

The convergent generalized central paths for linearly constrained convex programming

Qian, Xun; Liao, Lizhi; Sun, Jie; Zhu, Hong

Published in:
SIAM Journal on Optimization

DOI:
[10.1137/16M1104172](https://doi.org/10.1137/16M1104172)

Published: 01/04/2018

Document Version:
Publisher's PDF, also known as Version of record

[Link to publication](#)

Citation for published version (APA):
Qian, X., Liao, L., Sun, J., & Zhu, H. (2018). The convergent generalized central paths for linearly constrained convex programming. *SIAM Journal on Optimization*, 28(2), 1183-1204. <https://doi.org/10.1137/16M1104172>

General rights

Copyright and intellectual property rights for the publications made accessible in HKBU Scholars are retained by the authors and/or other copyright owners. In addition to the restrictions prescribed by the Copyright Ordinance of Hong Kong, all users and readers must also observe the following terms of use:

- Users may download and print one copy of any publication from HKBU Scholars for the purpose of private study or research
- Users cannot further distribute the material or use it for any profit-making activity or commercial gain
- To share publications in HKBU Scholars with others, users are welcome to freely distribute the permanent publication URLs

THE CONVERGENT GENERALIZED CENTRAL PATHS FOR LINEARLY CONSTRAINED CONVEX PROGRAMMING*

XUN QIAN[†], LI-ZHI LIAO[†], JIE SUN[‡], AND HONG ZHU[§]

Abstract. The convergence of central paths has been a focal point of research on interior point methods. Quite detailed analyses have been made for the linear case. However, when it comes to the convex case, even if the constraints remain linear, the problem is unsettled. In [*Math. Program.*, 103 (2005), pp. 63–94], Gilbert, Gonzaga, and Karas presented some examples in convex optimization, where the central path fails to converge. In this paper, we aim at finding some continuous trajectories which can converge for all linearly constrained convex optimization problems under some mild assumptions. We design and analyze a class of continuous trajectories, which are the solutions of certain ordinary differential equation (ODE) systems for solving linearly constrained smooth convex programming. The solutions of these ODE systems are named generalized central paths. By only assuming the existence of a finite optimal solution, we are able to show that, starting from any interior feasible point, (i) all of the generalized central paths are convergent, and (ii) the limit point(s) are indeed the optimal solution(s) of the original optimization problem. Furthermore, we illustrate that for the key example of Gilbert, Gonzaga, and Karas, our generalized central paths converge to the optimal solutions.

Key words. continuous trajectory, convex programming, interior point method, ordinary differential equation

AMS subject classifications. 90C25, 90C51, 37C75, 37N40

DOI. 10.1137/16M1104172

1. Introduction. Consider the following linearly constrained convex programming problem

$$(P) \quad \begin{array}{ll} \min & f(x) \\ \text{s.t.} & Ax = b, \ x_i \geq 0, \ i = 1, \dots, k, \end{array}$$

where $x \in \mathbb{R}^n$, $1 \leq k \leq n$, $f(x)$ is smooth and convex over the feasible set, and A is an m by n matrix with full row rank. As a blanket assumption, we assume that the optimal value of problem (P) is finite and attainable; therefore, we use \min rather than \inf in problem (P). In addition, the following notations are used in this paper:

$$\begin{aligned} \mathbb{R}_{k+}^n &= \{x \in \mathbb{R}^n \mid x_i \geq 0, \ 1 \leq i \leq k\}, \quad \mathbb{R}_{k++}^n = \{x \in \mathbb{R}^n \mid x_i > 0, \ 1 \leq i \leq k\}, \\ \mathcal{P}^+ &= \{x \in \mathbb{R}_{k+}^n \mid Ax = b\}, \quad \text{and} \quad \mathcal{P}^{++} = \{x \in \mathbb{R}_{k++}^n \mid Ax = b\}. \end{aligned}$$

It is conventional to assume that \mathcal{P}^{++} is nonempty in the analysis of interior point methods.

*Received by the editors November 18, 2016; accepted for publication (in revised form) November 21, 2017; published electronically April 24, 2018.

<http://www.siam.org/journals/siopt/28-2/M110417.html>

Funding: The work of the second author was supported in part by grants from Hong Kong Baptist University (FRG) and General Research Fund (GRF) of Hong Kong. The work of the third author was partially supported by Australia Council Research under grant DP160102819 and the National Science Foundation of China under grant-111 B16002.

[†]Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Kowloon, Hong Kong, People's Republic of China (13479784@life.hkbu.edu.hk, liliao@hkbu.edu.hk).

[‡]Department of Mathematics and Statistics, Curtin University, Perth, Australia (jie.sun@curtin.edu.au).

[§]Faculty of Science, Jiangsu University, Zhenjiang, Jiangsu, China (zhuhongmath@126.com).

The central path plays a vital role in interior point methods. The central path can be defined by setting up the corresponding barrier function. Among various barrier functions, the logarithm barrier function is mostly used, and the barrier function usually only involves the constraints (for example, if x_n is free in (P), then the barrier function usually does not contain x_n). However, the central path with this type barrier function may fail to converge for problem (P) when $k < n$. In [7], Gilbert, Gonzaga, and Karas constructed an example where the objective function is infinitely differentiable but the central path fails to converge with the logarithm barrier function. In [11], Iusem, Svaiter, and Neto proved that under certain hypotheses the central path defined by a general barrier function for a monotone variational inequality problem is well defined, bounded, continuous, and converges to the analytic center of the solution set. However, we note that there is no discussion in [11] on how to find a good barrier function such that the corresponding central path can converge.

Motivated by the counterexample in [7], in this paper, we aim at finding some continuous trajectories which can converge for all linearly constrained convex optimization problems under some mild assumptions. With several modifications on the original central path, we arrive at a class of new interior point continuous trajectories, which are the solutions of some ordinary differential equation (ODE) systems. For these ODE systems, we can prove that their solution trajectories will converge to some optimal solution(s) of problem (P) under quite mild conditions. Moreover, the initial point of these ODE systems can be any interior feasible point.

The first in-depth study of the central path is due to McLinden [15]. Then the central path was studied in more details for linear and convex quadratic programming, such as Adler and Monteiro [1], Güler [8], Megiddo [16], Monteiro and Tsuchiya [18], Vavasis and Ye [24]. In [1], the affine scaling continuous trajectories which contain the central path were studied for linear programming problems, and the optimality and convergence were obtained. The characterizations of the limit point and the dual estimates were also considered in [1]. In [17], Megiddo and Shub analyzed the limiting behavior of the affine scaling trajectories near the optimal vertex under the assumption of primal and dual nondegeneracy in linear programming. In [8], the limiting behavior of the weighted central path in linear programming at both $\mu = 0$ (μ is the parameter) and $\mu = +\infty$ was studied, and the convergence of the k th order derivatives was also obtained. In [16], the limiting behavior and the derivatives of the weighted central paths were studied and the behavior of trajectories near corners was also considered. In [18], the limiting behavior of the derivatives of certain trajectories which contain the central path associated with a monotone horizontal linear complementarity problem was studied and it was shown that each trajectory converges to the solution set along a unique and well-characterized direction. In [12], Kojima, Mizuno, and Noma studied the limiting behavior of some trajectories which are more general than the central path in [15] for monotone complementarity problems. In [24], it was shown that for linear programming, the central path composes at most n^2 alternating straight and curved segments. There are also several papers in the literature studying the central path for nonlinear programming, such as Drummond and Svaiter [6], Iusem, Svaiter, and Neto [11], McLinden [15], and Monteiro and Zhou [19]. For the central path with the logarithm barrier function and $k = n$ for problem (P), the convergence can be guaranteed under the strict complementarity condition [15], or the analyticity of $f(x)$ [19], or the condition that there exists a subspace W of \mathbb{R}^n such that $\text{Ker}(\nabla^2 f(x)) = W$ [6].

It should be mentioned that there have been some studies on the continuous trajectory for semidefinite optimization problems. Halická, de Klerk, and Roos [9]

studied the convergence of the central path and some properties of the limiting point in semidefinite optimization. Sim and Zhao [21] studied the underlying paths in interior point methods for the monotone semidefinite linear complementarity problem. They showed that each off-central path is a well-defined analytic curve with parameter μ ranging over $(0, \infty)$ and any accumulation point of the off-central path is a solution. Furthermore they also studied the analyticity of the off-central path through a simple example. Then they investigated the asymptotic behavior of off-central paths for general semidefinite linear complementarity problems (using the dual HKM direction) under the strict complementarity condition in [22].

Our main focus of this paper is the (interior point) continuous solution trajectory of the following ODE system

$$(1) \quad \frac{dx}{dt} = - [\gamma_1 I_n + tDP_{AD}D\nabla^2 f(x)]^{-1} DP_{AD}D\nabla f(x), \quad x(t_0) = x^0 \in \mathcal{P}^{++},$$

where

$$\begin{aligned} \gamma_1 > 0, \quad \frac{1}{2} \leq \gamma_2 < 1, \quad t_0 > 0, \quad t > 0, \quad x \in \mathbb{R}_{k^{++}}^n, \quad X = \text{diag}(x) \in \mathbb{R}^{n \times n}, \\ d \in \mathbb{R}^n, \quad d_i = x_i^{\gamma_2} \text{ for } i = 1, \dots, k, \quad d_i = 1 \text{ for } i = k + 1, \dots, n, \\ D = \text{diag}(d) \in \mathbb{R}^{n \times n}, \quad P_{AD} = I_n - DA^T(AD^2A^T)^{-1}AD, \end{aligned}$$

and I_n stands for the $n \times n$ identity matrix. For the ODE system (1), we sometimes use its equivalent implicit form

$$(2) \quad \frac{dx}{dt} = -\frac{1}{\gamma_1} DP_{AD}D \left[\nabla f(x) + t\nabla^2 f(x) \frac{dx}{dt} \right], \quad x(t_0) = x^0 \in \mathcal{P}^{++},$$

and we require $\nabla^2 f(x) \in C^1$ on $\mathbb{R}_{k^+}^n$. Note that the right-hand side of the ODE system (1) is defined on the open set $(0, +\infty) \times \mathbb{R}_{k^{++}}^n$. We need $t > 0$ to guarantee that the inverse of the matrix $\gamma_1 I_n + tDP_{AD}D\nabla^2 f(x)$ always exists (this will be shown by Lemma 5 later). The reason that x is required to be in $\mathbb{R}_{k^{++}}^n$ is that at the boundary of $\mathbb{R}_{k^{++}}^n$, the matrix P_{AD} may not exist. For example, if $k = 3, n = 4, b = (1, 1)^T$, and

$$A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix},$$

then at the point $(0, 0, 1, 0)^T, D = \text{diag}(0, 0, 1, 1)$ and

$$AD = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

which implies that the inverse of (AD^2A^T) does not exist. Hence P_{AD} does not exist at the point $(0, 0, 1, 0)^T$.

