Convergence Depth Control for Interior Point Methods

Weifeng Chen, Zhijiang Shao, Kexin Wang, and Xi Chen
Dept. of Control Science and Engineering, Zhejiang University, Hangzhou 310027, P.R. China

Lorenz T. Biegler
Dept. of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213

DOI 10.1002/aic.12225
Published online July 26, 2010 in Wiley Online Library (wileyonlinelibrary.com).

For practical applications, optimization algorithms may converge to the optimal solution unreasonably slowly because of factors such as the poor scaling, ill-conditioning, errors in calculation, and so on. Most improvements during the optimization procedure are made within a small part of the total computation time. To relieve the heavy computational burden, it is necessary to balance the calculation accuracy and computation cost. The traditional termination criteria based on the Karush-Kuhn-Tucker conditions cannot appropriately meet this requirement. Convergence depth control (CDC) strategy for Reduced Hessian Successive Quadratic Programming (RSQP) was presented as an alternative measure in a previous study. This work incorporates interior point methods with the modified CDC strategy, which was tested through AMPL interface and Aspen Open Solvers interface. Related properties are proved.

Keywords: convergence depth control, interior point methods, nonlinear optimization, convergence, termination criteria

Introduction

Almost all optimization algorithms converge to a stationary point by infinite iterations, even though some terminate in a finite number of steps with exact arithmetic.[1] It is important to design appropriate termination criteria to judge whether, in practical application, an adequate approximate solution to the optimum is obtained. The traditional termination criteria for nonlinear optimization are based on the Karush-Kuhn-Tucker (KKT) conditions. If any of the KKT conditions is not satisfied, the optimization procedure will be endless.

The convergence and CPU time for solving problems are crucial in many applications such as real time optimization (RTO),[2–4] nonlinear model predictive control (NMPC),[5,6] and so on. In practical applications, an adequate approximate solution to the optimum is always sufficient. Higher precision is unnecessary from the viewpoint of engineering. Sometimes, optimization algorithms may converge to the optimal solution unreasonably slowly due to poor scaling, ill-conditioning and errors in calculation such as round-off error or truncating error.

Some behaviors of optimization algorithm for solving problems with traditional termination criteria are discussed in this section. The interior point optimizer[7–10] (IPOPT, version 3.6.1) is investigated as it is a leading nonlinear programming solver, especially for large-scale problems. The selected examples are *dallasl* and *sineali* from the CUTE test set. From Figure 1, Problem *dallasl* shows that the algorithm is terminated at the expense of 235 iterations, but the iterates after the 20th iteration are of little difference compared with the final solution. Problem *sineali* shows that the constraint violation is zero during the entire optimization process and the objective function value changes slightly from the initial iterations. However, the dual feasibility (the norm of the first order derivative of Lagrange function) has not achieved the predefined tolerance. The iterations continue until the maximum number of iterations (3000) is

Correspondence concerning this article should be addressed to Z. Shao at zjshao@iipc.zju.edu.cn.

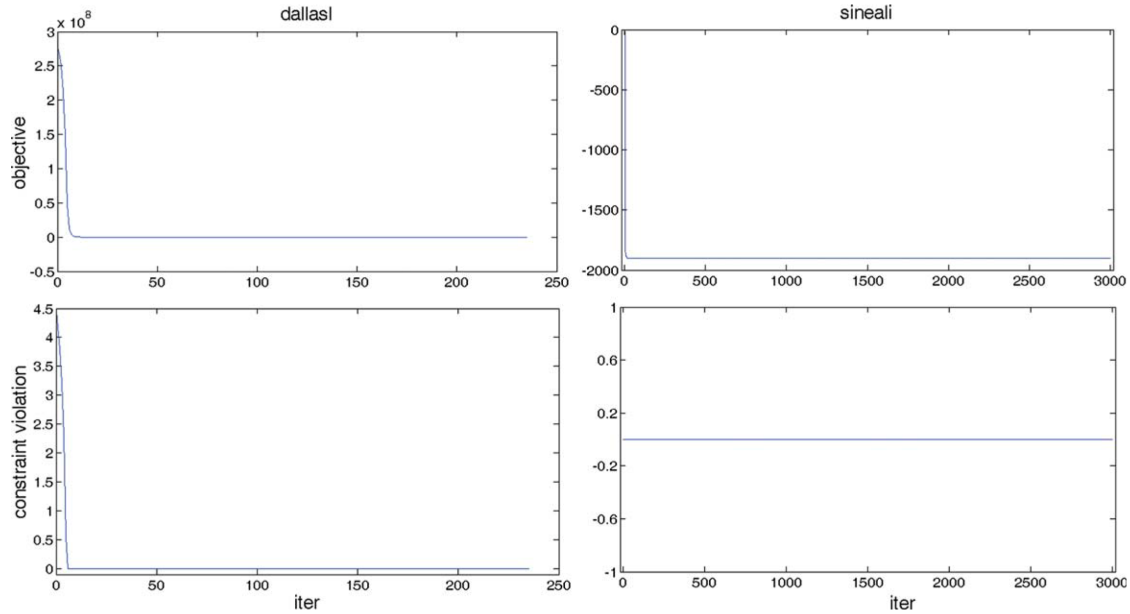

Figure 1. Behaviors of interior point methods with traditional termination criteria.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

exceeded. From both problems, an interesting phenomenon is observed: most of the objective and constraint violation improvements during the optimization procedure are made in a small part of the whole CPU time. This observed phenomenon is in accordance with the Pareto Principle (80/20 Rule), which states that roughly 80% of the effects come from 20% of the causes.

For process system optimization problems, the objective function is usually a profit function. There is no need to continue to iterate when the feasibility is satisfied and the objective function can only be slightly improved with further iteration. Here, it is necessary to balance the calculation accuracy and computation cost and to quickly detect unreasonably slow convergence or convergence failure in order to avoid unnecessary computational effort. The traditional termination criteria based on KKT conditions are inadequate for this purpose. As such, smart termination criteria are needed to halt the optimization algorithm when it reaches a suitable result.

The termination criteria for unconstrained optimization, linearly constrained optimization, and nonlinear constrained optimization have been described by Gill et al.[1] Recently, some studies have been conducted on the termination criteria for nonlinear optimization. Nocedal et al. provided a detailed description of the behavior of the gradient norm for the steepest descent method applied to quadratic objective functions and revealed that the gradient norm was not a reliable measure of successful termination.[11] Wang et al. put forward the convergence depth control (CDC) strategy[12] for reduced-Hessian successive quadratic programming (RSQP). It was designed to take advantage of the achievement estimation of the optimization process to discover the proper time to terminate the optimization algorithm.

In this article, the performance of CDC combined with interior point methods is studied. The CDC is modified according to the characteristics of interior point methods. We prove that the iteration sequence produced by interior point methods can be terminated by the modified CDC. We also prove that convergence depth indicates the degree of current iterate converging to the optimum. We test interior point methods with the modified CDC using AMPL interface[13] and Aspen Open Solvers interface.[14] This article is organized as follows. In the second section, we review interior point methods, including the solution for the barrier problem, and barrier parameter update strategies. The third section describes the CDC for interior point methods and gives the theoretical analysis. The fourth section presents the comparisons of performance between interior point methods with the modified CDC and interior point methods with termination criteria based on KKT conditions. The set of nonlinear programming (NLP) test problems consists of problems from the CUTE test set and problems formulated based on Aspen Plus, including optimization problems for depropanizer and debutanizer distillation column systems, data reconciliation for large-scale air separation systems, and large scale ethylene separation systems. The last section concludes this article.

Interior Point Methods

Interior point methods[7–10,15–17] transform bound constraints of NLP into a logarithmic barrier term and add it to the objective function. The solution for the original NLP is transformed into the solution of a sequence of barrier problems with a variable barrier parameter. As the barrier parameter tends to zero, the optimum of barrier problem converges to a local solution of the original NLP.[18] The following discussions are based on the work of Wächter and Biegler.[7–10]

Solving barrier problem

To facilitate the discussion, the formulation of NLP is considered as follows:

$$\begin{cases} \min_x & f(\boldsymbol{x}) \\ \text{s.t.} & \boldsymbol{c}(\boldsymbol{x}) = 0 \\ & \boldsymbol{x} \geq 0 \end{cases} \tag{1}$$

the bound constraints of (1) are replaced by a logarithmic barrier term that is added to objective function to give

$$\begin{cases} \min_x & \varphi_\mu(\boldsymbol{x}) = f(\boldsymbol{x}) - \mu \sum_{i=1}^{n} \ln(x^{(i)}) \\ \text{s.t.} & \boldsymbol{c}(\boldsymbol{x}) = 0 \end{cases} \tag{2}$$

with a barrier parameter $\mu > 0$. The KKT conditions for (2) are

$$\begin{cases} \nabla f(\boldsymbol{x}) + \nabla \boldsymbol{c}(\boldsymbol{x})\lambda - \boldsymbol{v} = 0 \\ \boldsymbol{c}(\boldsymbol{x}) = 0 \\ \boldsymbol{XVe} - \mu \boldsymbol{e} = 0 \end{cases} \tag{3}$$

where $\boldsymbol{e}$ is the vector of ones of appropriate dimension, $\lambda \in R^m$ are Lagrange multipliers, $\boldsymbol{v} \in R^n$ are dual variables corresponding to $\boldsymbol{x}$, $\boldsymbol{X} := diag(\boldsymbol{x})$, and $\boldsymbol{V} := diag(\boldsymbol{v})$. Newton's method is applied to (3). Search direction $(\boldsymbol{d}_i, \boldsymbol{d}_i^\lambda, \boldsymbol{d}_i^v)$ at iterate $(\boldsymbol{x}_i, \lambda_i, \boldsymbol{v}_i)$ is then obtained as a solution of the linearization of (3)

$$\begin{bmatrix} \boldsymbol{W}_i & \boldsymbol{A}_i & -\boldsymbol{I} \\ \boldsymbol{A}_i^T & 0 & 0 \\ \boldsymbol{V}_i & 0 & \boldsymbol{X}_i \end{bmatrix} \begin{pmatrix} \boldsymbol{d}_i \\ \boldsymbol{d}_i^\lambda \\ \boldsymbol{d}_i^v \end{pmatrix} = - \begin{pmatrix} \boldsymbol{g}_i + \boldsymbol{A}_i\lambda_i - \boldsymbol{v}_i \\ \boldsymbol{c}_i \\ \boldsymbol{X}_i\boldsymbol{v}_i - \mu\boldsymbol{e} \end{pmatrix} \tag{4}$$

We denote the identity matrix with $\boldsymbol{I}$, and the Hessian of the Lagrange function of (1) with $\boldsymbol{W}_i$. $\boldsymbol{A}_i$ denotes the transposition of the Jacobian of equality constraints, and $\boldsymbol{g}_i$ denotes the gradient of objective function at iteration i. The solution to (4) is obtained by first solving

$$\begin{bmatrix} \boldsymbol{H}_i & \boldsymbol{A}_i \\ \boldsymbol{A}_i^T & 0 \end{bmatrix} \begin{pmatrix} \boldsymbol{d}_i \\ \lambda_i^+ \end{pmatrix} = - \begin{pmatrix} \nabla \varphi_\mu(\boldsymbol{x}_i) \\ \boldsymbol{c}_i \end{pmatrix} \tag{5}$$

where $\boldsymbol{H}_i := \boldsymbol{W}_i + \varsigma_i$ with $\varsigma_i := \boldsymbol{X}_i^{-1}\boldsymbol{V}_i$ and then computing $\boldsymbol{d}_i^\lambda$, $\boldsymbol{d}_i^v$ from $\boldsymbol{d}_i^\lambda = \lambda_i^+ - \lambda_i$, $\boldsymbol{d}_i^v := \mu\boldsymbol{X}_i^{-1}\boldsymbol{e} - \boldsymbol{v}_i - \varsigma_i\boldsymbol{d}_i$. To analyze convergence behavior, a QR factorization of $\boldsymbol{A}_i$ can be computed to obtain matrices $\boldsymbol{Z}_i \in R^{n\times(n-m)}$ and $\boldsymbol{Y}_i \in R^{n\times m}$ so that the columns of $[\boldsymbol{Z}_i\ \boldsymbol{Y}_i]$ form an orthonormal basis of R^n, and the column of $\boldsymbol{Z}_i$ becomes a basis of the null space of $\boldsymbol{A}_i^T$. The search direction can then be decomposed into

$$\boldsymbol{d} = \boldsymbol{q}_i + \boldsymbol{p}_i \tag{6}$$

where $\boldsymbol{q}_i := \boldsymbol{Y}_i\tilde{\boldsymbol{q}}_i$ and $\boldsymbol{p}_i := \boldsymbol{Z}_i\tilde{\boldsymbol{p}}_i$ with

$$\tilde{\boldsymbol{q}}_i := -\left[\boldsymbol{A}_i^T\boldsymbol{Y}_i\right]^{-1}\boldsymbol{c}_i \tag{7}$$

$$\tilde{\boldsymbol{p}}_i := -\left[\boldsymbol{Z}_i^T\boldsymbol{H}_i\boldsymbol{Z}_i\right]^{-1}\boldsymbol{Z}_i^T(\boldsymbol{g}_i - \mu\boldsymbol{X}_i^{-1}\boldsymbol{e} + \boldsymbol{H}_i\boldsymbol{q}_i) \tag{8}$$

When the search direction is obtained, the new iterate is denoted by

$$\boldsymbol{x}_{i+1} = \boldsymbol{x}_i + a_i\boldsymbol{d}_i, \quad \lambda_{i+1} = \lambda_i + a_i\boldsymbol{d}_i^\lambda, \quad \boldsymbol{v}_{i+1} = \boldsymbol{v}_i + a_i^v\boldsymbol{d}_i^v \tag{9}$$

where a_i and a_i^v are obtained by linear search methods satisfying the fraction to the boundary rule

$$\boldsymbol{x}_i + a_i\boldsymbol{d}_i \geq (1-\tau)\boldsymbol{x}_i, \quad \boldsymbol{v}_i + a_i^v\boldsymbol{d}_i^v \geq (1-\tau)\boldsymbol{v}_i \tag{10}$$

where $\tau \in (0, 1)$ is some constant chosen close to one.

To ensure global convergence, a filter method framework based on linear search was proposed and analyzed by Wächter and Biegler.[8] Under the proper assumptions, all limit points of the iteration sequence generated by the proposed algorithm framework are feasible, and there exists a limit point which is a first-order optimal point for the barrier problem (2).

