Comput Optim Appl (2010) 47: 501–527
DOI 10.1007/s10589-008-9229-y

An aggregate deformation homotopy method for min-max-min problems with max-min constraints

Hui-juan Xiong · Bo Yu

Received: 1 May 2008 / Revised: 7 December 2008 / Published online: 30 December 2008

Abstract In this paper, the constrained min-max-min problem, which is an essentially nonsmooth and nonconvex problem, is considered. Based on a twice aggregate function with a modification, an aggregate deformation homotopy method is established. Under some suitable assumptions, a smooth path from a randomly given point to a solution of the generalized KKT system is proven to exist. By numerically tracing the smooth path, a globally convergent algorithm for some solution of the problem is given. Some numerical results are given to show the feasibility of the method.

Keywords Nonsmooth optimization · Min-max-min problem · Aggregate function · Homotopy method

1 Introduction

Consider the constrained min-max-min problem ($\mathrm{CM}^3\mathrm{P}$):

$$\begin{aligned} &\min_{x\in\mathbf{R}^n}\left\{f(x)=\max_{1\le i\le m}\ \min_{1\le j\le l_i}\{f_{ij}(x)\}\right\},\\ &\text{s.t.}\quad g(x)=\max_{1\le i\le k}\ \min_{1\le j\le p_i}\{g_{ij}(x)\}\le 0, \end{aligned} \tag{1.1}$$

where $f_{ij}, g_{ij}\in C^r\ (r>2)$.

The research was supported by the National Nature Science Foundation of China (10671029) and the Research Fund for the Doctoral Programme of Higher Education (20060141029).

H.-j. Xiong (✉) · B. Yu
Department of Mathematics, Dalian University of Technology, Dalian, China
e-mail: xhj98021135@163.com

B. Yu
e-mail: yubo@dlut.edu.cn

The CM^{3}P is a typical problem that appears in various fields like engineering design, circuit design, heat exchange, data mining and etc. (see [3, 17, 30, 31, 35, 36, 39, 40] for examples). Despite its wide applications, the literature on implementable and convenient algorithms for solving it is little because the max-min functions $f(x)$ and $g(x)$ are nonconvex and nonsmooth even if f_{ij} and g_{ij} are all linear functions. In [7], quasi-differentiability of a max-min function was discussed, and a first order necessary condition in terms of quasi-differential for optimization problems with nonsmooth max-min functions was given, however, it is not convenient to use general quasi-differential theory to construct practical algorithms. It's preferable to utilize its special structure to construct effective algorithms.

A simple and direct idea is to decompose a CM^{3}P into a collection of optimization problems with smooth inequality constraints and the solution of the CM^{3}P can be obtained by solving all subproblems in the collection. However, the collection of the CM^{3}P may be very large that makes the direct method exorbitantly expensive. Moreover, the feasible set of some subproblems may be empty and thus the direct method may not work (see [17]). In [40], S. Scholtes extended the SQP method to combinatorial nonlinear programming and applied it to nonlinear programming problems with max-min inequality constraints, where an active set strategy was given for the subproblem that can be seen as an improvement of this idea.

In [17], based on the fact that a set of real numbers contains a non-positive element if and only if its convex hull contains a non-positive element, C. Kirjner-Neto and E. Polak presented a transcription of the problem into a single smooth inequality constrained nonlinear programming problem. This method is convenient to utilize existing algorithms and softwares for standard optimization problems, however, because some auxiliary variables were introduced, the computation will be performed in highly dimensional spaces and hence is not efficient when m, k, p_i, l_i in (1.1) are big. Moreover, the correspondence between the minimizers of the original problem and the transformed problem need some assumptions. In [20], S.J. Li et al. proved the equivalence of the transformed standard nonlinear programming problem with the original problem under weaker assumptions.

Since Karmarkar published his famous paper [16], the interior point algorithm started from 1950's has been well studied and it has become a kind of very efficient algorithm for solving linear programming, nonlinear programming, semi-definite programming (SDP) and some other related problems. For linear programming and smooth convex nonlinear programming, many good theoretical results as well as excellent implementable techniques for interior point algorithms are presented, however, only a few results on interior point methods or homotopy methods are given for nonconvex and/or nonsmooth optimization problems.

In [9, 10], a combined homotopy interior point method (CHIP) for nonconvex programming with some inequality constraints was given. Under a so-called normal cone condition as well as some commonly used constraint qualification conditions, the existence and convergence of a smooth interior path, which is trivial for linear programming and convex programming, from a randomly given interior point to a solution of the KKT system, was proven. In [26], the combined homotopy method for nonconvex programming problem with both inequality constraints and equality constrains was given. In [29, 53], weaker versions of the normal cone condition were given, under which existence and convergence of smooth interior paths were also proven.

In [44], a different globally convergent homotopy method for nonconvex programming was presented by L.T. Watson. Under some conditions, a smooth curve emanating from a randomly initial point was proven to exist. Under some extra assumptions on the smooth curve, it was proven that the curve can approach to a KKT solution point of the original problem. Based on the idea by L.T. Watson and the existing combined homtopy idea, a new constraint shifting homotopy method for smooth nonconvex problems with inequality constraints was given by Y.F. Shang and B. Yu in [41]. By constructing some auxiliary constraint shifting functions, the existence and convergence of a smooth homotopy path, from an arbitrarily initial point to a solution of the KKT system, was proven only if the auxiliary constraint shifting functions satisfying some suitable conditions. The new constraint shifting homotopy method is better than the original CHIP. It's feasible for almost all initial point. Moreover, the normal cone condition of the original feasible set, that is essential for the original CHIP, is relaxed to that of the initial deforming feasible set. If choosing suitable constraint shifting functions such that the initial deforming feasible set is convex, the normal cone condition of the initial deforming feasible set holds naturally.

Some more general deformation ideas were discussed in [5, 13–15, 18] and etc. In [15], a deformation strategy was discussed by H.Th. Jongen and et al. for standard nonlinear programming problem in the sense that Lagrange multipliers corresponding to active inequality constraints were allowed to be negative. Respectively, a continuous deformation based on the notion of Kojima's strong stability was discussed by M. Kojima et al. in [18]. In spite of the difference, both of them were embedding the original problem into a parameterized one and focused on the discussion of the properties of the parameterized problem. In [5, 13, 14] and etc., J. Guddat et al. did much work on the idea. They constructed some detailed deformations such that the original problem can be deformed from a simple problem with an easy obtained solution, such as a problem with a quadratic convex objective function and a convex ball constraint. By some path-following strategies in parameterized optimization with jumps, they proved that a generalized critical point of the original problem can be traced. However, the focus of them is not on constructing effective deformation to determine a smooth path convenient to trace but on analyzing the possible singularities of the path determined by a given deformation. As a result, the strategy sometimes is difficult to numerically implement. In the worst case all connected components must be found, but the problem determining the connected components number is still open to solve.

Aggregate function approximation is an attractive smoothing technique for max-type nonsmooth optimization problems. It was derived based on the Jaynes' maximum entropy principles by X.S. Li in [21] and was sometimes referred to as an exponential penalty item (see [19]). Because of its good approximation property to the max-function, it was used extensively for many problems such as the nonlinear programming problem [21, 22, 51], the non-smooth min-max problem [23, 24, 36, 48], variational inequalities [27], generalized complementarity problems [8, 34, 37], mathematical programm with equilibrium constraints (MPEC) [4] etc.

In [36], based on the aggregate function, an adaptive smoothing method was given for the finite min-max problems. In [35], a similar adaptive smoothing strategy was given for an unconstrained min-max-min problem with the aggregate function smoothing the inner min-function.

In [51], an aggregate homotopy method for a non-convex programming problem with inequality constraints was presented by utilizing the aggregate function smoothing the unified single max-type inequality constraint. In [27, 28, 52], the method was extended to the constrained sequential max-min problems, where the CM^{3}P in this paper was seen as a special case and considered. In the papers, a first-order necessary optimality condition of the CM^{3}P based on Clarke's subdifferential was first analyzed and a generalized KKT system was established. By constructing the following twice aggregate function and its modified version,

$$f(x,\mu) = \mu \ln \sum_{i=1}^{m} \left(\sum_{j=1}^{l_i} \exp\left(-\frac{f_{ij}(x)}{\mu} \right) \right)^{-1}, \tag{1.2}$$

$$g(x,\mu) = \mu \ln \sum_{i=1}^{k} p_i \left(\sum_{j=1}^{p_i} \exp\left(-\frac{g_{ij}(x)}{\mu} \right) \right)^{-1}, \tag{1.3}$$

for smoothing $f(x)$ and $g(x)$ in (1.1) respectively, the generalized KKT system of the CM^{3}P can be embedded into the following aggregate homotopy equation,

$$H(\omega,\omega^0,\mu) \equiv \begin{pmatrix} (1-\mu)(\nabla_x f(x,\mu) + y\nabla_x g(x,\theta\mu)) + \mu(x-x^0) \\ yg(x,\theta\mu) - \mu y^0 g(x^0,\theta) \end{pmatrix} = 0, \tag{1.4}$$

where $\omega = (x,y) \in \mathbf{R}^{n+1}$, $\omega^0 = (x^0,y^0) \in \Omega^0 \times \mathbf{R}^1_{++}$, $\nabla_x f(x,\mu)$ and $\nabla_x g(x,\theta\mu)$ are the gradients of $f(x,\mu)$ and $g(x,\theta\mu)$ with respect to x, and $\theta \in (0,1]$ is given in advance. Under some conditions, the existence and convergence of a smooth interior path approaching to a generalized KKT point of the problem was proven. Compared to the existing methods for nonsmooth problems with max-min functions, the aggregate homotopy method doesn't require any auxiliary variables or solving any subproblems and is globally convergent. However, the method is deficient in that it requires a restrictive weak normal cone condition of the feasible set and the initial point must be in the interior of a closed subset of the feasible set.

In this paper, we try to establish a new aggregate deformation homotopy method for the CM^{3}P. To obtain this, we first construct a detailed deformation by giving some modifications on the twice aggregate function defined as (1.2) and (1.3), and adding an extra ball constraint to guarantee $\Omega(1)$ convex and $\Omega(t)$ bounded during the deformation. With this detailed constraint shifting deformation strategy, the generalized KKT system of the CM^{3}P can be embedded into an aggregate deformation homotopy equation. Under an extension of the enlarged MFCQ condition in [14], which is only on the original feasible set and unrelated to the constraint shifting functions, the regularity during deformation can be proven to be satisfied. As a result, the existence and convergence of a smooth regular path, starting from a randomly initial point and approaching to a generalized KKT solution of the CM^{3}P, is proven. Under some extra conditions, it's proven the obtained solution point is not a local maximum solution point.

The paper is organized as follows. In Sect. 2, some necessary preliminaries are given. Section 3 introduces the aggregate deformation homotopy method with a detailed predictor-corrector procedure tracing the homotopy solution curve. In Sect. 4, the comparison results of the numerical examples are given. The conclusions with some future work is given in the last section.

1.1 Notation

The following notations are used throughout the paper. We denote the feasible set of (1.1) as $\Omega = \{x \in \mathbf{R}^n \,|\, \max_{1\le i\le k} \min_{1\le j\le p_i}\{g_{ij}(x)\} \le 0\}$, the strict feasible set as $\Omega^0 = \{x \in \mathbf{R}^n \,|\, \max_{1\le i\le k} \min_{1\le j\le p_i}\{g_{ij}(x)\} < 0\}$, the boundary of the constraint set as $\partial\Omega = \Omega \setminus \Omega^0$. $f_i(x) = \min_{1\le j\le l_i}\{f_{ij}(x)\}$, $g_i(x) = \min_{1\le j\le p_i}\{g_{ij}(x)\}$, $g(x) = \max_{1\le i\le k} \min_{1\le j\le p_i}\{g_{ij}(x)\}$, $I(x) = \{i \in \{1,2,\dots,m\} : f(x) = f_i(x)\}$, $J_i(x) = \{j \in \{1,2,\dots,l_i\} : f_{ij}(x) = f_i(x)\}$, $II(x) = \{i \in \{1,2,\dots,k\} : g(x) = g_i(x)\}$, $JJ_i(x) = \{j \in \{1,2,\dots,p_i\} : g_i(x) = g_{ij}(x)\}$. Denote $|I(x)|$ as the number of $I(x)$, $|II(x)|$ as the number of $II(x)$. $\|\cdot\|$ is the Euclidean norm of a vector. The ball with center zero and radius p is denoted by $\mathcal{B}_p$.

