J Glob Optim (2010) 46:217–232
DOI 10.1007/s10898-009-9419-x

On the convergence properties of modified augmented Lagrangian methods for mathematical programming with complementarity constraints

H. Z. Luo · X. L. Sun · Y. F. Xu · H. X. Wu

Received: 3 December 2007 / Accepted: 20 March 2009 / Published online: 3 April 2009

Abstract In this paper, we present new convergence results of augmented Lagrangian methods for mathematical programs with complementarity constraints (MPCC). Modified augmented Lagrangian methods based on four different algorithmic strategies are considered for the constrained nonconvex optimization reformulation of MPCC. We show that the convergence to a global optimal solution of the problem can be ensured without requiring the boundedness condition of the multipliers.

Keywords Mathematical program with complementarity constraints · Modified augmented Lagrangian methods · Nonconvex constrained optimization · Convergence to global solution

H. Z. Luo · X. L. Sun (✉) · Y. F. Xu
Department of Management Science, School of Management, Fudan University,
200433 Shanghai, People's Republic of China
e-mail: xls@fudan.edu.cn

Y. F. Xu
e-mail: yfxu@fudan.edu.cn

H. Z. Luo
Department of Applied Mathematics, Zhejiang University of Technology, 310032 Hangzhou,
Zhejiang, People's Republic of China
e-mail: hzluo@zjut.edu.cn

H. X. Wu
Department of Mathematics, Hangzhou Dianzi University, 310018 Hangzhou,
Zhejiang, People's Republic of China
e-mail: hxwu@hdu.edu.cn

1 Introduction

Consider the following mathematical program with complementarity constraints (MPCC):

$$\begin{aligned}\min\ & f(x)\\ \text{s.t.}\ & G_i(x)\geq 0,\ H_i(x)\geq 0,\ G_i(x)H_i(x)=0,\quad i=1,\ldots,m_1,\\ & g_j(x)\leq 0,\quad j=1,\ldots,m_2,\\ & h_l(x)=0,\quad l=1,\ldots,m_3,\\ & x\in X,\end{aligned}$$

where f, G_i, H_i, g_j and $h_l : \mathbb{R}^n \to \mathbb{R}$ are all continuously differentiable functions, and X is a nonempty closed set in $\mathbb{R}^n$.

Mathematical program with complementarity constraints is an important class of optimization problems with wide applications in economics and engineering [20]. Due to the presence of the complementarity constraints, Mangasarian–Fromovitz constraint qualification (CQ) or the linear independence CQ is never satisfied at any feasible point of MPCC [32]. Traditional nonlinear programming methods are therefore not applicable to MPCC. It has been shown in [24] that the linear complementarity problem is equivalent to the linear integer 0–1 feasibility problem. In the case of linear complementarity constraints this makes the problem computationally more tractable. Some complementarity problems can be viewed as special cases of variational inequalities (see, e.g., [10,11]). In recent years, MPCC has attracted much attention in nonlinear optimization. A comprehensive and in-depth theoretical study of MPCC can be found in [8,32]. Various methods for MPCC have been proposed (see e.g., [7,9,13,30,31] and the references therein).

One of the approaches for MPCC is to reformulate it as a constrained optimization problem by NCP function. The reformulated problem can then be dealt with by methods developed in nonlinear programming. Recently, Yang and Huang [35] used the Fischer–Burmeister function to reformulate MPCC as a nonsmooth constrained optimization problem and applied Rockafellar's augmented Lagrangian method to solve the resulting constrained optimization problem. A partial augmented Lagrangian method was also proposed in [14] for solving MPCC. The convergence of the Lagrangian methods was analyzed in [14,35] in terms of their first and second order optimality conditions under the boundedness assumption of the multiplier sequence. Convergence properties of augmented Lagrangian methods for constrained nonconvex optimization have been investigated by many researchers [1–3,5,16,17,21,23,25,26,34].

The purpose of this paper is to study the convergence properties of augmented Lagrangian methods based on two classes of Lagrangian functions for MPCC. In particular, we are concerned with the following question: When does the augmented Lagrangian method converge to a *global* solution of MPCC if the Lagrangian relaxation problems can be solved globally? To answer this question, we propose four modified augmented Lagrangian methods that adopt safeguarding strategy, conditional multiplier updating rule, penalty parameter updating criteria and normalization of the multipliers, respectively. The convergence to a global solution is proved for these modified augmented Lagrangian methods without appealing to the boundedness assumption on the multipliers. The convergence results obtained in this paper provide theoretical foundations for the use of augmented Lagrangian methods for mathematical programs with complementarity constraints.

The paper is organized as follows. In Sect. 2, we introduce two classes of augmented Lagrangian functions for MPCC. In Sect. 3, we present the convergence properties of the basic augmented Lagrangian method for MPCC under standard assumptions. The modified augmented Lagrangian method with safeguarding is investigated in Sect. 4. In Sect. 5, we establish the convergence results of the augmented Lagrangian method with the conditional multiplier updating. The use of penalty parameter updating criteria and the normalization of multipliers are discussed in Sect. 6. Finally, some concluding remarks are given in Sect. 7.

2 Two classes of augmented Lagrangian functions

In this section, we describe two classes of augmented Lagrangian functions for MPCC. The augmented Lagrangian methods investigated in the subsequent sections are based on these two classes of functions.

A function $\phi : \mathbb{R}^2 \to \mathbb{R}$ is called an NCP function if it satisfies

$$\phi(a, b) = 0 \quad \text{if and only if} \quad a \geq 0, \quad b \geq 0, \quad ab = 0.$$

Examples of ϕ include the min-function $\phi^{\min}(a, b) = \frac{1}{2}\left(a + b - \sqrt{(a-b)^2}\right)$, the Fischer–Burmeister function $\phi^{\text{FB}}(a, b) = a + b - \sqrt{a^2 + b^2}$, and the Mangasarian function $\phi^{\text{M}}(a, b) = \theta(|a - b|) - \theta(a) - \theta(b)$, where $\theta : \mathbb{R} \to \mathbb{R}$ is a strictly increasing function with $\theta(0) = 0$ [22]. Observe that ϕ^{FB} and $\phi^{\min}$ are smooth on $\mathbb{R}^2$ except at $(0, 0)$ and the line $a = b$. Note also that ϕ^{M} is smooth when taking $\theta(t) = \frac{1}{2}t|t|$ or t^3. Recently, a new class of NCP-functions $\phi^{\text{KYF}}(a, b) = \psi_0(ab) - \psi_1(-a, -b)$ are given in [15], where $\psi_0 : \mathbb{R} \to \mathbb{R}_+$ and $\psi_1 : \mathbb{R}^2 \to \mathbb{R}_+$ satisfy

$$\psi_0(t) = 0 \Leftrightarrow t \leq 0, \quad \psi_1(a, b) = 0 \Leftrightarrow a \leq 0, \ b \leq 0.$$

Notice that when setting $\psi_0(t) = [\max\{0, t\}]^p$, $\psi_1(a, b) = [(\max\{0, a\})^2 + (\max\{0, b\})^2]^{p/2}$ or $\psi_1(a, b) = \left[\max\left\{0, a + b + \sqrt{a^2 + b^2}\right\}\right]^p$, where $p > 1$, ϕ^{KYF} is continuously differentiable up to $(p-1)$th order.

Now, let ϕ be an NCP function. Then MPCC can be reformulated as the following constrained optimization problem:

$$\begin{array}{lll}(P) & \min & f(x) \\ & \text{s.t.} & \phi_i(x) = \phi(G_i(x), H_i(x)) = 0, \quad i = 1, \ldots, m_1, \\ & & g_j(x) \leq 0, \quad j = 1, \ldots, m_2, \\ & & h_l(x) = 0, \quad l = 1, \ldots, m_3, \\ & & x \in X.\end{array}$$

Let $\Phi(x) = (\phi_1(x), \ldots, \phi_{m_1}(x))^T$, $g(x) = (g_1(x), \ldots, g_{m_2}(x))^T$, $h(x) = (h_1(x), \ldots, h_{m_3}(x))^T$. The augmented Lagrangian of Rockafellar and Wets associated with (P) [29, Chapt. 11, Sect. K^*] is defined as

$$\begin{aligned} L_1(x, \mu, \lambda, \nu, c) = f(x) - \mu^T \Phi(x) + c\sigma_1(\Phi(x)) - \nu^T h(x) + c\sigma_3(h(x)) \\ + \min_{g(x)+z \leq 0} \left[-\lambda^T z + c\sigma_2(z)\right], \end{aligned} \tag{1}$$

where $c > 0$, $x \in X$ and $(\mu, \lambda, \nu) \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \times \mathbb{R}^{m_3}$, and the function $\sigma_i : \mathbb{R}^{m_i} \to \mathbb{R}$ satisfies the following conditions $(i = 1, 2, 3)$:

(A1) σ_i is a strictly convex function on $\mathbb{R}^{m_i}$;
(A2) $\sigma_i(0) = 0$ and 0 is the unique minimizer of σ_i over $\mathbb{R}^{m_i}$;
(A3) $\lim_{\|z\| \to \infty} \frac{\sigma_i(z)}{\|z\|} = +\infty$.