Updating the barrier parameter

The other major ingredient of interior point methods is the procedure for choosing the barrier parameter. There are two strategies for updating the barrier parameter.[15,19,20] The first is the monotonic strategy, namely the Fiacco-McCormick approach. This approach fixes the barrier parameter until an approximate solution is obtained. The second approach, the adaptive strategy, updates the barrier parameter at every iteration.

Using interior point methods with Fiacco-McCormick approach fixes the barrier parameter and solves the barrier problem using line search strategy, trust region strategy, or filter strategy. This ensures that the iteration sequence globally converges to the optimal solution of barrier problem under certain conditions. The optimal iterate $(\boldsymbol{x}, \lambda, \boldsymbol{v})$ is measured by the following function:

$$F_\mu(\boldsymbol{x}, \lambda, \boldsymbol{v}) := \max\left\{\frac{\max\left\{\|\boldsymbol{g} + \boldsymbol{A}\lambda - \boldsymbol{v}\|_\infty, \|\boldsymbol{XVe} - \mu\boldsymbol{e}\|_\infty\right\}}{s_\mathrm{d}}, \frac{\|\boldsymbol{c}(\boldsymbol{x})\|_\infty}{s_\mathrm{c}}\right\} \tag{11}$$

$$s_\mathrm{d} = 1 + \frac{\|\lambda\|_1 + \|\boldsymbol{v}\|_1}{(m+n)}, \quad s_\mathrm{c} = 1 + \frac{\|\boldsymbol{x}\|_1}{n} \tag{12}$$

when $(\tilde{\boldsymbol{x}}_j, \tilde{\lambda}_i, \tilde{\boldsymbol{v}}_i)$ satisfies $F_{\mu_j}(\tilde{\boldsymbol{x}}_i, \tilde{\lambda}_i, \tilde{\boldsymbol{v}}_i) \leq \varepsilon_{\mu_j}$, $(\tilde{\boldsymbol{x}}_i, \tilde{\lambda}_i, \tilde{\boldsymbol{v}}_i)$ is regarded as the approximate solution to the barrier problem (2) with barrier parameter μ_i. ε_{μ_j} is denoted as

$$\varepsilon_{\mu_j} = \min\{\varepsilon_{\max}, \tau_\varepsilon \min\{\mu_j, \mu_j^{\theta_\varepsilon}\}\} \tag{13}$$

where $\varepsilon_{\max}$ and τ_ε are positive constants and $\theta_\varepsilon \in (1{,}2/\theta_\mu)$, $\theta_\mu \in (1{,}2)$. After solving the barrier problem with barrier parameter μ_i, the barrier parameter is renewed as

$$\mu_{j+1} = \min\{\tau_\mu\mu_j, \mu_j^{\theta_\mu}\} \tag{14}$$

where $\theta_\mu \in (1{,}2)$, $\tau_\mu \in (0{,}1)$. This optimization procedure continues to iterate until $(\boldsymbol{x}^*, \lambda^*, \boldsymbol{v}^*)$ satisfies $F_0(\boldsymbol{x}^*, \lambda^*\ \boldsymbol{v}^*) \leq \varepsilon_{\mathrm{tol}}$, where $\varepsilon_{\mathrm{tol}}$ is defined by the user.

Interior point methods with adaptive strategy[19] update the barrier parameter at every iteration. The barrier parameter is generally proportional to the complementarity value,

$$\mu_{j+1} = \sigma_j \frac{\boldsymbol{x}_j^T\boldsymbol{v}_j}{n} \tag{15}$$

where n denotes the number of variables and σ_j is a centering parameter that varies at every iteration.

σ_j can be calculated through the quality function (17). At the initial stage, the barrier parameter μ in (4) is substituted by Eq. 15. Let $\boldsymbol{x}(\sigma)$, $\lambda(\sigma)$, and $\boldsymbol{v}(\sigma)$ be defined as

$$\begin{aligned} \boldsymbol{x}(\sigma) &= \boldsymbol{x}_i + a_i^{\max}(\sigma)\boldsymbol{d}_i(\sigma), \quad \lambda(\sigma) = \lambda_i + a_i^{\max}(\sigma)\boldsymbol{d}_i^{\lambda}(\sigma) \\ \boldsymbol{v}(\sigma) &= \boldsymbol{v}_i + a_i^{v,\max}(\sigma)\boldsymbol{d}_i^{v}(\sigma) \end{aligned} \tag{16}$$

where $\boldsymbol{d}(\sigma)$ is the solution to (4). $a_j^{\max}(\sigma)$ and $a_j^{v,\max}(\sigma)$ are the maximum step lengths satisfying the fraction to the boundary rule (10). The quality function $q_N(\sigma)$ is denoted as

$$q_N(\sigma) = \|\boldsymbol{g}(\boldsymbol{x}(\sigma)) + \boldsymbol{A}(\boldsymbol{x}(\sigma))\lambda(\sigma) - \boldsymbol{v}(\sigma)\|^2 + \|\boldsymbol{X}(\sigma)\boldsymbol{V}(\sigma)\boldsymbol{e} - \mu\boldsymbol{e}\|^2 + \|\boldsymbol{c}(\boldsymbol{x}(\sigma))\|^2 \tag{17}$$

and σ_j is the value that minimizes $q_N(\sigma)$. Other methods for renewing σ_j include the LOQO rule[15] and Mehrotra's probing method.[20]

To ensure global convergence of the iterates of interior point methods with adaptive strategy, Nocedal et al.[19] proposed a new algorithm framework. $\phi(\boldsymbol{x},\lambda,\ \boldsymbol{v})$ measures the optimality of (1) as follows:

$$\phi(\boldsymbol{x},\lambda,\boldsymbol{v}) = \|\boldsymbol{g}(\boldsymbol{x}) + \boldsymbol{A}(\boldsymbol{x})\lambda - \boldsymbol{v}\|^2 + \|\boldsymbol{XVe}\|^2 + \|\boldsymbol{c}(\boldsymbol{x})\|^2 \tag{18}$$

when iterate $(\boldsymbol{x}_{i+1},\lambda_{i+1},\ v_{i+1})$ satisfies

$$\phi(\boldsymbol{x}_{i+1},\lambda_{i+1},\boldsymbol{v}_{i+1}) \le \kappa M_i \tag{19}$$

and $\kappa \in (0,1)$, $M_i = \max\{\phi_{i-l}, \phi_{i-l+1}, \ldots \phi_i\}$, and $l = \min\{i, l^{\max}\}$, the barrier parameter is renewed using Eq. 15. If (19) does not hold, the barrier parameter will be updated with the Fiacco-McCormick strategy to obtain a new iterate satisfying (19). The updating mode then returns to an adaptive strategy.

CDC Strategy and Related Property

The causes for the slow convergence of optimization algorithm, especially the phenomenon in accordance with Pareto's Principle, are diverse and case dependent. Basically, these causes could be classified as follows:

- Poor scaling of the optimization problem. The objective function is highly sensitive to changes for some components of $\boldsymbol{x}$, but insensitive to others. Moreover, the optimum lies in a narrow valley; thus, the contours of the objective near the optimum tend towards highly eccentric ellipses.[21] It is difficult for the solver to find the optimum.
- Ill-conditioning. The KKT matrix may become ill-conditioned in the course of optimization because of rank deficiency. In this sense, the Lagrange multipliers will be highly sensitive, and the resulting solution will likely improve slowly or be inaccurate.[11,22]
- Errors in calculation. Because of limited computer precision and the limitations of existing techniques, the Hessian matrix cannot always be evaluated exactly, particularly when the Hessian matrix is approximated by the finite differences of the gradient and there are small discontinuities in the computed gradient.[1] Therefore, the rate of convergence of iteration sequence generated by the optimization algorithm may not be superlinear or quadratic, as expected in theory.

These reasons cannot be completely avoided. Thus, unreasonably slow convergence, especially the phenomenon in accordance with Pareto's principle, may occur. Early termination then becomes appropriate and necessary.

CDC strategy

From the Introduction, we know that traditional termination criteria are sometimes too strict to be utilized for many practical applications. Such criteria detect the slow convergence or the convergence failure until the specified maximum iterations are exceeded or the specified maximum CPU time has run out. The convergence depth and degree of progress are introduced to quantitatively evaluate the status of current iterate, and the sigmoid function is used to soften the traditional termination criteria. We are then adequately able to control the termination of algorithm.

In light of Fiacco and McCormick,[18] as the barrier parameter tends to zero, the optimum of (2) converges to a local solution of (1). It is reasonable to consider the barrier parameter when the CDC strategy for interior point methods is established.

Feasibility and predictive improvement on the objective[12] at $\boldsymbol{x}^t$ are defined by

$$\delta^t_{\text{feasErr}} = \max\{\|\boldsymbol{c}(\boldsymbol{x}^t)\|; -x_i^t, i \in I\} \tag{20}$$

$$\delta^t_{\text{objErr}} = \left|\boldsymbol{g}^{\boldsymbol{T}}(\boldsymbol{x}^t)\boldsymbol{d}(\boldsymbol{x}^t)\right| \tag{21}$$

The progress of optimization procedure[12] is described by

$$\delta^t_{\text{feasChg}} = \left|\delta^t_{\text{feasErr}} - \delta^{t-1}_{\text{feasErr}}\right| \tag{22}$$

$$\delta^t_{\text{objChg}} = \left|f(\boldsymbol{x}^t) - f(\boldsymbol{x}^{t-1})\right| \tag{23}$$

Feasibility, predictive improvement, and the progress of optimization procedure defined above reflect the characteristics of iteration sequence of (2) with a fixed barrier parameter. In order to describe the whole iteration sequence produced by interior point methods, the barrier parameter is taken into consideration.

Convergence depth and degree of progress are defined according to (24)–(26).

$$\theta^t_{\text{convg}} = S(\max\{\delta^t_{\text{feasErr}}, \delta^t_{\text{objErr}}, \mu^t\}, \varepsilon_0) \tag{24}$$

$$\theta^t_{\text{prog}} = S(\max\{\delta^t_{\text{feasChg}}, \delta^t_{\text{objChg}}, \mu^t\}, \varepsilon_0) \tag{25}$$

where ε_0 is a given tolerance and S is a sigmoid function that is defined as

$$S(\delta^t, \varepsilon_0) = \frac{\tanh(\xi * (\log \delta^t / \log \varepsilon_0))}{\tanh(\xi)} \tag{26}$$

In this interval, $S(\delta^t,\ \varepsilon_0)$ has continuous values in the interval $[\varepsilon_0, 1/\varepsilon_0]$, and ξ influences the variation trend of $S(\delta^t,\ \varepsilon_0)$. The specific interpretation of ξ has been given by Wang et al.[12] The value for ξ is set to 1.5 in this article. With this value, rigid criteria are "softened" into flexible criteria as

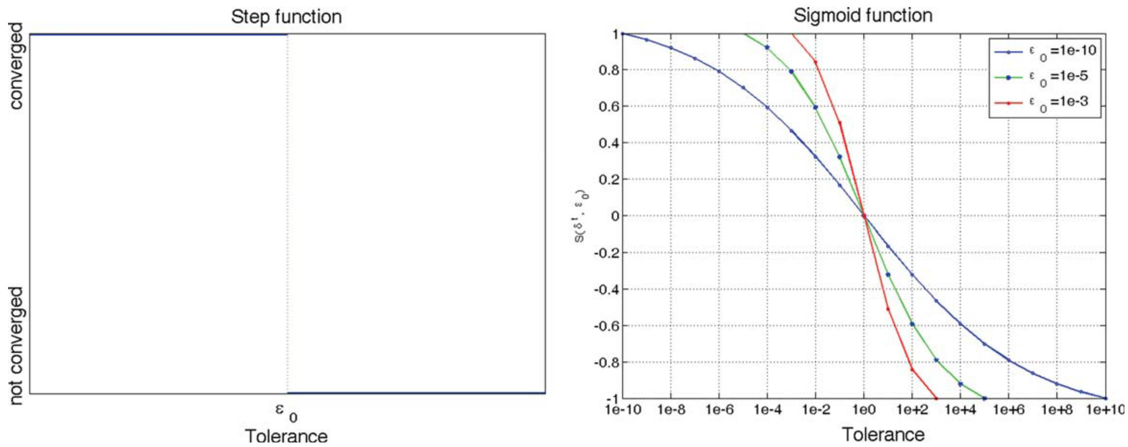

Figure 2. Transforming rigid criteria into flexible criteria.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

shown in Figure 2. This could be compared to the traditional termination criteria, with which the conclusion is ''converged'' or ''not converged'' only. From Figure 2, it is evident that $S(\delta^t, \varepsilon_0)$ decreases monotonically with the increase of δ^t.

In the implementation of CDC in Wang,[12] if the convergence depth of current iterate has reached a predefined acceptable value, then optimization is considered terminated. Otherwise, it proceeds unless the degree of improvement indicates that there is no more possibility of improvement. In Wang et al.,[12] an exception (*mancino*) occurs when the run with CDC obtains an underconverged result; however, the algorithm with traditional termination criteria is successfully terminated. The termination criteria should be well designed to avoid prematurely terminating the optimization procedure. In this article, the implementation of CDC is modified by strengthening the condition for declaring no improvement of the optimization procedure. The modified implementation of CDC is shown in Figure 3. The optimization procedure is considered to have ''no improvement'' possibility only if the degree of progress has reached a threshold value more than η (a predefined constant) times. The proposed modification for CDC is based on the principle that the convergence depth should reach as far as possible before the condition ''no improvement'' is activated.