Now we explain where the ODE system (1) comes from when $k = n$. The central path with barrier function $-\beta_1 \sum_{i=1}^n x_i^{\alpha_1}$ ($0 < \alpha_1 < 1, \beta_1 > 0$) can be defined as the homotopy solution of the KKT system of an auxiliary optimization problem as follows:

$$(3) \quad \begin{cases} \nabla f(x) - z - A^T y = 0, \\ Ax = b, \quad x > 0, \\ X^{1-\alpha_1} z = \mu \alpha_1 \beta_1 e, \quad z > 0, \end{cases}$$

where $\mu > 0$ is the barrier penalty parameter, $e = (1, \dots, 1)^T \in \mathbb{R}^n$, and $X^{1-\alpha_1}$ is the power matrix in the usual sense of matrix analysis. The study of the existence

and convergence of the central path for a more general convex program can be found in [19]. In a recent paper [5], the central path and the affine scaling trajectory are studied for both linear programming and semidefinite programming from a dynamical system perspective. Let $(x(\mu), y(\mu), z(\mu))$ be the solution of the system (3). If for any $\mu > 0$, $(x(\mu), y(\mu), z(\mu))$ exists, then we obtain a trajectory of $(x(\mu), y(\mu), z(\mu))$ in terms of μ . By taking the derivative with respect to μ in (3), we can arrive at

$$(4) \quad \begin{cases} \nabla^2 f(x) \frac{dx}{d\mu} - \frac{dz}{d\mu} - A^T \frac{dy}{d\mu} = 0, \\ A \frac{dx}{d\mu} = 0, \\ (1 - \alpha_1)X^{-\alpha_1}Z \frac{dx}{d\mu} + X^{1-\alpha_1} \frac{dz}{d\mu} = \alpha_1 \beta_1 e, \end{cases}$$

where $Z = \text{diag}(z_1, z_2, \dots, z_n)$. From the above system, we can obtain (for detailed derivations of the following ODE system and the ODE system (6), see the appendix)

$$(5) \quad \frac{dx}{d\mu} = \frac{1}{\mu^2} \left[\gamma_1 I_n + \frac{1}{\mu} DP_{AD} D \nabla^2 f(x) \right]^{-1} DP_{AD} D \nabla f(x), \quad x(\mu_0) = x^0 \in \mathcal{P}^{++},$$

where $D = X^{\gamma_2}$, $\gamma_1 = (1 - \alpha_1)\alpha_1\beta_1$, and $\gamma_2 = 1 - \frac{\alpha_1}{2}$. Let $t = \frac{1}{\mu}$ with $t_0 = \frac{1}{\mu_0}$ in the above equation (for a general case of the conversion, see section 2.2 in [13]), we have

$$(6) \quad \frac{dx}{dt} = - \left[\gamma_1 I_n + t DP_{AD} D \nabla^2 f(x) \right]^{-1} DP_{AD} D \nabla f(x), \quad x(t_0) = x^0 \in \mathcal{P}^{++},$$

which is just the ODE system (1) (x^0 may be different) with $k = n$, $\gamma_1 > 0$, and $\frac{1}{2} < \gamma_2 < 1$. By using the barrier function $\beta_2 \sum_{i=1}^n x_i \ln x_i$ ($\beta_2 > 0$), we can similarly obtain the ODE system (6) with $\gamma_1 = \beta_2$ and $\gamma_2 = \frac{1}{2}$. This explains where the ODE system (1) comes from for $k = n$. It should also be mentioned that the solution of the ODE system (6) defines the primal central path only if x^0 is on the primal central path. In fact, finding an (x^0, y^0, z^0) on the central path is not an easy task. However, for our ODE system (1), x^0 is only required in \mathcal{P}^{++} .

Next we explain where the ODE system (1) comes from when $k < n$. For $k = n$, the matrix X^{γ_2} is like a barrier to prevent the trajectory from going into the nonpositive region. Hence for $k < n$, it is natural that we replace $x_i^{\gamma_2}$ with 1 for $i = k + 1, \dots, n$. In fact, the ODE system (1) can also be generated by proposing some specific barrier function. For $\frac{1}{2} < \gamma_2 < 1$, the barrier function takes the form

$$(7) \quad -\beta_1 \sum_{i=1}^k (x_i^{\alpha_1} + p_i x_i) + \frac{\alpha_1 \beta_1 (1 - \alpha_1)}{2} \sum_{i=k+1}^n (x_i - \bar{x}_i^0)^2,$$

where p_i 's and \bar{x}_i^0 's may depend on x^0 , $\gamma_1 = (1 - \alpha_1)\alpha_1\beta_1$, and $\gamma_2 = 1 - \frac{\alpha_1}{2}$. For problem (P), if we use the above barrier function, then from the KKT condition, we have the following system

$$\begin{cases} (\nabla f(x))_i - \mu \alpha_1 \beta_1 x_i^{\alpha_1 - 1} - \mu \beta_1 p_i + (A^T y)_i = 0, & i = 1, \dots, k, \\ (\nabla f(x))_i + \mu \alpha_1 \beta_1 (1 - \alpha_1) (x_i - \bar{x}_i^0) + (A^T y)_i = 0, & i = k + 1, \dots, n, \\ Ax = b, \\ x_i > 0, & i = 1, \dots, k. \end{cases}$$

In the above system, we can choose parameters p_i 's and \bar{x}_i^0 's such that the initial point of the ODE system (1) satisfies the above KKT condition. By taking the derivative with respect to μ in the above system, we can get the ODE system (1) similarly.

For $\gamma_2 = \frac{1}{2}$, the barrier function takes the form

$$(8) \quad \beta_2 \sum_{i=1}^k (x_i \ln x_i + p_i x_i) + \frac{\beta_2}{2} \sum_{i=k+1}^n (x_i - \bar{x}_i^0)^2,$$

where p_i 's and \bar{x}_i^0 's may depend on x^0 , $\gamma_1 = \beta_2$, and $\gamma_2 = \frac{1}{2}$. The barrier function (7) may not satisfy the condition H12 in [11], i.e., the barrier function attains its minimum on \mathbb{R}_{k++}^n at some point \tilde{x} . However, after we obtain the existence, uniqueness, and boundedness of these solution trajectories, the convergence can be proved similarly by the method in [11] if x^0 is on the primal central path (but here we only require x^0 in \mathcal{P}^{++}). In order to distinguish this from the usual central path, we call the solution trajectories of the ODE system (1) (parameterized by γ_1 , γ_2 , and x^0) as the generalized central paths. At the end of section 4, we will illustrate the trajectory behaviors of the central path and our generalized central paths by using a C^2 example in [7].

For simplicity, in what follows, $\|\cdot\|$ denotes the 2-norm. C^k stands for the class of k th order continuously differentiable functions. Unless otherwise specified, x_j denotes the j th component of a vector x , e denotes the column vector of all ones, and e_i denotes the unity column vector whose i th component is 1, the dimension of e and e_i are clear from the context. For any index subset $J \subseteq \{1, \dots, n\}$, we denote by x_J the vector composed of those components of $x \in \mathbb{R}^n$ indexed by $j \in J$, $\text{rank}(Q)$ denotes the rank of the matrix Q .

The rest of this paper is organized as follows. In section 2, we (i) verify that the ODE system (1) has a unique solution in $[t_0, +\infty)$, and (ii) introduce a potential function for the ODE system (1). In section 3, we prove that every accumulation point of the continuous trajectory of the ODE system (1) is an optimal solution for problem (P). In section 4, we show the strong convergence of the continuous trajectory and verify that the limiting point has the maximal number of the positive components in $\{x_1, \dots, x_k\}$ among the optimal solutions. At the end of section 4, several figures are shown to illustrate the trajectory behaviors of the central path and our generalized central paths for an example in [7] and we also discuss the calculation of the direction in the ODE system (1) and the choice of parameters γ_1 and γ_2 briefly. Finally, some conclusions are drawn in section 5.

2. Properties of the generalized central paths. The following assumptions are made throughout this paper.

ASSUMPTION 1. \mathcal{P}^{++} is nonempty, and there exists a point $x^* \in \mathcal{P}^+$ such that $f(x^*)$ is the optimal value of problem (P).

ASSUMPTION 2. $f(x) \in C^3$ on \mathbb{R}_{k+}^n .

In this section, we first show that for any fixed γ_1 , γ_2 , and x^0 in the ODE system (1), the solution of the ODE system (1) is unique and can be extended to infinity. Then a potential function is introduced in assisting the convergence proof for the solution of the ODE system (1). We begin our discussion by revealing some smoothness properties for the right-hand side of the ODE system (1).

LEMMA 3. For any $\gamma_2 > 0$, $(AD^2A^T)^{-1} \in C^1$ on \mathbb{R}_{k++}^n with respect to x .

Proof. Since A has full row rank m , for any $x \in \mathbb{R}_{k++}^n$, it is not hard to see that $(AD^2A^T)^{-1}$ exists. Then since $AD^2A^T \in C^1$ on \mathbb{R}_{k++}^n with respect to x , and the inverse of a matrix is smooth on the open set of invertible matrices (see Chapter 2 in [20]), we know that $(AD^2A^T)^{-1} \in C^1$ on \mathbb{R}_{k++}^n with respect to x . \square

LEMMA 4. *If $A, B \in \mathbb{R}^{n \times n}$ are both symmetric and positive semidefinite, then all eigenvalues of AB are nonnegative.*

Proof. Since B is symmetric and positive semidefinite, then from Theorem 7.2.6 in [10], there exists a unique symmetric and positive semidefinite square root of B , say $B^{\frac{1}{2}}$. Then from Theorem 1.3.20 in [10], we know AB has the same eigenvalues as $B^{\frac{1}{2}}AB^{\frac{1}{2}}$. Since $B^{\frac{1}{2}}AB^{\frac{1}{2}}$ is symmetric and positive semidefinite, its eigenvalues are all nonnegative. Hence all eigenvalues of AB are nonnegative. \square

LEMMA 5. *For any $\gamma_1 > 0$ and $\gamma_2 > 0$, $(\gamma_1 I_n + tDP_{AD}D\nabla^2 f(x))^{-1} \in C^1$ on $(t, x) \in (0, +\infty) \times \mathbb{R}_{k++}^n$.*

Proof. For any $\gamma_1 > 0$, $\gamma_2 > 0$, and $x \in \mathbb{R}_{k++}^n$, from Lemma 3, we know that

$$P_{AD} = I - DA^T(AD^2A^T)^{-1}AD \in C^1.$$

From this and Assumption 2, it is easy to see $DP_{AD}D\nabla^2 f(x) \in C^1$, and for $(t, x) \in (0, +\infty) \times \mathbb{R}_{k++}^n$, $\gamma_1 I_n + tDP_{AD}D\nabla^2 f(x) \in C^1$.

Since $f(x)$ is convex, $\nabla^2 f(x)$ is positive semidefinite. Noticing $P_{AD}^2 = P_{AD}$, it is easy to see that for any $x \in \mathbb{R}_{k++}^n$, $DP_{AD}D$ is positive semidefinite. So all eigenvalues of $tDP_{AD}D\nabla^2 f(x)$ are nonnegative from Lemma 4 for any $t \geq 0$. Then $(\gamma_1 I_n + tDP_{AD}D\nabla^2 f(x))^{-1}$ always exists. Moreover, since the inverse of a matrix is smooth on the open set of invertible matrices (see Chapter 2 in [20]), then for $(t, x) \in (0, +\infty) \times \mathbb{R}_{k++}^n$,

$$[\gamma_1 I_n + tDP_{AD}D\nabla^2 f(x)]^{-1} \in C^1.$$

The proof is completed. \square

Theorems 6 and 7 below guarantee the existence, uniqueness, and feasibility for the solution of the ODE system (1).