Obviously, $\sigma(z) = \|z\|^\gamma$ with $\gamma > 1$ satisfies conditions (A1)–(A3). The augmented Lagrangian defined in (1) includes a wide class of augmented Lagrangian functions in the literature. When σ_i $(i = 1, 2, 3)$ are separable and twice continuously differentiable, L_1 reduces to the essentially quadratic augmented Lagrangian function [3,33]. In particular, when $\sigma_i(z) = \frac{1}{2}\|z\|_2^2$ for $i = 1, 2, 3$, L_1 becomes Rockafellar's augmented Lagrangian in [28]:

$$L_1^R(x, \mu, \lambda, \nu, c) = f(x) - \mu^T \Phi(x) + \frac{c}{2}\|\Phi(x))\|^2 - \nu^T h(x) + \frac{c}{2}\|h(x)\|^2 + \frac{1}{2c}\sum_{j=1}^{m_2}\left\{[\max(0, \lambda_j + cg_j(x))]^2 - \lambda_j^2\right\}. \quad (2)$$

Note that when setting $\phi = \phi^{\mathrm{FB}}$, the above function reduces to the augmented Lagrangian in [35].

Another general class of augmented Lagrangians for (P) is given by Mangasarian [21] as follows:

$$L_2(x, \mu, \lambda, \nu, c) = f(x) + \frac{1}{c}\sum_{i=1}^{m_1}[\theta(c\phi_i(x) + \mu_i) - \theta(\mu_i)] + \frac{1}{c}\sum_{j=1}^{m_2}\left[\theta(cg_j(x) + \lambda_j)_+ - \theta(\lambda_j)\right] + \frac{1}{c}\sum_{l=1}^{m_3}[\theta(ch_l(x) + \nu_l) - \theta(\nu_l)], \quad (3)$$

where $c > 0$, $x \in X$, $(\mu, \lambda, \nu) \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \times \mathbb{R}^{m_3}$, $\theta(s)_+ = \theta(s)$ if $s \geq 0$ and $\theta(s)_+ = 0$ otherwise, and the function $\theta : \mathbb{R} \to \mathbb{R}$ satisfies the following conditions:

(B1) θ is continuously differentiable and strictly convex on $\mathbb{R}$;
(B2) $\theta(0) = 0$, θ' maps$\mathbb{R}$ onto$\mathbb{R}$ and $\theta'(0) = 0$;
(B3) $\frac{\theta(s)}{|s|} \to \infty$, $(|s| \to \infty)$.

Note that $L_2(x, \mu, \lambda, \nu, c)$ is not concave with respect to (μ, λ, ν) [21, Remark 2.13]. Observe also that L_2 reduces to L_1^R when setting $\theta(s) = \frac{1}{2}s^2$. We see that if θ is twice differentiable and $\theta''(0) = 0$, then $L_2(x, \mu, \lambda, \nu, c)$ is twice differentiable with respect to x when ϕ, f and all g_i are twice differentiable. Examples of θ that satisfy conditions (B1)–(B3) and $\theta''(0) = 0$ include $\theta(s) = \frac{1}{\rho}|s|^\rho$ $(\rho > 2)$, $\theta(s) = \frac{1}{2}(e^s + e^{-s}) - \frac{1}{2}s^2 - 1$ and $\theta(s) = \frac{1}{2}\left[(e^s + e^{-s})/2 - 1\right]^2$.

3 Basic primal–dual scheme

In this section, we present the basic primal–dual scheme based on the two classes of augmented Lagrangians L_j $(j = 1, 2)$ in the previous section and discuss its convergence to a global optimal solution to MPCC under standard conditions.

The augmented Lagrangian relaxation problem associated with $L_j (j = 1, 2)$ is

$$d_c^j(\mu, \lambda, \nu) = \min_{x \in X} L_j(x, \mu, \lambda, \nu, c), \quad j = 1, 2. \quad (4)$$

Let

$$P(y, \lambda, c) = \min_{y+z \leq 0}\left[-\lambda^T z + c\sigma_2(z)\right]. \quad (5)$$

Condition (A3) for σ_2 ensures that $P(y, \lambda, c) > -\infty$ for any $y, \lambda \in \mathbb{R}^{m_2}$. Moreover, since $-\lambda^T z + c\sigma_2(z)$ is a strictly convex function of z, the minimization problem in (5) has a unique global optimal solution $\pi(y, \lambda, c)$. We have the following result.

Proposition 1 *The dual function d_c^1 is a concave function on $\mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \times \mathbb{R}^{m_3}$. Moreover, for any $(\bar{\mu}, \bar{\lambda}, \bar{\nu}) \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \times \mathbb{R}^{m_3}$ with $d_c^1(\bar{\mu}, \bar{\lambda}, \bar{\nu}) > -\infty$, $(-\Phi(\bar{x}), -\pi(g(\bar{x}), \bar{\lambda}, c), -h(\bar{x}))$ is an ϵ-subgradient of d_c^1 at $(\bar{\mu}, \bar{\lambda}, \bar{\nu})$, where $\bar{x}$ is an ϵ-approximate optimal solution to the relaxation problem* (4) *associated with L_1, i.e.,*

$$L_1(\bar{x}, \bar{\mu}, \bar{\lambda}, \bar{\nu}, c) \le L_1(x, \bar{\mu}, \bar{\lambda}, \bar{\nu}, c) + \epsilon, \quad \forall x \in X.$$

Proof From (5), $P(g(x), \lambda, c)$ is the minimum of linear functions of λ. Thus $P(g(x), \lambda, c)$ is a concave function of λ. Since

$$L_1(x, \mu, \lambda, \nu, c) = f(x) - \mu^T \Phi(x) + c\sigma_1(\Phi(x)) + P(g(x), \lambda, c) - \nu^T h(x) + c\sigma_3(h(x)),$$

we infer that $d_c^1(\mu, \lambda, \nu)$ is also a concave function of (μ, λ, ν). Moreover, from (5), it is easy to see that $-\pi(g(\bar{x}), \bar{\lambda}, c)$ is a subgradient of $P(g(\bar{x}), \lambda, c)$ at $\bar{\lambda}$. Thus, for any $(\mu, \lambda, \nu) \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \times \mathbb{R}^{m_3}$, we have

$$\begin{aligned}
d_c^1(\mu, \lambda, \nu) &= \min_{x \in X} L_1(x, \mu, \lambda, \nu, c) \\
&\le L_1(\bar{x}, \mu, \lambda, \nu, c) \\
&= f(\bar{x}) - \mu^T \Phi(\bar{x}) + c\sigma_1(\Phi(\bar{x})) + P(g(\bar{x}), \lambda, c) - \nu^T h(\bar{x}) + c\sigma_3(h(\bar{x})) \\
&\le f(\bar{x}) - \bar{\mu}^T \Phi(\bar{x}) + c\sigma_1(\Phi(\bar{x})) + P(g(\bar{x}), \bar{\lambda}, c) - \bar{\nu}^T h(\bar{x}) + c\sigma_3(h(\bar{x})) \\
&\quad - \Phi(\bar{x})^T(\mu - \bar{\mu}) - \pi(g(\bar{x}), \bar{\lambda}, c)^T(\lambda - \bar{\lambda}) - h(\bar{x})^T(\nu - \bar{\nu}) \\
&\le d_c^1(\bar{\mu}, \bar{\lambda}, \bar{\nu}) - \Phi(\bar{x})^T(\mu - \bar{\mu}) - \pi(g(\bar{x}), \bar{\lambda}, c)^T(\lambda - \bar{\lambda}) - h(\bar{x})^T(\nu - \bar{\nu}) + \epsilon,
\end{aligned}$$

which implies that $\left(-\Phi(\bar{x}), -\pi(g(\bar{x}), \bar{\lambda}, c), -h(\bar{x})\right)$ is an ϵ-subgradient of d_c^1 at $(\bar{\mu}, \bar{\lambda}, \bar{\nu})$. □