2 Preliminaries

We first recall the necessary optimality condition of the CM^{3}P from [27] and [28]. Our method will focus on obtaining a solution point satisfying the optimality condition.

Lemma 2.1 *Suppose that $f_{ij}(x)$ $(1 \le i \le m, 1 \le j \le l_i)$ and $g_{ij}(x)$ $(1 \le i \le k, 1 \le j \le p_i)$ are continuously differentiable functions, then $f : R^n \to R$ and $g : R^n \to R$ are locally Lipschitz at $x \in R^n$ and their subdifferential in the sense of Clarke can be computed as follows,*

$$\partial f(x) \subseteq \left\{ \sum_{i\in I(x)} \delta_i \sum_{j\in J_i(x)} \eta_{ij}\nabla f_{ij}(x) : \sum_{i\in I(x)} \delta_i = 1, \sum_{j\in J_i(x)} \eta_{ij} = 1, \delta_i \ge 0, \eta_{ij} \ge 0 \right\},$$

$$\partial g(x) \subseteq \left\{ \sum_{i\in II(x)} \lambda_i \sum_{j\in JJ_i(x)} \mu_{ij}\nabla f_{ij}(x) : \right.$$
$$\left. \sum_{i\in I(x)} \lambda_i = 1, \sum_{j\in J_i(x)} \mu_{ij} = 1, \lambda_i \ge 0, \mu_{ij} \ge 0 \right\},$$

where

$$I(x) = \{i \in \{1,2,\dots,m\} : f(x) = f_i(x)\}\},$$
$$J_i(x) = \{j \in \{1,2,\dots,l_i\} : f_i(x) = f_{ij}(x)\};$$
$$II(x) = \{i \in \{1,2,\dots,k\} : g(x) = g_i(x)\}\},$$
$$JJ_i(x) = \{j \in \{1,2,\dots,p_i\} : g_i(x) = g_{ij}(x)\}.$$

Definition 2.1 The CM^{3}P is called to be regular at $x \in \partial\Omega$, if $\nabla g_{ij}(x)$ $(i \in II(x), j \in JJ_i(x))$ are positive independent, that is,

$$\sum_{i \in II(x)} \alpha_i \sum_{j \in JJ_i(x)} \beta_{ij} \nabla g_{ij}(x) = 0, \quad \alpha \geq 0, \beta \geq 0 \quad \Rightarrow \quad \alpha = 0, \beta = 0.$$

Theorem 2.1 *Suppose $f_{ij}(x)$ $(1 \leq i \leq m, 1 \leq j \leq l_i)$ and $g_{ij}(x)$ $(1 \leq i \leq k, 1 \leq j \leq p_i)$ are continuously differential functions, x^* is a local minimum of the CM^{3}P and the regularity holds at x^*. Then there exists a $y^* \geq 0, \delta_i \geq 0, \eta_{ij} \geq 0, \lambda_i \geq 0, \mu_{ij} \geq 0$, such that*

$$\begin{aligned} &\sum_{i \in I(x^*)} \delta_i \sum_{j \in J_i(x^*)} \eta_{ij} \nabla f_{ij}(x^*) + y \sum_{i \in II(x^*)} \lambda_i \sum_{j \in JJ_i(x^*)} \mu_{ij} \nabla g_{ij}(x^*) = 0, \\ &yg(x^*) = 0, \quad y \geq 0, \ g(x^*) \leq 0, \\ &\sum_{i \in I(x^*)} \delta_i = 1, \quad \sum_{j \in J_i(x^*)} \eta_{ij} = 1, \quad \sum_{i \in II(x^*)} \lambda_i = 1, \quad \sum_{j \in JJ_i(x^*)} \mu_{ij} = 1 \end{aligned} \tag{2.1}$$

(2.1) *is said a generalized KKT system of* (1.1) *and the point* (x^*, y^*) *satisfying* (2.1) *is called a generalized KKT point.*

We use the following twice aggregate function

$$f(x,t) = t \ln\left(\sum_{i=1}^{m}\left(\sum_{j=1}^{l_i} \exp\left(-\frac{f_{ij}(x)}{t}\right)\right)^{-1}\right) \tag{2.2}$$

and

$$g(x,t) = t \ln\left(\frac{1}{k}\sum_{i=1}^{k}\left(\sum_{j=1}^{p_i} \exp\left(-\frac{g_{ij}(x)}{t}\right)\right)^{-1}\right) \tag{2.3}$$

to give smooth approximations of $f(x)$ and $g(x)$. The two functions have good approximate properties listed in the following proposition. The proves are similar with the proof in [27, 28] and omitted.

Proposition 2.1 *If $f_{ij}(x)$ and $g_{ij}(x)$ are in C^r $(r > 2)$, then*

(1) *$f(x,t)$ is smooth with respect to $t > 0$ for all $x \in \mathbf{R}^n$, and*

$$f(x) - t \ln\left(\max_{1 \leq i \leq m} \{l_i\}\right) \leq f(x,t) \leq f(x) + t \ln m;$$

as $t \to 0$, $f(x,t) \to f(x)$;

(2) *$g(x,t)$ is smooth with respect to $t > 0$ for all $x \in \mathbf{R}^n$, and*

$$g(x) - t \ln\left(k \max_{1 \leq i \leq k} \{p_i\}\right) \leq g(x,t) \leq g(x);$$

as $t \to 0$, $g(x,t) \to g(x)$;

(3) *$f(x,t)$ is r-time continuously differential with respect to x for all $t>0$, and*

$$\nabla_x f(x,t)=\sum_{i=1}^{m}\delta_i(x,t)\sum_{j=1}^{l_i}\eta_{ij}(x,t)\nabla f_{ij}(x),$$

where

$$\delta_i(x,t)=\frac{\left(\sum_{j=1}^{l_i}\exp(-\frac{f_{ij}(x)}{t})\right)^{-1}}{\sum_{i=1}^{m}\left(\sum_{j=1}^{l_i}\exp(-\frac{f_{ij}(x)}{t})\right)^{-1}},\qquad \eta_{ij}(x,t)=\frac{\exp(-\frac{f_{ij}(x)}{t})}{\sum_{j=1}^{l_i}\exp(-\frac{f_{ij}(x)}{t)})};$$

(4) *$g(x,t)$ is r-time continuously differential with respect to x for all $t>0$, and*

$$\nabla_x g(x,t)=\sum_{i=1}^{k}\lambda_i(x,t)\sum_{j=1}^{p_i}\mu_{ij}(x,t)\nabla g_{ij}(x),$$

where

$$\lambda_i(x,t)=\frac{\left(\sum_{j=1}^{p_i}\exp(-\frac{g_{ij}(x)}{t})\right)^{-1}}{\sum_{i=1}^{k}\left(\sum_{j=1}^{p_i}\exp(-\frac{g_{ij}(x)}{t})\right)^{-1}},\qquad \mu_{ij}(x,t)=\frac{\exp(-\frac{g_{ij}(x)}{t})}{\sum_{j=1}^{p_i}\exp(-\frac{g_{ij}(x)}{t})};$$

(5) *For any $0<t\le 1$, $x\in\mathbf{R}^n$, $\delta_i(x,t)$, $\eta_{ij}(x,t)$ defined in (3), $\lambda_i(x,t)$ and $\mu_{ij}(x,t)$ defined in (4), we have*

$$0\le\delta_i(x,t),\quad \eta_{ij}(x,t)\le 1,\quad \sum_{i=1}^{m}\delta_i(x,t)=1,\quad \sum_{j=1}^{l_i}\eta_{ij}(x,t)=1;$$

$$0\le\lambda_i(x,t),\quad \mu_{ij}(x,t)\le 1,\quad \sum_{i=1}^{k}\lambda_i(x,t)=1,\quad \sum_{j=1}^{p_i}\mu_{ij}(x,t)=1;$$

(6) *For any given $x\in\mathbf{R}^n$, $\delta_i(x,t)$, $\eta_{ij}(x,t)$ defined in (3), $\lambda_i(x,t)$ and $\mu_{ij}(x,t)$ defined in (4), we have as $t\to 0$, $\delta_i(x,t)\to 0$, for $i\notin I(x)$, $\delta_i(x,t)\to\frac{1}{|I(x)|}$, for $i\in I(x)$; $\eta_{ij}(x,t)\to 0$ for $j\notin J_i(x)$, $\eta_{ij}(x,t)\to\frac{1}{|J_i(x)|}$ for $j\in J_i(x)$. $\lambda_i(x,t)\to 0$, for $i\notin II(x)$; $\lambda_i(x,t)\to\frac{1}{|II(x)|}$, for $i\in II(x)$; $\mu_{ij}(x,t)\to 0$, for $j\notin JJ_i(x)$, $\mu_{ij}(x,t)\to\frac{1}{|JJ_i(x)|}$ for $j\in JJ_i(x)$.*

3 The aggregate deformation homotopy method and its convergence

In this section, we introduce the aggregate deformation homotopy method. For this, we first give an assumption,

Assupmtion A1 Ω^0 is nonempty and the solution set is bounded.

In many interior point methods for nonlinear programming problem such as the methods listed in [2, 50, 51], it's essential to assume that the feasible set is bounded or the iteration point sequences keep bounded. In [47], B. Yu et al. discussed the convergence of the combined homotopy method in an unbounded feasible set. Under an extra weak assumption that the problem has no solution at infinity, that is, for any given $x \in \Omega$, sequence $\{x^k\} \subset \Omega$, $\|x^k\| \to +\infty$, there exists $y^k \geq 0$ such that $\lim_{k\to+\infty}(x - x^k)^T(\nabla f(x^k) + \nabla g(x^k)y^k) < 0$, a smooth interior point homotopy path was proven to exist.

Assumption A1 is actually an alternative description of the assumption that the problem has no solution at infinity in [47]. If the Assumption A1 holds, then there exists a large enough constant p such that the ball $\mathcal{B}_p$ include all the solutions of Ω.

For a randomly given initial point $x^{(0)} \in \mathbf{R}^n$ and $x^{(0)} \in \mathcal{B}_p^0$, we define the following deforming function to give a smooth deformation of the original non-smooth constraint,

$$\begin{aligned} g_1(x,t) &= (1-t)^{\beta_1} g_\theta(x,t) - t^{\beta_2}\rho \leq 0, \\ g_2(x,t) &= x^T x - p \leq 0, \end{aligned} \tag{3.1}$$

where $g_\theta(x,t) = g(x,\theta t)$, $g(\cdot,\cdot)$ are defined as (2.3), $\theta \in (0,1]$ is a constant, $\beta_1, \beta_2, \rho > 0$.

Denote $\Omega_1(t) = \{x|g_1(x,t) \leq 0\}$, $\Omega_2(t) = \{x|g_2(x,t) \leq 0\}$, $\Omega(t) = \Omega_1(t) \cap \Omega_2(t)$, $I(x,t) = \{i \in \{1,2\} : g_i(x,t) = 0\}$. Followed from the construction of $g_1(x,t)$ and $g_2(x,t)$, it's obvious that as $t = 1$, $\Omega(1) = \mathcal{B}_p$ is a convex set and as $t \to 0$, $\Omega(t) \to \Omega \cap \mathcal{B}_p$.