THEOREM 6. *For the ODE system (1), there exists a unique solution $x(t)$ with a maximal existence interval $[t_0, \beta)$; in addition, $x_i(t) > 0$ for $i = 1, \dots, k$ on the existence interval.*

Proof. By Lemma 5, the right-hand side of the ODE system (1) is locally Lipschitz continuous on $(0, +\infty) \times \mathbb{R}_{k++}^n$. Then from Theorem IV.1.2 in [3], there exists a unique solution $x(t)$ of the ODE system (1) on the maximal existence interval $[t_0, \beta)$ for some $\beta > t_0$ or $\beta = +\infty$ such that $x(t) \in \mathbb{R}_{k++}^n$. Since $x(t) \in \mathbb{R}_{k++}^n$, $x_i(t) > 0$ for $i = 1, \dots, k$ on the existence interval. The proof is completed. \square

Later in this section, it will be shown that $\beta = +\infty$ (Theorem 15). To simplify the presentation, in the remainder of this paper, $x(t)$ (or $D(t)$) will be replaced by x (or D) whenever no confusion would occur. The next theorem shows that the solution trajectory is feasible for problem (P).

THEOREM 7. *Let $x(t)$ be the solution of the ODE system (1) in the maximal existence interval $[t_0, \beta)$. Then $Ax(t) = b \forall t \in [t_0, \beta)$.*

Proof. By the implicit form (2), we know that for any $t \in [t_0, \beta)$

$$x(t) = x^0 - \frac{1}{\gamma_1} \int_{t_0}^t \left(DP_{AD}D \left[\nabla f(x) + t\nabla^2 f(x) \frac{dx}{dt} \right] \Big|_{t=\tau} \right) d\tau.$$

Noticing

$$ADP_{AD} = AD - AD^2A^T(AD^2A^T)^{-1}AD \equiv 0,$$

we can get

$$Ax(t) = Ax^0 - \frac{1}{\gamma_1} \int_{t_0}^t \left(ADP_{AD}D \left[\nabla f(x) + t \nabla^2 f(x) \frac{dx}{dt} \right] \Big|_{t=\tau} \right) d\tau = b.$$

Thus the theorem is proved. □

The next four lemmas lay the foundation for the potential function which will be introduced in (9) below. The potential function will play a vital role in the proofs of the main results.

LEMMA 8. For any fixed $\gamma_2 \geq \frac{1}{2}$, if $0 < x_i \leq M$ for $i = 1, \dots, k$, and $|x_i| \leq M$ for $i = k + 1, \dots, n$, and $\nabla f(x) \in C^1$ on \mathbb{R}_{k+}^n , then there exists some positive constant \hat{M} such that $\|\nabla f(x)\| \leq \hat{M}$, moreover, every entry of $DP_{AD}D\nabla f(x)$ and $D^{1-\frac{1}{\gamma_2}}P_{AD}D\nabla f(x)$ is bounded, and the bound depends only on A, M, n , and \hat{M} .

Proof. If $0 < x_i \leq M$ for $i = 1, \dots, k$ and $|x_i| \leq M$ for $i = k + 1, \dots, n$, then x belongs to a closed bounded subset of \mathbb{R}_{k+}^n . Along with the assumption that $\nabla f(x) \in C^1$ on \mathbb{R}_{k+}^n , there must exist some positive constant \hat{M} such that $\|\nabla f(x)\| \leq \hat{M}$. From Lemma 3 and the remark in [23], we know that if $d > 0$, then every entry of $(AD^2A^T)^{-1}AD^2$ is bounded, and the bound depends only on A and n . So if $0 < x_i \leq M$ for $i = 1, \dots, k$, then every entry of

$$DP_{AD}D = D^2 - D^2A^T(AD^2A^T)^{-1}AD^2$$

is bounded, and the bound depends only on A, M , and n .

Hence, we know that every entry of $DP_{AD}D\nabla f(x)$ is bounded. Furthermore the bound depends only on A, M, n , and \hat{M} . Noticing that

$$D^{1-\frac{1}{\gamma_2}}P_{AD}D = D^{2-\frac{1}{\gamma_2}} - D^{2-\frac{1}{\gamma_2}}A^T(AD^2A^T)^{-1}AD^2$$

and $\gamma_2 \geq \frac{1}{2}$, we can similarly verify that every entry of $D^{1-\frac{1}{\gamma_2}}P_{AD}D$ is bounded, and the bound depends only on A, M , and n . Therefore, we know that every entry of $D^{1-\frac{1}{\gamma_2}}P_{AD}D\nabla f(x)$ is bounded, and the bound depends only on A, M, n , and \hat{M} . Thus the proof is completed. □

LEMMA 9 (see [4]). Suppose f is differentiable (i.e., its gradient ∇f exists at each point in $\text{dom}f$). Then f is convex if and only if $\text{dom}f$ is convex and

$$f(y) \geq f(x) + \nabla f(x)^T(y - x)$$

holds for all $x, y \in \text{dom}f$.

Proof. See section 3.1.3 in [4] for the proof. □

LEMMA 10. Let a be any positive constant and let $g(x) = x - a - a \cdot \ln \frac{x}{a}$. Then for any scalar $x > 0$, $g(x) \geq 0$ and $g(x) = 0$ if and only if $x = a$. Furthermore, $g(x) \rightarrow +\infty$ as $x \rightarrow 0^+$ or $x \rightarrow +\infty$.

Proof. The proof is obvious and omitted. □

LEMMA 11. Let a be any positive constant and $1 < r < 2$. Then for any scalar $x > 0$, $g(x) = \frac{1}{2-r}(x^{2-r} - a^{2-r}) - \frac{a}{1-r}(\frac{1}{x^{r-1}} - \frac{1}{a^{r-1}}) \geq 0$ and $g(x) = 0$ if and only if $x = a$. Furthermore, $g(x) \rightarrow +\infty$ as $x \rightarrow 0^+$ or $x \rightarrow +\infty$.

Proof. This can be verified by basic calculus and the proof is omitted. □

Now we introduce a potential function for the ODE system (1). With the help of this potential function, the boundedness of the optimal solution set is no longer needed in the convergence proof for the solution of the ODE system (1). Instead, only the weaker Assumption 1 is needed. In 1983, Losert and Akin [14] introduced a kind of potential function for both the discrete and continuous dynamical systems in a classical model of population genetics. Their potential function was originally designed for a single constraint. We extend their potential function to multiple constraints. In order to define the potential function, we first introduce some notations. For any $y \in \mathbb{R}_{k+}^n$, $B(y) = \{i \mid y_i > 0, i = 1, \dots, k\}$ and $N(y) = \{i \mid y_i = 0, i = 1, \dots, k\}$. Obviously, for any $y \in \mathbb{R}_{k+}^n$, $B(y) \cap N(y) = \emptyset$ and $B(y) \cup N(y) = \{1, \dots, k\}$. Then the potential function $V(t, x, y)$ for the ODE system (1) can be defined as follows:

$$(9) \quad V(t, x, y) = \gamma_1 I(x, y) + t [f(y) - f(x) + (x - y)^T \nabla f(x)],$$

where

$$(10) \quad I(x, y) = \sum_{i=k+1}^n \frac{1}{2} (x_i - y_i)^2$$

$$(11) \quad \begin{cases} \sum_{i=1}^k (x_i - y_i) - \sum_{i \in B(y)} y_i \cdot \ln \frac{x_i}{y_i} & \text{if } \gamma_2 = \frac{1}{2}, B(y) \subseteq B(x), \\ + \sum_{i=1}^k \frac{x_i^{2-2\gamma_2} - (y_i)^{2-2\gamma_2}}{2 - 2\gamma_2} \\ - \sum_{i \in B(y)} \frac{y_i}{1 - 2\gamma_2} \left(\frac{1}{x_i^{2\gamma_2-1}} - \frac{1}{y_i^{2\gamma_2-1}} \right) & \text{if } \frac{1}{2} < \gamma_2 < 1, B(y) \subseteq B(x), \\ + \infty & \text{if } B(y) \not\subseteq B(x), \end{cases}$$

$t > 0$ and $x \in \mathbb{R}_{k+}^n$ are the variables, $y \in \mathbb{R}_{k+}^n$ is a parameter.

With the introduction of function $V(t, x, y)$ in (9), some important results can be obtained in Theorems 15 and 16. But first, the following Lemmas 12, 13, and 14 are needed.

LEMMA 12. For any $(t, x) \in (0, +\infty) \times \mathbb{R}_{k++}^n$, the formula

$$- [\gamma_1 I_n + t DP_{AD} D \nabla^2 f(x)]^{-1} DP_{AD} D \nabla f(x)$$

in the ODE system (1) has the following equivalent form:

$$(13) \quad - DP_{AD} [\gamma_1 I_n + t P_{AD} D \nabla^2 f(x) DP_{AD}]^{-1} P_{AD} D \nabla f(x).$$

Proof. Since P_{AD} is a symmetric matrix and $P_{AD}^2 = P_{AD}$, we know the eigenvalues of P_{AD} are 0 or 1. From the eigenvalue decomposition of P_{AD} , there exist two matrices $Q \in \mathbb{R}^{n \times p}$ and $R \in \mathbb{R}^{n \times (n-p)}$ such that $P_{AD} = QQ^T$ and $[Q \ R]^T [Q \ R] = I_n$ with $p = \text{rank}(P_{AD})$. Then

$$\begin{aligned} \gamma_1 I_n + t P_{AD} D \nabla^2 f(x) DP_{AD} &= \gamma_1 I_n + t QQ^T D \nabla^2 f(x) D QQ^T \\ &= \gamma_1 QQ^T + \gamma_1 RR^T + t QQ^T D \nabla^2 f(x) D QQ^T \\ &= Q(\gamma_1 I_p + t Q^T D \nabla^2 f(x) D Q) Q^T + \gamma_1 RR^T \\ &= [Q \ R] \begin{pmatrix} \gamma_1 I_p + t Q^T D \nabla^2 f(x) D Q & 0 \\ 0 & \gamma_1 I_{n-p} \end{pmatrix} [Q \ R]^T, \end{aligned}$$

which implies

$$\begin{aligned} & [\gamma_1 I_n + tP_{AD}D\nabla^2 f(x)DP_{AD}]^{-1} \\ &= [Q \ R] \begin{pmatrix} [\gamma_1 I_p + tQ^T D\nabla^2 f(x)DQ]^{-1} & 0 \\ 0 & (\gamma_1 I_{n-p})^{-1} \end{pmatrix} [Q \ R]^T. \end{aligned}$$