We now describe the basic primal–dual method based on L_1 and L_2. Define $h^j(x, d, c) = (h_1^j(x, d, c), \ldots, h_{m_j}^j(x, d, c))^T$ $(j = 1, 2, 3)$ with

$$\begin{aligned}
h_i^1(x, \mu, c) &= \begin{cases} -\phi_i(x), & \text{for } L_1 \\ (1/c)\left[\theta'(c\phi_i(x) + \mu_i) - \theta'(\mu_i)\right], & \text{for } L_2 \end{cases} \\
h_i^2(x, \lambda, c) &= \begin{cases} -z_i, & \text{for } L_1 \\ (1/c)\left[\theta'(cg_i(x) + \lambda_i)_+ - \theta'(\lambda_i)\right], & \text{for } L_2 \end{cases} \\
h_i^3(x, \nu, c) &= \begin{cases} -h_i(x), & \text{for } L_1 \\ (1/c)\left[\theta'(ch_i(x) + \nu_i) - \theta'(\nu_i)\right], & \text{for } L_2 \end{cases}
\end{aligned} \tag{6}$$

where $z = (z_1, \ldots, z_{m_2})^T$ is the optimal solution of problem (5).

Algorithm 1 (Basic primal–dual method)

Step 0. (Initialization) Select two positive sequences $\{c_k\}_{k=0}^{\infty}$ and $\{\epsilon_k\}_{k=0}^{\infty}$ with $\epsilon_k \to 0$ as $k \to \infty$. Choose $\mu^0 \in \mathbb{R}^{m_1}$, $\lambda^0 \in \mathbb{R}^{m_2}$, $\nu^0 \in \mathbb{R}^{m_3}$. Set $k = 0$.

Step 1. (Relaxation problem) Compute an $x^k \in X$ such that

$$L_j(x^k, \mu^k, \lambda^k, \nu^k, c_k) \le \min_{x \in X} L_j(x, \mu^k, \lambda^k, \nu^k, c_k) + \epsilon_k \quad (j = 1, 2). \tag{7}$$

For L_1, compute the optimal solution $z^k = \pi(g(x^k), \lambda^k, c_k)$ to the following convex problem

$$\min\{-(\lambda^k)^T z + c_k \sigma_2(z) \mid g(x^k) + z \leq 0\}. \tag{8}$$

Step 2. (Multiplier updating) Compute

$$\begin{cases} \mu^{k+1} = \mu^k + c_k h^1(x^k, \mu^k, c_k), \\ \lambda^{k+1} = \lambda^k + c_k h^2(x^k, \lambda^k, c_k), \\ \nu^{k+1} = \nu^k + c_k h^3(x^k, \nu^k, c_k). \end{cases} \tag{9}$$

Step 3. Set $k := k + 1$, go to Step 1.

Remark 1 By Proposition 1 and Step 1, $(-\Phi(x^k), -z^k, -h(x^k))$ is an ϵ_k-subgradient of the dual function $d^1_{c^k}$ at $(\mu^k, \lambda^k, \nu^k)$. Thus, the multiplier update (9) for L_1 can be viewed as executing an ϵ_k-steepest ascent step for maximizing the dual function $d^1_{c_k}$ with step-size c_k. Also, note that

$$\frac{\partial L_2(x, \mu, \lambda, \nu, c)}{\partial \mu_i} = \frac{1}{c}[\theta'(c\phi_i(x) + \mu_i) - \theta'(\mu_i)],$$
$$\frac{\partial L_2(x, \mu, \lambda, \nu, c)}{\partial \lambda_i} = \frac{1}{c}[\theta'(cg_i(x) + \lambda_i)_+ - \theta'(\lambda_i)],$$
$$\frac{\partial L_2(x, \mu, \lambda, \nu, c)}{\partial \nu_i} = \frac{1}{c}[\theta'(ch_i(x) + \nu_i) - \theta'(\nu_i)].$$

Thus, the multiplier update (9) for L_2 is an ϵ_k-steepest ascent step for maximizing the dual function $d^2_{c_k}$ with stepsize c_k.

We need the following assumption:

Assumption 1 $\underline{f} = \inf_{x \in X} f(x) > -\infty$.

Assumption 1 and conditions (A1)–(A3) for $\sigma_i (i = 1, 2, 3)$ ensure that Step 1 is well defined. Obviously, Assumptions 1 is satisfied if X is a compact set. We point out that Assumption 1 is standard in the convergence analysis for augmented Lagrangian methods for constrained global optimization [18,19,33].

The following convergence result for Algorithm 1 can be proved by using similar arguments as in the proof of Proposition 2.1 in [3].

Theorem 1 *Assume that Assumption 1, $(A1)$–$(A3)$ for $\sigma_i (i = 1, 2, 3)$ and $(B1)$–$(B3)$ for θ are satisfied. Suppose that $\{(\mu^k, \lambda^k, \nu^k)\}$ is bounded. If $c_k \to \infty$ as $k \to \infty$, then each limit point of the sequence $\{x^k\}$ generated by Algorithm 1 is a global optimal solution to* MPCC.

We point out that the boundedness assumption for $\{(\mu^k, \lambda^k, \nu^k)\}$ is essential for ensuring the convergence result of Theorem 1, whereas the multiplier update of $\{(\mu^k, \lambda^k, \nu^k)\}$ in Step 3 does not affect the convergence. In the subsequent sections, we will propose four different algorithmic strategies to modify the basic primal–dual method so as to circumvent the boundedness condition of the multipliers in Theorem 1.

4 Modified augmented Lagrangian method using safeguarding

In this section, we use the safeguarding technique to modify the basic primal–dual algorithm [1,2,4,18]. This simple technique ensures the boundedness of the multiplier sequence by projecting the multipliers in Step 2 of Algorithm 1 onto suitable bounded intervals.

Algorithm 2 (Modified primal–dual method using safeguarding)

Step 0. (Initialization) Choose a positive sequence $\{\epsilon_k\}_{k=1}^{\infty}$ with $\epsilon_k \to 0$ as $k \to \infty$. Choose $\gamma > 1$, $\delta \in (0, 1)$, $c_1 > 0$, $-\infty < \alpha_i^j < \beta_i^j < \infty$ for $i = 1, \ldots, m_j$, $j \in \{1, 2, 3\}$. Let $T_j = \{z \in \mathbb{R}^{m_j} \mid \alpha_i^j \le z_i \le \beta_i^j,\ i = 1, \ldots, m_j\}$, $j = 1, 2, 3$. Choose $\bar{\mu}^1 \in T_1$, $\bar{\lambda}^1 \in T_2$ and $\bar{\nu}^1 \in T_3$. Set $\rho^0 = 1$ and $k = 1$.

Step 1. (Relaxation problem) Compute an $x^k \in X$ such that

$$L_j(x^k, \bar{\mu}^k, \bar{\lambda}^k, \bar{\nu}^k, c_k) \le \min_{x\in X} L_j(x, \bar{\mu}^k, \bar{\lambda}^k, \bar{\nu}^k, c_k) + \epsilon_k, \quad j = 1, 2. \tag{10}$$

For L_1, compute also a z^k such that

$$z^k = \arg\min\{-(\bar{\lambda}^k)^T z + c_k\sigma_2(z) \mid g(x^k) + z \le 0\}. \tag{11}$$

Step 2. (Multiplier updating) Compute

$$\begin{cases} \mu^{k+1} = \bar{\mu}^k + c_k h^1(x^k, \bar{\mu}^k, c_k), \\ \lambda^{k+1} = \bar{\lambda}^k + c_k h^2(x^k, \bar{\lambda}^k, c_k), \\ \nu^{k+1} = \bar{\nu}^k + c_k h^3(x^k, \bar{\nu}^k, c_k). \end{cases} \tag{12}$$

Step 3. (Safeguarding projection) Compute

$$\begin{cases} \bar{\mu}^{k+1} = \mathcal{P}_{T_1}(\mu^{k+1}), \\ \bar{\lambda}^{k+1} = \mathcal{P}_{T_2}(\lambda^{k+1}), \\ \bar{\nu}^{k+1} = \mathcal{P}_{T_3}(\nu^{k+1}), \end{cases} \tag{13}$$

where $\mathcal{P}_{T_j}(z)$ denotes the Euclidean projection of z on T_j.

