\ &+ \frac{b_0}{2}[\|h_{l_k}\|^2 - \|h_{l_k} + \nabla h_{l_k}^T \mu_k d_k\|^2].\end{aligned}$$

According to rule (2.31), we have $\rho_k \geq \sigma_k^2$. Hence

$$\begin{aligned}\frac{\sigma_k^2}{2} C_4 \mu_k \|h_{l_k}\|^2 \min\{\frac{\delta_{min}}{b_1}, \frac{\alpha_1(1-\gamma_1)C_3}{4C_4}, \alpha_1\} &\leq (\nabla_x \ell^s(x_k,\lambda_k))^T \mu_k d_k + \frac{1}{2}\mu_k^2 d_k^T \tilde{H}_k d_k \\ &+ \Delta\lambda_k^T(h_{l_k} + \nabla h_{l_k}^T \mu_k d_k) \\ &+ \frac{\sigma_k}{2}[\|W_k(g_u(x_k) + \nabla g_u(x_k)^T \mu_k d_k)\|^2 + \frac{b_0}{2}\|h_{l_k}\|^2.\end{aligned}$$

Then,

$$\begin{aligned}\frac{\sigma_k}{2} C_4 \mu_k \|h_{l_k}\|^2 \min\{\frac{\delta_{min}}{b_1}, \frac{\alpha_1(1-\gamma_1)C_3}{4C_4}, \alpha_1\} &\leq \frac{1}{\sigma_k}[(\nabla_x \ell^s(x_k,\lambda_k))^T \mu_k d_k + \frac{1}{2}\mu_k^2 d_k^T \tilde{H}_k d_k \\ &+ \Delta\lambda_k^T(h_{l_k} + \nabla h_{l_k}^T \mu_k d_k) + \frac{b_0}{2}\|h_{l_k}\|^2] \\ &+ \frac{1}{2}\|W_k(g_u(x_k) + \nabla g_u(x_k)^T \mu_k d_k)\|^2 \\ &\leq \frac{1}{\sigma_k}[|\nabla_x \ell^s(x_k,\lambda_k))^T d_k| + \frac{1}{2}|d_k^T \tilde{H}_k d_k| \\ &+ |\Delta\lambda_k^T(h_{l_k} + \nabla h_{l_k}^T \mu_k d_k)| + \frac{b_0}{2}\|h_{l_k}\|^2]\end{aligned}$$

$$+ \quad \frac{1}{2}\|W_k(g_u(x_k) + \nabla g_u(x_k)^T \mu_k d_k)\|^2,$$

where $\mu_k \leq 1$. Using the Cauchy-Schwarz inequality, assumptions A_3–A_5, and the fact that $\|d_k\| \leq \delta_{max}$, the proof is completed.

Lemma 3.15. *Under assumptions A_1–A_5. If $\sigma_k \to \infty$ and there exists an infinite subsequence $\{k_i\}$ of the iteration sequence at which ρ_k is increased, then*

$$\lim_{k_i \to \infty} \|h_{k_i}\| = 0. \tag{3.36}$$

Proof. The proof follows directly from $\lim_{k_i \to \infty} \mu_{k_i} = 1$, $\sigma_k \to \infty$, and Lemma (3.14).

Theorem 3.1. *Under assumptions A_1–A_5. If $\sigma_k \to \infty$, then*

$$\lim_{k \to \infty} \|h_{l_k}\| = 0. \tag{3.37}$$

Proof. The proof similar to the proof of Theorem 4.18 [19].

We notice from the way of updating σ that, the sequence $\{\sigma_k\}$ is unbounded only when there exist an infinite subsequence of indices $\{k_i\}$, at which

$$\frac{1}{2}Tpred_k(\mu_k \bar{d}_k^t) < \|\nabla g_u(x_k) W_k g_u(x_k)\| \min\{\|\nabla g_u(x_k) W_k g_u(x_k)\|, \Delta_k\}. \tag{3.38}$$

The following lemma shows that, if $\sigma_k \to \infty$ and $\limsup_{k \to \infty} \|W_k g_u(x_k)\| > 0$, then the iteration sequence generated by the algorithm ACBTR has a subsequence that satisfies IFJ conditions in the limit.

Lemma 3.16. *Under assumptions A_1–A_5. If $\sigma_k \to \infty$ and there exists a subsequence $\{k_j\}$ of indices indexing iterates that satisfy $\|W_k g_u(x_k)\| \geq \varepsilon > 0$ for all $k \in \{k_j\}$, then a subsequence of the iteration sequence indexed $\{k_j\}$ satisfies the IFJ conditions as $k \to \infty$.*

Proof. The proof is by contradiction. Let the subsequence $\{k_j\}$ be renamed to $\{k\}$ to simplify the notation. Suppose that there is no a subsequence of the sequence of iterates that satisfies IFJ conditions in the limit. Then we have $|\|W_k g_u(x_k)\|^2 - \|W_k(g_u(x_k) + \nabla g_u(x_k)^T Z_k \mu_k \bar{d}_k^t)\|^2| \geq \varepsilon_1 > 0$ from Lemma (3.1). Also we have $\|Z_k \nabla g_u(x_k) W_k g_u(x_k)\| \geq \varepsilon_2 > 0$ from (3.13). Since

$$\|Z_k^T \nabla g_u(x_k) W_k(g_u(x_k) + \nabla g_u(x_k)^T d_k^n)\| \quad \geq \quad \|Z_k^T \nabla g_u(x_k) W_k g_u(x_k)\| - \|Z_k^T \nabla g_u(x_k) W_k \nabla g_u(x_k)^T\| \|d_k^n\|,$$

and using (3.17), then we have

$$\|Z_k^T \nabla g_u(x_k) W_k(g_u(x_k) + \nabla g_u(x_k)^T d_k^n)\| \quad \geq \quad \varepsilon_2 - C_2 \|Z_k^T \nabla g_u(x_k) W_k \nabla g_u(x_k)^T\| \|h_{l_k}\|.$$

But $\{\|h_{l_k}\|\}$ convergence to zero and $\|Z_k^T \nabla g_u(x_k) W_k \nabla g_u(x_k)^T\|$ is bounded. Then $\|Z_k^T \nabla g_u(x_k) W_k(g_u(x_k) + \nabla g_u(x_k)^T d_k^n)\| \geq \frac{\varepsilon_2}{2}$ and therefore

$$\begin{aligned}
\|Z_k^T \nabla q_k(d_k^n)\| \quad &\geq \quad \sigma_k \|Z_k^T \nabla g_u(x_k) W_k(g_u(x_k) + \nabla g_u(x_k)^T d_k^n)\| - \|Z_k^T(\nabla_x \ell_k^s + \tilde{H}_k d_k^n)\| \\
&\geq \quad \sigma_k \frac{\varepsilon_2}{2} - \|Z_k^T(\nabla_x \ell_k^s + \tilde{H}_k d_k^n)\|.
\end{aligned}$$

Hence inequality (3.29) can be written as follows

$$T pred_k(\mu_k \bar{d}_k^t) \geq \frac{1}{2} C_5 \mu_k \sigma_k [\frac{\varepsilon_2}{2} - \frac{1}{\sigma_k} \|Z_k^T [\nabla_x \ell_k^s + \tilde{H}_k d_k^n]\| \min\{\Delta_k, \frac{\frac{\varepsilon_2}{2} - \frac{1}{\sigma_k}\|Z_k^T[\nabla_x \ell_k^s + \tilde{H}_k d_k^n]\|}{\|Z_k^T \nabla g_u(x_k) W_k \nabla g_u(x_k)^T Z_k\| + \frac{1}{\sigma_k}\|Z_k^T \tilde{H}_k Z_k\|}\}.$$

That is for k sufficiently large we have

$$T pred_k(\mu_k \bar{d}_k^t) \geq \frac{\varepsilon_2}{4} C_5 \mu_k \sigma_k \min\{\Delta_k, \frac{\varepsilon_2}{2\|Z_k^T \nabla g_u(x_k) W_k \nabla g_u(x_k)^T Z_k\|}\}.$$

Since $\sigma_k \to \infty$, then there exists infinite number of acceptable iterates at which (3.38) holds. That is, there exists a contradiction unless $\sigma_k \Delta_k$ is bounded. Hence $\Delta_k \to 0$ and therefore $\|d_k\| \to 0$. Now we will consider two cases:

Firstly, if $\|W_k g_u(x_k)\|^2 - \|W_k(g_u(x_k) + \nabla g_u(x_k)^T Z_k \mu_k \bar{d}_k^t)\|^2 > \varepsilon_1$, we have

$$\sigma_k\{\|W_k g_u(x_k)\|^2 - \|W_k(g_u(x_k) + \nabla g_u(x_k)^T Z_k \mu_k \bar{d}_k^t)\|^2|\} > \sigma_k \varepsilon_1 \to \infty. \tag{3.39}$$

Thus, from (2.19), (3.39), and using assumptions A_3–A_5, we have $T pred_k(\mu_k \bar{d}_k^t) \to \infty$. That is, the left hand side of inequality (3.38) goes to infinity while the right hand side of the same inequality goes to zero. That is, there exists a contradiction in this case.