Properties of CDC

Under proper assumptions, the optimization procedure can be finally terminated by the modified CDC for interior point methods. The convergence depth indicates the distance between the current iterate and the optimum. Proofs in this section are based mainly on the results by Wächter and Biegler[8] and Nocedal et al.[19]

Assumptions A. As the barrier parameter μ_i converges to zero, interior point methods with Fiacco-McCormick barrier strategy or adaptive barrier strategy can produce the iteration sequence $\boldsymbol{x}_{1,1}, \boldsymbol{x}_{2,1}, \ldots, \boldsymbol{x}_{m_1,1}, \boldsymbol{x}_{1,2}, \boldsymbol{x}_{2,2}, \ldots, \boldsymbol{x}_{m_2,2}, \boldsymbol{x}_{1,3}\ldots$, where $\boldsymbol{x}_{m_j,j}$ is the approximate solution of barrier problem (2) with barrier parameter μ_j, i.e., $F_{\mu_j}(\boldsymbol{x}_{m_j,j}, \lambda_{m_j,j}, \boldsymbol{v}_{m_j,j}) \le \varepsilon_{\mu_j}$. The subscripts i and j for $\boldsymbol{x}_{i,j}$ represent the i^{th} iteration for solving the barrier problem (2) with the barrier parameter μ_j. Theorem 1 and Theorem 2 in Wächter and Biegler[8] ensure that the iteration sequence of the barrier problem with barrier parameter μ_j converges to the optimal solution. Therefore, for every j, there exists m_j to satisfy $F_{\mu_j}(\boldsymbol{x}_{m_j,j}, \lambda_{m_j,j}, \boldsymbol{v}_{m_j,j}) \le \varepsilon_{\mu_j}$. It is assumed that the iteration sequence $\boldsymbol{x}_{1,1}, \boldsymbol{x}_{2,1}, \ldots, \boldsymbol{x}_{m_1,1}, \boldsymbol{x}_{1,2}, \boldsymbol{x}_{2,2}, \ldots, \boldsymbol{x}_{m_2,2}, \boldsymbol{x}_{1,3}\ldots$ satisfies the following nine assumptions:

(A1) There exists an open set $C \subseteq R^n$ with $[\boldsymbol{x}_{i,j}, \boldsymbol{x}_{i,j} + a_{i,j}^{\max}\boldsymbol{d}_{i,j}] \subseteq C$ for all $i \notin R_{\text{inc}}$ (R_{inc} is the set of those iteration counters, in which (5) is highly inconsistent) and j, so that the objective function f and constraints $\boldsymbol{c}$ are differentiable on C. Their function values, as well as their derivatives, are bounded and Lipschitz-continuous over C.

(A2) The matrices $\boldsymbol{W}_{i,j}$ approximating the Hessian of the Lagrangian of original problem (1) is uniformly bounded for all $i \notin R_{\text{inc}}$ and j, i.e., there exists a constant $M_{\text{W}} > 0$ with $\|\boldsymbol{W}_{i,j}\| < M_{\text{W}}$ at every iteration.

(A3) The minimal eigenvalue of $\boldsymbol{Z}_{i,j}^T \boldsymbol{H}_{i,j} \boldsymbol{Z}_{i,j}$ is larger than some positive constant, i.e., there exists a constant $M_{\text{H}} > 0$ so $\lambda_{\min}(\boldsymbol{Z}_{i,j}^T \boldsymbol{H}_{i,j} \boldsymbol{Z}_{i,j}) \ge M_{\text{H}}$ for all $i \notin R_{\text{inc}}$ and j, where columns of $\boldsymbol{Z}_{i,j} \in R^{n \times (n-m)}$ form an orthonormal basis matrix of null space of $\boldsymbol{A}_{i,j}^T$.

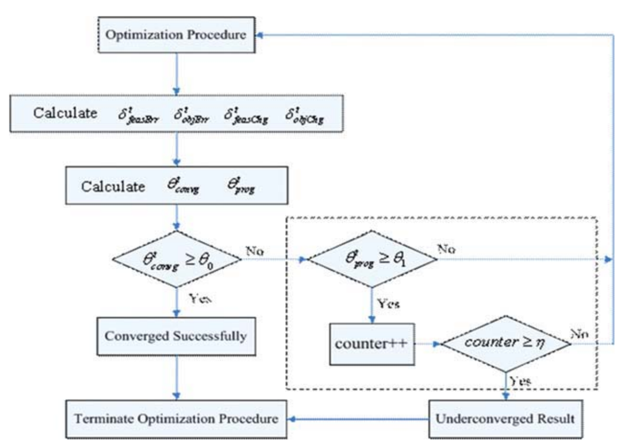

Figure 3. Implementation framework of the modified CDC.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

(A4) The minimal singular value of $A_{i,j}$ is larger than some positive constant., i.e., there exists a $M_A > 0$; therefore, $\sigma_{\min}(A_{i,j}) \geq M_A$ for all $i \notin R_{\text{inc}}$ and j.

(A5) There exists a constant $\theta_{\text{inc}} > 0$, so that $i \notin R_{\text{inc}}$ whenever $\theta(x_{i,j}) \leq \theta_{\text{inc}}$, where $\theta(x_{i,j}) = \|c(x_{i,j})\|$.

(A6) The iterates $\{x_{i,j}\}$ are bounded for all i and j, i.e., there exists a constant $M_x > 0$, and all the iterates satisfy $\|x_{i,j}\| \leq M_x$.

(A7) For every μ_j, at all feasible limit points $x^*_{\mu j}$ for (2), the gradients of the active constraints are linearly independent.

(A8) For every μ_j, there exist constants $\tilde{\delta}_\theta, \tilde{\delta}_x > 0$, so that whenever the restoration phase is called in an iteration $i \in R$ with $||c(x_{i,j})|| < \tilde{\delta}_\theta$, it returns a new iterate with $x^{(k)}_{i+1,j} \geq x^{(k)}_{i,j}$ for all components satisfying $x^{(k)}_{i,j} \leq \tilde{\delta}_x$.

(A9) The Mangasarian-Fromovitz constraint qualification (MFCQ) holds at optimal point x^* of (1).

Assumptions A correspond to the entire iteration sequence produced by interior point methods. Under Assumptions A, the lemmas and theorems of global convergence analysis by Wächter[8] still hold. In addition, the conditions of Theorem 3.12 in Forsgren[17] are also satisfied.

Lemma 1. $\{\varepsilon_{\mu_j}\}$ converges to zero as $j \to \infty$ (i.e., $\lim_{j\to\infty} \varepsilon_{\mu_j} = 0$).

Proof. From (14), we have $\mu_{j+1} = \min\{\tau_\mu \mu_j, \mu_j^{\theta_\mu}\} \leq \tau_\mu \mu_j$, where $\tau_\mu \in (0,1)$ is bounded. It implies that $0 \leq \mu_{j+1} \leq (\tau_\mu)^{j+1}\mu_0$. By applying the Squeeze Rule,* we obtain $\lim_{j\to\infty} \mu_j = 0$. From (13), we have $0 \leq \varepsilon_{\mu_j} \leq \tau_\varepsilon \min\{\mu_j, \mu_j^{\theta_\varepsilon}\} \leq \tau_\varepsilon \mu_j$, where τ_ε is a positive constant. Applying the Squeeze Rule once again, we conclude the following: $\lim_{j\to\infty} \varepsilon_{\mu_j} = 0$. ■

Lemma 2. Suppose Assumptions A hold, then $\left\|\left[A^T_{i,j} Y_{i,j}\right]^{-1}\right\|_2 \leq \frac{1}{M_A}$ for all $i \notin R_{inc}$.

Proof. As the columns of $[Z_{i,j} \;\; Y_{i,j}]$ form an orthonormal basis of R^n, and the columns of $Z_{i,j}$ become a basis matrix of the null space of $A^T_{i,j}$, it follows that

$$\begin{aligned}
\left|A_{i,j}A^{\mathrm{T}}_{i,j} - \lambda I_n\right| &= \left|\begin{bmatrix} Z^{\mathrm{T}}_{i,j} \\ Y^T_{i,j}\end{bmatrix}(A_{i,j}A^{\mathrm{T}}_{i,j} - \lambda I_n)[Z_{i,j} \quad Y_{i,j}]\right| \\
&= \left|\begin{bmatrix} Z^{\mathrm{T}}_{i,j}(A_{i,j}A^{\mathrm{T}}_{i,j} - \lambda I_n)Z_{i,j} & Z^{\mathrm{T}}_{i,j}(A_{i,j}A^{\mathrm{T}}_{i,j} - \lambda I_n)Y_{i,j} \\ Y^{\mathrm{T}}_{i,j}(A_{i,j}A^{\mathrm{T}}_{i,j} - \lambda I_n)Z_{i,j} & Y^T_{i,j}(A_{i,j}A^{\mathrm{T}}_{i,j} - \lambda I_n)Y_{i,j}\end{bmatrix}\right| \\
&= \left|\begin{bmatrix} -\lambda I_{n-m} & 0 \\ 0 & Y^{\mathrm{T}}_{i,j}(A_{i,j}A^{\mathrm{T}}_{i,j} - \lambda I_n)Y_{i,j}\end{bmatrix}\right| \\
&= \lambda^{n-m}\left|Y^{\mathrm{T}}_{i,j}A_{i,j}A^{\mathrm{T}}_{i,j}Y_{i,j} - \lambda I_m\right|
\end{aligned}$$

Thus, $A_{i,j}\ A^{\mathrm{T}}_{i,j}$ has the same non-zero eigenvalues as $Y^{\mathrm{T}}_{i,j}A_{i,j}A^{\mathrm{T}}_{i,j}Y_{i,j}$. From (A4), the minimum non-zero eigenvalue of $A_{i,j}\ A^T_{i,j}$ is larger than M^2_{A}. As $A^{\mathrm{T}}_{i,j}Y_{i,j}$ is invertible, all the eigenvalues of $Y^{\mathrm{T}}_{i,j}A_{i,j}A^{\mathrm{T}}_{i,j}Y_{i,j}$ are positive and larger than M^2_A. Therefore, $\lambda([Y^{\mathrm{T}}_{i,j}A_{i,j}A^{\mathrm{T}}_{i,j}Y_{i,j}]^{-1}) \leq \frac{1}{M^2_{\mathrm{A}}}$, indicating that $\left\|\left[A^{\mathrm{T}}_{i,j}Y_{i,j}\right]^{-1}\right\|_2 \leq \frac{1}{M_{\mathrm{A}}}$. ■

Lemma 3. Suppose Assumptions A hold for both interior point methods with the Fiacco-McCormick barrier strategy and with the adaptive barrier strategy. The norm of $d_{i,j}$ at iterate $x_{i,j}$ is bounded by the constraint violation, the dual infeasibility, the complementarity conditions, and $X_{i,j}$. In other words,

$$\begin{aligned}
\left\|d_{i,j}\right\|_2 &\leq \kappa_1\left\|c_{i,j}\right\|_\infty + \kappa_2\left\|g_{i,j} + A_{i,j}\lambda_{i,j} - v_{i,j}\right\|_\infty \\
&+ \kappa_3\left\|X^{-1}_{i,j}\right\| \cdot \left\|X_{i,j}V_{i,j}e - \mu_j e\right\|_\infty \\
&+ \kappa_4\left\|\mu_j X^{-2}_{i,j}e\right\| \cdot \left\|c_{i,j}\right\|_\infty
\end{aligned}$$

where κ_1, κ_2, κ_3, and κ_4 are positive constants.

Proof. Suppose $x_{i,j}$ is an iterate of the barrier problem with μ_j. From Lemma 2 and (A3), it follows that

$$\left\|\left[A^T_{i,j}Y_{i,j}\right]^{-1}\right\|_2 \leq \frac{1}{M_A}, \left\|\left[Z^T_{i,j}H_{i,j}Z_{i,j}\right]^{-1}\right\|_2 \leq \frac{1}{M_H} \tag{27}$$

(A2) and (A6) ensure that $X_{i,j}$ and $W_{i,j}$ are bounded,

$$\left\|X_{i,j}\right\| \leq M_x, \quad \left\|W_{i,j}\right\|_2 \leq M_W \tag{28}$$

From the equivalence of any vector norm, we have

$$\begin{aligned}
&\left\|c_{i,j}\right\|_2 < M_{\text{Con}}\left\|c_{i,j}\right\|_\infty \\
&\left\|g_{i,j} + A_{i,j}\lambda_{i,j} - v_{i,j}\right\|_2 < M_{Lag}\left\|g_{i,j} + A_{i,j}\lambda_{i,j} - v_{i,j}\right\|_\infty \\
&\left\|X_{i,j}V_{i,j}e - \mu_j e\right\|_2 < M_{Comp}\left\|X_{i,j}V_{i,j}e - \mu_j e\right\|_\infty
\end{aligned} \tag{29}$$

where M_{Con}, M_{Lag}, and M_{Comp} are constants. From (6), (7), (8), (27), (28), and (29), it follows that

$$\begin{aligned}
\left\|d_{i,j}\right\|_2 &= \left\|q_{i,j} + p_{i,j}\right\|_2 \\
&= \|Y_{i,j}[A^T_{i,j}Y_{i,j}]^{-1}c_{i,j} + Z_{ij}[Z^T_{i,j}H_{i,j}Z_{i,j}]^{-1}Z^T_{i,j}(g_{i,j} \\
&\quad -\mu_j X^{-1}_{i,j}e + H_{i,j}q_{i,j})\|_2 \\
&\leq \frac{1}{M_{\mathrm{A}}}\left\|c_{i,j}\right\|_2 + \frac{1}{M_{\mathrm{H}}}\left\|Z^T_{i,j}(g_{i,j} - \mu_j X^{-1}_{i,j}e + H_{i,j}q_{i,j})\right\|_2 \\
&\overset{A^T_{i,j}Z_{i,j}\equiv 0}{=} \frac{1}{M_{\mathrm{A}}}\left\|c_{i,j}\right\|_2 + \frac{1}{M_{\mathrm{H}}}\left\|Z^T_{i,j}(g_{i,j} + A_{i,j}\lambda_{i,j} - V_{i,j}e\right. \\
&\quad \left. + V_{i,j}e - \mu_j X^{-1}_{i,j}e + H_{i,j}q_{i,j})\right\|_2 \\
&\leq \frac{1}{M_{\mathrm{A}}}\left\|c_{i,j}\right\|_2 + \frac{1}{M_{\mathrm{H}}}\left\|g_{i,j} + A_{i,j}\lambda_{i,j} - v_{i,j}\right\|_2 \\
&\quad + \frac{1}{M_{\mathrm{H}}}\left\|V_{i,j}e - \mu_j X^{-1}_{i,j}e\right\|_2 + \frac{1}{M_{\mathrm{H}}M_{\mathrm{A}}}\left\|W_{i,j} + \mu_j X^{-2}_{i,j}e\right\| \cdot \left\|c_{i,j}\right\|_2 \\
&= \frac{(M_{\mathrm{H}} + M_{\mathrm{W}})M_{\text{Con}}}{M_{\mathrm{H}}M_{\mathrm{A}}}\left\|c_{i,j}\right\|_\infty + \frac{M_{\text{Lag}}}{M_{\mathrm{H}}}\left\|g_{i,j} + A_{i,j}\lambda_{i,j} - v_{i,j}\right\|_\infty \\
&\quad + \frac{M_{\text{Comp}}}{M_{\mathrm{H}}}\left\|X^{-1}_{i,j}\right\| \cdot \left\|X_{i,j}V_{i,j}e - \mu_j e\right\|_\infty + \frac{M_{\text{Con}}}{M_{\mathrm{H}}M_{\mathrm{A}}}\left\|\mu_j X^{-2}_{i,j}e\right\| \cdot \left\|c_{i,j}\right\|_\infty
\end{aligned} \tag{30}$$

Defining $\kappa_1 = \frac{(M_{\mathrm{H}}+M_{\mathrm{W}})M_{\text{Con}}}{M_{\mathrm{H}}M_{\mathrm{A}}}$, $\kappa_2 = \frac{M_{\text{Lag}}}{M_{\mathrm{H}}}$, $\kappa_3 = \frac{M_{\text{Comp}}}{M_{\mathrm{H}}}$, and $\kappa_4 = \frac{M_{\text{Con}}}{M_{\mathrm{H}}M_{\mathrm{A}}}$, we obtain the desired result. ■

Lemma 4. Suppose Assumptions A hold for interior point methods with the Fiacco-McCormick barrier strategy. There exists $J_1 > 0$, whenever $j > J_1$, the norm of $d_{m_j,j}$ at iterate $x_{m_j,j}$ is bounded by μ_j^{ϖ}, where ϖ is a constant, i.e., $||d_{m_j,j}||_2 \leq \rho\mu_j^{\varpi}$, where ρ is a positive constant.