Step 4. (Parameter updating) Let $\rho^k = \|(h^1(x^k, \bar{\mu}^k, c_k), h^2(x^k, \bar{\lambda}^k, c_k), h^3(x^k, \bar{\nu}^k, c_k))\|$. If

$$\rho^k \le \delta\rho^{k-1}, \tag{14}$$

set $c_{k+1} = c_k$, otherwise, set $c_{k+1} = \gamma c_k$. Set $k := k + 1$ and go to Step 1.

We first discuss the convergence of Algorithm 2 using L_1.

Theorem 2 *Assume that Assumption 1 and (A1)–(A3) for σ_i $(i = 1, 2, 3)$ are satisfied. Then, each limit point of the sequence $\{x^k\}$ generated by Algorithm 2 using L_1 is a global optimal solution to* MPCC.

Proof Let x^* be a global solution to MPCC. By (10) in Algorithm 2, we have

$$L_1(x^k, \bar{\mu}^k, \bar{\lambda}^k, \bar{\nu}^k, c_k) \le L_1(x^*, \bar{\mu}^k, \bar{\lambda}^k, \bar{\nu}^k, c_k) + \epsilon_k \tag{15}$$

Since $P(y, \lambda, c)$ defined by (5) is nondecreasing of y for fixed $\lambda \in \mathbb{R}^{m_2}$ and $c > 0$, $g(x^*) \le 0$ together with $\sigma_2(0) = 0$ implies

$$P(g(x^*), \bar{\lambda}^k, c_k) \le P(0, \bar{\lambda}^k, c_k) = \min_{z\le 0}\left[-(\bar{\lambda}^k)^T z + c_k\sigma_2(z)\right] \le -(\bar{\lambda}^k)^T \cdot 0 + c_k\sigma_2(0) = 0$$

for all k. Note that the feasibility of x^* implies $h(x^*) = 0$ and $\Phi(x^*) = 0$. Also, $\sigma_j(0) = 0$ for $j = 1, 3$. Therefore, by the definition of L_1, we have from (15) that

$$f(x^k) - (\bar{\mu}^k)^T \Phi(x^k) + c_k\sigma_1(\Phi(x^k)) - (\bar{\nu}^k)^T h(x^k) + c_k\sigma_3(h(x^k)) + \min_{g(x^k)+z\le 0}\left[-(\bar{\lambda}^k)^T z + c_k\sigma_2(z)\right] \le f(x^*) + \epsilon_k. \tag{16}$$

Since z^k is the optimal solution of problem (11), it follows from (11) and (16) that

$$g(x^k) + z^k \le 0, \quad \forall k, \tag{17}$$

$$f(x^k) - (\bar{\mu}^k)^T \Phi(x^k) + c_k \sigma_1(\Phi(x^k)) - (\bar{\nu}^k)^T h(x^k) + c_k \sigma_3(h(x^k)) \\ -(\bar{\lambda}^k)^T z^k + c_k \sigma_2(z^k) \le f(x^*) + \epsilon_k. \tag{18}$$

Case (i): $c_k \to \infty$ as $k \to \infty$. Define $\sigma(u, v, z) = \sigma_1(u) + \sigma_2(z) + \sigma_3(v)$. Let $\tau^k = (u^k, v^k, z^k)$, where $u^k = \Phi(x^k)$, $v^k = h(x^k)$. By Assumption 1, $\underline{f} = \inf_{x \in X} f(x) > -\infty$. It follows from (18) that

$$-\left(\frac{\bar{\mu}^k}{c_k}\right)^T u^k - \left(\frac{\bar{\nu}^k}{c_k}\right)^T v^k - \left(\frac{\bar{\lambda}^k}{c_k}\right)^T z^k + \sigma(\tau^k) \le \frac{1}{c_k}[f(x^*) - \underline{f} + \epsilon_k]. \tag{19}$$

We claim that $\{\tau^k\}$ is bounded. Otherwise, suppose, without loss of generality, that $\|\tau^k\| \to \infty$ $(k \to \infty)$. Since $\{(\bar{\mu}^k, \bar{\lambda}^k, \bar{\nu}^k)\} \subset T_1 \times T_2 \times T_3$ is bounded by Step 3 of the algorithm, we deduce from (19) that

$$\limsup_{k \to \infty} \frac{\sigma(\tau^k)}{\|\tau^k\|} \le 0. \tag{20}$$

Now, by conditions (A1)–(A3) for σ_j $(j = 1, 2, 3)$, σ is a strictly convex function with $\sigma(\tau) \ge 0$ on $\mathbb{R}^{m_1+m_2+m_3}$ and $\sigma(0) = 0$. Thus, using Corollary 3.27 in [29], we obtain $\liminf_{k \to \infty} \sigma(\tau^k)/\|\tau^k\| > 0$, a contradiction to (20).

Now assume, without loss of generality, that $\tau^k = (u^k, v^k, z^k) \to \bar{\tau} = (\bar{u}, \bar{v}, \bar{z})$ as $k \to \infty$. It follows from (19) that

$$\sigma(\bar{\tau}) \le \liminf_{k \to \infty} \sigma(\tau^k) \\ \le \liminf_{k \to \infty} \left\{ \frac{1}{c_k}[f(x^*) - \underline{f} + \epsilon_k] + \left(\frac{\bar{\mu}^k}{c_k}\right)^T u^k + \left(\frac{\bar{\nu}^k}{c_k}\right)^T v^k + \left(\frac{\bar{\lambda}^k}{c_k}\right)^T z^k \right\} = 0.$$

Since $\sigma(\tau) \ge 0$ and $\sigma(0) = 0$, we have $\bar{\tau} = (\bar{u}, \bar{v}, \bar{z}) = 0$. Thus, from (17) and the definitions of u^k and v^k, we have

$$\lim_{k \to \infty} \Phi(x^k) = 0, \quad \lim_{k \to \infty} h(x^k) = 0, \quad \limsup_{k \to \infty} g(x^k) \le 0. \tag{21}$$

On the other hand, it follows from (18) that

$$f(x^k) - (\bar{\mu}^k)^T \Phi(x^k) - (\bar{\nu}^k)^T h(x^k) - (\bar{\lambda}^k)^T z^k \le f(x^*) + \epsilon_k, \quad \text{for all } k. \tag{22}$$

Since $\{(\bar{\mu}^k, \bar{\lambda}^k, \bar{\nu}^k)\}$ is bounded, $z^k \to 0$ and $\epsilon_k \to 0$, taking limit in (22) gives rise to

$$\limsup_{k \to \infty} f(x^k) \le f(x^*), \tag{23}$$

which together with (21) and the closedness of X implies that any limit point $\bar{x}$ of $\{x^k\}$ is feasible for (P) and $f(\bar{x}) = f(x^*)$, i.e., $\bar{x}$ is a global optimal solution to (P). Since (P) and MPCC are equivalent, $\bar{x}$ is also a global optimal solution of MPCC.

Case (ii): $\{c_k\}$ is bounded. In this case, condition (14) in Step 4 must be satisfied at each iteration for sufficiently large k. Since $\delta \in (0, 1)$, we infer that $\rho^k \to 0$ $(k \to \infty)$. By (6), (17) and the definition of ρ^k in Step 4, we get (21). Moreover, since $\{(\bar{\mu}^k, \bar{\lambda}^k, \bar{\nu}^k)\}$ is bounded, $\sigma_j(0) = 0$ $(j = 1, 2, 3)$ and $\epsilon_k \to 0$, taking limits in (18), we also obtain (23). □

Lemma 1 ([21])

(i) $\theta'(s+t) - \theta'(t) = 0 \Rightarrow s = 0$;
(ii) $\theta'(s+t)_+ - \theta'(t) = 0 \Rightarrow s \le 0,\ t \ge 0,\ st = 0$.