Secondly, if $\|W_k g_u(x_k)\|^2 - \|W_k(g_u(x_k) + \nabla g_u(x_k)^T Z_k \mu_k \bar{d}_k^t)\|^2 < -\varepsilon_1$, then

$$\sigma_k\{\|W_k g_u(x_k)\|^2 - \|W_k(g_u(x_k) + \nabla g_u(x_k)^T Z_k \mu_k \bar{d}_k^t)\|^2|\} < -\sigma_k \varepsilon_1 \to -\infty,$$

where $\sigma_k \to \infty$ as $k \to \infty$. Similar to the above case, $T pred_k(\mu_k \bar{d}_k^t) \to -\infty$. This gives a contradiction in this case with $T pred_k(\mu_k \bar{d}_k^t) > 0$. This two contradictions prove the lemma.

The following lemma shows that if $\lim_{k\to\infty} \sigma_k \to \infty$ and $\liminf_{k\to\infty} \|W_k g_u(x_k)\| = 0$, then the iteration sequence generated by the algorithm ACBTR has a subsequence that satisfies FJ conditions in the limit.

Lemma 3.17. *Under assumptions A_1–A_5. Let $\{k_j\}$ be a subsequence of iterates that satisfy $\|W_k g_u(x_k)\| > 0$ for all $k \in \{k_j\}$ and $\lim_{k_j\to\infty} \|W_{k_j} g_{k_j}\| = 0$. If $\lim_{k\to\infty} \sigma_k = \infty$, then a subsequence of $\{k_j\}$ satisfies FJ conditions in the limit.*

Proof. The proof of this lemma is similar to the proof of Lemma 4.20 [19].

3.3. Global convergence theory when σ_k is bounded

In this section, we will continue our discussion assuming that the parameter σ_k is bounded. We mean that there exists an integer $\bar{k}$ such that for all $k \geq \bar{k}$, $\sigma_k = \bar{\sigma} < \infty$, and

$$\frac{1}{2} T pred_k(\mu_k \bar{d}_k^t) \geq \|\nabla g_u(x_k) W_k g_u(x_k)\| \min\{\|\nabla g_u(x_k) W_k g_u(x_k)\|, \Delta_k\}. \tag{3.40}$$

From assumptions A_3, A_5, and assumption (3.40), we can say that there exists a constant $0 < b_2$ such that for all $k \geq \bar{k}$

$$\|B_k\| \leq b_2, \quad \|Z_k^T B_k\| \leq b_2, \quad and \quad \|Z_k^T B_k Z_k\| \leq b_2, \tag{3.41}$$

where $B_k = \tilde{H}_k + \bar{\sigma} \nabla g_u(x_k) W_k \nabla g_u(x_k)^T$.

Lemma 3.18. *Under assumptions A_1–A_5, there exists a constant $C_8 > 0$ such that*

$$q_k(0) - q_k(\mu_k d_k^n) - \Delta\lambda_k^T(h_{I_k} + \nabla h_{I_k}^T \mu_k d_k) \geq -C_8 \mu_k \|h_{I_k}\|, \tag{3.42}$$

for all $k \geq \bar{k}$.

Proof. By using the definition (2.28), we have

$$\begin{aligned} q_k(0) - q_k(\mu_k d_k^n) &= -(\nabla_x \ell^s(x_k,\lambda_k))^T \mu_k d_k^n - \frac{1}{2}\mu_k^2 d_k^{n^T} \tilde{H}_k d_k^n \\ &\quad + \frac{\bar{\sigma}}{2}[\|W_k g_u(x_k)\|^2 - \|W_k(g_u(x_k) + \nabla g_u(x_k)^T \mu_k d_k^n)\|^2] \\ &= -(\nabla_x \ell^s(x_k,\lambda_k) + \bar{\sigma}\nabla g_u(x_k) W_k g_u(x_k))^T \mu_k d_k^n \\ &\quad - \frac{1}{2}\mu_k^2 d_k^{nT}(\tilde{H}_k + \bar{\sigma}\nabla g_u(x_k) W_k \nabla g_u(x_k)^T) d_k^n \\ &= -(\nabla_x \ell^s(x_k,\lambda_k) + \bar{\sigma}\nabla g_u(x_k) W_k g_u(x_k))^T \mu_k d_k^n - \frac{1}{2}\mu_k^2 d_k^{nT} B_k d_k^n. \end{aligned}$$

That is,

$$\begin{aligned} q_k(0) - q_k(\mu_k d_k^n) \; - \; & \Delta\lambda_k^T(h_{I_k} + \nabla h_{I_k}^T \mu_k d_k) = -(\nabla_x \ell^s(x_k,\lambda_k) + \bar{\sigma}\nabla g_u(x_k) W_k g_u(x_k))^T \mu_k d_k^n \\ & - \frac{1}{2}\mu_k^2 d_k^{nT} B_k d_k^n - \Delta\lambda_k^T(h_{I_k} + \nabla h_{I_k}^T \mu_k d_k) \\ \geq \; & -\mu_k \|\nabla_x \ell^s(x_k,\lambda_k)\|\|d_k^n\| - \bar{\sigma}\mu_k \|\nabla g_u(x_k) W_k g_u(x_k)\|\|d_k^n\| - \mu_k^2 \|B_k\|\|d_k^n\|^2 \\ & - \|\Delta\lambda_k\|\|h_{I_k} + \nabla h_{I_k}^T \mu_k d_k\| \\ \geq \; & -\mu_k[\|\nabla_x \ell^s(x_k,\lambda_k)\| + \bar{\sigma}\|\nabla g_u(x_k) W_k g_u(x_k)\| + \|B_k\|\|d_k^n\|]\|d_k^n\| \\ & - \mu_k \|\Delta\lambda_k\|\|\nabla h_{I_k}\|\|d_k^n\|. \end{aligned}$$

By using inequality (3.17), we can obtain the following inequality

$$\begin{aligned} q_k(0) - q_k(\mu_k d_k^n) - \Delta\lambda_k^T(h_{I_k} + \nabla h_{I_k}^T \mu_k d_k) \;\; \geq \;\; & -\mu_k[(\|\nabla_x \ell^s(x_k,\lambda_k)\| + \bar{\sigma}\|\nabla g_u(x_k) W_k g_u(x_k)\| \\ & + \|B_k\|\|d_k^n\| + \|\Delta\lambda_k\|\|\nabla h_{I_k}\|) C_2]\|h_{I_k}\|. \end{aligned}$$

From assumptions A_3–A_5, the fact that $\|d_k^n\| \leq \delta_{max}$, and using (3.41), then for all $k \geq \bar{k}$ there exists a constant $C_8 > 0$ such that inequality (3.42) hold. This completes the proof.

Lemma 3.19. *Under assumptions A_1–A_5, we have*

$$\begin{aligned} Pred_k \;\; \geq \;\; & \frac{1}{2} C_5 \mu_k \|Z_k^T \nabla q_k(d_k^n)\| \min\{\Delta_k, \frac{\|Z_k^T \nabla q_k(d_k^n)\|}{\|\bar{B}_k\|}\} \\ & + \|\nabla g_u(x_k) W_k g_u(x_k)\| \min\{\|\nabla g_u(x_k) W_k g_u(x_k)\|, \Delta_k\} \\ & - C_8 \mu_k \|h_{I_k}\| + \rho_k[\|h_{I_k}\|^2 - \|h_{I_k} + \nabla h_{I_k}^T \mu_k d_k\|^2], \end{aligned} \tag{3.43}$$

for all $k \geq \bar{k}$.

Proof. Since the definition of $Pred_k$ (2.27) can be written as follows

$$Pred_k \;\; = \;\; [q_k(\mu_k d_k^n) - q_k(\mu_k d_k)] + [q_k(0) - q_k(\mu_k d_k^n) - \Delta\lambda_k^T(h_{I_k} + \nabla h_{I_k}^T \mu_k d_k)]$$

$$+\rho_k[\|h_{l_k}\|^2 - \|h_{l_k} + \nabla h_{l_k}^T \mu_k d_k\|^2].$$

and by using (2.19), we have

$$\begin{aligned} Pred_k &= \frac{1}{2}Tpred_k(\mu_k \bar{d}_k^t) + \frac{1}{2}Tpred_k(\mu_k \bar{d}_k^t) \\ &\quad +[q_k(0) - q_k(\mu_k d_k^n) - \Delta\lambda_k^T(h_{l_k} + \nabla h_{l_k}^T \mu_k d_k)] \\ &\quad +\rho_k[\|h_{l_k}\|^2 - \|h_{l_k} + \nabla h_{l_k}^T \mu_k d_k\|^2]. \end{aligned}$$

Using inequalities (3.29), (3.40), and (3.42), we can obtain the desired result.