Proof. For every j, there exists m_i satisfying $F_{\mu_j}(x_{m_j,j}, \lambda_{m_j,j}, v_{m_j}, j) \leq \varepsilon_{\mu_j}$, thus

*The Squeeze Rule states that if $g(x) \leq f(x) \leq h(x)$ and $\lim_{x\to c} g(x) = \lim_{x\to c} h(x) = L$, then $\lim_{x\to c} f(x) = L$.

$$\frac{\left\|\boldsymbol{g}_{m_j,j}+\boldsymbol{A}_{m_j,j}\lambda_{m_j,j}-\boldsymbol{v}_{m_j,j}\right\|_\infty}{s_d}\le\varepsilon_{\mu_j},\quad\frac{\left\|\boldsymbol{c}(\boldsymbol{x}_{m_j,j})\right\|_\infty}{s_c}\le\varepsilon_{\mu_j},\frac{\left\|\boldsymbol{X}_{m_j,j}\boldsymbol{V}_{m_j,j}\boldsymbol{e}-\mu_j\boldsymbol{e}\right\|_\infty}{s_d}\le\varepsilon_{\mu_j}\tag{31}$$

The complementarity condition $\boldsymbol{XVe}-\mu\boldsymbol{e}=0$ is bilinear; therefore, the dual variable $\boldsymbol{v}$ is bounded, as well as primal variable x. Lemma 1 by Wächter[8] ensures that λ is uniformly bounded. From (12), we have $s_d\le M_{sd}$ and $s_c\le M_{sc}$, where M_{sd} and M_{sd} are positive constants. From (13), we have $\lim_{\mu_j\to0}\varepsilon_{\mu_j}=0$; thus, $\lim_{\mu_j\to0}F_{\mu_j}(\boldsymbol{x}_{m_j,j},\lambda_{m_j,j},\boldsymbol{v}_{m_j,j})=0$. Since $\varepsilon_{\mu_j}\le\tau_\varepsilon\mu_j^{\theta_\varepsilon}$ and $\theta_\varepsilon\in(1,2)$, similar to Lemma 3.13 by Forsgren,[17] we obtain $||\boldsymbol{x}_{m_j,j}-\boldsymbol{x}*||=O(\mu_j)$. $\boldsymbol{x}^*$ is the limit point of iteration sequence $\{\boldsymbol{x}_{m_j,\ j}\}$ and $F_0(\boldsymbol{x}^*,\lambda^*,\boldsymbol{v}^*)=0$ is satisfied. Let $\boldsymbol{A}(\boldsymbol{x}^*)$ denote the active set, and let $k\in\boldsymbol{A}(\boldsymbol{x}^*)$ represent that the k^{th} component of $\boldsymbol{x}^*$ is zero. If $\boldsymbol{A}(\boldsymbol{x}^*)$ is null, there exists $J^{(1)}>0$; whenever $j>J^{(1)}$, $||\boldsymbol{X}_{m_j,j}^{-1}||$ is uniformly bounded and $\mu_j<1$. According to Lemma 3, we obtain

$$\begin{aligned}\left\|\boldsymbol{d}_{m_j,j}\right\|_2&\le\kappa_1\left\|\boldsymbol{c}_{m_j,j}\right\|_\infty+\kappa_2\left\|\boldsymbol{g}_{m_j,j}+\boldsymbol{A}_{m_j,j}\lambda_{m_j,j}-\boldsymbol{v}_{m_j,j}\right\|_\infty\\&+\kappa_3\left\|\boldsymbol{X}_{m_j,j}^{-1}\right\|\cdot\left\|\boldsymbol{X}_{m_j,j}\boldsymbol{V}_{m_j,j}\boldsymbol{e}-\mu_j\boldsymbol{e}\right\|_\infty\\&+\kappa_4\left\|\mu_j\boldsymbol{X}_{m_j,j}^{-2}\boldsymbol{e}\right\|\cdot\left\|\boldsymbol{c}_{m_j,j}\right\|_\infty\\&\le\kappa_1M_{sc}\tau_\varepsilon\mu_j^{\theta_\varepsilon}+\kappa_2M_{sd}\tau_\varepsilon\mu_j^{\theta_\varepsilon}+\kappa_3M_{sd}\left\|\boldsymbol{X}_{m_j,j}^{-1}\right\|\tau_\varepsilon\mu_j^{\theta_\varepsilon}\\&+\kappa_4M_{sc}\left\|\boldsymbol{X}_{m_j,j}^{-2}\right\|\tau_\varepsilon\mu_j^{\theta_\varepsilon}\le\rho\mu_j^{\varpi}\end{aligned}$$

Define $\rho_1=\kappa_1M_{sc}\tau_\varepsilon+\kappa_2M_{sd}\tau_\varepsilon+\kappa_3M_{sd}||X_{m_j,j}^{-1}||\tau_\varepsilon+\kappa_4M_{sc}||X_{m_j,j}^{-2}||\tau_\varepsilon$ and $\varpi=\theta_\varepsilon-1$. If $\boldsymbol{A}(\boldsymbol{x}^*)$ is not null, we have $||\boldsymbol{X}_{m_j,j}^{-1}||=\frac{1}{O(\mu_j)}$. According to Lemma 3, we get

$$\begin{aligned}\left\|\boldsymbol{d}_{m_j,j}\right\|_2&\le\kappa_1\left\|\boldsymbol{c}_{m_j,j}\right\|_\infty+\kappa_2\left\|\boldsymbol{g}_{m_j,j}+\boldsymbol{A}_{m_j,j}\lambda_{m_j,j}-\boldsymbol{v}_{m_j},j\right\|_\infty\\&+\kappa_3\left\|\boldsymbol{X}_{m_j,j}^{-1}\right\|\cdot\left\|\boldsymbol{X}_{m_j,j}\boldsymbol{V}_{m_j,j}\boldsymbol{e}-\mu_{m_j,j}\boldsymbol{e}\right\|_\infty\\&+\kappa_4\left\|\mu_j\boldsymbol{X}_{m_j,j}^{-2}\boldsymbol{e}\right\|\cdot\left\|\boldsymbol{c}_{m_j,j}\right\|_\infty\\&\le\kappa_1M_{sc}\tau_\varepsilon\mu_j^{\theta_\varepsilon}+\kappa_2M_{sd}\tau_\varepsilon\mu_j^{\theta_\varepsilon}+\kappa_3M_{sd}\tau_\varepsilon\boldsymbol{O}(\mu_j^{\theta_\varepsilon-1})\\&+\kappa_4M_{sc}\tau_\varepsilon\boldsymbol{O}(\mu_j^{\theta_\varepsilon-1})\end{aligned}$$

There exists $J^{(2)}>0$, whenever $j>J^{(2)}$, $O(\mu_j^{\theta_\varepsilon-1})<M_{\theta_\varepsilon}\mu_j^{\theta_\varepsilon-1}$, M_{θ_ε} is a positive constant and $\mu_j<1$. Define $\rho=\max\{\rho_1,\kappa_1M_{sc}\tau_\varepsilon+\kappa_2M_{sd}\tau_\varepsilon+\kappa_3M_{sd}M_{\theta_\varepsilon}\tau_\varepsilon+\ \kappa_4M_{sc}M_{\theta_\varepsilon}\tau_\varepsilon\}$ and $\varpi=\theta_\varepsilon-1$. Let $J_1=\max\{J^{(1)},J^{(2)}\}$ then the lemma is proved. ■

Lemma 5. Suppose Assumptions A hold for interior point methods with the Fiacco-McCormick barrier strategy. For an arbitrarily given $\varepsilon>0$, there exists an integer J_2, $|f(\boldsymbol{x}_{m_j+1,j})-f(\boldsymbol{x}_{m_j,j})|<\varepsilon/2$ holds for all $j>J_2$.

Proof. From (A1), objective function $f(\boldsymbol{x})$ is continuous in $[\boldsymbol{x}_{m_j,j},\boldsymbol{x}_{m_j,j}+\alpha_{\max}d_{m_j,j}]$. For an arbitrarily given $\varepsilon>0$, there exists δ_{d_1}, if $||\alpha\boldsymbol{d}_{m_j,j}||_2<\delta_{d_1}$; therefore,

$$\left|f(\boldsymbol{x}_{m_j+1,j})-f(\boldsymbol{x}_{m_j,j})\right|=\left|f(\boldsymbol{x}_{m_j,j}+\alpha\boldsymbol{d}_{m_j,j})-f(\boldsymbol{x}_{m_j,j})\right|<\varepsilon/2\tag{32}$$

It follows from Lemma 4 that if $j>J_1$, $||\alpha\boldsymbol{d}_{m_j,j}||_2<\rho\mu_j^{\varpi}<\delta_{d_1}$, we have

$$\mu_j<\left(\frac{\delta_{d_1}}{\rho}\right)^{\frac{1}{\varpi}}\tag{33}$$

From Lemma 1, $\lim_{j\to\infty}\mu_j=0$, there is an integer $J^{(3)}$ such that if $j>J^{(3)}$, then (33) is satisfied. Let $J_2=\max\{J_1,J^{(3)}\}$,(32) is thus satisfied. ■

Lemma 6. Suppose Assumptions A hold for interior point methods with the Fiacco-McCormick barrier strategy. For an arbitrarily given $\varepsilon>0$, there exists an integer J_3, $|f(\boldsymbol{x}_{1,j+1})-f(\boldsymbol{x}_{m_j,j})|<\varepsilon$ holds for all $J>J_3$.

Proof. With barrier parameter μ_j and starting point $\boldsymbol{x}_{m_j,j}$, $\boldsymbol{x}_{m_j+1,j}$ is the next iterate. $\boldsymbol{x}_{1,j+1}$ is the next iterate with barrier parameter μ_{j+1} and the same starting point as $\boldsymbol{x}_{m_j+1,j}$. This is similar to the proof for Lemma 5 that for an arbitrarily given $\varepsilon>0$, there exists δ_{d_2}, if $||\boldsymbol{x}_{m_j+1,j}-\boldsymbol{x}_{1,j+1}||<\delta_{d_2}$, then

$$\left|f(\boldsymbol{x}_{m_j+1,j})-f(\boldsymbol{x}_{1,j+1})\right|<\varepsilon/2\tag{34}$$

From (6), (7), and (8), it follows that

$$\begin{aligned}\left\|\boldsymbol{x}_{m_j+1,j}-\boldsymbol{x}_{1,j+1}\right\|&=\left\|\alpha_{m_j,j}\boldsymbol{d}_{m_j,j}-\alpha_{1,j+1}\boldsymbol{d}_{1,j+1}\right\|\\&\le\left\|(\alpha_{m_j,j}-\alpha_{1,j+1})\boldsymbol{d}_{m_j,j}\right\|+\left\|\alpha_{1,j+1}(\boldsymbol{d}_{m_j,j}-\boldsymbol{d}_{1,j+1})\right\|\\&\overset{j>J_1}{\le}\left|(\alpha_{m_i,j}-\alpha_{1,j+1})\right|\rho\mu_j^{\varpi}\\&+\left|\alpha_{1,j+1}\right|\cdot\left\|\boldsymbol{Z}_{m_j,j}\left[\boldsymbol{Z}_{m_j,j}^T\boldsymbol{H}_{m_j,j}\boldsymbol{Z}_{m_j,j}\right]^{-1}\boldsymbol{Z}_{m_j,j}^T\boldsymbol{X}_{m_j,j}^{-1}\boldsymbol{e}(\mu_{j+1}-\mu_j)\right\|\end{aligned}\tag{35}$$

If active set $\boldsymbol{A}(\boldsymbol{x}^*)$ is empty, then $\left\|\boldsymbol{Z}_{m_j,j}\left[\boldsymbol{Z}_{m_j,j}^T\boldsymbol{H}_{m_j,j}\boldsymbol{Z}_{m_j,j}\right]^{-1}\boldsymbol{Z}_{m_j,j}^T\boldsymbol{X}_{m_j,j}^{-1}\boldsymbol{e}\right\|$ is, of course, bounded. If active set $\boldsymbol{A}(\boldsymbol{x}^*)$ is not null, $\lim_{j\to\infty}\left\|\boldsymbol{Z}_{m_j,j}\left[\boldsymbol{Z}_{m_j,j}^T\boldsymbol{H}_{m_j,j}\boldsymbol{Z}_{m_j,j}\right]^{-1}\boldsymbol{Z}_{m_j,j}^T\boldsymbol{X}_{m_j,j}^{-1}\boldsymbol{e}\right\|$ is also bounded because of the definition of $\boldsymbol{H}_{m_j,\ j}$ and $||\boldsymbol{x}_{m_j,j}-\boldsymbol{x}^*||=O(\mu_j)$. In either case, there exists $J^{(4)}>0$, whenever $j>J^{(4)}$, $\left\|\boldsymbol{Z}_{m_j,j}\left[\boldsymbol{Z}_{m_j,j}{}^T\boldsymbol{H}_{m_j,j}\boldsymbol{Z}_{m_j,j}\right]^{-1}\boldsymbol{Z}_{m_j,j}^T\boldsymbol{X}_{m_j,j}^{-1}\boldsymbol{e}\right\|<M_0$, M_0 is a positive constant and $\mu_j<1$. Let $J^{(5)}=\max\{J^{(4)},J_1\}$, when $j>J^{(5)}$. From (35), we get

$$\left\|\boldsymbol{x}_{m_j+1,j}-\boldsymbol{x}_{1,j+1}\right\|<(\rho+2M_0)\mu_j^{\varpi}<\delta_{d_2}\tag{36}$$

From Lemma 1, $\lim_{j\to\infty}\mu_j=0$; thus, there exists $J^{(6)}>0$, whenever $j>J^{(6)}$. Equations (36) and (34) are then satisfied. We define $J_3=\max\{J^{(5)},J^{(6)}\}$. From Lemma 5 and (34), whenever $j>J_3$, it follows that

$$\begin{aligned}&\left|f(\boldsymbol{x}_{1,j+1})-f(\boldsymbol{x}_{m_j,j})\right|\\&\le\left|f(\boldsymbol{x}_{m_j+1,j})-f(\boldsymbol{x}_{1,j+1})\right|+\left|f(\boldsymbol{x}_{m_j+1,j})-f(\boldsymbol{x}_{m_j,j})\right|<\varepsilon\end{aligned}\tag{37}$$

and we get the desired result. ■

Lemma 7. Suppose Assumptions A hold for interior point methods with the Fiacco-McCormick barrier strategy. Then, for an arbitrarily given $\varepsilon>0$, there exists an integer J_4, $\left|\left\|\boldsymbol{c}(\boldsymbol{x}_{1,j+1})\right\|_\infty-\left\|\boldsymbol{c}(\boldsymbol{x}_{m_j,j})\right\|_\infty\right|<\varepsilon$ holds for all $j>J_4$.