Next, we give the convergence result for Algorithm 2 when using L_2.

Theorem 3 *Assume that Assumption 1 and (B1)–(B3) for θ are satisfied. Then, each limit point of the sequence $\{x^k\}$ generated by Algorithm 2 using L_2 is a global optimal solution to* MPCC.

Proof Let x^* be a global solution to (*MPCC*). By (10), we have

$$L_2(x^k, \bar{\mu}^k, \bar{\lambda}^k, \bar{\nu}^k, c_k) \le L_2(x^*, \bar{\mu}^k, \bar{\lambda}^k, \bar{\nu}^k, c_k) + \epsilon_k. \tag{24}$$

By (B1)–(B2), $\theta(s) \ge 0$ for $s \in \mathbb{R}$ and θ' is strictly increasing on $\mathbb{R}$. Also $\theta'(0) = 0$ and $\theta'(s) \ge 0$ for $s \ge 0$ imply that $\theta'(s)_+ \ge 0$ for $s \in \mathbb{R}$. Thus, $\theta(\cdot)_+$ is monotonically increasing on R. Since $g_j(x^*) \le 0$, θ is nonnegative and $\theta(0) = 0$, it holds

$$\theta(c_k g_j(x^*) + \bar{\lambda}_j^k)_+ \le \theta(\bar{\lambda}_j^k)_+ \le \theta(\bar{\lambda}_j^k)$$

for all j and k. Also, the feasibility of x^* implies $\phi_i(x^*) = 0$ for all i and $h(x^*) = 0$. So, by the definition of L_2 and (24) we have

$$\begin{aligned} f(x^k) &+ \frac{1}{c_k}\sum_{i=1}^{m_1}[\theta(c_k\phi_i(x^k) + \bar{\mu}_i^k) - \theta(\bar{\mu}_i^k)] + \frac{1}{c_k}\sum_{j=1}^{m_2}[\theta(c_k g_j(x^k) + \bar{\lambda}_j^k)_+ - \theta(\bar{\lambda}_j^k)] \\ &+ \frac{1}{c_k}\sum_{l=1}^{m_3}[\theta(c_k h_l(x^k) + \bar{\nu}_l^k) - \theta(\bar{\nu}_l^k)] \le f(x^*) + \epsilon_k. \end{aligned} \tag{25}$$

Let $\bar{x}$ be a limit point of $\{x^k\}$, and $\mathcal{K} \subset \{1, 2, \ldots\}$ be such that $x^k \to \bar{x} (k \in \mathcal{K})$. The closedness of X implies $\bar{x} \in X$.

Case (i): $c_k \to \infty$ as $k \to \infty$. We first prove that $g_j(\bar{x}) \le 0$, $j = 1, \ldots, m_2$. Suppose that there exists j_0 such that $g_{j_0}(\bar{x}) > 0$. Let $\epsilon = g_{j_0}(\bar{x})/2$. Then there exists $k_0 > 0$ such that $g_{j_0}(x^k) \ge \epsilon$ for $k \ge k_0$, $k \in \mathcal{K}$. Thus, $c_k g_{j_0}(x^k) + \bar{\lambda}_{j_0}^k \ge c_k\epsilon + \bar{\lambda}_{j_0}^k$ for $k \ge k_0$, $k \in \mathcal{K}$. By the monotonicity of $\theta(\cdot)_+$, we have

$$\theta(c_k g_{j_0}(x^k) + \bar{\lambda}_{j_0}^k)_+ \ge \theta(c_k\epsilon + \bar{\lambda}_{j_0}^k)_+ \quad \text{for} \quad k \ge k_0,\ k \in \mathcal{K}. \tag{26}$$

By Step 3 of the algorithm, $\{(\bar{\mu}^k, \bar{\lambda}^k, \bar{\nu}^k)\} \subset T_1 \times T_2 \times T_3$ is bounded. Using (26) and Assumption 1, we obtain from (25) that

$$\begin{aligned} f(x^*) + \epsilon_k &\ge \underline{f} - \frac{1}{c_k}\sum_{i=1}^{m_1}\theta(\bar{\mu}_i^k) + \frac{1}{c_k}\theta(c_k g_{j_0}(x^k) + \bar{\lambda}_{j_0}^k)_+ - \frac{1}{c_k}\sum_{j=1}^{m_2}\theta(\bar{\lambda}_j^k) - \frac{1}{c_k}\sum_{l=1}^{m_3}\theta(\bar{\nu}_l^k) \\ &\ge \underline{f} - \frac{1}{c_k}\sum_{i=1}^{m_1}\theta(\bar{\mu}_i^k) + \frac{1}{c_k}\theta(c_k\epsilon + \bar{\lambda}_{j_0}^k)_+ - \frac{1}{c_k}\sum_{j=1}^{m_2}\theta(\bar{\lambda}_j^k) - \frac{1}{c_k}\sum_{l=1}^{m_3}\theta(\bar{\nu}_l^k) \\ &\to \infty\ (k \to \infty,\ k \in \mathcal{K}), \end{aligned}$$

where the third term in the last inequality tends to ∞ due to $c_k \to \infty$, condition (B3) and the boundedness of $\{\bar{\lambda}^k\}$, while the second, forth and fifth terms all tend to zero because of the continuity of θ and the boundedness of $\{(\bar{\mu}^k, \bar{\lambda}^k, \bar{\nu}^k)\}$. The above contradiction implies that

$g_j(\bar{x}) \le 0$, $j = 1, \ldots, m_2$. Similarly, we can prove $\phi_i(\bar{x}) = 0$, $i = 1, \ldots, m_1$ and $h(\bar{x}) = 0$, and therefore $\bar{x}$ is a feasible solution to MPCC.

On the other hand, it follows from (25) that

$$f(x^k) - \frac{1}{c_k}\sum_{i=1}^{m_1}\theta\left(\bar{\mu}_i^k\right) - \frac{1}{c_k}\sum_{j=1}^{m_2}\theta\left(\bar{\lambda}_j^k\right) - \frac{1}{c_k}\sum_{l=1}^{m_3}\theta\left(\bar{\nu}_l^k\right) \le f(x^*) + \epsilon_k. \tag{27}$$

Since $\{(\bar{\mu}^k, \bar{\lambda}^k, \bar{\nu}^k)\}$ is bounded, $\epsilon_k \to 0$ and θ is continuous, taking limit in (27) with respect to $k \in \mathcal{K}$ gives rise to

$$f(\bar{x}) \le f(x^*).$$

Hence, $f(\bar{x}) = f(x^*)$ and by the feasibility of $\bar{x}$, $\bar{x}$ is a global optimal solution to MPCC.

Case (ii): $\{c_k\}$ is bounded as $k \to \infty$. From Step 3, we have $\rho_k \to 0$ $(k \to \infty)$ and $c_k = c_{k_0}$ for k large enough. By the definition of ρ^k, we have $h^j(x^k, \bar{\mu}^k, c_k) \to 0$ as $k \to \infty$, $j = 1, 2, 3$, which, by the definition of h^j [cf.(6)], implies

$$\begin{cases} \lim_{k\to\infty} \frac{1}{c_k}\left[\theta'(c_k\phi_i(x^k) + \bar{\mu}_i^k) - \theta'\left(\bar{\mu}_i^k\right)\right] = 0, \quad i = 1, \ldots, m_1, \\ \lim_{k\to\infty} \frac{1}{c_k}\left[\theta'(c_k g_j(x^k) + \bar{\lambda}_j^k)_+ - \theta'\left(\bar{\lambda}_j^k\right)\right] = 0, \quad j = 1, \ldots, m_2, \\ \lim_{k\to\infty} \frac{1}{c_k}\left[\theta'(c_k h_l(x^k) + \bar{\nu}_l^k) - \theta'\left(\bar{\nu}_l^k\right)\right] = 0, \quad l = 1, \ldots, m_3. \end{cases}$$

Note that $\{(\bar{\mu}^k, \bar{\lambda}^k, \bar{\nu}^k)\}$ is bounded. Without loss of generality, we may assume that $(\bar{\mu}^k, \bar{\lambda}^k, \bar{\nu}^k) \to (\bar{\mu}, \bar{\lambda}, \bar{\nu})(k \to \infty, k \in \mathcal{K})$. Thus

$$\begin{cases} \theta'(c_{k_0}\phi_i(\bar{x}) + \bar{\mu}_i) - \theta'(\bar{\mu}_i) = 0, \quad i = 1, \ldots, m_1, \\ \theta'(c_{k_0}g_j(\bar{x}) + \bar{\lambda}_j)_+ - \theta'(\bar{\lambda}_j) = 0, \quad j = 1, \ldots, m_2, \\ \theta'(c_{k_0}h_l(\bar{x}) + \bar{\nu}_l) - \theta'(\bar{\nu}_l) = 0, \quad l = 1, \ldots, m_3. \end{cases} \tag{28}$$

By Lemma 1, we have

$$\begin{cases} \phi_i(\bar{x}) = 0, \ i = 1, \ldots, m_1, \quad h(\bar{x}) = 0, \\ g_j(\bar{x}) \le 0, \ \ \bar{\lambda}_j \ge 0, \quad \bar{\lambda}_j g_j(\bar{x}) = 0, \quad j = 1, \ldots, m_2. \end{cases} \tag{29}$$

Thus, $\bar{x}$ is feasible to MPCC.