To solve (1.1), we embed the KKT system of the smoothing parameterized problem into the following homotopy equation $H : \mathbf{R}^{n+2} \times (0,1] \to \mathbf{R}^{n+2}$,

$$\begin{aligned} H_{\omega^0}(\omega,t) &= \begin{pmatrix} (1-t)(\nabla_x f(x,t) + y_1\nabla_x g_1(x,t) + y_2\nabla_x g_2(x,t)) + t(x-x^0) \\ y_1 g_1(x,t) - t y_1^0 g_1(x^{(0)},1) \\ y_2 g_2(x,t) - t y_2^0 g_2(x^{(0)},1) \end{pmatrix} \\ &= 0, \end{aligned} \tag{3.2}$$

where $\omega = (x,y)$, $\omega^0 = (x^0,y^0) \in \Omega^0(1) \times \mathbf{R}^2_{++}$, $\nabla_x g_1(x,t) = (1-t)^{\beta_1}\nabla_x g_\theta(x,t)$, $\nabla_x g_\theta(x,t)$ and $\nabla_x f(x,t)$ are defined as Proposition 2.1, $\nabla_x g_2(x,t) = 2x$.

For a given $\omega^0 \in \Omega^0(1) \times \mathbf{R}^2_{++}$, we denote the zero-point set of H by $H^{-1}(0)$, that is,

$$H^{-1}(0) = \{(\omega,t) \in \Omega(t) \times \mathbf{R}^2_+ \times (0,1] : H_{\omega^0}(\omega,t) = 0\}.$$

In the following, we will prove $H^{-1}(0)$ includes a smooth solution curve, which starts from $(\omega^0, 1)$, and as $t \to 0$, approaches to a point whose (x, y_1) component is a solution of (2.1).

3.1 Convergence

To establish the convergence of the new method, the following assumption is necessary.

Assupmtion A2 For each $x \in \mathcal{B}_p$, there exists a $\xi \in \mathbf{R}^n$ such that

$$\begin{aligned} g(x) + \eta^T \xi &< 0, \quad g(x) \geq 0; \\ x^T \xi &< 0, \quad \text{if } x^T x - p = 0, \end{aligned}$$

holds for arbitrary $\eta \in \partial g(x)$. From the Lemma 2.1,

$$\eta = \sum_{i \in II(x)} \lambda_i \sum_{j \in JJ_i(x)} \mu_{ij} \nabla g_{ij}(x),$$

where λ_i and μ_{ij} satisfy $\lambda_i \geq 0$ and $\sum_{i \in II(x)} \lambda_i = 1$, $\mu_{ij} \geq 0$ and $\sum_{j \in JJ_i(x)} \mu_{ij} = 1$.

Assumption A2 is an extension of the Enlarged MFCQ in [13, 14]. By employing Gordan's theorem of the alternative, it can be rewritten into the following equivalent form:

Assupmtion A2* For each $x \in \mathcal{B}_p$,

$$\sum_{i \in II(x)} \alpha_i \sum_{j \in JJ_i(x)} \beta_{ij} \nabla g_{ij}(x) = 0, \quad \alpha_i \geq 0, \beta_{ij} \geq 0 \quad \Rightarrow \quad \alpha = 0, \beta = 0,$$

$g(x) \geq 0;$

$$s_1 \sum_{i \in II(x)} \lambda_i \sum_{j \in JJ_i(x)} \mu_{ij} \nabla g_{ij}(x) + 2 s_2 x = 0, \quad s_i \geq 0,\ i = 1, 2 \quad \Rightarrow \quad s_1 = s_2 = 0,$$

if $x^T x - p = 0$.

In the following, we prove the existence of the smooth path. We first give a lemma to show the regularity holds during deformation.

Lemma 3.2 *If the Assumption A2** *holds, then there exists a* $\theta \in (0, 1]$, *for arbitrary* $t \in [0, 1]$ *and* $x \in \partial\Omega(t)$, $\sum_{i \in I(x,t)} s_i \nabla_x g_i(x, t) = 0$ *has no semipositive solution, that is,*

$$\sum_{i \in I(x,t)} s_i \nabla_x g_i(x, t) = 0, \quad s = (s_1, s_2) \geq 0 \quad \Rightarrow \quad s = 0.$$

Proof (1) Obviously, for any $t \in [0, 1]$, $x \in \Omega(t)$, such that $g_2(x, t) = x^T x - p = 0$, we have $\nabla_x g_2(x, t) = 2x \neq 0$.

As $t = 1$, from $\Omega(1) = \mathcal{B}_p$, the conclusion is obvious. We only consider $t \in [0, 1)$.

(2) If $g_1(x, t) = 0$, that is, $I(x, t) = \{1\}$, then the conclusion holds. Otherwise, take a sequence $\theta_k \to 0$, $t_k \in [0, 1]$, $x^{(k)} \in \partial\Omega(t_k)$ such that

$$\begin{aligned} &g_1(x^{(k)}, t_k) = (1 - t_k)^{\beta_1} g(x^{(k)}, \theta_k t_k) - t_k^{\beta_2} \rho = 0, \\ &\nabla_x g_1(x^{(k)}, t_k) = (1 - t_k)^{\beta_1} \sum_{i=1}^{k} \lambda_i(x^{(k)}, \theta_k t_k) \sum_{j=1}^{p_i} \mu_{ij}(x^{(k)}, \theta_k t_k) \nabla g_{ij}(x^{(k)}) = 0. \end{aligned} \tag{3.3}$$

From the boundedness of t_k, $x^{(k)}$, $\lambda(x^{(k)},\theta_k t_k)$ and $\mu(x^{(k)},\theta_k t_k)$, there must exist convergent subsequences, still denoted as t_k and $x^{(k)}$, $\lambda(x^{(k)},\theta_k t_k)$ and $\mu(x^{(k)},\theta_k t_k)$, such that $t_k\to t$ and $x^{(k)}\to x$, $\lambda(x^{(k)},\theta_k t_k)\to\lambda^*$ and $\mu(x^{(k)},\theta_k t_k)\to\mu^*$. Moreover, from the definition of $\lambda(x^{(k)},\theta_k t_k)$ and $\mu_{ij}(x^{(k)},\theta_k t_k)$ in Proposition 2.1, we have $\lambda_i^*=0, i\notin II(x^*)$, $\mu_{ij}^*=0$, $j\notin JJ_i(x^*)$. Take limits of the two equalities in (3.3), we have

$$\begin{aligned}
&g_1(x,t)=(1-t)^{\beta_1}g(x)-t^{\beta_2}\rho=0 \quad\Rightarrow\quad g(x)=\frac{t^{\beta_2}\rho}{(1-t)^{\beta_1}}\geq 0,\\
&\nabla_x g_1(x,t)=(1-t)^{\beta_1}\sum_{i\in II(x)}\lambda_i^*\sum_{j\in JJ_i(x)}\mu_{ij}^*\nabla g_{ij}(x)=0,\\
&\sum_{i\in II(x)}\lambda_i^*=1,\qquad \sum_{j\in JJ_i(x)}\mu_{ij}^*=1,
\end{aligned}$$

which contradicts with the Assumption A2*.

(3) If $I(x,t)=\{1,2\}$, the conclusion also holds. Otherwise, take sequences $\theta_k\to 0$, $t_k\in[0,1)$, $x^{(k)}\in\partial\Omega(t_k)$ such that $t_k\to t$, $x^{(k)}\to x$, and

$$\begin{aligned}
&g_1(x^{(k)},t_k)=(1-t_k)^{\beta_1}g(x^{(k)},\theta_k t_k)-t_k^{\beta_2}\rho=0,\\
&g_2(x^{(k)},t_k)=x^{(k)T}x^{(k)}-p=0,\\
&s_1^k\nabla_x g_1(x^{(k)},t_k)+s_2^k\nabla_x g_2(x^{(k)},t_k)=0\\
&\quad\text{has a semipositive solution } s^{(k)}=(s_1^k,s_2^k).
\end{aligned}\tag{3.4}$$

From the boundedness of sequences t_k, $x^{(k)}$ and $s^{(k)}$, there must exist convergent subsequences, still denote as t_k, $x^{(k)}$ and $s^{(k)}$. Assume $t_k\to t$, $x^{(k)}\to x$ and $s^{(k)}\to s\neq 0$, take limits of the three equalities in (3.4), we have

$$\begin{aligned}
&g_1(x,t)=(1-t)^{\beta_1}g(x)-t^{\beta_2}\rho=0 \quad\Rightarrow\quad g(x)=\frac{t^{\beta_2}\rho}{(1-t)^{\beta_1}}\geq 0,\\
&g_2(x,t)=x^Tx-p=0,\\
&s_1(1-t)^{\beta_1}\sum_{i\in II(x)}\lambda_i^*\sum_{j\in JJ_i(x)}\mu_{ij}^*\nabla g_{ij}(x)+2s_2x=0,\qquad s=(s_1,s_2)\neq 0,
\end{aligned}$$

it also leads to a contradiction with Assumption A2*.

Conclude from (1), (2) and (3), the conclusion holds. □

Theorem 3.2 *If $f_{ij}(x)$ and $g_{ij}(x)$ are three times continuously differentiable functions, and Assumptions A1 and A2* hold, then the generalized KKT system (2.1) has at least a solution. Furthermore, for almost all $w^{(0)}\in\Omega^0(1)\times\mathbf{R}^2_{++}$, the zero-point set $H^{-1}(0)$ of homotopy equation (3.2) includes a smooth curve Γ, which starts from $(w^{(0)},1)$ and terminates in or approaches to the hyperplane $t=0$. Moreover, let $(x^*,y^*,0)$ be any limit point of Γ on the hyperplane at $t=0$, then y_2^* equals 0 and*

(x^*, y_1^*) *is a solution of* (2.1), *that is*, (x^*, y_1^*) *satisfying the following equations*,

$$\sum_{i\in I(x^*)} \delta_i^* \sum_{j\in J_i(x^*)} \eta_{ij}^* \nabla f_{ij}(x^*) + y_1^* \sum_{i\in II(x^*)} \lambda_i^* \sum_{j\in JJ_i(x^*)} \mu_{ij}^* \nabla g_{ij}(x^*) = 0,$$
$$y_1^* g(x^*) = 0, \quad y_1^* \geq 0,\ g(x^*) \leq 0, \tag{3.5}$$

where $\delta^* = (\delta_1^*, \ldots, \delta_m^*), \eta_i^* = (\eta_{i1}^*, \ldots, \eta_{il_i}^*), \lambda^* = (\lambda_1^*, \ldots, \lambda_k^*), \mu_i^* = (\mu_{i1}^*, \ldots, \mu_{ip_i}^*)$ *are accumulation points of* $\delta(x,t), \eta_i(x,t)$ $(i = 1, \ldots, m), \lambda(x,t), \mu_i(x,t)$ $(i = 1, \ldots, k)$, *satisfying* $\delta_i^* = 0$ $(i \notin I(x^*)), \eta_{ij}^* = 0$ $(j \notin J_i(x^*)), \lambda_i^* = 0$ $(i \notin II(x^*))$, $\mu_{ij}^* = 0$ $(j \notin JJ_i(x^*))$.

Proof Using the parameterized Sard theorem (see [1]), similar with the proof of Lemma 2.1 in [10], we can prove for almost all $w^{(0)} \in \Omega^0(1) \times \mathbf{R}_{++}^2$, 0 is a regular value of $H_{w^{(0)}}(\cdot)$. By the inverse image theorem and $H_{w^{(0)}}(w^{(0)}, 1) = 0$, we know $H^{-1}(0)$ includes a smooth solution curve Γ starting from $(w^{(0)}, 1)$. It will return to $(w^{(0)}, 1)$, terminates in or approaches to the boundary of $\Omega(t) \times \mathbf{R}_+^2 \times [0,1]$ or goes to infinity.