Thus from $Q^T R = 0$ and $Q^T Q = I_p$,

$$\begin{aligned} & -DP_{AD} [\gamma_1 I_n + tP_{AD}D\nabla^2 f(x)DP_{AD}]^{-1} P_{AD}D\nabla f(x) \\ &= -DQQ^T [Q \ R] \begin{pmatrix} [\gamma_1 I_p + tQ^T D\nabla^2 f(x)DQ]^{-1} & 0 \\ 0 & (\gamma_1 I_{n-p})^{-1} \end{pmatrix} [Q \ R]^T QQ^T D\nabla f(x) \\ &= -DQ [\gamma_1 I_p + tQ^T D\nabla^2 f(x)DQ]^{-1} Q^T D\nabla f(x). \end{aligned}$$

In addition,

$$\begin{aligned} & -[\gamma_1 I_n + tDP_{AD}D\nabla^2 f(x)]^{-1} DP_{AD}D\nabla f(x) \\ &= -D [\gamma_1 I_n + tP_{AD}D\nabla^2 f(x)D]^{-1} P_{AD}D\nabla f(x) \\ &= -D [\gamma_1 I_n + tQQ^T D\nabla^2 f(x)D]^{-1} QQ^T D\nabla f(x). \end{aligned}$$

Hence the lemma is proved, if we have the following equality

$$(14) \quad Q [\gamma_1 I_p + tQ^T D\nabla^2 f(x)DQ]^{-1} = [\gamma_1 I_n + tQQ^T D\nabla^2 f(x)D]^{-1} Q.$$

Noticing that

$$\begin{aligned} & [\gamma_1 I_n + tQQ^T D\nabla^2 f(x)D] Q [\gamma_1 I_p + tQ^T D\nabla^2 f(x)DQ]^{-1} \\ &= \gamma_1 Q [\gamma_1 I_p + tQ^T D\nabla^2 f(x)DQ]^{-1} \\ &\quad + tQQ^T D\nabla^2 f(x)DQ [\gamma_1 I_p + tQ^T D\nabla^2 f(x)DQ]^{-1} \\ &= Q [\gamma_1 I_p + tQ^T D\nabla^2 f(x)DQ] [\gamma_1 I_p + tQ^T D\nabla^2 f(x)DQ]^{-1} \\ &= Q, \end{aligned}$$

the equality (14) holds and the proof is completed. □

LEMMA 13. For any fixed $\gamma_1 > 0$ and $\gamma_2 \geq \frac{1}{2}$, if $0 < x_i \leq M$ for $i = 1, \dots, k$ and $|x_i| \leq M$ for $i = k + 1, \dots, n$ with $M > 0$, $\nabla^2 f(x) \in C^1$ on \mathbb{R}_{k+}^n , and $0 \leq t \leq \beta_1$ with $t_0 < \beta_1 < +\infty$, where $\beta_1 \leq \beta$ and β is defined in Theorem 6, then

- (i) there exists some positive constant \tilde{M} such that $\|\nabla^2 f(x)\| \leq \tilde{M}$;
- (ii) every entry of $[\gamma_1 I_n + tDP_{AD}D\nabla^2 f(x)]^{-1} DP_{AD}D\nabla f(x)$ is bounded, and the bound depends only on $A, M, n, \beta_1, \hat{M}$, and \tilde{M} , where \hat{M} is defined in Lemma 8;
- (iii) every entry of $D^{-\frac{1}{\gamma_2}} [\gamma_1 I_n + tDP_{AD}D\nabla^2 f(x)]^{-1} DP_{AD}D\nabla f(x)$ is bounded, and the bound depends only on $A, M, n, \beta_1, \hat{M}$, and \tilde{M} , where \hat{M} is defined in Lemma 8.

Proof. (i) If $0 < x_i \leq M$ for $i = 1, \dots, k$ and $|x_i| \leq M$ for $i = k + 1, \dots, n$ with $M > 0$, then x belongs to a closed bounded subset of \mathbb{R}_{k+}^n . Furthermore $\nabla^2 f(x) \in C^1$ on \mathbb{R}_{k+}^n , hence there exists some positive constant \tilde{M} such that $\|\nabla^2 f(x)\| \leq \tilde{M}$.

(ii) From Lemma 8, we know that when $0 < x_i \leq M$ for $i = 1, \dots, k$ and $|x_i| \leq M$ for $k + 1 \leq i \leq n$, every entry of $DP_{AD}D\nabla f(x)$ and $D^{1-\frac{1}{\gamma_2}}P_{AD}D\nabla f(x)$ is bounded, and the bound depends only on A, M, n , and \hat{M} .

From (i), we know that every entry of $DP_{AD}D\nabla^2 f(x)$ is bounded, and the bound depends only on A, M, n , and \tilde{M} . Therefore, when $0 < x_i \leq M$ for $i = 1, \dots, k$ and $|x_i| \leq M$ for $k + 1 \leq i \leq n$ and $0 \leq t \leq \beta_1$, every entry of $\gamma_1 I_n + tDP_{AD}D\nabla^2 f(x)$ is bounded, and every entry of its adjoint matrix is also bounded. Furthermore its determinant is not less than γ_1^n because the eigenvalues of $tDP_{AD}D\nabla^2 f(x)$ are all nonnegative from Lemma 4. Hence every entry of $[\gamma_1 I_n + tDP_{AD}D\nabla^2 f(x)]^{-1}$ is bounded, and the bound depends only on A, M, n, β_1 , and \tilde{M} .

Therefore we know that if $0 < x_i \leq M$ for $i = 1, \dots, k$ and $|x_i| \leq M$ for $k + 1 \leq i \leq n$ and $0 \leq t \leq \beta_1$, then every entry of

$$[\gamma_1 I_n + tDP_{AD}D\nabla^2 f(x)]^{-1} DP_{AD}D\nabla f(x)$$

is bounded, and the bound depends only on $A, M, n, \beta_1, \hat{M}$, and \tilde{M} .

(iii) Let

$$\tilde{d} = -[\gamma_1 I_n + tDP_{AD}D\nabla^2 f(x)]^{-1} DP_{AD}D\nabla f(x).$$

Then every entry of \tilde{d} is bounded, and we can write \tilde{d} in the following form,

$$\tilde{d} = -\frac{1}{\gamma_1} DP_{AD}D(\nabla f(x) + t\nabla^2 f(x)\tilde{d}),$$

then we have

$$D^{-\frac{1}{\gamma_2}}\tilde{d} = -\frac{1}{\gamma_1} D^{1-\frac{1}{\gamma_2}} P_{AD}D(\nabla f(x) + t\nabla^2 f(x)\tilde{d}).$$

From the proof of Lemma 8, we know that every entry of $D^{1-\frac{1}{\gamma_2}}P_{AD}D$ is bounded, and the bound depends only on A, M , and n . Furthermore it is easy to see that when $0 < x_i \leq M$ for $i = 1, \dots, k$ and $|x_i| \leq M$ for $k + 1 \leq i \leq n$ and $0 \leq t \leq \beta_1$, then every entry of $\nabla f(x) + t\nabla^2 f(x)\tilde{d}$ is bounded, and the bound depends only on $A, M, n, \beta_1, \hat{M}$, and \tilde{M} . Hence we know that every entry of

$$D^{-\frac{1}{\gamma_2}}\tilde{d} = D^{-\frac{1}{\gamma_2}} [\gamma_1 I_n + tDP_{AD}D\nabla^2 f(x)]^{-1} DP_{AD}D\nabla f(x)$$

is bounded, and the bound depends only on $A, M, n, \beta_1, \hat{M}$, and \tilde{M} . Thus the proof is completed. \square

LEMMA 14. *Let $x(t)$ be the solution of the ODE system (1) on the maximal existence interval $[t_0, \beta)$. Then there exists an $M > 0$ which depends only on t_0, x^0, x^* , and $f(x^*) - f(x^0) + (x^0 - x^*)^T \nabla f(x^0)$ such that $|x_i(t)| \leq M \forall t \in [t_0, \beta)$ for $i = 1, \dots, n$.*

Proof. We can define function $I_1(t, x)$ as follows:

$$I_1(t, x) = V(t, x, x^*) \quad \forall (t, x) \in (0, +\infty) \times \mathbb{R}_{k+}^n,$$

where $x^* \in \mathcal{P}^+$ according to Assumption 1.

From Theorem 6, $x(t) \in \mathbb{R}_{k+}^n$, so $I_1(t, x(t))$ is well defined. First, we have

$$\begin{aligned} \frac{dI_1(t, x(t))}{dt} &= \frac{\partial V(t, x, x^*)}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial V(t, x, x^*)}{\partial t} \cdot 1 \\ &= \gamma_1 \frac{\partial I(x, x^*)}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial [f(y) - f(x) + (x - y)^T \nabla f(x)] \cdot t}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial V(t, x, x^*)}{\partial t}. \end{aligned}$$

Since $x(t) \in \mathbb{R}_{k^{++}}^n$, $B(x(t)) = \{1, 2, \dots, k\}$. Then it can be verified for $k + 1 \leq i \leq n$,

$$\frac{\partial I(x, x^*)}{\partial x_i} = x_i - x_i^*.$$

For $i \in B(x^*)$ and $\gamma_2 = \frac{1}{2}$,

$$\frac{\partial I(x, x^*)}{\partial x_i} = 1 - \frac{x_i^*}{x_i} = (x_i - x_i^*)x_i^{-2\gamma_2}.$$

For $i \in B(x^*)$ and $\frac{1}{2} < \gamma_2 < 1$,

$$\frac{\partial I(x, x^*)}{\partial x_i} = x_i^{1-2\gamma_2} - x_i^* x_i^{-2\gamma_2} = (x_i - x_i^*)x_i^{-2\gamma_2}.$$

For $i \in N(x^*)$ and $\gamma_2 = \frac{1}{2}$, since $x_i^* = 0$, hence

$$\frac{\partial I(x, x^*)}{\partial x_i} = 1 = (x_i - 0)x_i^{-1} = (x_i - x_i^*)x_i^{-2\gamma_2}.$$

For $i \in N(x^*)$ and $\frac{1}{2} < \gamma_2 < 1$, since $x_i^* = 0$, hence

$$\frac{\partial I(x, x^*)}{\partial x_i} = x_i^{1-2\gamma_2} = (x_i - 0)x_i^{-2\gamma_2} = (x_i - x_i^*)x_i^{-2\gamma_2}.$$

Therefore $\frac{\partial I(x, x^*)}{\partial x} = (x - x^*)^T D^{-2}$. Moreover, from Theorem 7, we have

$$\begin{aligned} (x(t) - x^*)^T D^{-2} DP_{AD}D &= (x(t) - x^*)^T - (x(t) - x^*)^T A^T (AD^2 A^T)^{-1} AD^2 \\ &= (x(t) - x^*)^T, \end{aligned}$$

hence, from the implicit form (2), we have

$$\begin{aligned} \frac{dI_1(t, x(t))}{dt} &= [\gamma_1(x - x^*)^T D^{-2} + t(x - x^*)^T \nabla^2 f(x)] \frac{dx}{dt} \\ &\quad + f(x^*) - f(x) + (x - x^*)^T \nabla f(x) \\ &= (x^* - x)^T \left[\nabla f(x) + t \nabla^2 f(x) \frac{dx}{dt} \right] + t(x - x^*)^T \nabla^2 f(x) \frac{dx}{dt} \\ &\quad + f(x^*) - f(x) + (x - x^*)^T \nabla f(x) \\ &= (x^* - x)^T \nabla f(x) + f(x^*) - f(x) + (x - x^*)^T \nabla f(x) \\ &= f(x^*) - f(x) \\ &\leq 0. \end{aligned}$$