Lemma 3.20. *Under assumptions A_1–A_5. If ρ_k increased at iteration k, then there exists a constant $C_9 > 0$ such that*

$$\rho_k \mu_k \min\{\|h_{l_k}\|, \delta_k\} \le C_9. \tag{3.44}$$

Proof. Since ρ_k is increased at iteration k, then from (2.30) we have

$$\begin{aligned} \frac{\rho_k}{2}[\|h_{l_k}\|^2 - \|h_{l_k} + \nabla h_{l_k}^T \mu_k d_k\|^2] &= [q_k(\mu_k d_k) - q_k(\mu_k d_k^n)] + [q_k(\mu_k d_k^n) - q_k(0)] \\ &\quad +\Delta\lambda_k^T(h_{l_k} + \nabla h_{l_k}^T \mu_k d_k) \\ &\quad +\frac{b_0}{2}[\|h_{l_k}\|^2 - \|h_{l_k} + \nabla h_{l_k}^T \mu_k d_k\|^2] \\ &= -\frac{1}{2}Tpred_k(\mu_k \bar{d}_k^t) - \frac{1}{2}Tpred_k(\mu_k \bar{d}_k^t) \\ &\quad +[q_k(\mu_k d_k^n) - q_k(0) + \Delta\lambda_k^T(h_{l_k} + \nabla h_{l_k}^T \mu_k d_k)] \\ &\quad +\frac{b_0}{2}[\|h_{l_k}\|^2 - \|h_{l_k} + \nabla h_{l_k}^T \mu_k d_k\|^2]. \end{aligned}$$

Applying inequality (3.23) to the left hand side and inequalities (3.29), (3.40), and (3.42) to the right hand side, we obtain

$$\begin{aligned} \frac{\rho_k}{2}C_4 \mu_k \|h_{l_k}\| \min\{\delta_k, \|h_{l_k}\|\} &\le -\frac{C_5}{2}\mu_k \|Z_k^T \nabla q_k(d_k^n)\| \min\{\Delta_k, \frac{\|Z_k^T \nabla q_k(d_k^n)\|}{\|\bar{B}_k\|}\} \\ &\quad -\|\nabla g_u(x_k) W_k g_u(x_k)\| \min\{\|\nabla g_u(x_k) W_k g_u(x_k)\|, \Delta_k\} \\ &\quad +C_8 \mu_k \|h_{l_k}\| + \frac{b_0}{2}\|h_{l_k}\|^2 \\ &\le C_8 \mu_k \|h_{l_k}\| + \frac{b_0}{2}\|h_{l_k}\|^2. \end{aligned}$$

The rest of the proof follows using the fact that $\mu_k \le 1$ and assumption A_3.

Lemma 3.21. *Under assumptions A_1–A_5. If $\|Z_k^T(\nabla_x \ell^s(x_k, \lambda_k) + \bar{\sigma}\nabla g_u(x_k) W_k g_u(x_k))\| + \|\nabla g_u(x_k) W_k g_u(x_k)\| \ge \varepsilon > 0$ and $\|h_{l_k}\| \le \eta\delta_k$ where $\eta > 0$ is given by*

$$\eta \le \min\left\{\frac{\varepsilon}{6b_2 C_2 \delta_{max}}, \frac{\sqrt{3}}{2C_2}, \frac{C_5\varepsilon}{12C_8}\min\{\frac{2\varepsilon}{3\delta_{max}}, 1\}, \frac{\varepsilon}{4C_8}\min\{\frac{\varepsilon}{\delta_{max}}, 1\}\right\}. \tag{3.45}$$

then there exists a constant $C_{10} > 0$, such that

$$Pred_k \ge C_{10}\mu_k \delta_k + \rho_k[\|h_{l_k}\|^2 - \|h_{l_k} + \nabla h_{l_k}^T \mu d_k\|^2]. \tag{3.46}$$

Proof. Suppose that $\|Z_k^T(\nabla_x \ell^s(x_k,\lambda_k) + \bar{\sigma}\nabla g_u(x_k)W_k g_u(x_k))\| \geq \frac{\varepsilon}{2}$, then $\|\nabla g_u(x_k)W_k g_u(x_k)\| \geq \frac{\varepsilon}{2}$. From inequality (3.17) and using (3.41), we have

$$\begin{aligned}
\|Z_k^T(\nabla_x \ell^s(x_k,\lambda_k) + \bar{\sigma}\nabla g_u(x_k)W_k g_u(x_k) + B_k d_k^n)\| &\geq \|Z_k^T(\nabla_x \ell^s(x_k,\lambda_k) + \bar{\sigma}\nabla g_u(x_k)W_k g_u(x_k))\| \\
&\quad -\|Z_k^T B_k d_k^n\| \\
&\geq \|Z_k^T(\nabla_x \ell^s(x_k,\lambda_k) + \bar{\sigma}\nabla g_u(x_k)W_k g_u(x_k))\| \\
&\quad -b_2 C_2 \|h_{l_k}\| \\
&\geq \frac{\varepsilon}{2} - b_2 C_2 \eta \delta_k.
\end{aligned}$$

Since $\eta \leq \frac{\varepsilon}{6 b_2 C_2 \delta_{max}}$, then we have

$$\|Z_k^T(\nabla_x \ell^s(x_k,\lambda_k) + \bar{\sigma}\nabla g_u(x_k)W_k g_u(x_k) + B_k d_k^n)\| \geq \frac{\varepsilon}{2} - \frac{\varepsilon}{6} \geq \frac{\varepsilon}{3}. \tag{3.47}$$

Because $\Delta_k = \sqrt{\delta_k{}^2 - \|d_k^n\|^2}$ and $\|d_k^n\| \leq C_2\|h_{l_k}\| \leq C_2\eta\delta_k \leq C_2\frac{\sqrt{3}}{2C_2}\delta_k = \frac{\sqrt{3}}{2}\delta_k$, hence $\Delta_k^2 = \delta_k^2 - \|d_k^n\|^2 \geq \delta_k^2 - \frac{3}{4}\delta_k^2 = \frac{1}{4}\delta_k^2$. Thus,

$$\Delta_k \geq \frac{1}{2}\delta_k. \tag{3.48}$$

From inequalities (3.43), (3.47) and (3.48), we have

$$\begin{aligned}
Pred_k &\geq \frac{1}{2}C_5\mu_k\|Z_k^T(\nabla_x \ell^s(x_k,\lambda_k) + \bar{\sigma}\nabla g_u(x_k)W_k g_u(x_k) + B_k d_k^n)\| \\
&\quad \min\{\|Z_k^T(\nabla_x \ell^s(x_k,\lambda_k) + \bar{\sigma}\nabla g_u(x_k)W_k g_u(x_k) + B_k d_k^n)\|, \frac{1}{2}\delta_k\} \\
&\quad +\|\nabla g_u(x_k)W_k g_u(x_k)\| \min\{\|\nabla g_u(x_k)W_k g_u(x_k)\|, \frac{1}{2}\delta_k\} \\
&\quad -C_8\mu_k\|h_{l_k}\| + \rho_k[\|h_{l_k}\|^2 - \|h_{l_k} + \nabla h_{l_k}^T \mu d_k\|^2] \\
&\geq \frac{C_5\mu_k\varepsilon}{12}\delta_k \min\{\frac{2\varepsilon}{3\delta_{max}}, 1\} + \frac{\mu\varepsilon}{4}\min\{\frac{\varepsilon}{\delta_{\max}}, 1\}\delta_k \\
&\quad -\frac{1}{2}C_8\eta\mu_k\delta_k - \frac{1}{2}C_8\eta\mu_k\delta_k + \rho_k[\|h_{l_k}\|^2 - \|h_{l_k} + \nabla h_{l_k}^T \mu_k d_k\|^2].
\end{aligned}$$

Since $\eta \leq \min\left\{\frac{C_5\varepsilon}{12C_8}\min\{\frac{2\varepsilon}{3\delta_{max}}, 1\}, \frac{\varepsilon}{4C_8}\min\{\frac{\varepsilon}{\delta_{max}}, 1\}\right\}$, then we have

$$Pred_k \geq \frac{C_5\mu_k\varepsilon}{24}\min\{\frac{2\varepsilon}{3\delta_{max}}, 1\}\delta_k + \frac{\mu\varepsilon}{8}\min\{\frac{2\varepsilon}{\delta_{\max}}, 1\}\delta_k + \rho_k[\|h_{l_k}\|^2 - \|h_{l_k} + \nabla h_{l_k}^T \mu_k d_k\|^2].$$

The result follows if we take $C_{10} = \min\left\{\frac{C_5\varepsilon}{24}\min\{\frac{2\varepsilon}{3\delta_{max}}, 1\}, \frac{\varepsilon}{8}\min\{\frac{2\varepsilon}{\delta_{\max}}, 1\}\right\}$.