Notice that

$$\frac{\partial H_{w^{(0)}}(w^{(0)}, 1)}{\partial w} = \begin{pmatrix} I_n & 0 & 0 \\ y_1 \frac{\partial g_1(x^{(0)}, t)}{\partial x} & g_1(x^{(0)}, 1) & 0 \\ y_2 \frac{\partial g_2(x^{(0)}, t)}{\partial x} & 0 & g_2(x^{(0)}, 1) \end{pmatrix} \tag{3.6}$$

is nonsingular, so $w^{(0)}$ is not a multiple solution of $H_{w^{(0)}}(w, 1) = 0$, that is, Γ cannot return to $(w^{(0)}, 1)$.

Let $(\bar{x}, \bar{y}, \bar{t})$ be any limit of Γ other than $(x^{(0)}, y^{(0)}, 1)$, the following cases are possible:

(i) $(\bar{x}, \bar{y}, \bar{t}) \in \Omega(\bar{t}) \times \mathbf{R}_+^2 \times [0,1]$, $\|\bar{y}\| \to +\infty$;
(ii) $(\bar{x}, \bar{y}, \bar{t}) \in \Omega(1) \times \mathbf{R}_+^2 \times \{1\}$, $\|\bar{y}\| < +\infty$;
(iii) $(\bar{x}, \bar{y}, \bar{t}) \in \partial(\Omega(\bar{t})) \times \mathbf{R}_{++}^2 \times (0,1)$, $\|\bar{y}\| < +\infty$;
(iv) $(\bar{x}, \bar{y}, \bar{t}) \in \Omega(\bar{t}) \times \mathbf{R}_+^2 \times (0,1)$, $\|\bar{y}\| < +\infty$ and at least one component of $\bar{y}$ equals 0;
(v) $(\bar{x}, \bar{y}, \bar{t}) \in \Omega(0) \times \mathbf{R}_+^2 \times \{0\}$, $\|\bar{y}\| < +\infty$.

Because the equation $H_{w^{(0)}}(w, 1) = 0$ has only one single solution $(w^{(0)}, 1)$ in $\Omega(1) \times \mathbf{R}_+^2$, case (ii) is impossible.

From the continuity of Γ and the second and third equalities of homotopy equation (3.2), case (iii) and (iv) are impossible.

If case (i) holds, then there exists a sequence of points $\{(x^{(k)}, y^{(k)}, t_k)\}_{k=1}^\infty$ on Γ such that $t_k \to \bar{t}$, $x^{(k)} \to \bar{x}$ and $\|y^{(k)}\| \to +\infty$ as $k \to +\infty$. From the second and third equalities of the homotopy equation (3.2), we have

$$y_1^{(k)} g_1(x^{(k)}, t_k) = t_k y_1^{(0)} g_1(x^{(0)}, 1), \qquad y_2^{(k)} g_2(x^{(k)}, t_k) = t_k y_2^{(0)} g_2(x^{(0)}, 1). \tag{3.7}$$

Take the limits of the two equalities in (3.7), if $\bar{t} \in (0, 1]$, then $\bar{y}_i \to +\infty$ for $i \in I(\bar{x}, \bar{t})$, $0 < \bar{y}_i < +\infty$ for $i \notin I(\bar{x}, \bar{t})$; if $\bar{t} = 0$, then $\bar{y}_i = 0$ for $i \notin I(\bar{x}, \bar{t})$. From the first equality of (3.2), we have

$$\begin{aligned}&(1 - t_k)(\nabla_x f(x^{(k)}, t_k) + y_1^{(k)} \nabla g_1(x^{(k)}, t_k) + y_2^{(k)} \nabla g_2(x^{(k)}, t_k))\\&\quad + t_k(x^{(k)} - x^{(0)}) = 0.\end{aligned} \tag{3.8}$$

(1) If $\bar{t}=1$, from $\Omega(1)=\mathcal{B}_p$, we know $I(\bar{x}, \bar{t}) \neq \{1\}$. Denote $\tilde{y}_2 = \lim_{k\to\infty} y_2^{(k)}(1 - t_k) \geq 0$, following cases are possible:

- if $\tilde{y}_2 = \infty$, divide the (3.8) by $|y_2^{(k)}(1 - t_k)|$ and take the limit, we will have $\bar{x} = 0$ which contradicts with $g_2(\bar{x}, \bar{t}) = 0$;
- if $\tilde{y}_2 < \infty$, taking the limit of (3.8), we have $2\tilde{y}_2\bar{x} + \bar{x} = x^0$. If $\tilde{y}_2 = 0$, then $\bar{x} = x^0$, it contradicts with $g_2(x^{(0)}, 1) < 0$ and $g_2(\bar{x}, 1) = 0$; if $\tilde{y}_2 > 0$, we have $\bar{x}^T(x^0 - \bar{x}) = 2\tilde{y}_2\bar{x}^T\bar{x} > 0$ which contradicts with the convexity of $g_2(x, t)$.

So, as $k \to \infty$, it's impossible to have $\| y^{(k)} \| \to \infty$ for the case of $\bar{t} = 1$.

(2) If $\bar{t} \in (0, 1)$, rewrite the first equality of homotopy equation (3.2) as follows

$$\nabla_x f(x^{(k)}, t_k) + y_1^k \nabla_x g_1(x^{(k)}, t_k) + y_2^k \nabla_x g_2(x^{(k)}, t_k) + \frac{t_k}{1 - t_k}(x^{(k)} - x^{(0)}) = 0. \tag{3.9}$$

Divide (3.9) by $\|y^{(k)}\|$ and take the limit of it. Denote $\tilde{y} = \lim_{k\to\infty} \frac{y^{(k)}}{\|y^{(k)}\|}$,

- if $I(\bar{x}, \bar{t}) = \{1\}$, then we have $\tilde{y}_1 = 1$, $\tilde{y}_2 = 0$ and $\tilde{y}_1 \nabla g_1(\bar{x}, \bar{t}) = 0$, it contradicts with the conclusion of Lemma 3.2;
- if $I(\bar{x}, \bar{t}) = \{2\}$, then we have $\tilde{y}_1 = 0$, $\tilde{y}_2 = 1$ and $\tilde{y}_2 \nabla g_2(\bar{x}, \bar{t}) = 2\tilde{y}_2\bar{x} = 0$, so $\bar{x} = 0$ which contracts with $g_2(\bar{x}, \bar{t}) = \bar{x}^T\bar{x} - p = 0$;
- if $I(\bar{x}, \bar{t}) = \{1, 2\}$, we have $\tilde{y}_1 + \tilde{y}_2 = 1$, $\tilde{y}_1 > 0$, $\tilde{y}_2 > 0$, and $\tilde{y}_1 \nabla g_1(\bar{x}, \bar{t}) + \tilde{y}_2 \nabla g_2(\bar{x}, \bar{t}) = 0$, it also contracts with conclusion of Lemma 3.2.

(3) If $\bar{t} = 0$, from the definition of p, $I(\bar{x}, \bar{t}) \neq \{2\}$. Divide the first equality of the homotopy (3.2) by $\|y^{(k)}(1 - t_k)\|$ and take the limit, from the boundedness of λ and μ_i $(i = 1, \ldots, k)$, we have they have accumulation points λ^*, μ_i^* $(i = 1, \ldots, k)$ and satisfying $\lambda_i^* = 0, i \notin II(x)$, $\mu_{ij}^* = 0$, $j \notin JJ_i(x^*)$. As a result,

$$\sum_{i\in II(\bar{x})} \lambda_i^* \sum_{j\in JJ_i(\bar{x})} \mu_{ij}^* \nabla g_{ij}(\bar{x}) = 0,$$

which contradicts with the Assumption A2*.

Concluded from (1), (2) and (3), it's impossible that case (i) holds. As a result, case (v) is the only possible case. That is Γ must be with finite arclength and terminate in or approach to the hyperplane at $t = 0$.

From $I(\bar{x}, 0) \neq \{2\}$ and the third equality of (3.2), we have $\bar{y}_2 = 0$. From $g(\bar{x}) = \lim_{k\to+\infty} g_1(x^{(k)}, t_k) \leq 0$ and $H(\bar{x}, \bar{y}, \bar{t}) = 0$, we have if $\bar{x} \in \Omega^0$, $\bar{y}_1 = 0$, else $\bar{y}_1 \geq 0$. So we have $\bar{y}_1 g(\bar{x}) = 0$. As a result, $(\bar{x}, \bar{y}_1)$ satisfying (3.5) and is a solution of the generalized KKT system (2.1). □

Let s be the arclength of Γ, we can parameterize Γ with respect to s. That is, there exists continuously differentiable functions $x(s), y(s), t(s)$ such that

$$\begin{aligned}
&H(x(s), y(s), t(s)) = 0,\\
&\|(\dot{x}(s), \dot{y}(s), \dot{t}(s))\| = 1,\\
&x(0) = x^{(0)}, \quad y(0) = y^{(0)}, \quad t(0) = 1,\\
&\dot{t}(0) < 0.
\end{aligned} \tag{3.10}$$

Differentiating the first equality of (3.10), we can obtain the following theorem.

Theorem 3.3 *The homotopy path Γ is determined by the following initial value problem to the ordinary differential equation*

$$\begin{aligned}
&H'(x(s), y(s), t(s)) \begin{pmatrix} \dot{x}(s) \\ \dot{y}(s) \\ \dot{t}(s) \end{pmatrix} = 0,\\
&\|(\dot{x}(s), \dot{y}(s), \dot{t}(s))\| = 1,\\
&x(0) = x^{(0)}, \quad y(0) = y^{(0)}, \quad t(0) = 1,\\
&\dot{t}(0) < 0.
\end{aligned} \tag{3.11}$$

3.2 Predictor-corrector procedure

The path-following of the homotopy path Γ can be implemented by some predictor-corrector procedure, i.e., uses an explicit difference scheme (e.g. Euler method) for solving (3.11) to give a predictor point and then uses a locally fast algorithm (e.g., Newton method) to give a corrector point. Some detailed discussion on the predictor-corrector algorithm with the convergence can be seen from [1, 45] and etc. Here we only present a framework of a simple Euler predictor and a Newton corrector for the path-following. For simplicity, the sign of the predictor direction is chosen to maintain an acute angle with the former predictor tangent vector. The steplength for determining the predictor point is selected heuristically based on the former Newton iteration number.

When t is becoming small (we choose $\epsilon_c = 1e-3$ as a threshold here), we change t only in the predictor step with a small steplength h_ϵ and take the corrector step fixed at the hyperplane t equals to the predicted value. Such a disposal is beneficial to reduce the potential ill-conditioning coefficient matrix of the Newton iteration during the Newton corrector step. Detailed description of the algorithm is formulated as follows.

Algorithm 3.1 (Aggregate Deformation Homotopy Method for CM^{3}P)

Parameters. $\theta > 0$, initial steplength h_0, maximum steplength h_m, minimum steplength h_ϵ, stop tolerance $\varepsilon_1 > 0$ for procedure terminated, stop tolerance $\varepsilon_N > 0$ for Newton corrector stopped, stop tolerance $\varepsilon_c > 0$ for judging corrector plane, counter $N_i = 0$.

Data. $(x^{(0)}, y^{(0)}, t_0) \in \Omega^0(1) \times \mathbf{R}^2_{++} \times \{1\}$, $\omega^{(0)} = (x^{(0)}, y^{(0)})$.