Let $W(s,t) = \theta(s+t)_+ - \theta(t)$. By (B1)–(B3), we see that $W(s,t)$ is convex with respect to s for fixed $t \in \mathbb{R}$. Also, $W(0,t) \le 0$ for $t \in \mathbb{R}$ and $W_s'(s,t) = \theta'(s+t)_+$. Thus,

$$0 \ge W(0,t) \ge W(s,t) + (0-s)W_s'(s,t), \quad \forall s \in \mathbb{R},$$

this is

$$s\theta'(s+t)_+ \ge W(s,t), \quad \forall s \in \mathbb{R}, t \in \mathbb{R}. \tag{30}$$

Note also that

$$W(s,t) \ge \theta(t)_+ - \theta(t) + s\theta'(t)_+, \quad \forall s \in \mathbb{R}, t \in \mathbb{R}. \tag{31}$$

It then follows from (30)–(31) that for k large enough,

$$\frac{1}{c_k}\left[\theta\left(\bar{\lambda}_j^k\right)_+ - \theta\left(\bar{\lambda}_j^k\right)\right] + g_j(x^k)\theta'\left(\bar{\lambda}_j^k\right)_+ \le \frac{1}{c_k}W\left(c_k g_j(x^k), \bar{\lambda}_j^k\right)$$
$$\le g_j(x^k)\theta'\left(c_k g_j(x^k) + \bar{\lambda}_j^k\right)_+$$

for all j. Taking limits in the above inequality and using (28)–(29) and $\theta'(0) = 0$, we obtain

$$\lim_{k\to\infty, k\in\mathcal{K}} \frac{1}{c_k} W\left(c_k g_j(x^k), \bar{\lambda}_j^k\right) = g_j(\bar{x})\theta'(\bar{\lambda}_j) = 0, \quad j = 1, \ldots, m_2.$$

Using similar arguments, we can prove that

$$\lim_{k\to\infty, k\in\mathcal{K}} \frac{1}{c_k}\left[\theta\left(c_k\phi_i(x^k) + \bar{\mu}_i^k\right) - \theta\left(\bar{\mu}_i^k\right)\right] = 0, \; i = 1, \ldots, m_1,$$
$$\lim_{k\to\infty, k\in\mathcal{K}} \frac{1}{c_k}\left[\theta\left(c_k h_l(x^k) + \bar{\nu}_l^k\right) - \theta\left(\bar{\nu}_l^k\right)\right] = 0, \; l = 1, \ldots, m_3.$$

Hence, we obtain from (25) that $f(\bar{x}) \leq f(x^*)$, which combined with the feasibility of $\bar{x}$ implies that $\bar{x}$ is a global optimal solution to MPCC. □

5 Modified augmented Lagrangian method with conditional multiplier updating

In this section, we discuss an alternative strategy to modify the basic primal–dual algorithm using L_1 for solving MPCC. The underlying idea is to modify Step 3 of Algorithm 1 so that the multipliers remain unchanged unless certain reduction of the norm of the subgradient of the dual function d_c^1 is achieved. Similar idea for multiplier updating was used in [5] for smooth equality constrained optimization problem.

Algorithm 3 (Modified primal–dual method with conditional multiplier updating)

Step 0. Choose initial multiplier vectors λ^0, μ^0, ν^0 and the constants $c_0 > 1$, $u_0 > 0$, $v_0 > 0$, $\tau > 1$, $\gamma_1 \in (0, 1)$, $0 \leq \epsilon \ll 1$, $\alpha_\eta > 0.5$, $\beta_\eta > 0$, $\alpha_\omega > 0$, $\beta_\omega > 0$. Set $\alpha_0 = \min(\frac{1}{c_0}, \gamma_1)$, $\epsilon_0 = v_0\,(\alpha_0)^{\alpha_\omega}$, and $\eta_0 = u_0\,(\alpha_0)^{\alpha_\eta}$, and $k = 0$.

Step 1. Find an x^k satisfying

$$L_1(x^k, \mu^k, \lambda^k, \nu^k, c_k) \leq \min_{x\in X} L_1(x, \mu^k, \lambda^k, \nu^k, c_k) + \epsilon_k, \tag{32}$$

Step 2. Compute the optimal solution z^k to the following convex problem

$$\min\{-(\lambda^k)^T z + c_k\sigma_2(z) \mid g(x^k) + z \leq 0\}.$$

Let $\rho^k = \|(\Phi(x^k)^T, (z^k)^T, h(x^k)^T)\|$. If

$$\rho^k \leq \eta_k, \tag{33}$$

go to Step 3. Otherwise, go to Step 4.

Step 3. If $\rho^k \leq \epsilon$, stop. Otherwise, set

$$\begin{cases} \mu^{k+1} = \mu^k - c_k\Phi(x^k), \\ \lambda^{k+1} = \lambda^k - c_k z^k, \\ \nu^{k+1} = \nu^k - c_k h(x^k), \\ c_{k+1} = c_k, \\ \alpha_{k+1} = \min\left(\frac{1}{c_{k+1}}, \gamma_1\right), \\ \epsilon_{k+1} = \epsilon_k(\alpha_{k+1})^{\beta_\omega}, \\ \eta_{k+1} = \eta_k\,(\alpha_{k+1})^{\beta_\eta}, \end{cases} \tag{34}$$

Set $k := k + 1$, go to Step 1.

Step 4. Set

$$\begin{cases} \mu^{k+1} = \mu^k, \quad \lambda^{k+1} = \lambda^k, \quad \nu^{k+1} = \nu^k, \\ c_{k+1} = \tau c_k, \\ \alpha_{k+1} = \min\left(\frac{1}{c_{k+1}}, \gamma_1\right), \\ \epsilon_{k+1} = v_0(\alpha_{k+1})^{\alpha_\omega}, \\ \eta_{k+1} = u_0\,(\alpha_{k+1})^{\alpha_\eta}. \end{cases} \tag{35}$$

Set $k := k+1$, go to Step 1.

Let $\epsilon = 0$ in Algorithm 3 and $\omega^k = (\mu^k, \lambda^k, \nu^k)$. The following lemmas can be proved by using the similar arguments as in the proofs of Lemma 4.1 in [5] and Lemma 2 in [18].

Lemma 2 *If $c_k \to \infty$ when Algorithm 3 is executed, then* $\lim_{k\to\infty} \frac{\omega^k}{\sqrt{c_k}} = 0$.

Lemma 3 *If $\{c_k\}$ is bounded when Algorithm 3 is executed, then $\{\omega^k\}$ is convergent.*

We now present the convergence results for Algorithm 3. We assume that the condition (A3) for σ_j $(j = 1, 2, 3)$ is replaced by the following condition:

(A3′) There exist $\xi_j > 0$ $(j = 1, 2, 3)$ such that

$$\sigma_j(z) \geq \xi_j \|z\|^2, \ \ j = 1, 2, 3, \quad \forall z \in \mathbb{R}^{m_j}. \tag{36}$$

Note that condition (A3′) implies condition (A3).