As the proof of this lemma is similar to the proof for Lemma 6, we omit it for the sake of brevity. ■

In the following, we prove that if the optimization procedure has no improvement after a number of iterations, it can be terminated by the modified CDC for any predefined number η given in the Section on CDC strategy.

Theorem 1. Suppose Assumptions A hold for interior point methods with the Fiacco-McCormick barrier strategy. For an arbitrarily given positive integer N and arbitrarily given $\varepsilon > 0$, there exists an integer T_1. Whenever $t \geq T_1$, the number of t that satisfies

$$\theta^t_{\text{prog}} = S(\max\left\{\delta^t_{\text{feasChg}}, \delta^t_{\text{objChg}}\right\}, \varepsilon_0) \geq S(\varepsilon, \varepsilon_0), \tag{38}$$

is equal to or greater than N.

Proof. As the starting point for interior point methods is strictly larger than zero and the iterate in the process of line search satisfies fraction to the boundary rule:

$$a^{\max}_{i,j} := \max\left\{a \in (0, 1] : \boldsymbol{x}_{i,j} + a\boldsymbol{d}_{i,j} \geq (1-\tau)\boldsymbol{x}_{i,j}\right\} \tag{39}$$

If $\boldsymbol{x}_{i,j} > 0$, from (39), we have

$$\boldsymbol{x}_{i+1,j} = \boldsymbol{x}_{i,j} + a\boldsymbol{d}_{i,j} \geq (1-\tau)\boldsymbol{x}_{i,j} > 0 \text{ or } \boldsymbol{x}_{1,j+1} = \boldsymbol{x}_{m_j,j} + a\boldsymbol{d}_{m_j,j} \geq (1-\tau)\boldsymbol{x}_{m_j,j} > 0 \tag{40}$$

As $\lim_{j\to\infty} \mu_j = 0$, there exists $J^{(7)}$, whenever $j > J^{(7)}$,

$$\mu_j < \varepsilon \tag{41}$$

Let $t(j) = 1 + \sum_{l=1}^{j} m_l$ and define $J^{(8)} = \max\{J^{(7)}, J_3, J_4\}$. From Lemma 6, whenever $j \geq J^{(8)}$, we have $\delta^{t(j)}_{\text{objChg}} = |f(\boldsymbol{x}_{1,j+1}) - f(\boldsymbol{x}_{m_j,j})| < \varepsilon$. From Lemma 7 and (40), whenever $j \geq J^{(8)}$, we have $\delta^{t(j)}_{\text{feasChg}} < \varepsilon$. Whenever $j \geq J^{(8)}$, (41) is also satisfied. Finally, we define $T_1 = t(J^{(8)} + N)$. Whenever $t > T_1$, the value of t that satisfies $\theta^{t(j)}_{\text{prog}} = S(\max\{\delta^{t(j)}_{\text{feasChg}}, \delta^{t(j)}_{\text{objChg}}\}, \varepsilon) \geq S(\varepsilon, \varepsilon_0)$ is equal to or greater than N. If $\varepsilon < \varepsilon_0$, we have $\theta^{t(j)}_{\text{prog}} \geq 1$ ■

Remark 1. In the solution process of the barrier problem (2), when the norm of the search direction becomes very small, (38) can also be satisfied before the tolerance (13) for this barrier problem is accessed. Theorem 1 ensures that the optimization procedure can be finally terminated by the modified CDC.

Assumptions B. Interior point methods with adaptive barrier strategy proposed by Nocedal et al.[19] satisfies the following conditions:

(B1) Centering parameter σ_j is uniformly bounded, i.e., there exists a positive constant $M_\sigma > 0$, with $0 < \sigma_j < M_\sigma$ for all iterations.

(B2) There exists a sufficiently small ϕ_0, whenever $\phi(\boldsymbol{x}_{i+1}, \lambda_{i+1}, v_{i+1}) < \phi_0$, the updating strategy of the barrier parameter will never involve the Fiacco-McCormick strategy.

Theorem 2. Suppose Assumptions A and Assumptions B hold for interior point methods with the adaptive barrier strategy.[19] For an arbitrarily given positive integer N and the arbitrarily given $\varepsilon > 0$, there exists an integer T_2. Whenever $t \geq T_2$, and the value of t that satisfies

$$\theta^t_{\text{prog}} = S(\max\left\{\delta^t_{\text{feasChg}}, \delta^t_{\text{objChg}}\right\}, \varepsilon_0) \geq S(\varepsilon, \varepsilon_0) \tag{42}$$

is equal to or greater than N.

Proof. From the section on Updating the Barrier Parameter, the iteration sequence produced by interior point methods with adaptive barrier strategy satisfies

$$\lim_{j\to\infty} \phi(\boldsymbol{x}_{m_j,j}, \lambda_{m_j,j}, \boldsymbol{v}_{m_j,j}) = 0 \tag{43}$$

Thus, from (18), we have

$$\begin{aligned} &\lim_{j\to\infty} \left\|\boldsymbol{g}_{m_j,j} + \boldsymbol{A}_{m_j,j}\lambda_{m_j,j} - \boldsymbol{v}_{\boldsymbol{m}_j,j}\right\|_\infty = 0 \\ &\lim_{j\to\infty} \left\|\boldsymbol{X}_{m_j,j}\boldsymbol{V}_{m_j,j}\boldsymbol{e}\right\|_\infty = 0, \lim_{j\to\infty} \left\|\boldsymbol{c}_{m_j,j}\right\|_\infty = 0 \end{aligned} \tag{44}$$

Under the Fiacco-McCormick barrier strategy, the iterate $(\boldsymbol{x}_{m_j,\ j}, \lambda_{m_j,\ j}, \boldsymbol{v}_{m_j,\ j})$ denotes the approximate solution of the barrier problem with the barrier parameter μ_j. Under the adaptive barrier parameter updating strategy, $(\boldsymbol{x}_{m_j,\ j}, \lambda_{m_j,\ j}, \boldsymbol{v}_{m_j,\ j})$ denotes that the iterate after one iteration and the subscript m_j is a constant equal to one.

For ϕ_0 in (B2), from (43), there exists integer $J^{(9)} > 0$, whenever $j > J^{(9)}$, such that $\phi_j < \phi_0$. From (B2), if $j > J^{(9)}$, we have that m_j is a constant equal to one. From (15), (44), and (B1), it follows that

$$\begin{aligned} &\lim_{j\to\infty} \mu_j = 0 \\ &\lim_{j\to\infty} \boldsymbol{x}_{1,j} = \boldsymbol{x}* \end{aligned} \tag{45}$$

From (A1), the objective function $f(\boldsymbol{x})$ is continuous in the open set C containing interval $[\boldsymbol{x}_{1,j}, \boldsymbol{x}_{1,j+1}]$. Therefore, we have $\lim_{j\to\infty} f(\boldsymbol{x}_{1,j}) = f(\boldsymbol{x}*)$. For an arbitrarily given $\varepsilon > 0$, there exists $J^{(10)} > 0$, whenever $j > J^{(10)}$. It follows that

$$\begin{aligned} \left|f(\boldsymbol{x}_{1,j+1}) - f(\boldsymbol{x}_{1,j})\right| &= \left|f(\boldsymbol{x}_{1,j+1}) - f(\boldsymbol{x}*) + f(\boldsymbol{x}*) - f(\boldsymbol{x}_{1,j})\right| \\ &\leq \left|f(\boldsymbol{x}_{1,j+1}) - f(\boldsymbol{x}*)\right| + \left|f(\boldsymbol{x}*) - f(\boldsymbol{x}_{1,j})\right| < \varepsilon \end{aligned} \tag{46}$$

We define $t(j) = 1 + \sum_{l=1}^{j} m_l$, whenever $j > J^{(10)}$, $\delta^{t(j)}_{\text{feasChg}} = |f(\boldsymbol{x}_{1,j+1}) - f(\boldsymbol{x}_{1,j})| < \varepsilon$ is satisfied. From (44), there exists $J^{(11)} > 0$, and $\delta^{t(j)}_{\text{feasChg}} = \left|\left\|\boldsymbol{c}_{1,j+1}\right\|_\infty - \left\|\boldsymbol{c}_{1,j}\right\|_\infty\right| < \varepsilon$ holds for all $j > J^{(11)}$. Let $J^{(12)} = \max(J^{(9)}, J^{(10)}, J^{(11)})$. We define $T_2 = t(J^{(12)} + N)$. Whenever $t > T_2$, the number of t that satisfies $\theta^{t(j)}_{\text{prog}} = S(\max\{\delta^{t(j)}_{\text{feasChg}}, \delta^{t(j)}_{\text{objChg}}\}, \varepsilon_0) \geq S(\varepsilon, \varepsilon_0)$ is equal to or greater than N. If $\varepsilon < \varepsilon_0$, we have $\theta^{t(j)}_{\text{prog}} \geq 1$. ■

In Wächter,[8] the criticality measure $\chi(\boldsymbol{x}_{i,j}) := ||\tilde{\boldsymbol{p}}_{i,j}||_2$ and constraint violation $\theta(\boldsymbol{x}_{i,j}) := ||\boldsymbol{c}(\boldsymbol{x}_{i,j})||_\infty$ are used to measure the optimality of the iterates of barrier problem with barrier parameter μ_j. In this study, we prove that convergence depth $\theta^{i,j}_{\text{convg}}$ can be used to measure the optimality of iterates. The monotonicity of sigmoid function (26) decreases on $\delta^{i,j}$ and it is continuous; therefore, there exists the inverse function S^{-1},

$$\delta^{i,j} = S^{-1}(\theta^{i,j}_{\text{convg}}, \varepsilon_0) \tag{47}$$

Theorem 3. Suppose Assumptions A hold for both interior point methods with the Fiacco-McCormick barrier strategy and with adaptive barrier strategy. Then, every iterate of iteration sequence $\boldsymbol{x}_{1,1}, \boldsymbol{x}_{2,1}, \ldots, \boldsymbol{x}_{m_1,1}, \boldsymbol{x}_{1,2}, \boldsymbol{x}_{2,2}, \ldots, \boldsymbol{x}_{m_2,2}, \boldsymbol{x}_{1,3} \ldots$ produced by interior point methods satisfies the following conditions,

$$\theta(\boldsymbol{x}_{i,j}) \leq \delta^{i,j} \tag{48}$$

$$\chi^2(\boldsymbol{x}_{i,j}) \leq \vartheta \delta^{i,j} \tag{49}$$

where ϑ is a positive constant.

Proof. Equation (47) implies that

$$\delta^{i,j}_{\text{objErr}} \leq \delta^{i,j}, \delta^{i,j}_{\text{feasErr}} \leq \delta^{i,j}, \mu_j \leq \delta^{i,j} \tag{50}$$

From (20), (40), and (47), we immediately get (48). Assumptions A indicate that $\chi_{i,j} = ||\tilde{\boldsymbol{p}}_{i,j}||_2$ is bounded by M_{p} (M_{p} is a positive constant). Similar to Lemma 6, we have $\left\|\boldsymbol{Z}_{i,j}[\boldsymbol{Z}^T_{i,j}\boldsymbol{H}_{i,j}\boldsymbol{Z}_{i,j}]^{-1}\boldsymbol{Z}^T_{i,j}\boldsymbol{X}^{-1}_{i,j}e\right\| < M_1$ and $\left\|\boldsymbol{Z}_{i,j}[\boldsymbol{Z}^T_{i,j}\boldsymbol{H}_{i,j}\boldsymbol{Z}_{i,j}]^{-1}\boldsymbol{Z}^T_{i,j}\boldsymbol{H}_{i,j}\right\| < M_2$. $\boldsymbol{d}_{i,j}$ is bounded by M_{d} (M_{d} is a positive constant). From (6) and (8), we get

$$\begin{aligned}
\left\|\boldsymbol{d}^T_{i,j}\boldsymbol{Z}_{i,j}\tilde{\boldsymbol{p}}_{i,j}\right\| &= \left\|\boldsymbol{d}^T_{i,j}\boldsymbol{Z}_{i,j}\left[\boldsymbol{Z}^T_{i,j}\boldsymbol{H}_{i,j}\boldsymbol{Z}_{i,j}\right]^{-1}\boldsymbol{Z}^T_{i,j}(\boldsymbol{g}_{i,j} - \mu_j \boldsymbol{X}^{-1}_{i,j}\boldsymbol{e} + \boldsymbol{H}_{i,j}\boldsymbol{q}_{i,j})\right\| \\
&\leq \frac{1}{M_{\text{H}}}\left|\boldsymbol{g}^T_{i,j}\boldsymbol{d}_{i,j}\right| + M_d M_1 \mu_j + \frac{M_{\text{d}}M_2}{M_{\text{A}}}\left\|\boldsymbol{c}(\boldsymbol{x}_{i,j})\right\| \\
&\leq \frac{1}{M_{\text{H}}}\delta^{i,j}_{\text{objErr}} + M_{\text{d}}M_1\mu_j + \frac{M_{\text{d}}M_2}{M_{\text{A}}}\delta^{i,j}_{\text{feasErr}} \\
&\leq \left(\frac{1}{M_{\text{H}}} + M_{\text{d}}M_1 + \frac{M_{\text{d}}M_2}{M_{\text{A}}}\right)\delta_{i,j}
\end{aligned} \tag{51}$$

$$\left\|\boldsymbol{d}^T_{i,j}\boldsymbol{Z}_{i,j}\tilde{\boldsymbol{p}}_{i,j}\right\| = \left\|(\boldsymbol{q}_{i,j} + \boldsymbol{p}_{i,j})^T\boldsymbol{Z}_{i,j}\tilde{\boldsymbol{p}}_{i,j}\right\| \geq \chi^2(\boldsymbol{x}_{i,j}) - \frac{M_{\text{p}}}{M_{\text{A}}}\boldsymbol{c}(\boldsymbol{x}_{i,j}) \tag{52}$$