Step 0. $h = h_0$, $d_0 = [0; -1]$, $k = 0$;

Step 1. Compute a predictor point $(\bar{\omega}^{(k)}, \bar{t}_k)$

(1.1) Solve a linear equation

$$\begin{bmatrix} DH(\omega^{(k)}, t_k) \\ d_0^T \end{bmatrix} d_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

to obtain a unit tangent vector $d_1 = \frac{d_1}{\|d_1\|}$;

(1.2) $h = \min(h_m, h)$; $(\bar{\omega}^{(k)}, \bar{t}_k) = (\omega^{(k)}, t_k) + hd_1$;

(1.3) If $\bar{t}_k < 0$ or $\bar{t}_k > 1$, $h = \frac{h}{2}$, return to (1.2); else, go to Step 2;

Step 2. Compute a corrector point $(\omega^{(k+1)}, t_{k+1})$

(2.1) If $\bar{t}_k < \epsilon_c$, take $d = [0; 1]$ and $h = \min\{h, h_\epsilon\}$; else, take $d = d_1$. Go to (2.2);

(2.2) Solve equation

$$\begin{pmatrix} H(\omega, t) \\ d^T[\omega - \bar{\omega}^k; t - \bar{t}_k] \end{pmatrix} = 0$$

by Newton method with the stopping criteria ϵ_N and go to (2.3);

(2.3) If Newton corrector fail, $h = 0.7\ h$, go to (1.2); else, denote the solution as (ω^{k+1}, t_{k+1}), go to (2.4);

(2.4) If $t_{k+1} > 1$ or $t_{k+1} < 0$, $h = \frac{h}{2}$, go to (1.2); else , go to (2.5);

(2.5) If $t_k < \epsilon_1$, stop; else $d_0 = d$, $k = k + 1$, go to (2.6);

(2.6) $N_i = N_i + 1$. If $N_i > 3$ and the iteration number of Newton corrector is less than 3, go to (2.7); else , go to (1.1);

(2.7) $N_i = 0$, $h = 1.5h$, go to (1.1).

4 Numerical results

We now present some numerical comparisons to illustrate the efficiency of our method. The computations are performed on a computer running the software Matlab 6.0 on Microsoft Windows XP Professional with Pentium 4 2.00 GHz processor and 512 megabytes of memory.

We focus on the comparisons of some existing smoothing methods for the CM^{3}P. We make comparisons of our method (ADH) with the methods by Polak in [17] (Polak^1) and [35] (Polak^2) for unconstrained min-max-min problem and the aggregate homotopy method in [27] and [28] (AH). For the subproblem arised in Polak^1, we solve it with the Matlab function *fmincon* in optimization toolbox. If the objective function is nonsmooth for Polak^1, an extra variable is introduced to transform the original problem into a CM^{3}P with a smooth objective function and a max-min constraint. As to the *quadratic* subproblem in Polak^2, we solve it with the Matlab function *quadprog* in the optimization toolbox. For the aggregate homotopy method in [27] and [28], the same path-following strategy as our method is taken. If the initial point isn't in the interior of the feasible set, a two-phase strategy similar with the primal-dual interior point method for linear programming is taken.

To avoid the overflow, we use some alternative expressions of the twice aggregate function during the computation. For example, we take the following equivalent formulation in place of the twice aggregate function for smoothing the objective function,

$$f(x,t) = f(x) + t\ln\left(\sum_{i=1}^{m}\exp\left(\frac{f_i(x)-f(x)}{t}\right)\left(\sum_{1\le j\le l_i}\exp\left(\frac{f_i(x)-f_{ij}(x)}{t}\right)\right)^{-1}\right). \tag{4.1}$$

Six examples are given. The first three are simple. The fourth is a variant of the semi-infinite unconstrained min-max-min example from [35], we introduce an extra variable to convert it into a constrained example with changeable function number. The fifth one is constructed with changeable dimension. The last example is constructed that both the dimension and the function number are changeable.

During the computation, we take $\theta = 0.1$, $h_0 = 0.1$, $h_m = 1$, $h_\epsilon = 1e-3$, $\varepsilon_1 = 1e-6$, $\varepsilon_N = 1e-3$, $\varepsilon_c = 1e-3$. The results listed in the following tables, where x^* denotes the final solution point, f^* is the optimal value, g^* is the constraint value at x^*.

Example 1 Let $n=2$,

$$\begin{aligned}
f(x) &= x_1^2 + (x_2-2)^2,\\
g_{11}(x) &= x_1^2 + x_2^2 - 4,\\
g_{21}(x) &= -x_1 - x_2 - 2,\\
g_{22}(x) &= x_1.
\end{aligned}$$

Example 2 Let $n=2$, $a=[1;1]$, $b=0.5$,

$$\begin{aligned}
f(x) &= 100(x_2-x_1^2)^2 + (1-x_1)^2,\\
g_{11}(x) &= 1.5b + \frac{(x_1-a_1)^2}{2b} - (x_2-a_2),\\
g_{12}(x) &= -0.5b - \frac{(x_1-a_1)^2}{2b} - (x_2-a_2),\\
g_{13}(x) &= 1.5b - \frac{(x_2-a_2-b)^2}{2b} - (x_1-a_1),\\
g_{14}(x) &= 1.5b - \frac{(x_2-a_2-b)^2}{2b} + (x_1-a_1),\\
g_{21}(x) &= b + (x_2-a_2) + \frac{(x_1-a_1)^2}{b},
\end{aligned}$$

Table 1 Comparison results of Examples 1, 2 and 3

Ex.	$x^{(0)}$	Method	x^*	f^*	g^*	Time (sec.)
1	(0, 0)	ADH	(0.0000, 1.9988)	0.0000	−0.0049	0.4310
		AH	(−0.0000, 1.9999)	0.0000	−0.0028	**0.2600**
		Polak[1]	(0.0000, 1.8708)	0.0167	−0.5001	0.4840
2	(2, 2)	ADH	(1.1999, 1.4402)	0.0400	−0.0000	**0.4120**
		AH	(1.1999, 1.4402)	0.0400	−0.0000	0.4380
		Polak[1]	(1.1999, 1.4402)	0.0400	−0.0000	0.6250
3	(0.9, 0.9)	ADH	(0.7500, 0.0000)	0.0625	−1.7500	**0.2340**
		AH	(0.7500, 0.0000)	0.0625	−1.7500	0.3750
		Polak[1]	(0.7500, 0.0000)	0.0625	−1.7500	0.4530

$\beta_1 = 1$, $\beta_2 = 1$, $\rho = \mathrm{abs}(g^0) + 1$, $p = x^{0T} x^0 + 10$

$$g_{22}(x) = 2b - (x_2 - a_2) - 2(x_1 - a_1) - \frac{2}{b}(x_1 - a_1)^2,$$
$$g_{23}(x) = 2b - (x_2 - a_2) + 2(x_1 - a_1) - \frac{2}{b}(x_1 - a_1)^2.$$

Example 3 Let $n = 2$,

$$\begin{aligned}
f(x) &= \max\{(x_1 - 1)^2 + x_2^2, (x_1 - 0.5)^2 + x_2^2\},\\
g_{11}(x) &= 1.5 + 0.5x_1^2 - x_2,\\
g_{12}(x) &= 0.5 + 0.5x_1^2 - x_2,\\
g_{13}(x) &= -1.5 + (x_2 - 1)^2/2 - x_1,\\
g_{14}(x) &= -1.5 + (x_2 - 1)^2/2 + x_1,\\
g_{21}(x) &= -5 + x_2 + x_1^2,\\
g_{22}(x) &= 2 - x_2 + 2x_1 + 2x_1^2,\\
g_{23}(x) &= -4 + x_2 - 2x_1 + 2x_1^2.
\end{aligned}$$

Example 4 Let $n = 4$, $m = 50, 100, 500, 1000, 5000$,

$$\begin{aligned}
f(x) &= x_4,\\
g_{i1}(x) &= \sum_{k=1}^{3}\left(x_k + 1 + \frac{i}{m}\right)^2 - \frac{i}{m} - x_4,\\
g_{i2}(x) &= \sum_{k=1}^{3}\left(x_k - 1 + \frac{i}{m}\right)^2 - \frac{i}{m} - x_4, \quad i = 1, \dots, m.
\end{aligned}$$

Table 2 Comparison results of Example 4 with $x^{(0)} = (1, 1, 1, 2.1)$

m	Method	x^*	f^*	g^*	Time (sec.)
50	ADH	(0.6567, 0.6567, 0.6567, 0.2936)	0.2936	−1.9906e-7	**1.2510**
	AH	(0.6567, 0.6567, 0.6567, 0.2936)	0.2936	−6.1784e-8	1.5310
	Polak[1]		fail		
	Polak[2]	(0.6567, 0.6567, 0.6567)	0.2936		2.1880
100	ADH	(0.6617, 0.6617, 0.6617, 0.3134)	0.3134	−3.2656e-7	3.6070
	AH	(0.6617, 0.6617, 0.6617, 0.3134)	0.3134	−2.4045e-8	**3.2660**
	Polak[1]		fail		
	Polak[2]	(0.6617, 0.6617, 0.6617)	0.3134		9.0310
500	ADH	(0.6657, 0.6657, 0.6657, 0.3293)	0.3293	−1.6942e-7	**21.3440**
	AH	(0.6657, 0.6657, 0.6657, 0.3293)	0.3293	−1.1784e-8	53.5310
	Polak[1]		fail		
	Polak[2]	(0.6657, 0.6657, 0.6657)	0.3293		1706.5470
1000	ADH	(0.6662, 0.6662, 0.6662, 0.3313)	0.3313	−4.4848e-7	**45.2960**
	AH	(0.6662, 0.6662, 0.6662, 0.3313)	0.3313	−2.2504e-8	93.3120
	Polak[1]		fail		
	Polak[2]		*		
5000	ADH	(0.6666, 0.6666, 0.6666, 0.3329)	0.3329	−1.4554e-7	**269.2970**
	AH	(0.6666, 0.6666, 0.6666, 0.3329)	0.3329	−3.3042e-8	416.3210
	Polak[1]		fail		
	Polak[2]		*		

$\beta_1 = 1,\ \beta_2 = 2,\ \rho = \text{abs}(g^0) + 1,\ p = x^{0T} x^0 + 100$

*denotes the algorithm is stopped because the time is larger than 3000 seconds

Example 5 $n = 10, 50, 100, 500, 800,$

$$f(x) = \max\left\{\sum_{i=1}^{n}(2 - x_i)^2, 2\exp\left(x_n - \sum_{i=1}^{n-1} x_i\right)\right\},$$

$$g_{11}(x) = x^T x - 14000,$$

$$g_{21}(x) = -\sum_{i=2}^{n}(x_i - 10)^2 - 40(x_1 + 4)^2 + 500,$$

$$g_{31}(x) = x_1 + 10 - \sum_{i=2}^{n} x_i^2,$$

$$g_{32}(x) = -x_1 + \sum_{i=2}^{n} x_i^2 - 5,$$

Table 3 Comparison results of Example 5 with $x^{(0)} = (0.01n, 0.01n, \dots, 0.01n)$

n	Method	x^*	f^*	g^*	Time (sec.)
10	ADH	(2.5890, 0.9183, ..., 0.9183)	10.8782	−1.0550e-6	**0.3910**
	AH	(2.5890, 0.9183, ..., 0.9183)	10.8782	−9.7469e-7	0.4220
	Polak[1]	(2.5890, 0.9183, ..., 0.9183)	10.8782	3.2237e-11	1.4070
50	ADH	(2.0000, 2.0000, ..., 2.0000)	7.4128e-12	−183.9999	**0.6970**
	AH	(2.0000, 2.0000, ..., 2.0000)	3.2390e-11	−183.9998	0.7190
	Polak[1]	(0.8944, 0.8944, ..., 0.8944)	61.1146	1.1017e-9	0.7500
100	ADH	(2.0000, 2.0000, ..., 2.0000)	3.5828e-12	−383.9999	1.4170
	AH	(2.0000, 2.0000, ..., 2.0000)	2.1912e-12	−384.0000	**1.3190**
	Polak[1]	(0.6325, 0.6325, ..., 0.6325)	187.0178	9.3765e-9	1.4590
500	ADH	(2.0000, 2.0000, ..., 2.0000)	9.0358e-9	−1.9840e+3	**111.3520**
	AH	(2.0000, 2.0000, ..., 2.0000)	1.0054e-8	−1.9840e+3	114.4060
	Polak[1]			fail	
800	ADH	(2.0000, 2.0000, ..., 2.0000)	4.2708e-8	−3.1840e+3	1.3723e+3
	AH			fail	
	Polak[1]			fail	