Theorem 4 *Assume that Assumption 1,* (A1)–(A2) *and* (A3′) *for σ_j $(j = 1, 2, 3)$ are satisfied. Then, each limit point of the sequence $\{x_k\}$ generated by Algorithm 3 is a global optimal solution to* MPCC.

Proof Let x^* be a global solution to MPCC. Similar to the proof of Theorem 2, we have

$$g(x^k) + z^k \leq 0, \quad \forall k, \tag{37}$$

$$f(x^k) - (\mu^k)^T \Phi(x^k) + c_k\sigma_1(\Phi(x^k)) - (\nu^k)^T h(x^k) + c_k\sigma_3(h(x^k)) \\ -(\lambda^k)^T z^k + c_k\sigma_2(z^k) \leq f(x^*) + \epsilon_k. \tag{38}$$

Suppose that $x^k \to \bar{x}$ as $k \to \infty$ and $k \in \mathcal{K} \subseteq \{1, 2, \ldots\}$. The closedness of X implies $\bar{x} \in X$. We consider the following two cases.

Case (i): $c_k \to \infty$ when the algorithm is executed. By Lemma 2, it holds

$$\lim_{k\to\infty} \frac{\omega^k}{\sqrt{c_k}} = 0, \tag{39}$$

which implies $\lim_{k\to\infty} \frac{\omega^k}{c_k} = 0$, where $\omega^k = (\mu^k, \lambda^k, \nu^k)$. Using the similar arguments as in the proof of Theorem 2, we can infer from (37) and (38) that $\bar{x}$ is a feasible solution to MPCC.

On the other hand, using Cauchy–Schwartz inequality, it follows from (36) and (38) that

$$\begin{aligned} f(x^*) + \epsilon_k &\geq f(x^k) - (\mu^k)^T \Phi(x^k) + c_k\sigma_1(\Phi(x^k)) \\ &\quad -(\nu^k)^T h(x^k) + c_k\sigma_3(h(x^k)) - (\lambda^k)^T z^k + c_k\sigma_2(z^k) \\ &\geq f(x^k) - \|\mu^k\|\|\Phi(x^k)\| + c_k\xi_1\|\Phi(x^k)\|^2 \\ &\quad -\|\nu^k\|\|h(x^k)\| + c_k\xi_3\|h(x^k)\|^2 - \|\lambda^k\|\|z^k\| + \xi_2 c_k\|z^k\|^2 \\ &\geq f(x^k) - \frac{\|\mu^k\|^2}{4\xi_1 c_k} - \frac{\|\lambda^k\|^2}{4\xi_2 c_k} - \frac{\|\nu^k\|^2}{4\xi_3 c_k}, \end{aligned}$$

where the last inequality follows from the fact that the minimum of the convex function $-bt + at^2$ $(a > 0)$ over $\mathbb{R}$ is $-\frac{b^2}{4a}$. Since $\epsilon_k \to 0$ as $k \to \infty$, taking limits in the above inequality and using (39) give rise to $f(\bar{x}) \leq f(x^*)$. Thus, $f(\bar{x}) = f(x^*)$ and by the feasibility of $\bar{x}$, $\bar{x}$ is a global optimal solution to MPCC.

Case (ii): $\{c_k\}$ is bounded when the algorithm is executed. Then Step 2 must be executed at each iteration for k sufficiently large. This implies that (33) is always satisfied for k large enough. Hence $\eta_k \to 0$ and $\rho^k \to 0$. By Lemma 3, $\{(\mu^k, \lambda^k, \nu^k)\}$ converges. Also, from the algorithm, we have $\epsilon_k \to 0$. Similar to Case (ii) in the proof of Theorem 2, we can infer from (37) and (38) that $\bar{x}$ is a global optimal solution to MPCC. □

6 Penalty parameter updating and normalization of multipliers

In this section, we investigate the use of two other strategies in modifying the basic augmented Lagrangian algorithm when using L_1 or L_2.

We first investigate the strategy of updating the penalty parameter c_k using the information of multiplier [12,27,34].

Theorem 5 *Assume that Assumption 1,* $(A1)$–$(A2)$ *and* $(A3')$ *for* σ_j $(j = 1, 2, 3)$ *and* $(B1)$–$(B3)$ *for* θ *are satisfied. Let* c_k *in Algorithm 1 be updated by the following formulations:*

$$c_{k+1} = \begin{cases} c_k \left[\max\left\{\gamma, \sum_{i=1}^{m_1} \left|\mu_i^{k+1}\right|, \sum_{j=1}^{m_2} \left|\lambda_j^{k+1}\right|, \sum_{l=1}^{m_3} \left|\nu_l^{k+1}\right|\right\}\right]^2, & \text{for } L_1 \\ c_k \max\left\{\gamma, \sum_{i=1}^{m_1} \theta\left(\mu_i^{k+1}\right), \sum_{j=1}^{m_2} \theta\left(\lambda_j^{k+1}\right), \sum_{l=1}^{m_3} \theta\left(\nu_l^{k+1}\right)\right\}, & \text{for } L_2 \end{cases} \tag{40}$$

where $\gamma > 1$. *Then each limit point of the sequence* $\{x_k\}$ *generated by the modified Algorithm 1 is a global optimal solution to MPCC.*

Proof We first prove the theorem for L_1. Let x^* be a global solution to (P). Similar to the proof of Theorem 2, we have (37) and (38). Since $\gamma > 1$, we see from (40) that $c_k \to \infty$ as $k \to \infty$. Again, from (40), we have

$$\frac{1}{\sqrt{c_k}} \geq \frac{\sum_{i=1}^{m_1} |\mu_i^{k+1}|}{\sqrt{c_{k+1}}} \geq 0.$$

Thus

$$\lim_{k\to\infty} \frac{\mu^k}{\sqrt{c_k}} = 0. \tag{41}$$

Similarly, we have

$$\lim_{k\to\infty} \frac{\lambda^k}{\sqrt{c_k}} = 0, \quad \lim_{k\to\infty} \frac{\nu^k}{\sqrt{c_k}} = 0. \tag{42}$$

Using the similar arguments as in Case (i) of the proof of Theorem 4, we can prove from (37), (38), (41) and (42), that each limit point of the sequence $\{x^k\}$ is a global optimal solution to MPCC.

Next, we prove the case for L_2. From (40), we have

$$\lim_{k\to\infty} \frac{1}{c_k} \sum_{i=1}^{m_1} \theta\left(\mu_i^k\right) = 0, \quad \lim_{k\to\infty} \frac{1}{c_k} \sum_{i=1}^{m_2} \theta\left(\lambda_i^k\right) = 0, \quad \lim_{k\to\infty} \frac{1}{c_k} \sum_{i=1}^{m_3} \theta\left(\nu_i^k\right) = 0. \tag{43}$$

Similar to Case (i) in the proof of Theorem 3, using (43), we can deduce from (25) (with $(\bar{\mu}^k, \bar{\lambda}^k, \bar{\nu}^k)$ replaced by $(\mu^k, \lambda^k, \nu^k)$) that each limit point of the sequence $\{x^k\}$ is a global optimal solution to MPCC. □

Next, we consider another approach to guarantee the boundedness of multipliers in the basic augmented Lagrangian method. The idea is to normalize the multipliers in the augmented Lagrangians $L_j (j = 1, 2)$. Similar idea was used in [6] for constructing another type of augmented Lagrangian function. Let

$$\bar{\mu}^k = \frac{\mu^k}{1 + \|\mu^k\|}, \quad \bar{\lambda}^k = \frac{\lambda^k}{1 + \|\lambda^k\|}, \quad \bar{\nu}^k = \frac{\nu^k}{1 + \|\nu^k\|}. \tag{44}$$

Replacing $(\mu^k, \lambda^k, \nu^k)$ in the definition of L_j $(j = 1, 2)$ with $(\bar{\mu}^k, \bar{\lambda}^k, \bar{\nu}^k)$ results in the following normalized Lagrangian functions:

$$\tilde{L}_j(x, \mu^k, \lambda^k, \nu^k, c) = L_j(x, \bar{\mu}^k, \bar{\lambda}^k, \bar{\nu}^k, c), \quad j = 1, 2. \tag{45}$$

Theorem 6 *Let* L_j $(j = 1, 2)$ *in* (7) and λ^k in subproblem (8) in Algorithm 1 be replaced by $\tilde{L}_j$ $(j = 1, 2)$ and $\bar{\lambda}^k$, respectively. Also, let the multiplier updating in Step 2 of Algorithm 1 be replaced by

$$\begin{cases} \mu^{k+1} = \bar{\mu}^k + c_k h^1(x^k, \mu^k, c_k), \\ \lambda^{k+1} = \bar{\lambda}^k + c_k h^2(x^k, \lambda^k, c_k), \\ \nu^{k+1} = \bar{\nu}^k + c_k h^3(x^k, \nu^k, c_k). \end{cases}$$

Suppose that Assumption 1, (A1)–(A3) for $\sigma_i (i = 1, 2, 3)$ and (B1)–(B3) for θ are satisfied. If $c_k \to \infty$ as $k \to \infty$, then each limit point of the sequence $\{x^k\}$ generated by the modified Algorithm 1 is a global optimal solution to MPCC.