Considering (51) and (52), we get

$$\begin{aligned}
\chi^2(\boldsymbol{x}_{i,j}) &\leq \frac{M_{\text{p}}}{M_{\text{A}}}\boldsymbol{c}(\boldsymbol{x}_{i,j}) + \left(\frac{1}{M_{\text{H}}} + M_d M_1 + \frac{M_{\text{d}}M_2}{M_{\text{A}}}\right)\delta_{i,j} \\
&\leq \left(\frac{1}{M_{\text{H}}} + M_{\text{d}}M_1 + \frac{M_{\text{d}}M_2 + M_{\text{p}}}{M_{\text{A}}}\right)\delta_{i,j}
\end{aligned} \tag{53}$$

where $\vartheta = \frac{1}{M_{\text{H}}} + M_{\text{d}}M_1 + \frac{M_{\text{d}}M_2 + M_{\text{p}}}{M_{\text{A}}}$. The theorem is therefore proved. ■

For both barrier parameter updating strategies (Fiacco-McCormick barrier strategy and adaptive barrier strategy), under the proper assumptions, we prove that the optimization procedure can be finally terminated (Theorem 1 and Theorem 2). Theorem 3 illustrates the relationships among $\delta_{i,j}$, constraint violation $\theta(\boldsymbol{x}_{i,j})$, and criticality measure $\chi(\boldsymbol{x}_{i,j})$. It shows that if $\{\boldsymbol{x}_{i,j}\}$ (or its subsequence) with $\delta(\boldsymbol{x}_{i,j}) \rightarrow 0$ converges to a feasible limit point $\boldsymbol{x}^*$, then $\chi(\boldsymbol{x}_{i,j}) \rightarrow \chi(\boldsymbol{x}^*) = 0$ is also satisfied. Therefore, $\boldsymbol{x}^*$ corresponds to a KKT solution.[8] The convergence depth $\theta^{i,j}_{\text{convg}}$ is the sigmoid transformation of $\delta^{i,j}$, which intuitively measures the optimality of the current iterate.

Table 1. Basic Information of Examples from CUTE

Problem	#Var	#Bdd	#Eq	#Ineq	#nzJac	#nzHes
avion2	49	49	15	0	43	123
bt8	5	0	2	0	6	5
dallasl	837	837	598	0	1674	821
dallasm	164	164	119	0	328	158
dallass	44	44	29	0	88	39
flosp2hh	650	0	0	0	0	9582
orthrds2	203	0	100	0	500	906
sineali	20	20	0	0	0	39

Numerical Results

In this section, the interior point method algorithm used is the IPOPT software (version 3.6.1)[23]. The linear solver utilized is MA57. IPOPT implements interior point methods both with the Fiacco-McCormick barrier strategy and with the adaptive barrier strategy. We present two numerical experiments to compare the performance of interior point methods with the traditional termination criteria and the modified CDC. The examples of the first numerical experiment are from the CUTE collection. The barrier parameter updating strategy of IPOPT used for the first experiment is the Fiacco-McCormick barrier strategy. The examples of the second experiment are formulated based on Aspen Plus, including the optimization problem of depropanizer and debutanizer distillation column system, the data reconciliation problem of large-scale air separation system, and the optimization problem of large-scale ethylene separation system. In the second experiment, the adaptive barrier strategy is used for IPOPT.

For IPOPT with traditional termination criteria, the desired convergence tolerance is 1.0×10^{-6}, the desired threshold for dual infeasibility is 1.0×10^{-4}, the desired threshold for constraint violation is 1.0×10^{-4}, and the desired threshold for complementarity conditions is 1.0×10^{-4}. We set the maximum number of iterations as 3000. For IPOPT with the modified CDC, the specified threshold ε_0 is set to 1.0×10^{-6}. The acceptable convergence depth θ_0 is set to 0.95 in the first numerical experiment. In the second experiment, we set $\theta_0 =$

Table 2. Comparison of IPOPT with Traditional Criteria and Modified CDC on CUTE

			Obj		Iter		CPU(s)		Feas		Terminated	
Problem	θ_{convg}	θ_{prog}	CDC	Trad	CDC	Trad	CDC	Trad	CDC	Trad	CDC	Trad
avion2	0.954	0.453	9.468×10^7	9.468×10^7	60	95	0.125	0.219	2.8×10^{-13}	2.8×10^{-13}	√	√
bt8	0.985	0.974	1.000	1.000	10	27	0.016	0.078	9.5×10^{-7}	0.00	√	√
dallasl	0.952	0.810	-2.026×10^5	-2.026×10^5	79	235	0.735	2.156	2.8×10^{-14}	1.7×10^{-14}	√	√
dallasm	0.952	0.809	-4.820×10^4	-4.820×10^4	23	280	0.063	0.875	7.1×10^{-15}	3.6×10^{-15}	√	√
dallass	0.964	0.581	-3.239×10^4	-3.239×10^4	21	418	0.031	0.797	3.6×10^{-15}	1.4×10^{-14}	√	√
flosp2hh	1.040	0.760	38.87	38.87	4	3000	0.281	136.94	0.0	0.0	√	X
orthrds2	0.956	0.956	1.544×10^3	1.544×10^3	35	43	0.219	0.328	2.5×10^{-15}	2.2×10^{-16}	√	√
sineali	0.986	0.986	-1.901×10^3	-1.901×10^3	29	3000	0.047	6.11	0.00	0.00	√	X

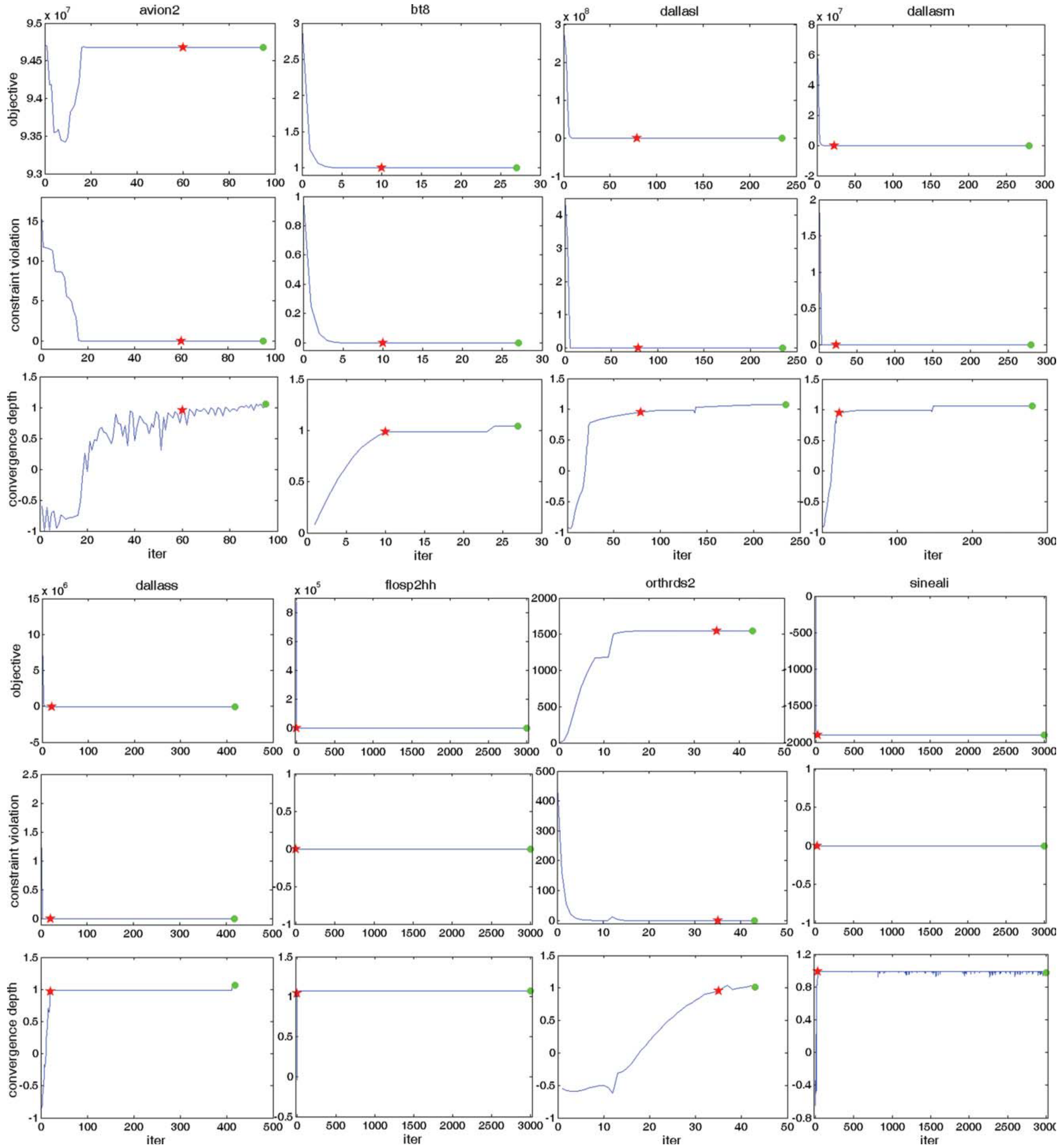

Figure 4. Comparison of IPOPT with modified CDC and traditional criteria for problems from CUTE.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

1.0 to obtain deeper convergence. In order to minimize the risk of incorrectly declaring "no improvement," the predefined number η for declaring "no improvement" is set to 10.

All the numerical experiments were performed on a Dell Precision T3400 running Windows XP and with Intel(R) Core (TM)2 Duo CPU, 2.66GHz and 2.0G RAM memory.

Numerical experiments on the CUTE

The examples were selected representatively from CUTE collection. Table 1 gives the detailed description of these

Table 3. Comparison of Feasibility of Solutions for Depropanizer and Debutanizer Distillation Column System

Model	Standard Algorithm (DMO)	IPOPT with CDC	IPOPT with Traditional Termination Criteria
Original model	-2.09×10^{-5}	-2.1912×10^{-8}	-1.1176×10^{-8}
First scaled model	–	-2.1912×10^{-8}	-1.1176×10^{-8}
Second scaled model	5.30×10^{-7}	–	–

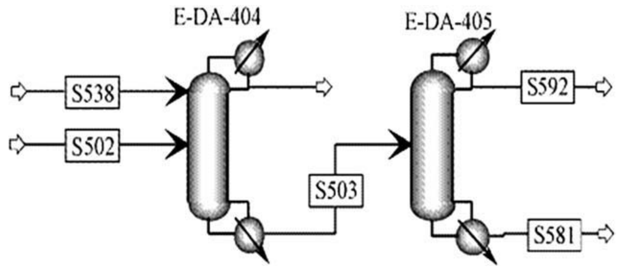

Figure 5. Flowsheet of depropanizer and debutanizer distillation column system.

examples. The columns of Table 1 represent the number of variables (#var), bounds on variables (#bdd), equality constraints (#eq), inequality constraints (#ineq), non-zero elements of Jacobian matrix of constraints (#nzJac), and non-zero elements of Hessian matrix of Lagrange function (#nzHes). Table 2 shows the numerical results of these examples using interior point methods with the traditional termination criteria and with modified CDC. The columns of Table 2 represent the number of iterations (iter), computation time cost (CPU), the value of objective function (obj), the constraint violation (feas), and termination status of the algorithm ("$\surd$" means terminated successfully while "X" means unreasonably slow convergence or convergence failure). CDC represents the CDC strategy, and Trad represents the traditional termination criteria.

Whenever the convergence depth is greater than 0.95, the CDC strategy successfully terminates the algorithm. If $\theta_{\text{prog}} \geq 1$ is satisfied for more than 10 times, an underconverged result is achieved. From the θ_{convg} column, the runs with CDC for all the problems terminate successfully, while from the *iter* and the *terminated* columns, the runs with traditional termination criteria for *flosp2hh* and *sineali* fail as the number of iterations exceeds the maximum number of iterations (3000). With nearly the same objective value and constraint violation, the runs with CDC reduce the number of iterations substantially when compared with the runs with traditional termination criteria. Furthermore, the CDC strategy adequately balances the computation cost and acceptable solution. We can see the advantage of CDC intuitively from Figure 4, where the termination points of the runs with CDC and traditional termination criteria are marked by "★" and "●," respectively.

Numerical experiments on Aspen Plus

The rigorous model obtained from Aspen Plus through the Aspen Open Solvers interface is properly and automatically scaled based on the physical meaning by Aspen Plus. We call it the first-scaled model, on which IPOPT termination is based. It may also be scaled by the optimization algorithm again (second-scaled model). To verify whether the solutions returned by IPOPT are feasible, we compare them with the optimization solution produced by standard algorithm (DMO, of which the termination is based on the second-scaled model) in Aspen Plus. The feasibility comparisons based on the original rigorous model (using command "print largest equations" in Aspen Plus to print the constraint violations) are shown for each example below (Tables 3, 5, and 7).

The optimization problems formulated with Aspen Plus are practical engineering problems that generally have two kinds of objective function: dimensional (e.g., the economic objective function) and dimensionless (e.g., the data reconciliation objective function). For the dimensional case, the calculations of predictive improvement (21) and objective change (23) are influenced by the unit of measure. In order to eliminate the impact of units, the predictive improvement and the objective change should be scaled by the objective function. Equations (21) and (23) are replaced with (54) and (55) for dimensional objective function.