$\beta_1 = 1$, $\beta_2 = 2.5$, $\rho = \text{abs}(g^0) + 50$, $p = x^{0T} x^0 + 14000$

$$g_{41}(x) = 10 - x_1 - \sum_{i=2}^{n} (x_i + 14)^2,$$

$$g_{51}(x) = 100 - (x_1 - 2)^2 - \sum_{i=2}^{n} (x_i + 20)^2.$$

Example 6 Let $x = (T, u, w) \in \mathbf{R}^{1+N+N}$, $N = 40, 60, 80, 100$. The example is defined as

$$\begin{aligned} f(x) &= T^2 + C_1 \sum_{i=1}^{N} u_i^2 + C_2 \sum_{i=1}^{N} w_i^2, \\ g_{i1}(x) &= x1(i+1) - \alpha, \\ g_{i2}(x) &= \beta - x2(i+1), \\ g_{N+i,1}(x) &= -x1(i+1) - \alpha, \\ g_{N+i,2}(x) &= \beta - x2(i+1), \quad i = 1, \dots, N, \end{aligned}$$

where $C_1 = 0.1$, $C_2 = 0.5$, $\alpha = 5$, $\beta = 0.3$,

$$\begin{aligned} x1(1) &= 0.5, \\ x2(1) &= 0.5, \end{aligned}$$

Table 4 Comparison results of Example 6 with $x^{(0)} = (2.5, 2.5, \ldots, 2.5)$

N	n	m	Method	f^*	g^*	Time (sec.)
40	81	80	ADH	8.7930e-9	−1.8223e-6	**29.1880**
			AH	3.2402e-11	−1.7640e-7	75.1720
			Polak[1]	3.1014e-8	3.4849e-20	36.2030
60	121	120	ADH	2.0642e-8	−1.8590e-6	**1.1056e+2**
			AH	1.4292e-12	−3.0253e-8	1.5375e+2
			Polak[1]	1.6795e-12	4.3300e-23	2.0166e+2
80	161	160	ADH	7.4599e-8	−2.5927e-6	**2.5473e+2**
			AH	1.4735e-9	−8.4129e-7	3.7611e+2
			Polak[1]	5.8976e-12	−5.4070e-19	3.4604e+2
100	201	200	ADCH	9.9949e-9	−8.2071e-7	**6.3173e+2**
			AH	8.4396e-9	−1.8008e-6	7.1610e+2
			Polak[1]	3.5781e-8	−4.7971e-20	7.6942e+2

$\beta_1 = 1$, $\beta_2 = 0.1$, $\rho = \mathrm{abs}(g^0) + 0.1$, $p = x^{0T} x^0 + 100$

$$
\begin{aligned}
&\vartheta(1) = 0,\\
&\theta(1) = 0,\\
&x1(i+1) = x1(i) + \frac{T}{N}\vartheta(i)\cos(\theta(i)),\\
&x2(i+1) = x2(i) + \frac{T}{N}\vartheta(i)\sin(\theta(i)),\\
&\vartheta(i+1) = \vartheta(i) + \frac{T}{N}u_i,\\
&\theta(i+1) = \theta(i) + \frac{T}{N}w_i, \quad i = 1, \ldots, N.
\end{aligned}
$$

From the results listed in the tables, we find that the aggregate deformation homotopy method (ADH) is comparable to the other methods. The aggregate homotopy method (AH) is sensitive to the non-interior initial point. For example, during the computation of Example 4 by the aggregate homotpy method, as $m \geq 500$, the initial point isn't in the interior of the aggregate feasible set, a phase-two strategy must be taken that leads to a much longer time. For Example 5, as $n = 800$, the initial point also isn't in the interior of the feasible set. However, the phase-two strategy fails to generate an interior initial point for this example. The method Polak[1] maybe leads to a non-optimal solution such as Example 1 or fails to find a feasible solution (see Table 2 and 3). The reason is that some extra variables must be introduced to the problem that changes the original feasible set and maybe leads to an infeasible problem. In Example 4, we compare the method Polak[2] that is presented in [35] for unconstrained min-max-min problem. We find that it is sensitive to the min-function number m. If m is large, it acts very slowly that is because the method must solve a

Table 5 The results of Example 3 for different parameters value with $x^{(0)} = (0.9; 0.9)$

	Para value $(\beta_1, \beta_2, \rho, p)$	x^*	f^*	N_1	N_2	Time
β_1 changeable	$(1, 1, 10 + a, 100 + b)$	(0.7500, 0.0000)	0.0625	22	52	0.1788
	$(2, 1, 10 + a, 100 + b)$	(0.7500, 0.0000)	0.0625	23	56	0.1882
	$(3, 1, 10 + a, 100 + b)$	(0.7500, 0.0000)	0.0625	21	49	0.1677
	$(0.5, 1, 10 + a, 100 + b)$	(0.7500, 0.0000)	0.0625	23	59	0.1900
	$(0.1, 1, 10 + a, 100 + b)$	(0.7500, 0.0000)	0.0625	25	64	0.1971
β_2 changeable	$(1, 1, 10 + a, 100 + b)$	(0.7500, 0.0000)	0.0625	22	52	0.1788
	$(1, 2, 10 + a, 100 + b)$	(0.7500, 0.0000)	0.0625	22	65	0.1899
	$(1, 3, 10 + a, 100 + b)$	(0.7500, 0.0000)	0.0625	**35**	**107**	**0.2735**
	$(1, 0.5, 10 + a, 100 + b)$	(0.7500, 0.0000)	0.0625	19	46	0.1622
	$(1, 0.1, 10 + a, 100 + b)$	(0.7500, 0.0000)	0.0625	22	47	0.1833
ρ changeable	$(1, 1, 10 + a, 100 + b)$	(0.7500, 0.0000)	0.0625	22	52	0.1788
	$(1, 1, 1 + a, 100 + b)$	(0.7500, 0.0000)	0.0625	21	51	0.1735
	$(1, 1, 100 + a, 100 + b)$	(0.7500, 0.0000)	0.0625	**31**	**98**	**0.2688**
p changeable	$(1, 1, 10 + a, 100 + b)$	(0.7500, 0.0000)	0.0625	22	52	0.1788
	$(1, 1, 10 + a, 10 + b)$	(0.7500, 0.0000)	0.0625	22	67	0.1834
	$(1, 1, 10 + a, 500 + b)$	(0.7500, 0.0000)	0.0625	19	54	0.1754

$a = |g(x^{(0)})|$, $b = x^{(0)^T} x^{(0)}$, N_1-predictor step number, N_2-Newton corrector step number

quadratic subproblem with dimension m in every iteration. While m is big, the time on the quadratic subproblem is large.

In our algorithm, some deformation parameters are introduced to construct deformation. They are chosen manually. To show the influence of the value of them on the efficiency of our algorithm, a test on Example 3 for different parameter values is given, the test result is listed in the following Table 5.

From the result in Table 5, we find that the performance of our algorithm moderately depends on the parameters value. Among the parameters, the value of parameter β_2 and ρ are more important. Some more efficient parameters determination strategy may be beneficial for the improvement of the algorithm.

Remark 1 Homotopy methods usually focus on obtaining a stationary solution point when used for solving optimization problems, so does the aggregate deformation homotopy method in this paper. A question will arise naturally: can the homotopy methods approach to a minimum? How to obtain this?

For the smooth problems, the problem obtaining a local minimum by homotopy methods were solved. For the smooth unconstrained problem $\min\{f(x) : x \in \mathbf{R}^n\}$, it was proven that the linear homotopy equation $H(x, t) = (1 - t)\nabla f(x) + t(x - x^0)$ can determine a smooth path approaching to a non-maximum stationary point x^* only if $\nabla^2 f(x^*)$ is nonsingular. For the smooth constrained optimization problems

$\min\{f(x) : g_i(x) \leq 0, i = 1, \ldots, m\}$, the smooth homotopy path, determined by the combined homotopy equation

$$H(x, y, t) = \begin{pmatrix} (1-t)(\nabla f(x) + \nabla g(x)y) + t(x - x^0) \\ Yg(x) - tY^0 g(x^0) \end{pmatrix} = 0, \tag{4.2}$$

can also approach to a non-maximum stationary point (x^*, y^*) only if $\frac{\partial H}{\partial(x,y)}(x^*, y^*, 0)$ is nonsingular. During the computation, an extra judging step can be included into the original path-following procedure, that judges whether x^* is a saddle point by computing the sequential principal minor determinant of $\nabla^2 f(x^*)$. If x^* is a saddle point, then choose a new initial point and repeat the path-following procedure until a local minimum solution can be obtained. More details of them can be seen from [42].

However, for the CM^{3}P in this paper, it's much more complex to determine a minimum by the path-following procedure because of its nonsmoothness. Though all the computed solution for the given examples are local minima up to now, we cannot prove theoretically the determined homotopy path approaches to a minimum. We can only prove now under some assumptions, the homotopy path can determine a direction d, for the feasible point $x(t) = x^* + td$ along the direction, it has $f(x(t)) \geq f(x^*)$, as a result, the obtained generalized KKT solution point x^* is not a maximum. The detailed proof is given as follows,

Proposition 4.2 *For the given aggregate deformation homtopy equation* (1.4), *denote* $\omega = (x, y_1, y_2)$, *denote* $\omega^* = (x^*, y_1^*, 0)$ *as the obtained solution by path-following the homotopy equation. If* $\frac{\partial H}{\partial \omega}(\omega^*, 0)$ *is nonsingular, then it must exist a neighborhood* N_ϵ *of* $t^* = 0$ *and a direction* $\bar{d} \in \mathbf{R}^{n+2}$, *such that in the neighborhood, it has* $\omega(t) = \omega^* + t\bar{d}$ *for any* $(\omega(t), t) \in H^{-1}(0)$. *Take* $d = \bar{d}(1:n)$, *then* $x(t) = x^* + td$. *If defining* $G(x) = \max\{(\alpha_1(f(x) - f(x^*)) + \alpha_2 g(x) : \alpha \geq 0, \alpha_1 + \alpha_2 = 1\}$, *then we have* $G(x^*) = 0$ *and* $G'(x^*; d) \geq 0$. *When* $G'(x^*; d) > 0$, *or* $G'(x^*; d) = 0$ *and* d *satisfying*

$$d^T \Bigg(\sum_{i \in I(x^*)} \delta_i^* \sum_{j \in J_i(x^*)} \eta_{ij}^* \nabla^2 f_{ij}(x^*) + y_1^* \sum_{i \in II(x^*)} \lambda_i^* \sum_{j \in JJ_i(x^*)} \mu_{ij}^* \nabla^2 g_{ij}(x^*) \Bigg) d > 0, \tag{4.3}$$

x^* *isn't a maximum solution.*

Proof If $\frac{\partial H}{\partial \omega}(\omega^*, 0)$ is nonsingular, from the implicit function theorem, there exists a neighborhood $N_{\epsilon'}$ of $t = 0$ such that in the neighborhood, for any $(\omega(t), t) \in H^{-1}(0)$, the gradient of $\omega(t)$ is given by ($t \in N_{\epsilon'}$),

$$\nabla \omega(t) = -[\nabla_\omega H(\omega(t), t)]^{-1} \nabla_t H(\omega(t), t).$$

Denote $d(t) = -\frac{[\nabla_\omega H(\omega(t),t)]^{-1} \nabla_t H(\omega(t),t)}{\|[\nabla_\omega H(\omega(t),t)]^{-1} \nabla_t H(\omega(t),t)\|}$, $d(t)$ is a unit vector in the neighborhood $N_{\epsilon'}$, it must have a limit point $\bar{d}$. As a result, there exists a neighborhood $N_\epsilon \subset N_{\epsilon'}$ such that $\omega(t) = \omega^* + t\bar{d}$, $t \in N_\epsilon$. Take $d = \bar{d}(1:n)$, then $x(t) = x^* + td$.