Proof We only prove the theorem for $\tilde{L}_1$. The case for $\tilde{L}_2$ can be proved similarly. By Step 1 of Algorithm 1 and (45), we have

$$L_1(x^k, \bar{\mu}^k, \bar{\lambda}^k, \bar{\nu}^k, c_k) \le \min_{x \in X} L_1(x, \bar{\mu}^k, \bar{\lambda}^k, \bar{\nu}^k, c_k) + \epsilon_k.$$

Note from (44) that $\{(\bar{\mu}^k, \bar{\lambda}^k, \bar{\nu}^k)\}$ is bounded. The reminder of the proof is similar to Case (i) of the proof of Theorem 2. □

7 Concluding remarks

We have presented some new convergence properties for modified augmented Lagrangian methods based on two classes of augmented Lagrangian functions for MPCC. The main contribution of the paper is an investigation of different strategies for modifying the basic primal–dual method so that the convergence to a global optimal solution can be achieved without appealing to the restrictive assumption on the boundedness of the multipliers. The results obtained in this paper may help to understand the global behaviors of augmented Lagrangian methods for MPCC.

Acknowledgments This work was supported by the Natural Science Foundation of Zhejiang Province grant Y7080184, the National Natural Science Foundation of China grant 70832002 and 70771028, and Shanghai Leading Academic Discipline Project (B210).

References

1. Andreani, R., Birgin, E.G., Martínez, J.M., Schuverdt, M.L.: On augmented Lagrangian methods with general lower-level constraints. SIAM J. Optim. **18**, 1286–1309 (2007)
2. Andreani, R., Birgin, E.G., Martínez, J.M., Schuverdt, M.L.: Augmented Lagrangian methods under the constant positive linear dependence constraint qualification. Math. Program. **111**, 5–32 (2008)
3. Bertsekas, D.P.: Constrained Optimization and Lagrangian Multiplier Methods. Academic Press, New York (1982)
4. Birgin, E.G., Castillo, R.A., Martínez, J.M.: Numerical comparison of augmented Lagrangian algorithms for nonconvex problems. Comput. Optim. Appl. **31**, 31–55 (2005)
5. Conn, A.R., Gould, N.I.M., Toint, P.L.: A globally convergent augmented Lagrangian algorithm for optimization with general constraints and simple bounds. SIAM J. Numer. Anal. **28**, 545–572 (1991)
6. Di Pillo, G., Lucidi, S.: An augmented Lagrangian function with improved exactness properties. SIAM J. Optim. **12**, 376–406 (2001)
7. Facchinei, F., Jiang, H., Qi, L.: A smoothing method of mathematical programs with equilibrium constraints. Math. Program. **85**, 107–134 (1999)
8. Fukushima, M., Pang, J.S.: Some feasibility issues in mathematical programs with equilibrium constraints. SIAM J. Optim. **8**, 673–681 (1998)
9. Fukushima, M., Tseng, P.: An implementable active-set algorithm for computing a B-stationary point of a mathematical program with linear complementarity constraints. SIAM J. Optim. **12**, 724–739 (2001)
10. Giannessi, F., Maugeri, A., Pardalos, P.M. (eds.): Equilibrium Problems and Variational Models. Kluwer, Dordrecht (2001)
11. Giannessi, F., Pardalos, P.M., Rapcsak, T. (eds.): New Trends in Equilibrium Systems. Kluwer, Dordrecht (2001)
12. Griva, I., Polyak, R.: A primal–dual nonlinear rescaling method with dynamic scaling parameter update. Math. Program. **106**, 237–259 (2006)
13. Hu, X.M., Ralph, D.: Convergence of a penalty method for mathematical programming with complementarity constraints. J. Optim. Theory Appl. **123**, 365–390 (2004)
14. Huang, X.X., Yang, X.Q., Teo, K.L.: Partial augmented Lagrangian method and mathematical programs with complementarity constraints. J. Global Optim. **35**, 235–254 (2006)
15. Kanzow, C., Yamashita, N., Fukushima, M.: New NCP-functions and thier properties. J. Optim. Theory Appl. **94**, 115–135 (1997)
16. Lewis, R.M., Torczon, V.: A globally convergent augmented Lagrangian pattern search algorithm for optimization with general constraints and simple bounds. SIAM J. Optim. **12**, 1075–1089 (2002)
17. Luo, H.Z.: Studies on the Augmented Lagrangian Methods and Convexification Methods for Constrained Global Optimization. PhD thesis, Shanghai University (2007)
18. Luo, H.Z., Sun, X.L., Li, D.: On the convergence of augmented Lagrangian methods for constrained global optimization. SIAM J. Optim. **18**, 1209–1230 (2007)
19. Luo, H.Z., Sun, X.L., Wu, H.X.: Convergence properties of augmented Lagrangian methods for constrained global optimization. Optim. Methods Softw. **23**, 763–778 (2008)
20. Luo, Z.Q., Pang, J.S., Ralph, D.: Mathematical Programs with Equilibrium Constraints. Cambridge University Press, Cambridge (1996)
21. Mangasarian, O.L.: Unconstrained Lagrangians in nonlinear programming. SIAM J. Control Optim. **13**, 772–791 (1975)
22. Mangasarian, O.L.: Equivalence of the complementarity problem to a system of nonlinear equations. SIAM Appl. Math. **31**, 89–92 (1976)
23. Nguyen, V.H., Strodiot, J.J.: On the convergence rate of a penalty function method of exponential type. J. Optim. Theory Appl. **27**, 495–508 (1979)
24. Pardalos, P.M., Rosen, J.B.: Global optimization approach to the linear complementarity problem. SIAM J. Sci. Stat. Comput. **9**, 341–353 (1988)
25. Polyak, R.: Modified barrier functions: theory and methods. Math. Program. **54**, 177–222 (1992)
26. Polyak, R.: Log-Sigmoid multipliers method in constrained optimization. Ann. Oper. Res. **101**, 427–460 (2001)
27. Polyak, R.: Nonlinear rescaling vs. smoothing technique in convex optimization. Math. Program. **92**, 197–235 (2002)
28. Rockafellar, R.T.: Lagrange multipliers and optimality. SIAM Rev. **35**, 183–238 (1993)
29. Rockafellar, R.T., Wets, R.J.-B.: Variational Analysis. Springer, Berlin (1988)
30. Scheel, H., Scholtes, S.: Exact penalization of mathematical programs with equilibrium constraints. SIAM J. Control Optim. **37**, 617–652 (1999)

31. Scholtes, S.: Convergence properties of a regularization scheme for mathematical programs with complementarity constraints. SIAM J. Optim. **11**, 918–936 (2001)
32. Scholtes, S., Stohr, M.: Mathematical programs with complementarity constraints: stationarity, optimality and sensitivity. Math. Oper. Res. **25**, 1–22 (2000)
33. Sun, X.L., Li, D., McKinnon, K.I.M.: On saddle points of augmented Lagrangians for constrained nonconvex optimization. SIAM J. Optim. **15**, 1128–1146 (2005)
34. Tseng, P., Bertsekas, D.P.: On the convergence of the exponential multiplier method for convex programming. Math. Program. **60**, 1–19 (1993)
35. Yang, X.Q., Huang, X.X.: Convergence analysis of an augmented Lagrangian methods for mathematical programs with complementarity constraints. Nonlinear Anal. **63**, 2247–2256 (2005)