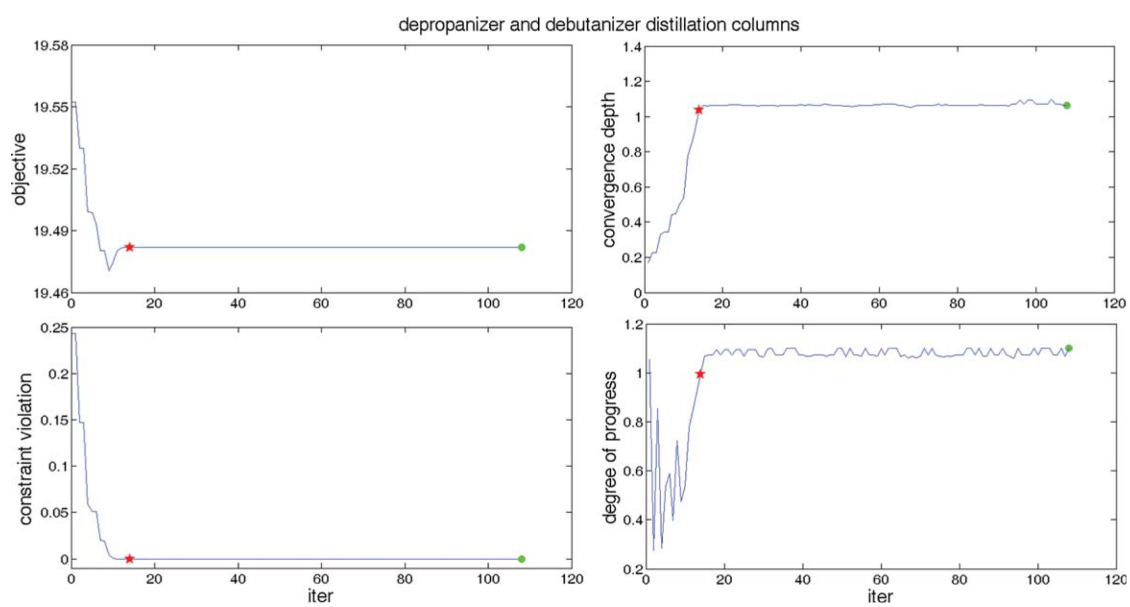

Figure 6. Comparison of IPOPT with traditional criteria and modified CDC for depropanizer and debutanizer distillation column system.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Table 4. Comparison of IPOPT with Traditional Termination Criteria and Modified CDC for Depropanizer and Debutanizer Distillation Column System

θ_{convg}	θprog	Convergence Depth Control	Traditional Termination Criteria
1.0350	0.9940		
Objective (RMB/h)		1.9481946227×10^5	1.9481946260×10^5
Iteration		14	108
CPU(s)		8.06	92.90
Feasibility		$-2.1\,912 \times 10^{-8}$	$-1.1\,176 \times 10^{-8}$
Terminated		√	√

$$\delta^t_{\text{objErr}} = |\boldsymbol{g}^T(\boldsymbol{x}^t)\boldsymbol{d}(\boldsymbol{x}^t)| / |f(\boldsymbol{x}_0)| \tag{54}$$

$$\delta^t_{\text{objChg}} = |f(\boldsymbol{x}^t) - f(\boldsymbol{x}^{t-1})| / |f(\boldsymbol{x}_0)| \tag{55}$$

where $\boldsymbol{x}_0$ is the starting point that can be obtained through the sequential modular method in Aspen Plus. For the dimensionless case, the calculations of predictive improvement and objective change are still based on (21) and (23)

In the following tests, the objective functions for depropanizer and debutanizer distillation columns system and ethylene separation system are dimensional, so the calculations of predictive improvement and objective function are based on (54) and (55). The objective function for reconciliation problem of air separation system is dimensionless, so (21) and (23) are utilized.

Remark 2. $f(\boldsymbol{x}_0)$ is bounded; therefore, all the lemmas and theories analyzed in the section properties of CDC still hold.

Optimization of depropanizer and debutanizer distillation column system

The depropanizer and debutanizer distillation column system is an essential part in ethylene production. We can refer to Jiang[24] for the flowsheet of the column system, as shown in Figure 5. The objective is defined as

$$\max : F = (278.88 \times S1\_bw + 227.752 \times S1\_dw)/10000 \tag{56}$$

where unit of F is 10,000 RMB/h*. The values *S1_bw* and *S1_dw* represent the flow rates of the top production of depropanizer column and debutanizer column, respectively, in terms of kmol/h. Considering the quality of the productions, the optimization problem is formulated as follows:

$$\begin{aligned} &\min : -F \\ &\text{s.t. mesh equations of the distillation columns} \\ &\quad \text{connection equations of the distillatin columns} \\ &\quad \text{Percentage of } C_3H_6 \geq 0.93 \\ &\quad \text{Percentage of } C_5 \leq 0.0095 \end{aligned} \tag{57}$$

For this model, the number of variables is 3889, the number of equality constraints is 3887, there are two degrees of freedom, and the number of non-zero elements of Jacobian of constraints is 76,197 (This model information is based on the output file of IPOPT).

Table 3 indicates that the solutions for optimization problem of the depropanizer and debutanizer distillation column system produced by IPOPT with traditional termination criteria and the IPOPT with CDC are both feasible, as compared with the results returned by standard algorithm in Aspen Plus.

Figure 6 shows the trends of all the performance indicators in the optimization procedure. From the objective and constraint violation subgraphs, we see that the objective value and constraint violation stay nearly the same after ~15 iterations. The subgraph convergence depth (θ_{convg}) shows that, during the first 20 iterations, the convergence depth rises dramatically until it reaches the predefined threshold of 1.0. It is almost the same in following iterations (i.e., the optimality of the solution is seldom improved). The degree of progress (θ_{prog}) subgraph shows the improvement of objective and feasibility.

The results in Table 4 show that the runs with modified CDC and traditional termination criteria are both successful. The run with CDC makes the profit of the products 1.9481946227×10^5 RMB/h. There is only a 3.3×10^{-4} RMB/h difference between the CDC and the traditional termination criteria. The difference is negligible from practical point of view. However, the iteration number is reduced from 108 to 14 (an 87.04% reduction), and the time cost is reduced from 92.90 to 8.06 s (a 91.32% reduction). We reduce the computation time cost substantially at the expense of little profit loss.

Reconciliation problem of large-scale air separation system

The flowsheet of large-scale air separation system is shown in Figure 7. The detailed information on this system can be found in Zhu[25] and Zhang.[26] The measurement data of the process system reflect the run status of equipment and serve as the basis for process simulation and optimization. As data measurement errors are inevitable, we need to reconcile the data before they are applied for real time optimization. In this section, we try to reconcile the measured data based on the rigorous model of air separation system, which is constructed as follows:

$$\begin{aligned} &\min \sum_{i=1}^{N} \left[(\hat{u}_i - \tilde{u}_i)^2 / d^2 \right] \\ &\text{s.t.} \quad c(\hat{\boldsymbol{u}}, \boldsymbol{x}) = 0 \\ &\qquad \boldsymbol{u}_L \leq \hat{\boldsymbol{u}} \leq \boldsymbol{u}_U \\ &\qquad \boldsymbol{x}_L \leq \boldsymbol{x} \leq \boldsymbol{x}_U \end{aligned} \tag{58}$$

where $\tilde{\boldsymbol{u}}$ are the measurement data, $\hat{\boldsymbol{u}}$ are the reconciliation data, d^2 is the square deviation of measurement data $\tilde{\boldsymbol{u}}$, and $\boldsymbol{x}$ are the unmeasured data. Constraints $\boldsymbol{c}(\tilde{\boldsymbol{u}}, \boldsymbol{x})$ are the rigorous models of the air separation system, which have mass balance equations, energy balance equations, and so on. For this reconciliation model, the number of variables is 7184, the number of equality constraints is 7168, with 16 degrees of freedom, and the number of non-zero elements of Jacobian of constraints is 40,034 (This model information is based on the output file of IPOPT).

Table 5 indicates that for this system the solutions produced by IPOPT with traditional termination criteria and IPOPT with CDC are both feasible when compared with the results returned by standard algorithm in Aspen Plus.

Figure 8 shows the trends of all the performance indicators in the optimization procedure.

*10,000 RMB/h ≅ 1465.03 USD/h.

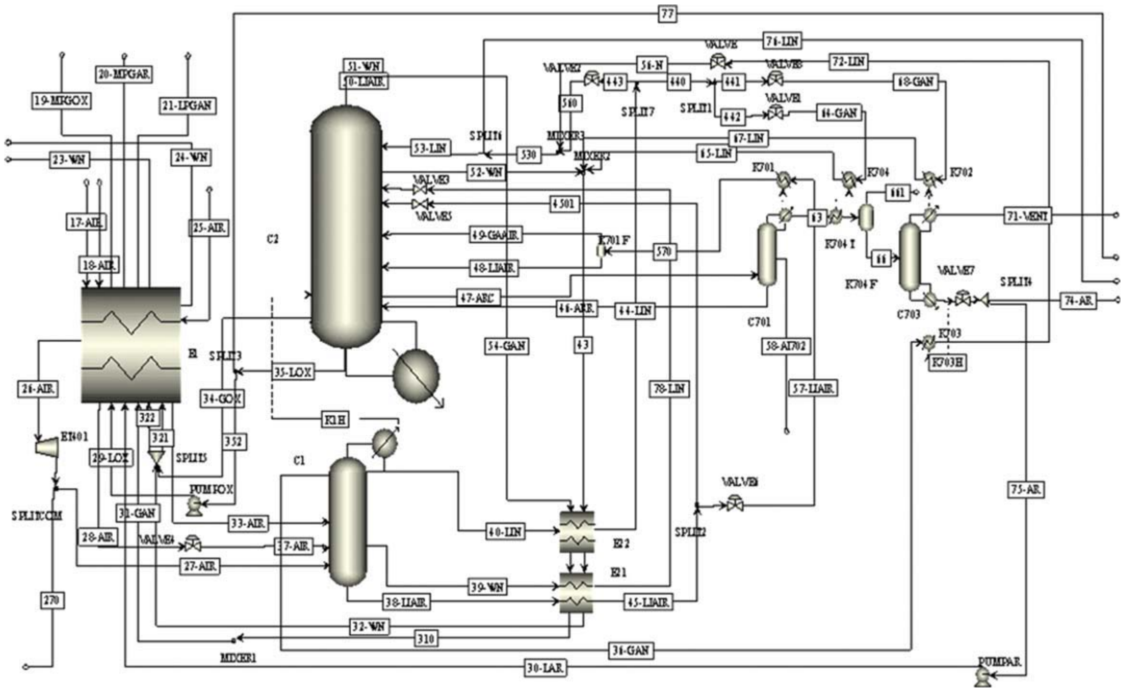

Figure 7. Flowsheet of large-scale air separation system.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Table 5. Comparison of Feasibility of Solutions for Air Separation System

Model	Standard Algorithm (DMO)	IPOPT with CDC	IPOPT with Traditional Termination Criteria
Original model	1.846×10^{-3}	2.5313×10^{-8}	1.3039×10^{-8}
First scaled model	–	2.5313×10^{-8}	1.3039×10^{-8}
Second scaled model	1.397×10^{-11}	–	–

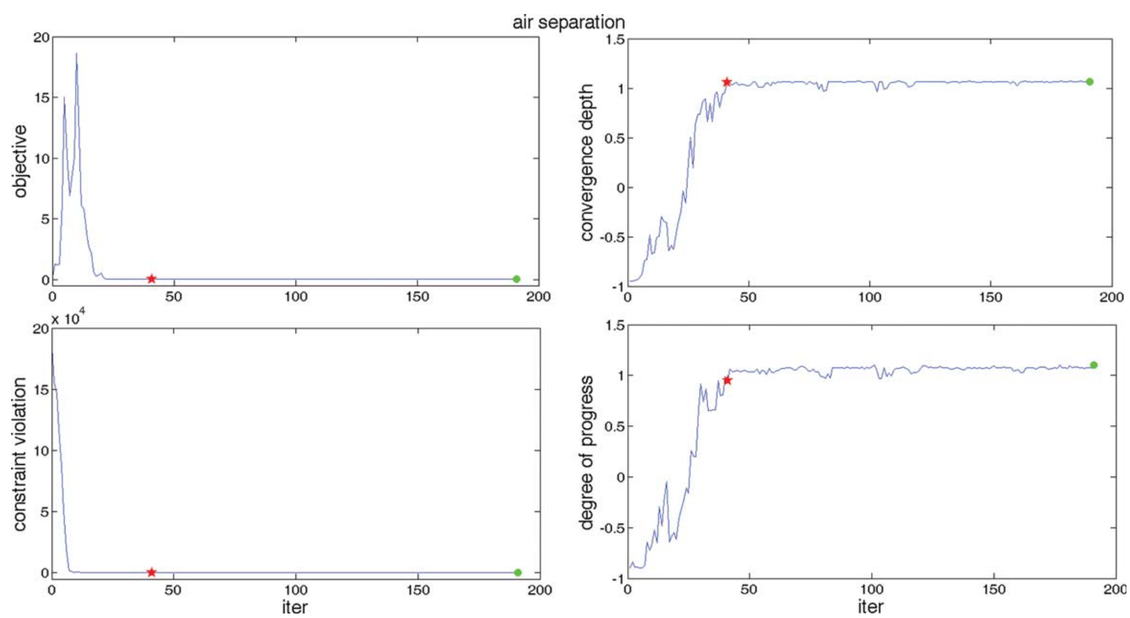

Figure 8. Comparison of IPOPT with traditional criteria and modified CDC for air separation.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Table 6. Comparison of IPOPT with Traditional Termination Criteria and Modified CDC for Air Separation System

θ_{convg} 1.0564	θprog 0.9510	Convergence Depth Control	Traditional Termination Criteria
Objective (RMB/h)		4.0051×10^{-2}	4.0051×10^{-2}
Iteration		41	191
CPU(s)		16.47	92.83
Feasibility		2.5313×10^{-8}	1.3039×10^{-8}
Terminated		√	√

The results given in Table 6 indicate that when comparing with the run under traditional termination criteria, the run with CDC reduces the number of iterations from 191 to 41 (a 78.53% reduction), while the computation cost is reduced from 92.83 to 16.47 s (an 82.23% reduction). Based on such results, CDC substantially saves on computation time.

Optimization problem of large-scale ethylene separation system

The flowsheet of a large-scale ethylene separation system is shown in Figure 9. The system of ethylene separation is composed of the following: demethanizer, deethanizer, acetylene hydrogenation system, ethylene column, ethylene product storage system, depropanizer, debutanizer and, allene, and methylacetylene hydrogenation system. The main products are ethylene, ethane, butane, and propylene. See Fang[27] for detailed information regarding this flowsheet.