Define

$$G(x) = \max\left\{\alpha_1(f(x) - f(x^*)) + \alpha_2 g(x) : \alpha \geq 0, \sum_{i=1}^{2} \alpha_i = 1\right\},$$

it obviously has $G(x^*) = 0$. Moreover, from the discussions in [33], we know f and g are semismooth functions, as a result $G(x)$ is a semismooth function (Lemma 2, [6]) and for any given x and u, $G'(x; u)$ exists. In the following, we prove $G'(x^*; d) \geq 0$ for the direction d.

From (x^*, y_1^*) is a generalized KKT solution, we know for any direction u such that $g'(x^*; u) \leq 0$, it has $f'(x^*; u) \geq 0$. From the construction of the homotopy equation (1.4), we know for the direction d, it has $g(x(t)) = g(x^* + td) < 0$ and $\lim_{t\downarrow 0} g(x(t)) \leq 0$. As a result, $G(x(t)) = f(x(t)) - f(x^*)$ and $G'(x^*; d) = f'(x^*; d) \geq 0$.

If $G'(x^*; d) > 0$, then $G(x(t)) = f(x(t)) - f(x^*) > 0$, where $x(t) = x^* + td$, $t \in N_\epsilon$, that is to say $f(x(t)) > f(x^*)$, as a result, x^* cannot be a maximum solution.

If $G'(x^*; d) = 0$, from $G'(x^*; \cdot)$ is upper semicontinuous with respect to direction $u \in \mathbf{R}^n$, for arbitrary sequence $x_k \to x^*$, $d_k \to d$ and $v_k \in \partial G(x_k)$, $v_k \to 0$, it has $v_k^T d \leq 0$. Similar with the description of theorem 2 in [6], if x^* is a local maximum solution of (1.1), we have

$$L''_+(x^*, y_1^*, 0, d) \leq 0, \tag{4.4}$$

where $L''_+(x^*, y_1^*, 0, d)$ is the generalized second-order directional derivative of the Lagrange function $L(x, y) = f(x) + yg(x)$ in the sense of Chaney [6], that is defined to be the supremum of all numbers

$$\limsup[L(x_k, y_1^*) - L(x^*, y_1^*) - v^T(x_k - x^*)]/t_k^2,$$

taken over all triples of sequence $\{x_k\}$, $\{v_k\}$ and $\{t_k\}$ such that

(a) $t_k \downarrow 0$, and $\{x_k\}$ converges to x^*, $\{(x_k - x^*)/t_k\}$ converges to d;
(b) $\{v_k\}$ converges to v and $v_k \in \partial_x L(x_k, y_1^*)$.

From the computation formula of the Chaney's generalized second-order directional derivative in [32], we can compute the Chaney's generalized second-order directional derivative of the Lagrange function $L(x, y)$ of (1.1) as follows,

$$L''_+(x^*, y_1^*, 0, d) = \max\left\{\sum_{i\in I(x^*)} \sum_{j\in J_i(x^*)} a_{1ij}\nabla^2 f_{ij}(x^*) \right.$$
$$\left. + y_1^* \sum_{i\in II(x^*} \sum_{j\in JJ_i(x^*)} a_{2ij}\nabla^2 g_{ij}(x^*) : a \in T_d(L, x^*, y_1^*, 0)\right\}$$

where

$$
\begin{aligned}
&T_d(L, x^*, y_1^*, 0)\\
&\quad= \Bigg\{ a = (a_1, a_2) : a_i \in R^{S_i}, i = 1, 2, \text{ and there exist } \{x^l\}, \{a^l\}, \{v^l\} \text{ such that}\\
&\qquad (1)\ x^l \to^d x^*, v^l \to 0, v^l \in \partial_x L(x^l, y_1^*);\\
&\qquad (2)\ a^l \to a, \text{ for each } i = 1, 2, a_i^l \in E_{S_i} \text{ and}\\
&\qquad v^l = \sum_{i \in K_1(x^l)} \sum_{j \in K_{1i}(x^l)} a_{1ij}^l \nabla f_{ij}(x^l) + y_1^* \sum_{i \in K_2(x^l)} \sum_{j \in K_{2i}(x^l)} a_{2ij}^l \nabla g_{ij}(x^l) \Bigg\},\\
&\quad a_{ijk}^l = 0, \text{ if } j \notin K_i(x^l) \text{ or } k \notin K_{ij}(x^l),
\end{aligned}
\tag{4.5}
$$

$K_1 = m$, $K_2 = k$, $K_{1i} = l_i$, $K_{2i} = p_i$, $S_i = \sum_{j \in K_i} K_{ij}$, $E_{S_i} = \{a_i \in \mathbf{R}^{S_i} : a_i \geq 0, \sum_{j=1}^{K_i} \sum_{s=1}^{K_{ij}} a_{ijs} = 1\}$, $K_1(x^l) = I(x^l)$, $K_2(x^l) = II(x^l)$, $K_{1i}(x^l) = J_i(x^l)$, $K_{2i}(x^l) = JJ_i(x^l)$. Denote $s_1^* = (\delta_1^* \eta_{11}^*, \ldots, \delta_1^* \eta_{1l_1}^*, \ldots, \delta_m^* \eta_{ml_m}^*)$, $s_2 = (\lambda_1^* \mu_{11}^*, \ldots, \lambda_1^* \mu_{1p_1}^*, \ldots, \lambda_k^* \mu_{kp_k}^*)$, then $s = (s_1, s_2) \in T_d(L, x^*, y_1^*, 0)$. If d satisfies (4.3), then $L''_+(x^*, y_1^*, 0, d) > 0$ that contradicts with (4.4), as a result, x^* cannot be a maximum solution. □

The conclusion of Proposition 4.2 is only a preparation work for obtaining a minimum of (1.1). To design the algorithm for obtaining a local minimum, more further work must be done in the future such as the proof that the inequality (4.3) holds naturally for the determined direction d under the nonsingular assumption, and the consideration on the non-regular circumstance, that is the circumstance $\frac{\partial H}{\partial \omega}(\omega^*, 0)$ is singular.

Remark 2 Ill-conditioning of the Hessian is an important numerical issue when using the second-order information of the aggregate function or twice aggregate function constructing smoothing algorithms. For the Hessian computation of the twice aggregate function in this paper, we have the following conclusion.

Proposition 4.3 *For a given $x \in \mathbf{R}^n$ and $t > 0$, the Hessian of $f(x, t)$ with respect to x have the following form*:

$$
\begin{aligned}
\nabla_x^2 f(x, t) = &\sum_{i=1}^m \delta_i(x, t) \sum_{j=1}^{l_i} \eta_{ij}(x, t) \nabla^2 f_{ij}(x)\\
&+ \frac{2}{t} \sum_{i=1}^m \delta_i(x, t) \left(\sum_{j=1}^{l_i} \eta_{ij}(x, t) \nabla f_{ij}(x) \right)\\
&\times \left(\sum_{j=1}^{l_i} \eta_{ij}(x, t) \nabla f_{ij}(x) \right)^T
\end{aligned}
$$

$$-\frac{1}{t}\sum_{i=1}^{m}\delta_i(x,t)\sum_{j=1}^{l_i}\eta_{ij}(x,t)\nabla f_{ij}(x)\nabla f_{ij}(x)^T$$
$$-\frac{1}{t}\nabla_x f(x,t)\nabla_x f(x,t)^T. \tag{4.6}$$

Moreover, for any given $x^* \in \mathbf{R}^n$ $(n \geq 2)$, *if there exists a neighborhood* O_{x^*} *of* x^*, *such that* $f(x)$ *is smooth in* $O_{x^*} \setminus \{x^*\}$, *then*

$$\lim_{t\downarrow 0}\nabla_x^2 f(x^*,t) = \frac{1}{|I(x^*)|}\sum_{i\in I(x^*)}\frac{1}{|J_i(x^*)|}\sum_{j\in J_i(x^*)}\nabla^2 f_{ij}(x^*). \tag{4.7}$$

Proof From the definition of $f(x,t)$, (4.6) can be directly computed. From the theorem 1 in [38], if $f(x)$ is smooth in $O_{x^*} \setminus \{x^*\}$, then $f(x)$ is smooth at O_{x^*}, and $\nabla f_{ij}(x^*) = \nabla f(x^*)$ for all $i \in I(x^*)$, $j \in J_i(x^*)$. As a result, from $\lim_{t\downarrow 0}\frac{1}{t}\exp(-\frac{a}{t}) = 0$ $(a > 0)$ and (4.6), we have

$$\lim_{t\downarrow 0}\frac{\eta_{ij}(x^*,t)}{t} = \lim_{t\downarrow 0}\frac{\exp(\frac{f_{ij}(x^*)-f_i(x^*)}{t})}{t\sum_{j\in J_i}\exp(\frac{f_{ij}(x^*)-f_i(x^*)}{t})} = 0, \quad j \notin J_i(x^*),$$

$$\lim_{t\downarrow 0}\frac{\delta_i(x^*,t)}{t} = 0, \quad i \notin I(x^*),$$

$$\lim_{t\downarrow 0}\frac{1-\sum_{j\in J_i(x^*)}\eta_{ij}(x^*,t)}{t} = \lim_{t\downarrow 0}\frac{\sum_{j\notin J_i(x^*)}\exp(\frac{f_{ij}(x^*)-f_i(x^*)}{t})}{t\sum_{j\in J_i}\exp(\frac{f_{ij}(x^*)-f_i(x^*)}{t})} = 0,$$

$$\lim_{t\downarrow 0}\frac{1-\sum_{i\in I(x^*)}\delta_i(x^*,t)}{t} = 0,$$

$$\begin{aligned}
&\lim_{t\downarrow 0}\Bigg(\frac{2}{t}\sum_{i\in I(x^*)}\delta_i(x^*,t)\Bigg(\sum_{j\in J_i(x^*)}\eta_{ij}(x^*,t)\nabla f_{ij}(x^*)\Bigg)\Bigg(\sum_{j\in J_i(x^*)}\eta_{ij}(x^*,t)\nabla f_{ij}(x^*)\Bigg)^T\\
&\quad -\frac{1}{t}\sum_{i\in I(x^*)}\delta_i(x^*,t)\sum_{j\in J_i(x^*)}\eta_{ij}(x^*,t)\nabla f_{ij}(x^*)\nabla f_{ij}(x^*)^T\\
&\quad -\frac{1}{t}\nabla_x f(x^*,t)\nabla_x f(x^*,t)^T\Bigg)\\
&= \lim_{t\downarrow 0}\Bigg(\frac{1}{t}\sum_{i\in I(x^*)}\delta_i(x^*,t)\Bigg(\sum_{j\in J_i(x^*)}\eta_{ij}(x^*)\Bigg(\sum_{j\in J_i(x^*)}\eta_{ij}(x^*)-1\Bigg)\Bigg)\\
&\quad +\frac{1}{t}\sum_{i\in I(x^*)}\delta_i(x^*,t)\Bigg(1-\sum_{i\in I(x^*)}\delta_i(x^*,t)\Bigg)\Bigg)\nabla f(x^*)\nabla f(x^*)^T = 0,
\end{aligned}$$

as a result, $\lim_{t\downarrow 0}\nabla_x^2 f(x^*,t) = \frac{1}{|I(x^*)|}\sum_{i\in I(x^*)}\frac{1}{|J_i(x^*)|}\sum_{j\in J_i(x^*)}\nabla^2 f_{ij}(x^*)$. □

From above Proposition 4.3, only in some special case, it can be proven that $\nabla_x^2 f(x,t)$ keeps finite as $t \to 0$, otherwise, it maybe approaches to infinite. Such a phenomenon was also considered by many other researchers. For example, in [43], the author discussed the potential ill-conditioning of the aggregate function $f(x,p) = p\ln(\sum_{i=1}^m \exp(\frac{1}{p} f_i(x)))$ for smoothing the max-type function $f(x) = \max_{1\le i\le m}\{f_i(x)\}$. He illustrated that the speed of the Hessian going to infinite depended not only on the parameter p but also on $\|\nabla f_i(x) - \nabla f_j(x)\|$, where $i, j \in I(x) = \{i \in \{1,\dots,m\} : f_i(x) = f(x)\}$.