We can easily see from lemma 3.21 that, at any iteration at which $\|Z_k^T(\nabla_x \ell^s(x_k,\lambda_k) + \bar{\sigma}\nabla g_u(x_k)W_k g_u(x_k)) + \nabla g_u(x_k)W_k g_u(x_k)\| \geq \varepsilon$ and $\|h_{l_k}\| \leq \eta\delta_k$, where η is given by (3.45), there is no need to increase the value of ρ_k. It is only increased when $\|h_{l_k}\| \geq \eta\delta_k$.

Lemma 3.22. *Under assumptions A_1–A_5. If ρ_{k^j} increased at the j^{th} trial iterate of any iteration k,then*

$$\rho_{k^j}\mu_{k^j}\|h_{l_k}\| \leq C_{11}, \tag{3.49}$$

where $C_{11} > 0$ is a constant.

Proof. The proof of this lemma follows directly from inequalities (3.31) and (3.44).

Lemma 3.23. *Under assumptions A_1–A_5. If $\rho_k \to \infty$, then*

$$\lim_{k_i \to \infty} \|h_{l_{k_i}}\| = 0, \tag{3.50}$$

where $\{k_i\}$ is a subsequence of iterates at which the penalty parameter is increased.

Proof. The proof of this lemma follows directly from Lemma 3.22 and $\lim_{k\to\infty} \mu_k = 1$.

3.4. Main global convergence theory

In this section, we will prove the main global convergence theorems for the proposed algorithm ACBTR.

Theorem 3.2. *Assume that assumptions A_1–A_5 hold, then the sequence of iterates generated by ACBTR algorithm satisfies*

$$\lim_{k\to\infty} \|h_{l_k}\| = 0. \tag{3.51}$$

Proof. Suppose that $\limsup_{k\to\infty} \|h_{l_k}\| \geq \varepsilon$, where $\varepsilon > 0$ is a constant. Then there exists an infinite subsequence of indices $\{k_j\}$ indexing iterates that satisfy $\|h_{k_j}\| \geq \frac{\varepsilon}{2}$. From Lemma (3.10), we know that there exists an infinite sequence of acceptable steps, so to simplify, we assume that all members of the sequence $\{k_j\}$ are acceptable iterates. Now we will consider two cases:
Firstly, we consider that, if $\{\rho_k\}$ is unbounded. Then there exists an infinite number of iterates $\{k_i\}$ at which ρ_k is increased. From Lemma (3.23) and for k sufficiently large, we can say $\{k_i\} \bigcap \{k_j\} = \emptyset$. Let k_{ζ_1} and k_{ζ_2} be two consecutive iterates at which ρ_k is increased and $k_{\zeta_1} < k < k_{\zeta_2}$, for any $k \in \{k_j\}$. Notice that, ρ_k is the same for all iterates between k_{ζ_1} and k_{ζ_2}. Since all the iterates of $\{k_j\}$ are acceptable, then

$$\Phi_k - \Phi_{k+1} = Ared_k \geq \gamma_1 Pred_k,$$

for all $k \in \{k_j\}$. Using inequality (3.24), we have

$$\frac{\Phi_k - \Phi_{k+1}}{\rho_k} \geq \frac{\gamma_1 C_4 \mu_k}{2} \|h_{l_k}\| \min\{\|h_{l_k}\|, \delta_k\}.$$

Summing over all acceptable iterates that lie between k_{ζ_1} and k_{ζ_2}, we have

$$\sum_{k=k_{\zeta_1}}^{k_{\zeta_2}-1} \frac{\Phi_k - \Phi_{k+1}}{\rho_k} \geq \frac{\gamma_1 C_4 \mu_k \varepsilon}{4} \min\{\hat{C}_6, \frac{\varepsilon}{2}\},$$

where $\hat{C}_6$ is as C_6 in (3.34), with ε is replaced by $\frac{\varepsilon}{2}$. Hence,

$$\frac{\ell^s(x_{k_{\zeta_1}}, \mu_{k_{\zeta_1}}; \bar{\sigma}) - \ell^s(x_{k_{\zeta_2}}, \mu_{k_{\zeta_2}}; \bar{\sigma})}{\rho_{k_{\zeta_1}}} + [\|h_{l_{k_{\zeta_1}}}\|^2 - \|h_{k_{\zeta_2}}\|^2] \geq \frac{\gamma_1 C_4 \varepsilon}{4} \min\{\hat{C}_6, \frac{\varepsilon}{2}\}.$$

Since $\rho_k \to \infty$, then for k_{ζ_1} sufficiently large, we have

$$\frac{|\ \ell^s(x_{k_{\zeta_1}}, \lambda_{k_{\zeta_1}}; \bar{\sigma}) - \ell^s(x_{k_{\zeta_2}}, \lambda_{k_{\zeta_2}}; \bar{\sigma})\ |}{\rho_{k_{\zeta_1}}} < \frac{\gamma_1 C_4 \varepsilon}{8} \min\{\hat{C}_6, \frac{\varepsilon}{2}\}.$$

Therefore,

$$\|h_{l_{k_{\zeta_1}}}\|^2 - \|h_{l_{k_{\zeta_2}}}\|^2 \geq \frac{\gamma_1 C_4 \varepsilon}{8} \min\{\hat{C}_6, \frac{\varepsilon}{2}\}.$$

But this leads to a contradiction with Lemma (3.23) unless $\varepsilon = 0$.
Secondly, if $\{\rho_k\}$ is bounded, then there exists an integer $\tilde{k}$ such that for all $k \geq \tilde{k}$, $\rho_k = \tilde{\rho}$. Hence from inequality (3.24), we have for any $\hat{k} \in \{k_j\}$ and $\hat{k} \geq \tilde{k}$

$$Pred_{\hat{k}} \;\geq\; \frac{\tilde{\rho} C_4 \mu_{\hat{k}}}{2} \|h_{l_{\hat{k}}}\| \min\{\delta_{\hat{k}}, \|h_{l_{\hat{k}}}\|\} \geq \frac{\varepsilon \tilde{\rho} C_4 \mu_{\hat{k}}}{4} \min\{\frac{\varepsilon}{2\delta_{max}}, 1\}\delta_{\hat{k}}. \tag{3.52}$$

Since all the iterates of $\{k_j\}$ are acceptable, then for any $\hat{k} \in \{k_j\}$, we have

$$\Phi_{\hat{k}} - \Phi_{\hat{k}+1} = Ared_{\hat{k}} \geq \gamma_1 Pred_{\hat{k}}.$$

Using inequality (3.52), we have

$$\Phi_{\hat{k}} - \Phi_{\hat{k}+1} \geq \frac{\gamma_1 \varepsilon \tilde{\rho} C_4 \mu_{\hat{k}}}{4} \min\{\frac{\varepsilon}{2\delta_{max}}, 1\}\delta_{\hat{k}}.$$

Using Lemma (3.13), we have

$$\Phi_{\hat{k}} - \Phi_{\hat{k}+1} \geq \frac{\gamma_1 \varepsilon \tilde{\rho} C_4 \mu_{\hat{k}}}{4} \min\{\frac{\varepsilon}{2\delta_{max}}, 1\}\hat{C}_6 > 0.$$

Thus there exists a contradiction with the fact that $\{\Phi_k\}$ is bounded when the sequence of the penalty parameter $\{\rho_k\}$ is bounded. Hence, in both cases the supposition is not correct and the theorem is proved.