From Lemma 9, we have

$$(15) \quad f(x^*) - f(x) + (x - x^*)^T \nabla f(x) \geq 0.$$

From Lemmas 10, 11, and inequality (15), we know that for any $(t, x) \in (0, +\infty) \times \mathbb{R}_{k^{++}}^n$, if there exists some $i \in \{1, \dots, n\}$ such that $|x_i| \rightarrow +\infty$, then $I_1(t, x) \rightarrow +\infty$. But $I_1(t, x(t)) \leq I_1(t_0, x(t_0)) \forall t \in [t_0, \beta)$. Since $I_1(t_0, x^0) = \gamma_1 I(x^0, x^*) + t_0[f(x^*) - f(x^0) + (x^0 - x^*)^T \nabla f(x^0)]$, we know that there exists an $M > 0$ which depends only on t_0, x^0, x^* , and $f(x^*) - f(x^0) + (x^0 - x^*)^T \nabla f(x^0)$ such that $|x_i(t)| \leq M \forall t \in [t_0, \beta)$ for $i = 1, \dots, n$. \square

THEOREM 15. *Let the maximal existence interval of the solution $x(t)$ of the ODE system (1) be $[t_0, \beta)$. Then $\beta = +\infty$.*

Proof. Assume $\beta \neq +\infty$. Noticing that

$$\frac{dx}{dt} = D^{\frac{1}{\gamma_2}} D^{-\frac{1}{\gamma_2}} [\gamma_1 I_n + tDP_{AD}D\nabla^2 f(x)]^{-1} DP_{AD}D\nabla f(x),$$

we have for $i \in \{1, \dots, k\}$,

$$\frac{dx_i}{dt} = x_i \cdot \left\{ D^{-\frac{1}{\gamma_2}} [\gamma_1 I_n + tDP_{AD}D\nabla^2 f(x)]^{-1} DP_{AD}D\nabla f(x) \right\}_i.$$

Hence from Lemmas 13 and 14, we know that there exists an $L > 0$ such that for every $i \in \{1, \dots, k\}$, we have

$$(16) \quad \left| \frac{dx_i}{dt} \right| \leq Lx_i \quad \forall t \in [t_0, \beta),$$

and for every $i \in \{k + 1, \dots, n\}$, we have

$$(17) \quad \left| \frac{dx_i}{dt} \right| \leq L \quad \forall t \in [t_0, \beta);$$

furthermore this L depends only on $A, n, \beta, t_0, x^0, x^*, \hat{M}, \tilde{M}$, and $f(x^*) - f(x^0) + (x^0 - x^*)^T \nabla f(x^0)$, where \hat{M} is defined in Lemma 8 and \tilde{M} is defined in Lemma 13.

For every $i \in \{1, \dots, n\}$, from inequalities (16), (17), and $0 < |x_i(t)| \leq M \forall t \in [t_0, \beta)$, we know that (without loss of generality we assume $M \geq 1$)

$$(18) \quad \left| \frac{dx_i}{dt} \right| \leq LM \quad \forall t \in [t_0, \beta);$$

furthermore, $x(t)$ is continuous on $[t_0, \beta)$, and it is not hard to see that $\lim_{t \rightarrow \beta^-} x(t)$ exists. We denote this limit as $x(\beta)$. Evidently $x(\beta) \in \mathbb{R}_{k+}^n$. According to the extension theorem in section 2.5 in [2], we know that $(t, x(t))$ will go to the boundary of the open set $(0, +\infty) \times \mathbb{R}_{k++}^n$. However, because of the hypothesis, $\beta \neq +\infty$, so there must exist at least one $i \in \{1, \dots, k\}$ such that $x_i(\beta) = 0$. From inequality (16), we know that if $t \in [t_0, \beta)$,

$$\frac{dx_i}{x_i} \geq -Ldt.$$

Integrating the inequality above, we have for every $t \in [t_0, \beta)$

$$\ln x_i(t) - \ln x_i(t_0) \geq -L(t - t_0).$$

Since $x_i(t) \rightarrow x_i(\beta) = 0$ as $t \rightarrow \beta^-$, $\ln x_i(t) - \ln x_i(t_0) \rightarrow -\infty$ as $t \rightarrow \beta^-$, but $-L(t - t_0)$ is bounded since $-L(t - t_0) \geq -L(\beta - t_0)$. This is a contradiction. Thus $\beta = +\infty$, and the proof is completed. \square

From Theorem 15, we can define the limit set for the solution of the ODE system (1). Let $x(t)$ be the solution of the ODE system (1), then its limit set is defined as

$$\Omega^1(x^0) = \left\{ p \in \mathcal{R}^n \mid \exists \{t_k\}_{k=0}^{+\infty} \text{ with } \lim_{k \rightarrow +\infty} t_k = +\infty \text{ such that } \lim_{k \rightarrow +\infty} x(t_k) = p \right\}.$$

Note that in fact, set $\Omega^1(x^0)$ also depends on parameters γ_1 and γ_2 in the ODE system (1). But to simplify our notation, γ_1 and γ_2 are omitted here.

THEOREM 16. *The limit set $\Omega^1(x^0)$ is nonempty and contained in \mathcal{P}^+ .*

Proof. From Theorems 6, 7, and 15, we know that the limit set $\Omega^1(x^0)$ is contained in \mathcal{P}^+ . From Lemma 14, we know that the solution $x(t)$ is contained in a bounded closed set. Hence the limit set $\Omega^1(x^0)$ is also nonempty. \square

The next theorem reveals some fundamental properties for the solutions of the ODE system (1).

THEOREM 17. *Let $x(t)$ be the solution of the ODE system (1), then $f(x(t))$ is a nonincreasing function on $[t_0, +\infty)$. Furthermore, if $x^0 \in \mathcal{P}^{++}$ is an optimal solution for problem (P), then $x(t) \equiv x^0$ on $[t_0, +\infty)$; otherwise $f(x(t))$ is a strictly decreasing function on $[t_0, +\infty)$.*

Proof. From the equivalent form (13), we have for $t \geq t_0$,

$$(19) \quad \frac{df(x(t))}{dt} = - \left\| [\gamma_1 I_n + tP_{AD}D\nabla^2 f(x)DP_{AD}]^{-\frac{1}{2}} P_{AD}D\nabla f(x) \right\|^2 \leq 0.$$

Hence we know that $f(x(t))$ is a nonincreasing function on $[t_0, +\infty)$.

The KKT conditions for problem (P) can be stated as follows:

$$(20) \quad \begin{cases} Ax = b, & x \in \mathbb{R}_{k+}^n, \\ Xz = 0, & z \in \mathbb{R}_{k+}^n, \\ A^T y + z = \nabla f(x), & y \in \mathbb{R}^m, \\ z_i = 0 & \text{for } k + 1 \leq i \leq n. \end{cases}$$

If $x \in \mathcal{P}^{++}$ is an optimal solution, there must exist corresponding (y, z) such that the system (20) holds, then

$$z = 0 \quad \text{and} \quad A^T y = \nabla f(x),$$

thus, it is easy to see that

$$P_{AD}D\nabla f(x) = P_{AD}DA^T y = 0.$$

Hence if $x^0 \in \mathcal{P}^{++}$ is an optimal solution, we know that the right-hand side of the ODE system (1) equals zero at $x = x_0$, therefore $x(t) \equiv x^0$ for $t \geq t_0$ is a solution of the ODE system (1). Because of the uniqueness of the solution, we know that $x(t) \equiv x^0$ on $[t_0, +\infty)$.

If $x^0 \in \mathcal{P}^{++}$ is not an optimal solution, we show that $f(x(t))$ is a strictly decreasing function on $[t_0, +\infty)$. Suppose that $f(x(t))$ is not a strictly decreasing function on $[t_0, +\infty)$, then there must exist t_1 and t_2 with $t_0 \leq t_1 < t_2$ such that $f(x(t_1)) = f(x(t_2))$. Since $\frac{df(x(t))}{dt} \leq 0$, we know that when $t_1 \leq t \leq t_2$, $\frac{df(x(t))}{dt} \equiv 0$. Then from (19), we have that when $t_1 \leq t \leq t_2$,

$$[\gamma_1 I_n + tP_{AD}D\nabla^2 f(x)DP_{AD}]^{-\frac{1}{2}} P_{AD}D\nabla f(x) \equiv 0.$$

For $0 < t_1 \leq t \leq t_2$, the matrix $\gamma_1 I_n + tP_{AD}D\nabla^2 f(x)DP_{AD}$ is symmetric and positive definite, thus $[\gamma_1 I_n + tP_{AD}D\nabla^2 f(x)DP_{AD}]^{-\frac{1}{2}}$ is always invertible, which implies that when $t_1 \leq t \leq t_2$,

$$P_{AD}D\nabla f(x) \equiv 0.$$

In particular, at the point $x(t_1) = x^1 \in \mathcal{P}^{++}$, we have

$$(21) \quad P_{AD}D\nabla f(x) = 0.$$

Consider the following ODE system

$$(22) \quad \frac{dx}{dt} = - [\gamma_1 I_n + tDP_{AD}D\nabla^2 f(x)]^{-1} DP_{AD}D\nabla f(x), \quad x(t_1) = x^1 \in \mathcal{P}^{++},$$

where the ODE is the same as the one in system (1). In order to distinguish the two solutions of the ODE systems (1) and (22), we denote the solution of the ODE system (22) by $\tilde{x}(t)$. From (21), we know that when $x = x^1$ for any $t > 0$, $\frac{dx}{dt} = 0$ in the ODE system (22). Hence $\tilde{x}(t) \equiv x^1$ on $[t_0, +\infty)$ is a solution of the ODE system (22). However, the solution $x(t)$ of the ODE system (1) also satisfies $x(t_1) = x^1$ and satisfies the ODE in the ODE system (22), which implies that $\tilde{x}(t) = x(t)$ on $[t_0, +\infty)$ is also a solution of the ODE system (22). According to the uniqueness of the solution for the ODE system (22), we have $x(t) \equiv x^1$ on $[t_0, +\infty)$. In particular, we have $x^0 = x(t_0) = x^1$. Thus from (21), we have $P_{AD}D\nabla f(x)|_{x=x^0} = 0$. Moreover, $x^0 \in \mathcal{P}^{++}$, D is invertible at x^0 , hence, from

$$P_{AD}D\nabla f(x)|_{x=x^0} = D [I_n - A^T(AD^2A^T)^{-1}AD^2] \nabla f(x)|_{x=x^0} = 0,$$

we have $\nabla f(x^0) = A^T(AD^2A^T)^{-1}AD^2\nabla f(x)|_{x=x^0}$.