Some other beneficial strategies for coping with the ill-conditioning of the aggregate function also exist. For example, in [36], Polak, E. et al. presented some feedback precision-adjustment rule to control the parameter. With this strategy, though the ill-conditioning still cannot be avoided, it can be reduced. From another point of view, by giving an improvement of the aggregate function, Yang, Q.Z. presented an adjustable aggregate function to avoid the ill-conditioning in [49]. Respectively, a similar cross entropy function was presented by X.S. Li et al. [25]. With the new technique, the smoothing functions can approximate to the original max-function when the control parameter t is larger than zero. However, as it was pointed out in [25], the new approximate functions are unstable in numerical computation. They just provided a new idea for coping with the ill-conditioning of the aggregate function.

In this paper, we take an adjustable corrector plane strategy in the algorithm, that is, as t is becoming small, we change t only in the predictor step and take the Newton corrector step fixed at a plane where t equals to the predicted value. Though the strategy is elementary, it's beneficial to avoid the possible ill-conditioning coefficient matrix during the Newton corrector and make t reduce to zero not too fast and as a result, is good at reducing the ill-conditioning.

Besides these, some other work on ill-conditioning arising from the interior-point methods by M.H. Wight et al. [11, 12, 46] are also beneficial for reference. It's one of our on-going work based on these existing work to make more further consideration on the potential ill-conditioning of the twice aggregate function.

5 Conclusion and future work

In this paper, an aggregate deformation homotopy method for the CM^{3}P is presented. The convergence analysis and numerical results show that the new method is feasible for the CM^{3}P. Compared to other existing methods, the new method is a non-interior-point path-following method. Though some deformation parameters need to be adjusted manually, the algorithm is comparatively stable to moderate changes of the parameters' value. Moreover, it has many benefits. For example, it doesn't require a restrictive weak normal cone condition and is feasible for almost all initial point, it's unnecessary to introduce any auxiliary variables or solve any extra subproblems with higher dimension. As a result, it's competitive to the existing methods. Our future work will include more effective twice aggregate function evaluation with the gradient and Hessian, the efficient linear equation solver for the predictor direction and Newton corrector, the applications to some problems arising in the field of data mining, the consideration on the degenerate or non-regular circumstance and some more efficient strategy on the ill-conditioning as well.

Acknowledgements The authors like to express our thanks to anonymous reviewers, whose precise and substantial remarks led to an improved version of the paper.

References

1. Allgower, E.L., Georg, K.: In: Numerical Continuation Methods: An Introduction, pp. 9–23. Springer, Berlin/New York (1990)
2. André, L.T., Andreas, W., Sasan, B., et al.: A primal-dual interior-point method for nonlinear programming with strong global and local convergence properties. SIAM J. Optim. **14**, 173–199 (2003)
3. Astorino, A., Gaudioso, M.: Polyhedral separability through successive LP. J. Optim. **112**, 265–293 (2002)
4. Birbil, S., Fang, S.C., Han, J.: Entropic regularization approach for mathematical programs with equilibrium constraints. Technical Report, Industrial Engineering and Operations Research, Carolina, USA (2002)
5. Bofill, W.G.: Properties of an interior embedding for solving nonlinear optimization problems. Math. Program. **86**, 649–659 (1999)
6. Chaney, R.W.: Second-order necessary conditions in constrained semismooth optimization. SIAM J. Control Optim. **25**, 1072–1081 (1987)
7. Demyanov, V.F., Vasilief, L.V.: Nondifferentiable Optimization. Nauka, Moscow (1981). (In Russian)
8. Fang, S.C., Han, J., Huang, Z.H., Birbil, S.: On the finite termination of an entropy function based smoothing newton method for vertical linear complementarity problems. Technical Report, Industrial Engineering and Operations Research, Carolina, USA (2002)
9. Feng, G.C., Yu, B.: Combined homotopy interior point method for nonlinear programming problems. In: Fujita, H., Yamaguti, M. (eds.) Advances Notes in Numerical and Applied Analysis. Proceedings of the Second Japan-China Seminar on Numerical Mathematics, Tokyo. Lecture Notes in Numerical and Applied Analysis, vol. 14, pp. 9–16. Springer, Berlin (1995)
10. Feng, G.C., Lin, Z.H., Yu, B.: Existence of an interior pathway to a Kraus-Kuhn-Tucker point of a nonlinear programming problem. Nonlinear Anal. **32**, 761–768 (1998)
11. Forsgren, A., Gill, P.E., Wright, M.H.: Interior methods for nonlinear optimization. SIAM Rev. **44**, 525–597 (2002)
12. Forsgren, A., Gill, P.E., Griffin, J.D.: Iterative solution of augmented systems arising in interior method. SIAM J. Optim. **18**, 666–690 (2007)
13. Guddat, J., Vazquez, F.G., Nowack, D., Rückmann, J.-J.: A modified standard embedding with jumps in nonlinear optimization. Eur. J. Oper. Res. **169**, 1185–1206 (2006)
14. Guddat, J., Vazquez, F.G., Nowack, D.: On the role of the Mangasarian-Fromovitz constraint qualification for penalty, exact penalty and Lagrange multiplier methods. In: Fiacco, A.V. (ed.) Mathematical Programming with Data Perturbation, pp. 159–183. Marcel Dekker, New York (1997)
15. Jongen, H.T., Jonker, P., Twilt, F.: On deformation in optimization. Methods Oper. Res. **37**, 171–184 (1980)
16. Karmarkar, N.: A new polynomial-time algorithm for linear programming. Combinatorica **4**, 373–395 (1984)
17. Kirjner-Neto, C., Polak, E.: On the conversion of optimization problems with max-min constraints to standard optimization problems. SIAM J. Optim. **8**, 887–915 (1998)
18. Kojima, M., Hirabayashi, R.: Continuous deformations of nonlinear programs. Math. Program. Study **21**, 150–198 (1984)
19. Kort, B.W., Bertsekas, D.P.: A new penalty function algorithm for constrained minimization. In: Proceedings of the 1972 IEEE Conference on Decision and Control, New Orleans, Louisiana (1972)
20. Li, S.J., Yang, X.Q., Teo, K.L.: On the conversion of optimization problems with max-min constraints to standard optimization problems. J. Optim. **109**, 691–698 (2001)
21. Li, X.S.: An aggregate function method for nonlinear programming. Sci. China (Ser. A) **12**, 1283–1288 (1991)
22. Li, X.S.: An aggregate constraint method for nonlinear programming. J. Oper. Res. Soc. **42**, 1003–1010 (1991)
23. Li, X.S.: An entropy-based aggregate method for minimax optimization. Eng. Optim. **18**, 227–285 (1992)
24. Li, X.S., Fang, S.C.: On the entropic regularization method for solving min-max problems with applications. Math. Methods Oper. Res. **46**, 119–130 (1997)

25. Li, X.S., Pan, S.: Solving the finite min-max problem via an exponential penalty method. Comput. Technol. **8**, 3–15 (2003)
26. Lin, Z.H., Li, Y., Yu, B.: A combined homotopy interior method for general nonlinear programming problem. Appl. Math. Comput. **80**, 209–224 (1996)
27. Liu, G.X.: Aggregate homotopy methods for solving sequential max-min Problems, complementarity problems and variational inequalities. Ph.D. thesis, pp. 14–42. Department of Mathematics, Jilin University, Jilin, China (2003)
28. Liu, G.X., Feng, G.C., Yu, B.: Aggregate homotopy method for sequential max-min problems. J. JiLin Univ. (Sci. edn.) **2**, 155–156 (2003)
29. Liu, Q.H., Yu, B., Feng, G.C.: An interior point path-following method for nonconvex programming with quasi normal cone condition. Adv. Math. **29**, 381–382 (2000)
30. Mangasarian, O.L., Wild, E.W.: Multiple instance classification via successive linear programming. J. Optim. **137**, 555–568 (2008)
31. Mangasarian, O.L., Rosen, J.B., Thompson, M.E.: Nonconvex piecewise-quadratic underestimation for global minimization. J. Glob. Optim. **34**, 475–488 (2006)
32. Meng, F.W., Xia, Z.Q., Hao, Y.: Generalized second-order directional derivatives for a class of nondifferentiable functions. Oper. Res. Trans. **3**, 40–46 (1999)
33. Mifflin, R.: Semismooth and semoconvex functions in constrained optimization. SIAM J. Control Optim. **15**, 959–972 (1977)
34. Peng, J.M., Lin, Z.: A non-interior continuation method for generalized linear complementarity problems. Math. Program. **86**, 533–563 (1999)
35. Polak, E., Royset, J.O.: Algorithms for finite and semi-infinite min-max-min problems using adaptive smoothing techniques. J. Optim. **119**, 421–457 (2003)
36. Polak, E., Royset, J.O., Womersley, R.S.: Algorithms with adaptive smoothing for finite minimax problems. J. Optim. **119**, 459–484 (2003)
37. Qi, H.D., Liao, L.Z., Lin, Z.: Regularized smoothing approximations to vertical nonlinear complementarity problems. Math. Anal. Appl. **230**, 261–276 (1999)
38. Rockafellar, R.T.: A property of piecewise smooth functions. Comput. Optim. Appl. **25**, 247–250 (2003)
39. Royset, J.O., Polak, E., Kiureghian, A.D.: Adaptive approximations and exact penalization for the solution of generalized semi-infinite min-max problems. SIAM J. Optim. **14**, 1–34 (2003)
40. Scholtes, S.: Nonconvex structures in nonlinear programming. Oper. Res. **52**, 368–383 (2004)
41. Shang, Y.F.: Constraint shifting combined homotopy method for nonlinear programming, equilibrium programming and variational inequalities. Ph.D. thesis, pp. 14–52. Department of Mathematics, Jilin University, Jilin, China (2006)
42. Sun, W.J.: Study on homotopy methods for a local minimum of nonconvex programming problem, M.S. thesis, JinLin University, China (2006)
43. Wang, Y.: The computer implementation of the homotopy method for optimization problems, pp. 44–64. Dalian Maritime University Press, China (1996)
44. Watson, L.T.: Theory of globally convergent probability-one homotopies for nonlinear programming. SIAM J. Optim. **11**, 761–780 (2000)
45. Watson, L.T., Billups, S.C., Morgan, A.P.: Algorithm 652 hompack a suite of codes for globally convergent homotopy algorithms. ACM Trans. Math. Softw. **13**, 281–310 (1987)
46. Wright, M.H.: Ill-conditioning and computational error in interior methods for nonlinear programming. SIAM J. Optim. **9**, 84–111 (1998)
47. Xu, Q., Yu, B.: Homotopy method for non-convex programming in unbounded set. Northeast. Math. **21**, 25–31 (2005)
48. Xu, S.: Smoothing method for minimax problems. Comput. Optim. Appl. **20**, 267–279 (2001)
49. Yang, Q.Z., Yang, D.Z., Zhang, M.H.: Adjustable entropy function method. Math. Numer. Sinica **23**, 81–86 (2001)
50. Yamashita, H., Yabe, H.: An interior point method with a primal-dual quadratic barrier penalty function for nonlinear optimization. SIAM J. Optim. **14**, 479–499 (2003)
51. Yu, B., Feng, G.C., Zheng, S.L.: The aggregate constraint homotopy method for nonconvex nonlinear programming. Nonlinear. Anal. **45**, 839–847 (2001)
52. Yu, B., Liu, G.X., Feng, G.C., Li, Y.: The aggregate homotopy method for constrained sequential max-min problems. Northeast. Math. **19**, 287–290 (2003)
53. Yu, B., Liu, Q.H., Feng, G.C.: A combined homotopy interior point method for nonconvex programming with pseudo cone condition. Northeast. Math. **16**, 383–386 (2000)