Theorem 3.3. *Under assumptions A_1–A_5, the sequence of iterates generated by ACBTR algorithm satisfies*

$$\liminf_{k\to\infty} \; [\; \|Z_k^T \nabla_x \ell_k^s\| + \|\nabla g_u(x_k) W_k g_u(x_k)\| \;] \;=\; 0. \tag{3.53}$$

Proof. To prove this theorem we will prove

$$\liminf_{k\to\infty} [\|Z_k^T (\nabla_x \ell_k^s + \bar{\sigma} \nabla g_u(x_k) W_k g_u(x_k))\| + \|\nabla g_u(x_k) W_k g_u(x_k)\; \|\;] \;=\; 0, \tag{3.54}$$

by contradiction. That is, we assume $\| \, Z_k^T (\nabla_x \ell_k^s + \bar{\sigma} \nabla g_u(x_k) W_k g_u(x_k)) \, \| + \|\nabla g_u(x_k) W_k g_u(x_k)\| > \varepsilon$ and there exists an infinite subsequence $\{k_i\}$ of the iteration sequence such that $\|h_{l_{k_i}}\| > \eta \delta_{k_i}$. Since $\|h_{l_{k_i}}\| \to 0$ as $k_i \to 0$, then

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

Let k^j be any iteration in $\{k_i\}$. Then we will consider two cases:
Firstly, if $\{\rho_k\}$ is unbounded and the trial step $j-1$ of iteration k is rejected. Thus $\|h_{l_k}\| > \eta \delta_{k^j} = \alpha_1 \eta \|d_{k^{j-1}}\|$. Hence, from inequalities (3.24), (3.19), and $d_{k^{j-1}}$ was rejected, we have

$$\begin{aligned} (1-\gamma_1) \;&\leq\; \frac{|Ared_{k^{j-1}} - Pred_{k^{j-1}}|}{Pred_{k^{j-1}}} \\ &\leq\; \frac{[2\kappa_1 \|d_{k^{j-1}}\| + 2\kappa_2 \rho_{k^{j-1}} \|d_{k^{j-1}}\| \|\|h_{l_k}\| + 2\kappa_3 \rho_{k^{j-1}} \|d_{k^{j-1}}\|^2]}{\rho_{k^{j-1}} C_4 \min(\alpha_1 \eta, 1) \|h_{l_k}\|} \end{aligned}$$

$$\leq \frac{2\kappa_1}{\rho_{k^{j-1}} C_4 \alpha_1 \eta \min(\alpha_1 \eta, 1)} + \frac{2\kappa_2 + 2\kappa_3 \alpha_1 \eta}{C_4 \alpha_1 \eta \min(\alpha_1 \eta, 1)} \|d_{k^{j-1}}\|.$$

Since $\{\rho_k\}$ is unbounded, then there exists an iterate $\hat{k}$ sufficiently large such that for all $k \geq \hat{k}$, we have

$$\rho_{k^{j-1}} < \frac{4\kappa_1}{C_4 \alpha_1 \eta \min(\alpha_1 \eta, 1)(1 - \gamma_1)}.$$

and

$$\|d_{k^{j-1}}\| \geq \frac{C_4 \alpha_1 \eta \min(\alpha_1 \eta, 1)(1 - \gamma_1)}{4(\kappa_2 + \kappa_3 \alpha_1 \eta)}.$$

From the way of updating the radius of the trust region, we have

$$\delta_{k^j} = \alpha_1 \|d_{k^{j-1}}\| \geq \frac{C_4 \alpha_1^2 \eta \min(\alpha_1 \eta, 1)(1 - \gamma_1)}{4(\kappa_2 + \kappa_3 \alpha_1 \eta)}.$$

But this is a contradiction and this means that δ_{k^j} can not go to zero in this case.
Secondly, if $\{\rho_k\}$ is bounded and there exists an integer $\bar{k}$ and a constant $\bar{\rho}$ such that for all $k \geq \bar{k}$, $\rho_k = \bar{\rho}$. Let j be a trial step of iteration k at which $\|h_k\| > \eta \delta_{k^j}$. Now we will consider the following two cases:
I). If $j = 1$, then from our way of updating the radius of the trust-region, we have $\delta_{k^j} \geq \delta_{\min}$. That is, δ_{k^j} is bounded in this case.
II). If $j > 1$ and $\|h_{l_k}\| > \eta \delta_{k^l}$ for all $l = 1, \cdots, j$, then for all rejected trial steps $l = 1, \cdots, j-1$ of iteration k, we have

$$(1 - \gamma_1) \leq \frac{|Ared_{k^l} - Pred_{k^l}|}{Pred_{k^l}} \leq \frac{2C_3 \|d_{k^l}\|}{C_4 \min(\eta, 1) \|h_{l_k}\|}.$$

That is

$$\begin{aligned} \delta_{k^j} = \alpha_1 \|d_{k^{j-1}}\| &\geq \frac{\alpha_1 C_4 \min(\eta, 1)(1 - \gamma_1) \|h_{l_k}\|}{2C_3} \geq \frac{\alpha_1 C_4 \min(\eta, 1)(1 - \gamma_1)\eta}{2C_3} \delta_{k^1} \\ &\geq \frac{\alpha_1 C_4 \min(\eta, 1)(1 - \gamma_1)\eta}{2C_3} \delta_{\min}. \end{aligned}$$

This means that, δ_{k^j} is bounded.
Otherwise, if $j > 1$ and $\|h_{l_k}\| > \eta \delta_{k^l}$ holds for some l, then there exists an integer β_1 such that $\|h_{l_k}\| > \eta \delta_{k^l}$ holds for $l = \beta_1 + 1, ..., j$ and $\|h_{l_k}\| \leq \eta \delta_{k^l}$ for $l = 1, ..., \beta_1$. As in the above case, we can write

$$\delta_{k^j} \geq \frac{\alpha_1 C_4 \min(\alpha, 1)(1 - \gamma_1)}{2C_3} \|h_{l_k}\| \geq \frac{\alpha_1 C_4 \min(\eta, 1)(1 - \gamma_1)\eta}{2C_3} \delta_{k^{\beta_1 + 1}}. \tag{3.55}$$

But from the way of updating the radius of the trust-region, we have

$$\delta_{k^{\beta_1 + 1}} \geq \alpha_1 \|d_{k^{\beta_1}}\|. \tag{3.56}$$

Since $\|h_{l_k}\| \leq \eta \delta_{k^l}$ for $l = 1, ..., \beta_1$, then from Lemma (3.21) and the fact that $d_{k^{\beta_1}}$ is rejected, we have

$$(1 - \gamma_1) \leq \frac{|Ared_{k^{\beta_1}} - Pred_{k^{\beta_1}}|}{Pred_{k^{\beta_1}}} \leq \frac{2C_3 \bar{\rho} \|d_{k^{\beta_1}}\|}{C_{10}}.$$

This implies

$$\|d_{k^{\beta_1}}\| \geq \frac{C_{10}(1-\gamma_1)}{2C_3\bar{\rho}}.$$

This implies that, $\|d_{k^{\beta_1}}\|$ is bounded. Hence, δ_{k^j} is bounded in this case too. But this is a contradiction. That is $\|h_{l_k}\| \leq \eta\delta_{k^j}$ for all k^j sufficiently large.

Letting $k^j \geq \bar{k}$ and using Lemma (3.21), we have

$$\Phi_{k^j} - \Phi_{k^j+1} = Ared_{k^j} \geq \gamma_1 Pred_{k^j} \geq \gamma_1 C_{10}\delta_{k^j}.$$

As $k \to \infty$, then

$$\lim_{k\to\infty} \delta_{k^j} = 0. \tag{3.57}$$

That is δ_{k^j} is not bounded below. But this leads to a contradiction and to prove this contradiction we will consider the following two cases:

i). If $k^j > \bar{k}$ and the step was accepted at $j = 1$, then $\delta_k \geq \delta_{\min}$. Hence δ_{k^j} is bounded in this case.

ii). If $j > 1$ and there exists at least one rejected trial step $d_{k^{j-1}}$. Then from Lemmas (3.7) and (3.21), we have

$$(1-\gamma_1) < \frac{\bar{\rho}C_3\|d_{k^{j-1}}\|^2}{C_{10}\delta_{k^{j-1}}}.$$

From the way of updating δ_{k^j} we have

$$\delta_{k^j} = \alpha_1\|d_{k^{j-1}}\| > \frac{\alpha_1 C_{10}(1-\gamma_1)}{\bar{\rho}C_3}.$$

Hence δ_{k^j} is bounded in this case too. But this contradicts (3.57). This means that, the supposition is incorrect. Hence,

$$\liminf_{k\to\infty} \ [\ \|Z_k^T(\nabla_x \ell_k^s + \bar{\sigma}\nabla g_u(x_k)W_k g_u(x_k))\| + \|\nabla g_u(x_k)W_k g_u(x_k)\|\ ] \ = \ 0.$$

But this also implies (3.53). This completes the proof of the theorem.

From the above two theorems, we conclude that, given any $\varepsilon > 0$, the algorithm terminates because $\|Z_k^T\nabla_x\ell_k^s\| + \|\nabla g_u(x_k)W_k g_u(x_k)\| + \|h_{l_k}\| < \varepsilon$, for some finite k.