Let $y^0 = (AD^2A^T)^{-1}AD^2\nabla f(x)|_{x=x^0}$ and $z^0 = 0 \in \mathbb{R}^n$. Then (x^0, y^0, z^0) satisfies the KKT system (20), which implies that x^0 is an optimal solution for problem (P). This is a contradiction. Hence $f(x(t))$ is a strictly decreasing function on $[t_0, +\infty)$. Thus the proof is completed. \square

3. Optimality of the cluster point(s). In this section, we will show that every accumulation point of the solution of the ODE system (1) is an optimal solution for problem (P).

THEOREM 18. *Every point in $\Omega^1(x^0)$ is an optimal solution for problem (P).*

Proof. When $i \in N(x^*)$, we have $x_i^* = 0$. From Lemmas 9, 10, and 11, it is easy to see that $I_1(t, x) \geq 0$ for all $x \in \mathbb{R}_{k^{++}}^n$ and $t > 0$. So for all $t \in [t_0, +\infty)$, $I_1(t, x(t))$ is bounded below. This along with the fact that $\frac{dI_1(t, x(t))}{dt} \leq 0$ implies that $I_1(t, x(t))$ has a finite limit as $t \rightarrow +\infty$.

Let $x^{(1)} \in \Omega^1(x^0)$. From Theorem 17 and Assumption 1, we know $f(x(t))$ has a finite limit as $t \rightarrow +\infty$, and for any $x^{(1)} \in \Omega^1(x^0)$, we have

$$\lim_{t \rightarrow +\infty} f(x(t)) = f(x^{(1)}) \geq f(x^*).$$

Assume $f(x^{(1)}) > f(x^*)$, then since $\frac{dI_1(t, x(t))}{dt} = f(x^*) - f(x)$ (see the proof of Lemma 14), we know for any $t \geq t_0$,

$$\frac{dI_1(t, x(t))}{dt} < f(x^*) - f(x^{(1)}) < 0,$$

which contradicts the fact that $I_1(t, x(t))$ has a finite limit as $t \rightarrow +\infty$. Hence $x^{(1)}$ is an optimal solution for problem (P). \square

4. Convergence of the generalized central paths. Now, it comes to a key result of the paper.

THEOREM 19. *The limit set $\Omega^1(x^0)$ only contains a single point.*

Proof. From Theorem 16, we know that $\Omega^1(x^0)$ is not empty. So we can choose a point $\tilde{x} \in \Omega^1(x^0)$, and evidently $\tilde{x} \in \mathcal{P}^+$. From (9), we can define $V_1(t, x)$ as follows:

$$V_1(t, x) = V(t, x, \tilde{x}) \quad \forall (t, x) \in (0, +\infty) \times \mathbb{R}_{k^+}^n.$$

Let $x(t)$ be the solution of the ODE system (1); similarly to the proof in Lemma 14, we can have

$$(23) \quad \frac{dV_1(t, x(t))}{dt} = f(\tilde{x}) - f(x(t)) \leq 0.$$

From Lemmas 9, 10, and 11, it is straightforward to see that $I(x, \tilde{x}) \geq 0$ and $V_1(t, x) \geq 0$ for any $t \geq t_0 > 0$ and $x \in \mathbb{R}_{k+}^n$. In addition, $I(x, \tilde{x}) = 0 \iff x = \tilde{x}$. Noticing from Lemma 9 that $f(\tilde{x}) - f(x) + (x - \tilde{x})^T \nabla f(x) \geq 0$ for any $x \in \mathbb{R}_{k+}^n$, if we can prove

$$\lim_{t \rightarrow +\infty} V_1(t, x(t)) = 0,$$

then the limit set $\Omega^1(x^0)$ only contains a single point.

Now we prove $\lim_{t \rightarrow +\infty} V_1(t, x(t)) = 0$. For any $T > t_0$ and $x(T) \in \mathbb{R}_{k++}^n$ (guaranteed by Theorems 6 and 15), we can define $V_2(t, x)$ in (9) as follows:

$$V_2(t, x) = V(t, x, x(T)).$$

Then we have

$$\frac{dV_2(t, x(t))}{dt} = f(x(T)) - f(x(t)).$$

This, (23), and Theorem 18 imply

$$\frac{d(V_1(t, x(t)) - V_2(t, x(t)))}{dt} = f(\tilde{x}) - f(x(T)) \leq 0.$$

Noticing $V_2(T, x(T)) = V(T, x(T), x(T)) = 0$, we have

$$V_1(T, x(T)) - V_2(T, x(T)) = V_1(T, x(T)) \leq V_1(t_0, x(t_0)) - V_2(t_0, x(t_0)).$$

Next we show that $V_1(t_0, x(t_0)) - V_2(t_0, x(t_0))$ can be sufficiently small for some T big enough. Since $x(T) \in \mathbb{R}_{k++}^n$, we know $B(x(T)) = \{1, \dots, k\}$ and $N(x(T)) = \emptyset$. If $\gamma_2 = \frac{1}{2}$,

$$\begin{aligned} & V_1(t_0, x(t_0)) - V_2(t_0, x(t_0)) \\ &= \frac{\gamma_1}{2} \sum_{i=k+1}^n [(x_i^0 - \tilde{x}_i)^2 - (x_i^0 - x(T)_i)^2] + t_0 [f(\tilde{x}) - f(x(T)) + (x(T) - \tilde{x})^T \nabla f(x^0)] \\ &+ \sum_{i=1}^k \gamma_1 (x(T)_i - \tilde{x}_i) + \sum_{i \in N(\tilde{x})} \gamma_1 x(T)_i \ln \frac{x_i^0}{x(T)_i} - \sum_{i \in B(\tilde{x})} \gamma_1 \left(\tilde{x}_i \ln \frac{x_i^0}{\tilde{x}_i} - x(T)_i \ln \frac{x_i^0}{x(T)_i} \right). \end{aligned}$$

For $i \in B(\tilde{x})$, $\tilde{x}_i > 0$, and for $i \in N(\tilde{x})$, $\tilde{x}_i = 0$. Noticing

$$\lim_{a \rightarrow 0} a \ln \frac{x_i^0}{a} = 0$$

for any $x_i^0 > 0$, $i = 1, \dots, k$, therefore $V_1(t_0, x(t_0)) - V_2(t_0, x(t_0))$ can be sufficiently small if $\|x(T) - \tilde{x}\|$ is sufficiently small. However $\tilde{x} \in \Omega^1(x^0)$, hence, we can choose some T big enough such that $\|x(T) - \tilde{x}\|$ is sufficiently small. So for any $\epsilon > 0$, we can choose some T big enough such that

$$V_1(t_0, x(t_0)) - V_2(t_0, x(t_0)) = V(t_0, x(t_0), \tilde{x}) - V(t_0, x(t_0), x(T)) \leq \epsilon.$$

If $\frac{1}{2} < \gamma_2 < 1$, we can get this similarly. Then we get $V_1(T, x(T)) \leq \epsilon$. But $V_1(t, x(t))$ is nonincreasing from (23), hence we can get

$$(24) \quad \lim_{t \rightarrow +\infty} V_1(t, x(t)) = 0.$$

Thus the theorem is proved. □

Interestingly, we can show that the limit point of the solution of the ODE system (1) has the maximal number of the positive components in $\{x_1, \dots, x_k\}$ among the optimal solutions of problem (P).

THEOREM 20. *The limit point of the solution of the ODE system (1) has the maximal number of the positive components in $\{x_1, \dots, x_k\}$ among the optimal solutions of problem (P).*

Proof. Let Ω^* denote the optimal solution set of problem (P), and define

$$B^* = \{i \mid \exists x \in \Omega^* \text{ such that } x_i > 0, 1 \leq i \leq k\}.$$

Evidently, for any $x \in \Omega^*$, $B(x) \subseteq B^*$. If $B^* = \emptyset$, then the theorem is evident. Thus in the following, we assume $B^* \neq \emptyset$ and use $|B^*|$ to denote the number of elements in B^* . Then there must exist some $\bar{x} \in \Omega^*$ such that $B(\bar{x}) = B^*$. In fact, for every $i \in B^* \subseteq \{1, \dots, k\}$, according to the definition of B^* , there exists a point $x^{(i)} \in \Omega^*$ such that $x_i^{(i)} > 0$, hence, for

$$\bar{x} = \sum_{i \in B^*} \frac{1}{|B^*|} x^{(i)},$$

$B(\bar{x}) = B^*$, and it is easy to see that $\bar{x} \in \Omega^*$. Now we can see the maximum number of the positive components in $\{x_1, \dots, x_k\}$ among the optimal solutions of problem (P) is actually $|B^*|$. Furthermore, for any optimal solution \tilde{x} which has the maximal number of the positive components in $\{x_1, \dots, x_k\}$ among the optimal solutions, since $B(\tilde{x}) \subseteq B^*$ and the number of the elements in $B(\tilde{x})$ is $|B^*|$, we have $B(\tilde{x}) = B^*$.

Without loss of generality, we can assume that the optimal solution x^* in Assumption 1 has the maximal number of the positive components in $\{x_1, \dots, x_k\}$ among the optimal solutions. Then we have $B(x^*) = B^*$. We consider the function $I_1(t, x(t))$ in Lemma 14. From the proof of Lemma 14, we have for any $t \in [t_0, +\infty)$,

$$(25) \quad I_1(t, x(t)) \leq I_1(t_0, x(t_0)),$$

where $x(t)$ is the solution of the ODE system (1). Noticing that for $i \in N(x^*)$, $x_i^* = 0$, we can rewrite $I_1(t, x(t))$ as follows:

$$I_1(t, x(t)) = \gamma_1 \sum_{i=k+1}^n \frac{1}{2} (x_i - x_i^*)^2 + t [f(x^*) - f(x) + (x - x^*)^T \nabla f(x)]$$

$$+ \gamma_1 \cdot \begin{cases} \sum_{i \in N(x^*)} x_i + \sum_{i \in B(x^*)} \left(x_i - x_i^* - x_i^* \cdot \ln \frac{x_i}{x_i^*} \right) & \text{if } \gamma_2 = \frac{1}{2}, \\ \sum_{i \in N(x^*)} \frac{x_i^{2-2\gamma_2}}{2-2\gamma_2} + \sum_{i \in B(x^*)} \left[\frac{x_i^{2-2\gamma_2} - (x_i^*)^{2-2\gamma_2}}{2-2\gamma_2} \right. \\ \left. - \frac{x_i^*}{1-2\gamma_2} \left(\frac{1}{x_i^{2\gamma_2-1}} - \frac{1}{(x_i^*)^{2\gamma_2-1}} \right) \right] & \text{if } \frac{1}{2} < \gamma_2 < 1. \end{cases}$$