Table 7. Comparison of Feasibility of Solutions for Ethylene Separation System

Model	Standard Algorithm (DMO)	IPOPT with CDC	IPOPT with Traditional Termination Criteria
Original model	−1.502	5.4539×10^{-5}	5.2399×10^{-5}
First scaled model	–	5.4539×10^{-5}	5.2399×10^{-5}
Second scaled model	5.732×10^{-9}	–	–

We constructed the objective function as the economic profit of these four main products, which is defined as follows[27]:

$$J = \left(\sum_{i=1}^{5} w_i S_i\right) \Big/ 10000 \qquad (59)$$

where S_1, S_2, and S_3 are the molar flow rates (kmol/h) of the products ethylene, ethane, and butane, respectively. S_4 and S_5 are the molar flow rates (kmol/h) of the propylene at ports S549 and S550, respectively. The terms w_i, $i = 1,\ldots,5$ represent the prices (RMB/kmol) of the product flow rates. J is the total profit (unit = 10,000 RMB/h).

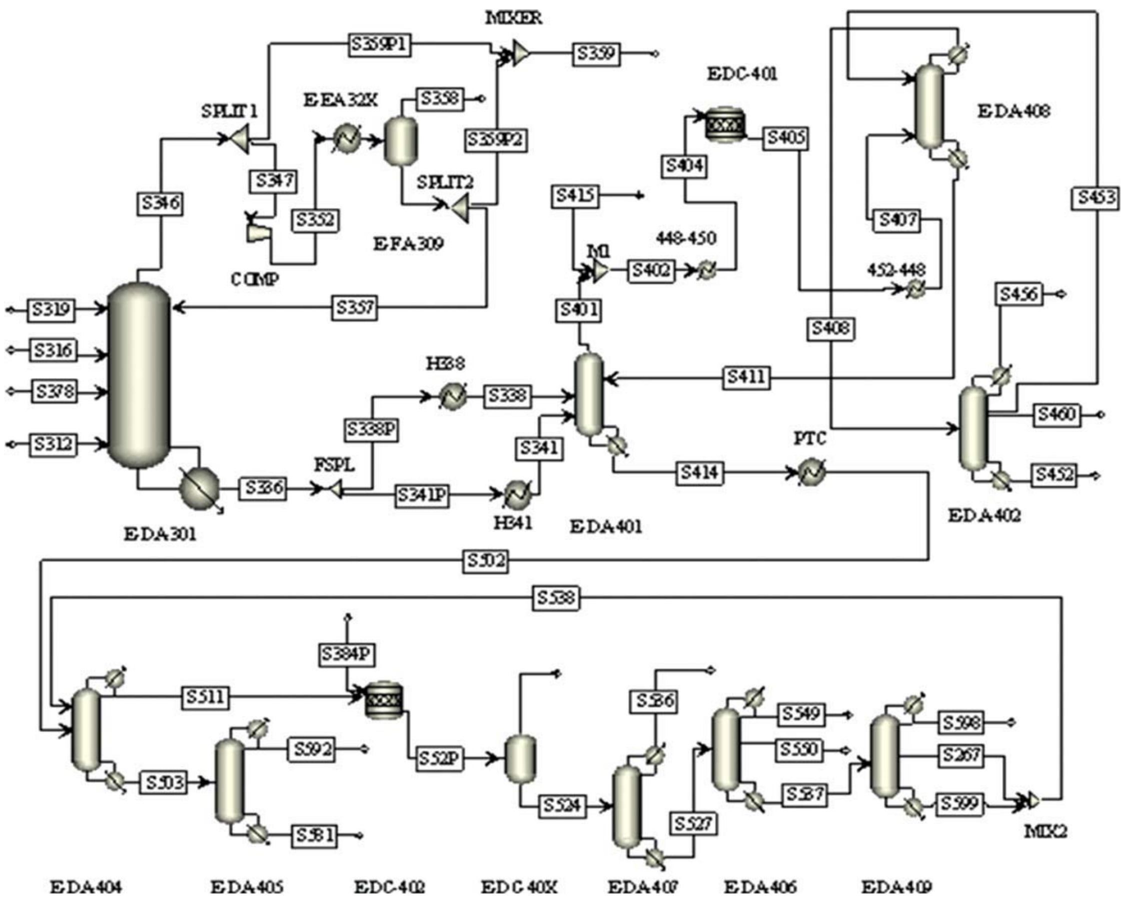

Figure 9. Flowsheet of ethylene separation system.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

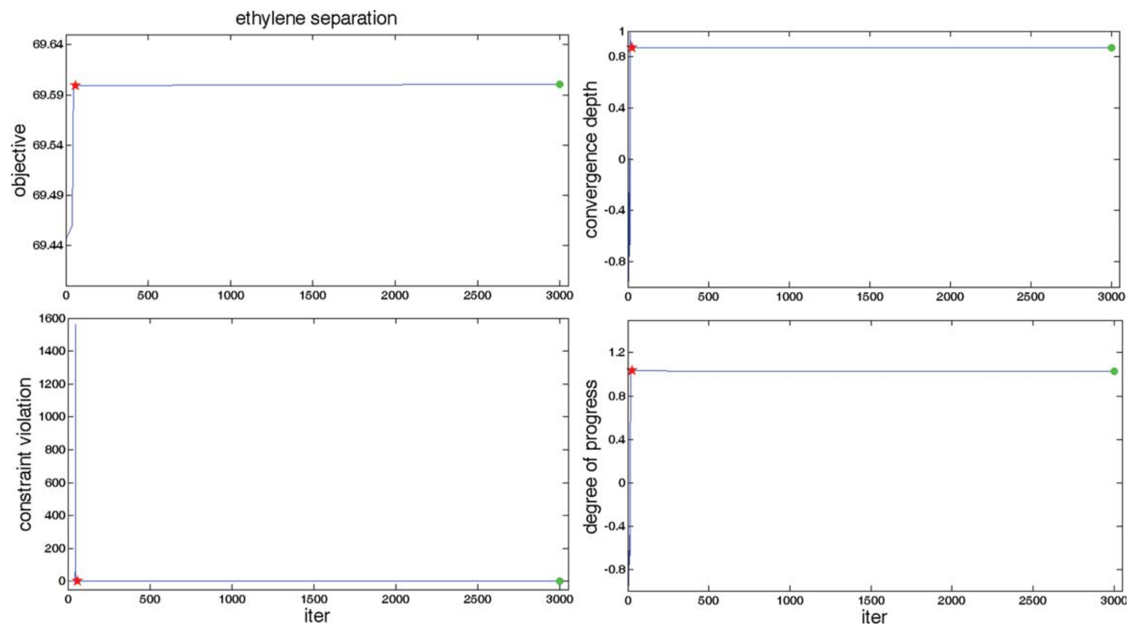

Figure 10. Comparison of IPOPT with traditional criteria and modified CDC for ethylene separation.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

The optimization problem is formulated as follows:

$$\begin{cases} \min -J \\ \text{s.t.} \quad \text{mesh equations (mass equation, energy equation, etc.)} \\ \qquad \text{connection equations} \\ \qquad \text{Purity constraints of the products} \end{cases} \tag{60}$$

For this optimization model, the number of variables is 23,288, the number of equality constraints is 23,280, the degrees of freedom are 8, and the number of non-zero elements of Jacobian of constraints is 409,426. This model information is based on the output file of IPOPT.

Table 7 indicates that for this system the solutions produced by IPOPT with traditional termination criteria and IPOPT with CDC are both feasible. This optimization problem is solved successfully by DMO, and the optimal objective is 6.850×10^5 RMB/h.

From the subgraphs objective and constraint in Figure 10, we see that the objective value increases slowly and constraint violation is nearly the same after a few iterations.

The convergence depth (θ_{convg}) subgraph shows that the convergence depth remains unchanged after ~30 iterations; however, it does not reach the predefined threshold θ_0. The degree of progress (θ_{prog}) subgraph shows that the degree of improvement fluctuates frequently after ~30 iterations; however, the extent of fluctuation is very small. $\theta_{\text{prog}} \geq 1$ is satisfied at nearly all the iterations after ~30 iterations; thus, the algorithm is terminated because of the unreasonably slow convergence. An underconverged solution is obtained.

The results in Table 8 show that both the runs with CDC and traditional termination criteria do not converge to the optima. The run with CDC detects the unreasonably slow convergence and terminates in the early stage of the optimization procedure, while the run with traditional termination criteria is stopped when the number of maximum iterations is exceeded.

Conclusions

Optimization procedures usually continue to iterate when feasibility is satisfied; however, the objective either seldom improves or the problems converge to the optima in an unreasonably slow way. In these cases, the CDC strategy can obtain a good compromise.

In this article, we incorporated interior point methods with the modified CDC. The implementation framework is modified based on the principle that, preferentially, the threshold for convergence depth should be accessed as far as possible. This is to avoid the optimization procedure being prematurely terminated by the CDC strategy. We have proved that with proper assumptions, the optimization procedure can be terminated by the modified CDC for interior point methods and that the convergence depth indicates the distance between the current iterate and the optimum.

Interior point methods using CDC strategy were tested with examples from the CUTE collection and examples

Table 8. Comparison of IPOPT with Traditional Termination Criteria and Modified CDC for Ethylene Separation System

θ_{convg}	θ_{prog}		
0.8704	1.0316	Convergence Depth Control	Traditional Termination Criteria
Objective (RMB/h)		6.8177592×10^5	6.8220972×10^5
Iteration		27	3000
CPU(s)		161.97	17,988.33
Feasibility		5.4539×10^{-5}	5.2399×10^{-5}
Terminated		X	X

formulated based on Aspen Plus, including the optimization problem of depropanizer and debutanizer distillation column system, the data reconciliation problem of large-scale air separation system, and the optimization problem of large-scale ethylene separation system through AMPL interface and Aspen Open Solvers interface. The modified CDC has advantages in dealing with the cases referred above. The CDC strategy obtains the approximate solution with acceptable optimality and significantly reduces the computation cost when compared with traditional termination criteria. Furthermore, it can detect the convergence failure much earlier as compared to traditional termination criteria, thus avoiding the consumption of unnecessary computing cost. Studies on the application of CDC to RTO and NMPC are recommended as future work.

Acknowledgments

This work was supported by the 973 Program of China (2009CB320603) and the 863 Program of China (2007A04Z192, 2008AA042902).

Literature Cited

1. Gill PE, Murray W, Wright MH. *Practical Optimization*. London: Academic Press, 1982.
2. Daniel JL. Real-time optimization in nonlinear chemical processes: Need for global optimizer. *AIChE J*. 2003;11:2980–2983.
3. Naysmith MR, Douglas PL. Review of real time optimization in the chemical process industries. *Dev Chem Eng Miner Process*. 1995;3: 67–87.
4. Yip WS, Marlin TE. The effect of model fidelity on real-time optimization performance. *Comp Chem Eng*. 2004;28:267–280.
5. Nagy ZK, Braatz RD. Robust nonlinear model predictive control of batch processes. *AIChE J*. 2003;49:1776–1786.
6. Findeisen R, Allgoewer F. *An introduction to nonlinear model predictive control*. In: 21st Benelux Meeting on Systems and Control, Veldhoven, 2002.
7. Wächter A. An Interior Point Algorithm for Large-Scale Nonlinear Optimization with Applications in Process Engineering. PhD Dissertation. Carnegie Mellon University, Pittsburgh, PA, 2002.
8. Wächter A, Biegler LT. Line search filter methods for nonlinear programming: motivation and global convergence. *SIAM J Optimization*. 2005;16:1–31.
9. Wächter A, Biegler LT. Line search filter methods for nonlinear programming: local convergence. *SIAM J Optimization*. 2005;16:32–48.
10. Wächter A, Biegler LT. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Math Program*. 2006;106:25–57.
11. Nocedal J, Sartenaer A, Zhu C. On the behavior of the gradient norm in the steepest descent method. *Comput Optimization Appl*. 2002;22:5–35.
12. Wang KX, Shao ZJ, Zhang ZJ, Chen ZQ, Fang XY, Zhou Z, Chen X, Qian JX. Convergence depth control for process system optimization. *Ind Eng Chem Res*. 2007;46:7729–7738.
13. Fourer R, Gay DM, Kernighan BW. A modeling language for mathematical programming. *Manage Sci*. 1990;36:519–554.
14. Aspen Technology, Inc. *Aspen Open Solvers 12.1 User Guide*. Aspen Technology, Inc., 2003.
15. Vanderbei RJ, Shanno DF. An interior point algorithm for nonconvex nonlinear programming. *Comput Optimization Appl*. 1999;13: 231–252.
16. Benson HY, Vanderbei RJ, Shanno DF. Interior Point Methods for Nonconvex Nonlinear Programming: Filter Methods and Merit Functions. *Comput Optimization Appl*. 2002;23:257–272.
17. Forsgren A, Gill PE, Wright MH. Interior methods for nonlinear optimization. *SIAM Rev*. 2002;44:525–597.
18. Fiacco AV, McCormick GP. The sequential unconstrained minimization technique for nonlinear programming, a primal-dual method. *Manage Sci*. 1964;10:360–366.
19. Nocedal J, Wächter A, Waltz RA. Adaptive barrier strategies for nonlinear interior methods. Tech. Rep., RC 23563, IBM T J. Watson Research Center, Yorktown Heights, NY, 2005.
20. Mehrotra S. On the implementation of a primal-dual interior point method. *SIAM J Optimization*. 1992;2:575–601.
21. Nocedal J, Wright S. *Numerical Optimization*. New York: Springer, 1999.
22. Health MT. *Scientific Computing: An Introductory Survey*. New York: McGraw-Hill Higher Education, 2001.
23. IPOPT web page, 2009. Available at: http://www.coin-or.org/Ipopt/. Accessed June 10, 2009.
24. Jiang AP, Shao ZJ, Chen X, Fang XY, Geng DZ, Zheng XQ, Qian JX. Simulation and optimization of distillation column sequence in large-scale ethylene production. *J Chem Ind Eng (China)*. 2006;57: 2128–2135.
25. Zhu LY. Study on simulation and optimization of complex separation process. PhD thesis. Zhejiang University, China, 2009.
26. Zhang ZJ, Shao ZJ, Chen X, Wang KX, Qian JX. Quasi-weighted least squares estimator for data reconciliation. *Comp Chem Eng*. 2010;34:154–162.
27. Fang XY, Shao ZJ, Wang ZQ, Chen WF, Wang KX, Zhang ZJ, Zhou Z, Chen X, Qian JX. Mnemonic enhancement optimization (MEO) for real-time optimization of industrial processes. *Ind Eng Chem Res*. 2009;48:499–509.

Manuscript received Oct. 14, 2009, and revision received Feb. 13, 2010.