4. Numerical results

Algorithm ACBTR was implemented as a MATLAB code and run under MATLAB version 8.2.701 (R2013b) 64-bit(win64). We begin by a starting point $x_0 \in F^+$ and the following parameter setting is used: $\delta_{min} = 10^{-4}$, $\delta_0 = max(\|d_0^{cp}\|, \delta_{min})$, $\delta_{max} = 10^4\delta_0$, $\gamma_1 = 10^{-4}$, $\gamma_2 = 0.75$, $\alpha_1 = 0.5$, $\alpha_2 = 2$ and $\varepsilon = 10^{-8}$.

Secondly, an extensive variety of possible numeric NBLP problems are introduced to clarify the effectiveness of the proposed ACBTR algorithm.

For each test problem, 10 independent runs with different initial starting point are proceeded to observe the matchmaking of the results. Statistical results of all test problems are summarized in Table 1. The results in Table 1 show that the resuls by the ACBTR Algorithm (2.5) are approximate or equal to those by the compared algorithms in the literature.

In Table 1, we adding the average of number of iterations (iter),the average of number of function evaluations (nfunc), the average of value of CPU time (CPUs) per seconds.

For comparison, we have included the corresponding results of the avarge value of CPU time (CPUs) which are obtained by Methods in [34] (Table 2), [29] (Table 3), and [44] (Table 4) respectively. It is obviously from the results that our algorithm ACBTR is qualified for treating NBLP problems even the upper and the lower levels are convex or not and the results converge to the optimal solution which is similarly or approximate to the optimal that reported in literature. Finally, it is obviously from the comparison between the solutions obtained by using ACBTR algorithm with literature, that ACBTR is able to find the optimal solution of all problems by a small number of iterations, small number of function evaluations, and less time.

Problem 1 [34]:

$$\begin{aligned} \min_t \quad & f_u = v_1^2 + v_2^2 + t^2 - 4t \\ s.t. \quad & 0 \le t \le 2, \\ \min_v \quad & f_l = v_1^2 + 0.5v_2^2 + v_1 v_2 + \\ & (1-3t)v_1 + (1+t)v_2, \\ s.t. \quad & 2v_1 + v_2 - 2t \le 1, \\ & v_1 \ge 0, \quad v_2 \ge 0. \end{aligned}$$

Problem 2 [34]:

$$\begin{aligned} \min_t \quad & f_u = v_1^2 + v_3^2 - v_1 v_3 - 4v_2 - 7t_1 + 4t_2 \\ s.t. \quad & t_1 + t_2 \le 1, \\ & t_1 \ge 0, \quad t_2 \ge 0 \\ \min_v \quad & f_l = v_1^2 + 0.5v_2^2 + 0.5v_3^2 + v_1 v_2 + \\ & (1-3t_1)v_1 + (1+t_2)v_2, \\ s.t. \quad & 2v_1 + v_2 - v_3 + t_1 - 2t_2 + 2 \le 0, \\ & v_1 \ge 0; \quad v_2 \ge 0 \quad v_3 \ge 0. \end{aligned}$$

Problem 3 [34]:

$$\begin{aligned} \min_t \quad & f_u = 0.1(t_1^2 + t_2^2) - 3v_1 - 4v_2 + 0.5(v_1^2 + v_2^2) \\ s.t. \quad & \\ \min_v \quad & f_l = 0.5(v_1^2 + 5v_2^2) - 2v_1 v_2 - t_1 v_1 - t_2 v_2, \\ s.t. \quad & -0.333v_1 + v_2 - 2 \le 0, \\ & v_1 - 0.333v_2 - 2 \le 0, \\ & v_1 \ge 0, \quad v_2 \ge 0, \end{aligned}$$

Problem 4 [34]:

$$\begin{aligned} \min_t \quad & f_u = t_1^2 - 2t_1 + t_2^2 - 2t_2 + v_1^2 + v_2^2 \\ s.t. \quad & t_1 \ge 0, \quad t_2 \ge 0 \\ \min_v \quad & f_l = (v_1 - t_1)^2 + (v_2 - t_2)^2, \\ s.t. \quad & 0.5 \le v_1 \le 1.5, \\ & 0.5 \le v_2 \le 1.5, \end{aligned}$$

Problem 5 [34]:

$$\begin{aligned} \min_t \quad & f_u = t^2 + (v-10)^2 \\ s.t. \quad & -t + v \le 0, \\ & 0 \le t \le 15, \\ \min_v \quad & f_l = (t + 2v - 30)^2, \\ s.t. \quad & t + v \le 20, \\ & 0 \le v \le 20, \end{aligned}$$

Problem 6 [34]:

$$\begin{aligned} \min_t \quad & f_u = (t-1)^2 + 2v_1^2 - 2t \\ s.t. \quad & t \ge 0, \\ \min_v \quad & f_l = (2v_1 - 4)^2 + (2v_2 - 1)^2 + tv_1, \\ s.t. \quad & 4t + 5v_1 + 4v_2 \le 12, \\ & -4t - 5v_1 + 4v_2 \le -4, \\ & 4t - 4v_1 + 5v_2 \le 4, \\ & -4t + 4v_1 + 5v_2 \le 4, \\ & v_1 \ge 0, \quad v_2 \ge 0, \end{aligned}$$

Problem 7 [34]:

$$\begin{aligned} \min_t \quad & f_u = (t-5)^2 + (2v+1)^2 \\ s.t. \quad & t \geq 0, \\ \min_v \quad & f_l = (2v-1)^2 - 1.5tv, \\ s.t. \quad & -3t + v \leq -3, \\ & t - 0.5v \leq 4, \\ & t + v \leq 7, \\ & v \geq 0. \end{aligned}$$

Problem 8 [34]:

$$\begin{aligned} \min_t \quad & f_u = t_1^2 - 3t_1 + t_2^2 - 3t_2 + v_1^2 + v_2^2 \\ s.t. \quad & t_1 \geq 0, \quad t_2 \geq 0, \\ \min_v \quad & f_l = (v_1 - t_1)^2 + (v_2 - t_2)^2, \\ s.t. \quad & 0.5 \leq v_1 \leq 1.5, \\ & 0.5 \leq v_2 \leq 1.5, \end{aligned}$$

Problem 9 [29]:

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

Problem 10 [29]:

$$\begin{aligned} \min_t \quad & f_u = t^3 v_1 + v_2 \\ s.t. \quad & 0 \leq t \leq 1, \\ \min_v \quad & f_l = -v_2 \\ s.t. \quad & t v_1 \leq 10, \\ & v_1^2 + t v_2 \leq 1, \\ & v_2 \geq 0. \end{aligned}$$

Problem 11 [44]:

$$\begin{aligned} \min_t \quad & f_u = -8t_1 - 4t_2 + 4v_1 - 40v_2 - 4v_3 \\ s.t. \quad & t_1 \geq 0, \quad t_2 \geq 0 \\ \min_v \quad & f_l = \frac{1 + t_1 + t_2 + 2v_1 - v_2 + v_3}{6 + 2t_1 + v_1 + v_2 - 3v_3}, \\ s.t. \quad & -v_1 + v_2 + v_3 + v_4 = 1, \\ & 2t_1 - v_1 + 2v_2 - 0.5v_3 + v_5 = 1, \\ & 2t_2 + 2v_1 - v_2 - 0.5v_3 + v_6 = 1, \\ & v_i \geq 0, \quad i = 1, ..., 6. \end{aligned}$$

Problem 12 [29]:

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

Problem 13 [44]:

$$\begin{aligned} \min_t \quad & f_u = -t_1^2 - 3t_2^2 - 4v_1 + v_2^2 \\ s.t. \quad & t_1^2 + 2t_2 \leq 4, \\ & t_1 \geq 0, \quad t_2 \geq 0, \\ \min_v \quad & f_l = 2t_1^2 + v_1^2 - 5v_2, \\ s.t. \quad & t_1^2 - 2t_1 + 2t_2^2 - 2v_1 + v_2 \geq -3, \\ & t_2 + 3v_1 - 4v_2 \geq 4, \\ & v_1 \geq 0, \quad v_2 \geq 0. \end{aligned}$$

Problem 14 [44]:

$$\begin{aligned} \min_t \quad & f_u = (t-1)^2 + (v-1)^2 \\ s.t. \quad & t \geq 0, \\ \min_v \quad & f_l = 0.5v^2 + 500v - 50tv \\ s.t. \quad & v \geq 0. \end{aligned}$$

Table 1. Comparisons of the results by ACBTR Algorithm 2.5 and Methods in reference.