For $k + 1 \leq i \leq n$, we have $\frac{1}{2}(x_i - x_i^*)^2 \geq 0$. From Lemma 9, $f(x^*) - f(x) + (x - x^*)^T \nabla f(x) \geq 0$. For $i \in N(x^*)$, we have $x_i > 0$ and $\frac{x_i^{2-2\gamma_2}}{2-2\gamma_2} > 0$. For $i \in B(x^*)$, from Lemmas 10 and 11, we know $(x_i - x_i^* - x_i^* \cdot \ln \frac{x_i}{x_i^*})$ and

$$\frac{x_i^{2-2\gamma_2} - (x_i^*)^{2-2\gamma_2}}{2 - 2\gamma_2} - \frac{x_i^*}{1 - 2\gamma_2} \left(\frac{1}{x_i^{2\gamma_2-1}} - \frac{1}{(x_i^*)^{2\gamma_2-1}} \right)$$

are both nonnegative and will go to $+\infty$ if $x_i \rightarrow 0$. Hence, from (25), we can see that for each $i \in B(x^*)$, the i th component of the solution of the ODE system (1) $x_i(t)$ is bounded below by some positive constant c_i . Therefore, the limit point \hat{x} of the solution of the ODE system (1) must satisfy $B(x^*) \subseteq B(\hat{x})$. On the other hand, $B(\hat{x}) \subseteq B^* = B(x^*)$. Hence $B(\hat{x}) = B(x^*) = B^*$. Thus the proof is completed. \square

Next we present an example in [7] to show the trajectories and limiting behaviors of the central path and our generalized central paths. The examples in [7] have the following form,

$$\begin{aligned} \min \quad & F(x, y) \\ \text{s.t.} \quad & y \geq 0, \end{aligned}$$

where $x \in \mathbb{R}$ and $y \in \mathbb{R}$ are variables and the solutions are on the x -axis $\{(x, 0) | x \in \mathbb{R}\}$. We choose a class C^2 example which is described in section 5.2 of [7]. Since the parameter ϵ_K (we use K instead of k , since k is already used in defining problem (P)) needs to satisfy $0 < |\epsilon_K| \leq \min\{c_K/4, |\epsilon_{K-1}|\}$ (see (15) in [7]), we choose $|\epsilon_K| = \min\{3c_K/32, |\epsilon_{K-1}|\}$. In this example, $z = (x, y)^T$ and for $K = 0, 1, \dots, y_K = 2^{-K}, a_K = y_K^2, b_K = \frac{y_K^3}{3}$,

$$r_K = \frac{1}{2} \min\{b_K - (b_{K-1} + a_{K-1}(y_K - y_{K-1})), b_{K-1} - (b_K + a_K(y_{K-1} - y_K))\}.$$

$c_K = \frac{1}{4} \min\{r_K, r_{K+1}\}$, and $\epsilon_K = (-1)^K |\epsilon_K|$. Furthermore, for $K = 0, 1, \dots$,

$$f_K(x, y) = a_K(y - y_K) + b_K + c_K g_K(z) + \epsilon_K x,$$

where $g_K(z) = (\frac{g_K^0(z)}{2})^3$ and

$$g_K^0(z) = \frac{1}{2} \left(\sqrt{(x+1)^2 + (y-y_K)^2} + \sqrt{(x-1)^2 + (y-y_K)^2} - 2 \right).$$

In section 5 of [7], many interesting properties for this example are given. For $y = y_K = 2^{-K}$ ($K \geq 0$) and $F(\cdot, y_K) = f_K(\cdot, y_K)$, a direct calculation yields

$$\arg \min_{x \in \mathbb{R}} F(x, y_K) = \{(-1)^{K+1}(1 + \tau_K)\},$$

where $0 < \tau_K = (8|\epsilon_K|/3c_K)^{\frac{1}{2}} \leq 0.5$. Hence the central path will be a zigzag path. We plot the central path (see Figure 1(a)) and our generalized central paths with different initial points (see Figures 1 and 2). For the generalized central paths, we let $\gamma_1 = 1, \gamma_2 = 0.75$, and $t_0 = 1$. The matrix $[\gamma_1 I_n + t DP_{AD} D \nabla^2 f(x)]$ in the ODE system (1) is invertible everywhere, a MATLAB ODE solver ode23s is used to compute the trajectories of the ODE system (1). For the central path, a MATLAB code provided by Karas is used. In Figure 1(a), cp represents the central path while

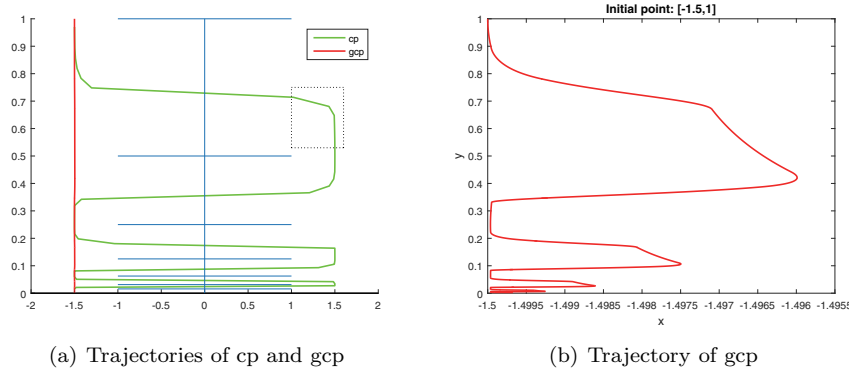


FIG. 1. Trajectories of the central path (cp) and generalized central paths (gcp).

gcp represents the generalized central paths, both paths having the same initial point $(-1.5, 1)$. Figure 1(b) is just a much magnified display of the generalized central path in Figure 1(a). Figure 2 depicts the trajectories of the generalized central paths with 6 different initial points.

Figures 1 and 2 clearly show that the central path is a zigzag path with a large loop and the generalized central paths do converge. Even though the generalized central paths swing back and forth, the magnitude of the swing is quite small and becomes smaller and smaller as t goes to infinity. This phenomenon can be explained partly by the potential function. Let $z^* = (x^*, 0)$ be an accumulation point of the generalized central paths. Then the potential function $V(t, z, z^*)$ has the following form

$$V(t, z, z^*) = \frac{1}{2}(x - x^*)^2 + 2\sqrt{y} + t [f(z^*) - f(z) + (z - z^*)^T \nabla f(z)].$$

Since

$$\frac{dV(t, z(t), z^*)}{dt} = f(z^*) - f(z) < 0,$$

from Lemma 9, we know $f(z^*) - f(z) + (z - z^*)^T \nabla f(z) \geq 0$, hence

$$\frac{1}{2}(x - x^*)^2 + 2\sqrt{y} \leq V(t_0, z^0, z^*) + \int_{t_0}^t (f(z^*) - f(z(\tau))) d\tau$$

for all $t \geq t_0$. In fact, from the proof of Theorem 19, we have $V(t, z(t), z^*) \rightarrow 0$ as $t \rightarrow +\infty$. Thus the magnitude of the swing of the generalized central paths cannot be very large and should become smaller and smaller. We think the reason that the central path does not converge in this example is that the barrier function only contains y , hence does not have any restriction on x . However, for the barrier function $-\frac{1}{\alpha_1}y^{\alpha_1}$ ($0 < \alpha_1 < 1$), if the term $\frac{1-\alpha_1}{2}(x-\bar{x}_0)^2$ is added, where \bar{x}_0 is any fixed number, to the barrier function, the corresponding ODE system can be described by our ODE system (1). The added term $\frac{1-\alpha_1}{2}(x-\bar{x}_0)^2$ can be regarded as a restriction on x , hence the resulting path may converge.

At the end of this section, we address some concerns about the cost of computing the right-hand side in the ODE system (1) and the choice of parameters γ_1 and γ_2 . First, the inverse of AD^2A^T is required in the ODE system (1), and the resulting computing cost is $O(m^2n + m^3)$. Also the inverse of an $n \times n$ matrix is involved in the

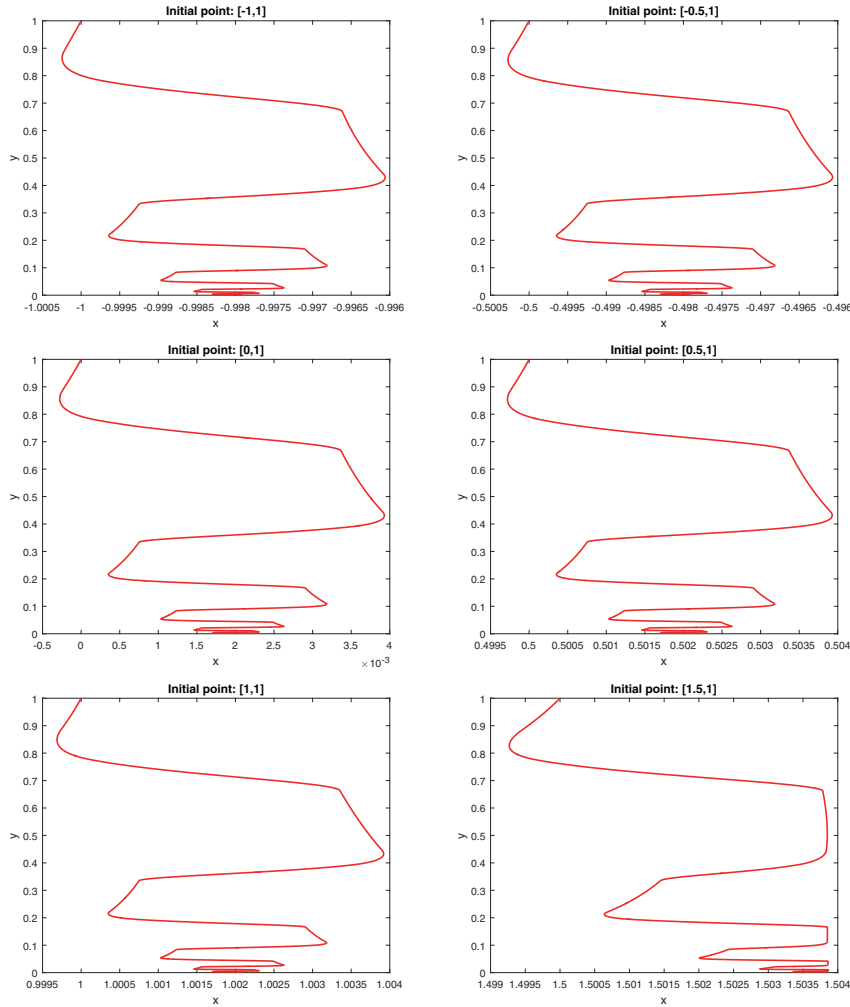


FIG. 2. Trajectories of the generalized central paths with different initial points.

ODE system (1), and thus we need to solve an $n \times n$ linear system. In any solution scheme for the ODE system (1), it appears that the solution of the linear system must be exact so that the next point would remain feasible. Therefore, the resulting cost for such an exact solution would be very high, but this is not true. Noticing the equivalent form (13) and $ADP_{AD} \equiv 0$, the equality constraint of the next iterate is always guaranteed even if the $n \times n$ linear system is solved approximately.

Next we study the impact of parameters γ_1 and γ_2 in the ODE system (1) by the integral of the residual in the objective function. According to (23), we have

$$\int_{t_0}^t (f(x(\tau)) - f(\tilde{x}))d\tau = V_1(t_0, x^0) - V_1(t, x(t)),$$

and along with (24), we can obtain

$$(26) \quad \int_{t_0}^{+\infty} (f(x(t)) - f(\tilde{x}))dt = V_1(t_0, x^0) = V(t_0, x^0, \tilde{x});$$

this describes the integral of the residual in the objective function. In the ODE system (1), γ_1 should not be too close to 0 since the condition number of the $n \times n$ matrix may go to infinity as $\gamma_1 \rightarrow 0$. If $\frac{1}{2} < \gamma_2 < 1$, from (9), (11), and (26), we can see if $\tilde{x}_i = 0$ for some $1 \leq i \leq k$, then $\tilde{V}(t_0, x^0, \tilde{x}) \rightarrow +\infty$ as $\gamma_2 \rightarrow 1$, which indicates that γ_2 should stay away from 1.