Problem name	(t_*, v_*) ACBTR	f_u^* f_l^* ACBTR	iter nfunc ACBTR	CPUs time ACBTR	(t_*, v_*) Ref.	f_u^* f_l^* Ref.
prob1 [34]	(0.8438, 0.7657, 1.121e-8)	-2.0769 -0.5863	14 16	1.77	(0.8438, 0.7657, 0)	-2.0769 -0.5863
prob2 [34]	(0.609, 0.391, 0, 0,1.828)	0.6086 1.6713	12 15	2.1	(0.609, 0.391, 0, 0, 1.828)	0.6426 1.6708
prob3 [34]	(0.97, 3.14, 2.6, 1.8)	-8.92 -6.05	9 10	3.09	(0.97, 3.14, 2.6, 1.8)	-8.92 -6.05
prob4 [34]	(.5,.5,.5,.5)	-1 0	13 15	1.87	(0.5, 0.5, 0.5, 0.5)	-1 0
prob5 [34]	(10.03, 9.9691)	100.58 0.0012	6 8	1.8	(10.03, 9.969)	100.58 0.001
prob6 [34]	(1.6879, 0.8805,0)	-1.3519 7.4991	8 12	4.5	NA	3.57 2.4
prob7 [34]	(1, 0)	17 1	10 11	2.05	(1, 0)	17 1
prob8 [34]	(0.75,0.75, 0.75, 0.75)	-2.25 0	9 11	1.05	($\sqrt{3}/2$, $\sqrt{3}/2$, $\sqrt{3}/2$, $\sqrt{3}/2$)	-2.1962 0
prob9 [29]	(11.138,5)	2209.8 222.52	11 13	1.85	(11.25,5)	2250 197.753
prob10 [29]	(1,0,7.6287e-08)	7.6287e-08 -7.6287e-08	7 9	3.34	(1,0,1)	1 -1
prob11 [44]	(0,0.9,0,0.6,0.4,0,0,0)	-29.2 0.3148	8 11	42.311	(0,0.9,0,0.6,0.4,0,0,0)	-29.2 0.3148
prob12 [29]	(3,5)	9 0	10 14	2.23	(3,5)	9 0
prob13 [44]	(0,1.7405, 1.8497,0.9692)	-15.548 -1.4247	6 7	2.5	(0,2,1.875,0.9063)	-12.68 -1.016
prob14 [44]	(10.016,0.81967)	81.328 -0.3359	8 11	2.15	(10.04,0.1429)	82.44 0.271

Table 2. Comparisons of the results by ACBTR (2.5) and Method [34].

Problem name	(t_*, v_*) ACBTR	f_u^* f_l^* ACBTR	CPUs ACBTR	(t_*, v_*) method [34]	f_u^* f_l^* method [34]	CPUs method [34].
prob1	(0.8438, 0.7657, 1.121e-8)	-2.0769 -0.5863	1.77	(0.8462,0.769 2,0)	-2.0769 -0.5917	1.734
prob2	(0.609, 0.391, 0, 0,1.828)	0.6086 1.6713	2.1	(0.6111, 0.3889,0, 0, 1.8333)	0.6389 1.6806	2.375
prob3	(0.97, 3.14, 2.6, 1.8)	-8.92 -6.05	3.9	(1.031 6, 3.097 8, 2.597 0, 1.792 9)	-8.9172 -6.137 0	3.315
prob4	(0.5,0.5,0.5,0.5)	-1 0	1.87	(0.5,0.5,0.5,0.5)	-1 0	1.576
prob5	(10.03, 9.9691)	100.58 0.0012	1.8	(10, 10)	100 0	1.825
prob6	(1.6879, 0.8805,0)	-1.3519 7.4991	4.5	(1.8889, 0.8889,0)	-1.2099 7.6173	4.689
prob7	(1, 0)	17 1	2.05	(1,0)	17 1	1.769
prob8	(0.75,0.75, 0.75, 0.75)	-2.25 0	1.05	(0.75,0.75, 0.75, 0.75)	-2.25 0	1.124

Table 3. Comparisons of the results by ACBTR (2.5) and Method [29].

Problem name	(t_*, v_*) ACBTR	f_u^* f_l^* ACBTR	CPUs ACBTR	(t_*, v_*) method [29]	f_u^* f_l^* method [29]	CPUs method [29].
prob9	(11.138,5)	2209.8 222.52	1.85	(11.25, 5)	2250 197.753	2.21
prob10	(1,0,7.6287e-08)	7.6287e-08 -7.6287e-08	3.34	(1,0,-1)	-1 1	3.38
prob12	(3,5)	9 0	2.23	(3,5)	9 0	-

Table 4. Comparisons of the results by ACBTR (2.5) and Method [44].

Problem name	(t_*, y_*) ACBTR	f_u^* f_l^* ACBTR	CPUs ACBTR	(t_*, y_*) method [44]	f_u^* f_l^* method [44]	CPUs method [44].
prob3	(0.97, 3.14, 2.6, 1.8)	-8.92 -6.05	3.9	(1.03, 3.097, 2.59,1.79	-8.92 -6.14	11.854
prob5	(10.03, 9.9691)	100.58 0.0012	1.8	(10,10)	100.014 4.93e-7	5.888
prob6	(1.6879, 0.8805,0)	-1.3519 7.4991	4.5	(1.8888,0.888)	-1.2091 7.6145	25.332
prob11	(0,0.9,0,0.6,0.4,0,0,0)	-29.2 0.3148	42.311	(0,0.9,0,0.6,0.4,0,0,0)	-29.2 0.3148	107.55
prob13	(0,1.7405, 1.8497,0.9692)	-15.548 -1.4247	2.5	(4.4e-7,2, 1.875,0.9063)	-12.65 -1.021	14.42
prob14	(10.016,0.81967)	81.328 -0.3359	2.15	(10.0164,0.8197)	18.3279 -0.3359	4.218

5. Conclusions

In this paper, we introduce an effective solution algorithm to solve NBLP problem with positive variables. This algorithm based on using KKT condition with Fischer-Burmeister function to transform NBLP problem into an equivalent smooth SONP problem. An active-set strategy with barrier method and the trust-region mechanism is used to ensure global convergence from any starting point. ACBTR algorithm can reduce the number of iteration and the number of function evaluation. The projected Hessian mechanism is used in ACBTR algorithm to overcome the difficulty of having an infeasible trust region subproblem. A global convergence theory of ACBTR algorithm is studied under five standard assumptions.

Preliminary numerical experiment on the algorithm is presented. The performance of the algorithm is reported. The numerical results show that our approach is of value and merit further investigation. For future work, there are many question should be answered

- Our approach used to transform problem 1.2 which is not smooth to smooth problem.
- Using the interior-point method guarantees the converges quadratically to a stationary point.

Conflict of interest

The authors declare that there is no conflict of interest in this paper.

References

1. D. Aksen, S. Akca, N. Aras, A bilevel partial interdiction problem with capacitated facilities and demand outsourcing, *Comput. Oper. Res.*, **41** (2014), 346–358. https://doi.org/10.1016/j.cor.2012.08.013

2. Y. Abo-Elnaga, M. El-Shorbagy, Multi-sine cosine algorithm for solving nonlinear bilevel programming problems, *Int. J. Comput. Int. Sys.*, **13** (2020), 421–432. https://doi.org/10.2991/ijcis.d.200411.001
3. A. Burgard, P. Pharkya, C. Maranas, Optknock: A bilevel programming framework for identifying gene knockout strategies formicrobial strain optimization, *Biotechnol. Bioeng.*, **84** (2003), 647–657. https://doi.org/10.1002/bit.10803
4. M. Bazaraa, H. Sherali, C. Shetty, *Nonlinear programming theory and algorithms*, Hoboken: John Wiley and Sons, 2006.
5. O. Ben-Ayed, O. Blair, Computational difficulty of bilevel linear programming, *Oper. Res.*, **38** (1990), 556–560. https://doi.org/10.1287/opre.38.3.556
6. R. Byrd, Omojokun, Robust trust-region methods for nonlinearly constrained optimization, *The second SIAM conference on optimization*, 1987.
7. R. Byrd, J. Gilbert, J. Nocedal, A trust region method based on interior point techniques for nonlinear programming, *Math. Program.*, **89** (2000), 149–185. https://doi.org/10.1007/PL00011391
8. J. Chen, The semismooth-related properties of a merit function and a descent method for the nonlinear complementarity problem, *J. Glob. Optim.*, **36** (2006), 565–580. https://doi.org/10.1007/s10898-006-9027-y
9. J. Chen, On some NCP-functions based on the generalized Fischer-Burmeister function, *Asia Pac. J. Oper. Res.*, **24** (2007), 401–420. https://doi.org/10.1142/S0217595907001292
10. J. Chen, S. Pan, A family of NCP-functions and a descent method for the nonlinear complementarity problem, *Comput. Optim. Appl.*, **40** (2008), 389–404. https://doi.org/10.1007/s10589-007-9086-0
11. J. Dennis, M. Heinkenschloss, L. Vicente. Trust-region interior-point SQP algorithms for a class of nonlinear programming problems. *SIAM J. Control Optim.*, **36** (1998), 1750–1794. https://doi.org/10.1137/S036012995279031
12. J. Dennis, M. El-Alem, K. Williamson, A trust-region approach to nonlinear systems of equalities and inequalities, *SIAM J. Optim.*, **9** (1999), 291–315. https://doi.org/10.1137/S1052623494276208
13. S. Dempe, *Foundation of bilevel programming*, London: Kluwer Academic Publishers, 2002.
14. T. Edmunds, J. Bard, Algorithms for nonlinear bilevel mathematical programs, *IEEE T. Syst. Man Cy.*, **21** (1991), 83–89. https://doi.org/10.1109/21.101139
15. B. El-Sobky, A robust trust-region algorithm for general nonlinear constrained optimization problems, PhD thesis, Alexandria University, 1998.
16. B. El-Sobky, A global convergence theory for an active trust region algorithm for solving the general nonlinear programming problem, *Appl. Math. Comput.*, **144** (2003), 127–157. https://doi.org/10.1016/S0096-3003(02)00397-1
17. B. El-Sobky, A Multiplier active-set trust-region algorithm for solving constrained optimization problem, *Appl. Math. Comput.*, **219** (2012), 928–946. https://doi.org/10.1016/j.amc.2012.06.072
18. B. El-Sobky, An interior-point penalty active-set trust-region algorithm, *J. Egypt. Math. Soc.*, **24** (2016), 672–680. https://doi.org/10.1016/j.joems.2016.04.003
19. B. El-Sobky, An active-set interior-point trust-region algorithm, *Pac. J. Optim.*, **14** (2018), 125–159.

20. B. El-Sobky, A. Abotahoun, An active-set algorithm and a trust-region approach in constrained minimax problem, *Comp. Appl. Math.*, **37** (2018), 2605–2631. https://doi.org/10.1007/s40314-017-0468-3
21. B. El-Sobky, A. Abotahoun, A trust-region algorithm for solving mini-max problem, *J. Comput. Math.*, **36** (2018), 776–791. https://doi.org/10.4208/jcm.1705-m2016-0735
22. B. El-Sobky, Y. Abouel-Naga, Multi-objective optimal load flow problem with interior-point trust-region strategy, *Electr. Pow. Syst. Res.*, **148** (2017), 127–135. https://doi.org/10.1016/j.epsr.2017.03.014
23. B. El-Sobky, Y. Abouel-Naga, A penalty method with trust-region mechanism for nonlinear bilevel optimization problem, *J. Comput. Appl. Math.*, **340** (2018), 360–374. https://doi.org/10.1016/j.cam.2018.03.004
24. B. El-Sobky, Y.Abo-Elnaga, A. Mousa, A. El-Shorbagy, Trust-region based penalty barrier algorithm for constrained nonlinear programming problems: An application of design of minimum cost canal sections, *Mathematics*, **9** (2021), 1551. https://doi.org/10.3390/math9131551
25. B. El-Sobky, G. Ashry, An interior-point trust-region algorithm to solve a nonlinear bilevel programming problem, *AIMS Mathematics*, **7** (2022), 5534–5562. https://doi.org/10.3934/math.2022307
26. A. Fiacco, G. McCormick. *Nonlinear programming: Sequential unconstrained minimization techniques*, New York: John Wiley and Sons, 1968.
27. J. Falk, J. M. Liu, On bilevel programming, Part I: General nonlinear cases, *Math. Program.*, **70** (1995), 47–72. https://doi.org/10.1007/BF01585928
28. F. Facchinei, H. Y. Jiang, L. Q. Qi, A smoothing method for mathematical programming with equilibrium constraints, *Math. Program.*, **85** (1999), 107–134. https://doi.org/10.1007/s10107990015a
29. H. Gumus, A. Flouda, Global optimization of nonlinear bilevel programming problems, *J. Global Optim.*, **20** (2001), 1–31. https://doi.org/10.1023/A:1011268113791
30. J. L. Gonzalez Velarde, J. F. Camacho-Vallejo, G. Pinto Serranoo, A scatter search algorithm for solving a bilevel optimization model for determining highway tolls, *Comput. Syst.*, **19** (2015), 5–16. https://doi.org/10.13053/CyS-19-1-1916
31. M. Hestenes, Muliplier and gradient methods, *J. Optim. Theorey Appl.*, **4** (1969), 303–320. https://doi.org/10.1007/BF00927673
32. G. Hibino, M. Kainuma, Y. Matsuoka, Two-level mathematical programming for analyzing subsidy options to reduce greenhouse-gas emissions, *IIASA Working Paper*, 1996.
33. D. Kouri, M. Heinkenschloss, D. Ridzal, B. van Bloemen Waanders, A trust-region algorithm with adaptive stochastic collocation for PDE optimization under uncertainty, *SIAM J. Sci. Comput.*, **35** (2013), A1847–A1879. https://doi.org/10.1137/120892362
34. H. Li, Y. C. Jiao, L. Zhang, Orthogonal genetic algorithm for solving quadratic bilevel programming problems. *J. Syst. Eng. Electron.*, **21** (2010), 763–770. https://doi.org/10.3969/j.issn.1004-4132.2010.05.008
35. Y. B. Lv, T. S. Hu, G. M. Wang, Z. P. Wan, A neural network approach for solving nonlinear bilevel programming problem, *Comput. Math. Appl.*, **55** (2008), 2823–2829. https://doi.org/10.1016/j.camwa.2007.09.010

36. N. N. Li, D. Xue, W. Y. Sun, J. Wang, A stochastic trust-region method for unconstrained optimization problems, *Math. Probl. Eng.*, **2019** (2019), 8095054. https://doi.org/10.1155/2019/8095054
37. L. M. Ma, G. M. Wang, A Solving algorithm for nonlinear bilevel programing problems based on human evolutionary model, *Algorithms*, **13** (2020), 260. https://doi.org/10.3390/a13100260
38. E. Omojokun, Trust-region strategies for optimization with nonlinear equality and inequality constraints, PhD thesis, University of Colorado, 1989.
39. T. Steihaug, The conjugate gradient method and trust-region in large scale optimization, *SIAM J. Numer. Anal.*, **20** (1983), 626–637. https://doi.org/10.1137/0720042
40. G. Savard, J. Gauvin, The steepest descent direction for the nonlinear bilevel programming problem, *Oper. Res. Lett.*, **15** (1994), 265–272. https://doi.org/10.1016/0167-6377(94)90086-8
41. S. Sadatrasou, M. Gholamian, K. Shahanaghi, An application of data mining classification and bi-level programming for optimal credit allocation, *Decis. Sci. Lett.*, **4** (2015), 35–50. https://doi.org/10.5267/j.dsl.2014.9.005
42. N. Thoai, Y. Yamamoto, A. Yoshise, Global optimization method for solving mathematical programs with linear complementarity constraints, *Mathematical programs with complementarity*, 2002.
43. M. Ulbrich, S. Ulbrich, L. N. Vicente, A globally convergent primal-dual interior-point filter method for nonlinear programming, *Math. Program.*, **100** (2004), 379–410. https://doi.org/10.1007/s10107-003-0477-4
44. Y. L. Wang, Y. C. Jiao, H. Li, An evolutionary algorithm for solving nonlinear bilevel programming based on a new constraint-handling scheme, *IEEE T. Syst. Man Cy. C*, **35** (2005), 221–232. https://doi.org/10.1109/TSMCC.2004.841908
45. X. Wang, Y. X. Yuan, A trust region method based on a new affine scaling technique for simple bounded optimization, *Optim. Method. Softw.*, **28** (2013), 871–888. https://doi.org/10.1080/10556788.2011.622378
46. X. Wang, Y. X. Yuan, An augmented Lagrangian trust region method for equality constrained optimization, *Optim. Method. Softw.*, **30** (2015), 559–582. https://doi.org/10.1080/10556788.2014.940947
47. Y. X. Yuan, Recent advances in trust region algorithms. *Math. Program.*, **151** (2015), 249–281. https://doi.org/10.1007/s10107-015-0893-2
48. M. Zeng, Q. Ni, A new trust region method for nonlinear equations involving fractional mode. *Pac. J. Optim.*, **15** (2019), 317–329.