5. Concluding remarks. As illustrated by the linearly constrained convex programming examples in [7], the central path may fail to converge. To circumvent this drawback, in this paper, the generalized central paths are introduced for linearly constrained convex programming. Under a very mild assumption on the existence of an optimal solution, we have proved that the generalized central paths always converge to the optimal solution(s) of the optimization problem from any starting interior feasible point. With the introduction of the generalized central paths, especially their representation as the solutions of the ODE system (1) with any starting point x^0 in \mathcal{P}^{++} , some efficient numerical algorithms might be developed. This is a future research topic.

Appendix. Derivations of ODE systems (5) and (6): We restate the system (4) as follows:

$$(27) \quad \nabla^2 f(x) \frac{dx}{d\mu} - \frac{dz}{d\mu} - A^T \frac{dy}{d\mu} = 0,$$

$$(28) \quad A \frac{dx}{d\mu} = 0,$$

$$(29) \quad (1 - \alpha_1) X^{-\alpha_1} Z \frac{dx}{d\mu} + X^{1-\alpha_1} \frac{dz}{d\mu} = \alpha_1 \beta_1 e,$$

where $Z = \text{diag}(z_1, z_2, \dots, z_n)$. From (29), we have

$$\frac{dz}{d\mu} = -(1 - \alpha_1) X^{-1} Z \frac{dx}{d\mu} + \alpha_1 \beta_1 X^{\alpha_1 - 1} e.$$

Then from (27), we have

$$(30) \quad A^T \frac{dy}{d\mu} = \nabla^2 f(x) \frac{dx}{d\mu} + (1 - \alpha_1) X^{-1} Z \frac{dx}{d\mu} - \alpha_1 \beta_1 X^{\alpha_1 - 1} e.$$

Multiplying AXZ^{-1} on both sides of (30) and using (28), we have

$$AXZ^{-1} A^T \frac{dy}{d\mu} = AXZ^{-1} \nabla^2 f(x) \frac{dx}{d\mu} - \alpha_1 \beta_1 AX^{\alpha_1} Z^{-1} e,$$

which implies

$$(31) \quad \frac{dy}{d\mu} = (AXZ^{-1} A^T)^{-1} \left[AXZ^{-1} \nabla^2 f(x) \frac{dx}{d\mu} - \alpha_1 \beta_1 AX^{\alpha_1} Z^{-1} e \right].$$

From (30) and (31), we have

$$\begin{aligned} & [A^T (AXZ^{-1} A^T)^{-1} AXZ^{-1} \nabla^2 f(x) - \nabla^2 f(x) - (1 - \alpha_1) X^{-1} Z] \frac{dx}{d\mu} \\ &= -\alpha_1 \beta_1 X^{\alpha_1 - 1} e + \alpha_1 \beta_1 A^T (AXZ^{-1} A^T)^{-1} AX^{\alpha_1} Z^{-1} e. \end{aligned}$$

Moreover, from the third equation in system (3), we have $Z = \mu\alpha_1\beta_1X^{\alpha_1-1}$. Hence we have

$$\begin{aligned} & \left[A^T(AX^{2-\alpha_1}A^T)^{-1}AX^{2-\alpha_1}\nabla^2 f(x) - \nabla^2 f(x) - \mu(1-\alpha_1)\alpha_1\beta_1X^{\alpha_1-2} \right] \frac{dx}{d\mu} \\ &= -\alpha_1\beta_1X^{\alpha_1-1}e + \alpha_1\beta_1A^T(AX^{2-\alpha_1}A^T)^{-1}AXe. \end{aligned}$$

From the first equation and the third equation in system (3), we have

$$\begin{aligned} & -\alpha_1\beta_1X^{\alpha_1-1}e + \alpha_1\beta_1A^T(AX^{2-\alpha_1}A^T)^{-1}AXe \\ &= -\frac{1}{\mu}z + \frac{1}{\mu}A^T(AX^{2-\alpha_1}A^T)^{-1}AX^{2-\alpha_1}z \\ &= -\frac{1}{\mu} \left[I - A^T(AX^{2-\alpha_1}A^T)^{-1}AX^{2-\alpha_1} \right] (\nabla f(x) - A^T y) \\ &= -\frac{1}{\mu} \left[I - A^T(AX^{2-\alpha_1}A^T)^{-1}AX^{2-\alpha_1} \right] \nabla f(x). \end{aligned}$$

Hence

$$\begin{aligned} & \left[A^T(AX^{2-\alpha_1}A^T)^{-1}AX^{2-\alpha_1}\nabla^2 f(x) - \nabla^2 f(x) - \mu(1-\alpha_1)\alpha_1\beta_1X^{\alpha_1-2} \right] \frac{dx}{d\mu} \\ &= -\frac{1}{\mu} \left[I - A^T(AX^{2-\alpha_1}A^T)^{-1}AX^{2-\alpha_1} \right] \nabla f(x). \end{aligned}$$

Multiplying $\frac{1}{\mu}X^{2-\alpha_1}$ on both sides of the above equation, we can have

$$\frac{dx}{d\mu} = \frac{1}{\mu^2} \left[\gamma_1 I_n + \frac{1}{\mu} DP_{AD} D \nabla^2 f(x) \right]^{-1} DP_{AD} D \nabla f(x),$$

where $D = X^{\gamma_2}$, $\gamma_1 = (1-\alpha_1)\alpha_1\beta_1$, and $\gamma_2 = 1 - \frac{\alpha_1}{2}$. Thus we get the ODE system (5). Let $t = \frac{1}{\mu}$ in the above equation. Then we have

$$\frac{dx}{dt} = \frac{dx}{d\mu} \cdot \frac{d\mu}{dt} = -\mu^2 \frac{dx}{d\mu} = - \left[\gamma_1 I_n + t DP_{AD} D \nabla^2 f(x) \right]^{-1} DP_{AD} D \nabla f(x).$$

Therefore we get the ODE system (6).

Acknowledgments. The authors would like to thank Professor E. Karas for providing the MATLAB code of the central path in Figure 1(a). The authors are also in debt to the two anonymous referees for their constructive comments and suggestions on the earlier version of this paper. In particular, the simpler proof of Lemma 12 was suggested by a referee. We really appreciate their valuable inputs.

REFERENCES

- [1] I. ADLER AND R. D. C. MONTEIRO, *Limiting behavior of the affine scaling continuous trajectories for linear programming problems*, Math. Program., 50 (1991), pp. 29–51.
- [2] D. V. ANOSOV, S. KH. ARANSON, V. I. ARNOLD, I. U. BRONSHTEIN, V. Z. GRINES, AND YU. S. IL'YASHENKO, *Ordinary Differential Equations and Smooth Dynamical Systems*, Springer, Berlin, 1988.
- [3] N. BOURBAKI, *Functions of a Real Variable*, Springer, Berlin, 2004.
- [4] S. BOYD AND L. VANDENBERGHE, *Convex Optimization*, Cambridge University Press, Cambridge, 2004.

- [5] M. T. CHU AND M. M. LIN, *Dynamical system characterization of the central path and its variants—A revisit*, SIAM J. Appl. Dyn. Syst., 10 (2011), pp. 887–905.
- [6] L. M. G. DRUMMOND AND B. F. SVAITER, *On well definedness of the central path*, J. Optim. Theory Appl., 102 (1999), pp. 223–237.
- [7] J. C. GILBERT, C. C. GONZAGA, AND E. KARAS, *Examples of ill-behaved central paths in convex optimization*, Math. Program., 103 (2005), pp. 63–94.
- [8] O. GÜLER, *Limiting behavior of the weighted central paths in linear programming*, Math. Program., 65 (1994), pp. 347–363.
- [9] M. HALICKÁ, E. DE KLERK, AND C. ROOS, *On the convergence of the central path in semidefinite optimization*, SIAM J. Optim., 12 (2002), pp. 1090–1099.
- [10] R. A. HORN AND C. R. JOHNSON, *Matrix Analysis*, Cambridge University Press, New York, 1985.
- [11] A. N. IUSEM, B. F. SVAITER, AND J. X. DA CRUZ NETO, *Central paths, generalized proximal point methods, and Cauchy trajectories in Riemannian manifolds*, SIAM J. Control Optim., 37 (1999), pp. 566–588.
- [12] M. KOJIMA, S. MIZUNO, AND T. NOMA, *Limiting behavior of trajectories by a continuation method for monotone complementarity problems*, Math. Oper. Res., 15 (1990), pp. 662–675.
- [13] L.-Z. LIAO, *A study of the dual affine scaling continuous trajectories for linear programming*, J. Optim. Theory Appl., 163 (2014), pp. 548–568.
- [14] V. LOSERT AND E. AKIN, *Dynamics of games and genes: Discrete versus continuous time*, J. Math. Biology, 17 (1983), pp. 241–251.
- [15] L. MCLINDEN, *An analogue of Moreau’s proximation theorem, with application to the nonlinear complementarity problem*, Pacific J. Math., 88 (1980), pp. 101–161.
- [16] N. MEGIDDO, *Pathways to the optimal set in linear programming*, in Progress in Mathematical Programming, Springer, New York, 1989, pp. 131–158.
- [17] N. MEGIDDO AND M. SHUB, *Boundary behavior of interior point algorithms for linear programming*, Math. Oper. Res., 14 (1989), pp. 97–146.
- [18] R. D. C. MONTEIRO AND T. TSUCHIYA, *Limiting behavior of the derivatives of certain trajectories associated with a monotone horizontal linear complementarity problem*, Math. Oper. Res., 21 (1996), pp. 793–814.
- [19] R. D. C. MONTEIRO AND F. J. ZHOU, *On the existence and convergence of the central path for convex programming and some duality results*, Comput. Optim. Appl., 10 (1998), pp. 51–77.
- [20] K. B. PETERSEN AND M. S. PEDERSEN, *The Matrix Cookbook*, Technical report, Technical University of Denmark, Kongens Lyngby, Denmark, 2008.
- [21] C.-K. SIM AND G. ZHAO, *Underlying paths in interior point methods for the monotone semidefinite linear complementarity problem*, Math. Program., 110 (2007), pp. 475–499.
- [22] C.-K. SIM AND G. ZHAO, *Asymptotic behavior of HKM paths in interior point methods for monotone semidefinite linear complementarity problems: General theory*, J. Optim. Theory Appl., 137 (2008), pp. 11–25.
- [23] J. SUN, *A convergence proof for an affine scaling algorithm for convex quadratic programming without nondegeneracy assumptions*, Math. Program., 60 (1993), pp. 69–79.
- [24] S. A. VAVASIS AND Y. YE, *A primal-dual interior point method whose running time depends only on the constraint matrix*, Math. Program., 74 (1996), pp. 79–120.