Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 20, No. 4, pp. 601–627, 2024

CONVERGENCE ANALYSIS OF FOUR INTERIOR POINT CONTINUOUS TRAJECTORIES FOR CONVEX SEMIDEFINITE PROGRAMMING

Xun Qian, Li-Zhi Liao* and Jie Sun†

This paper is dedicated to Professor Masao Fukushima in celebration of his 75th birthday.‡

Abstract: We study four interior continuous trajectories defined by matrix differential equations, aiming at developing new interior point methods for convex semidefinite programming (SDP). By assuming the boundedness of the level set, we establish the optimality and convergence of the first and the second trajectories for linear SDP. For the convex case, we show that, starting from any interior feasible point, the third trajectory converges to an optimal solution that has the maximal rank among all optimal solutions, under the assumptions that an optimal solution exists and the maximal rank of optimal solutions is one. Finally, we obtain the strongest result under the weakest assumption for the fourth trajectory, namely, by only assuming the existence of an optimal solution, we show that the trajectory converges to an optimal solution that has the maximal rank among all optimal solutions.

Key words: *continuous trajectory, interior point method, convex semidefinite programming, ordinary differential equation*

Mathematics Subject Classification: *34D23, 90C22, 90C51*

1 Introduction

We study four interior point continuous trajectories for convex semidefinite programming. First we introduce some notations. Let $\mathcal{S}^n$ denote the vector space of real symmetric $n \times n$ matrices. The standard inner product on $\mathcal{S}^n$ is

$$A \bullet B = \operatorname{tr}(AB) = \sum_{i,j} A_{ij} B_{ij}, \quad \operatorname{tr}(\cdot) = \operatorname{trace}(\cdot).$$

*Corresponding author. The work of L.-Z. Liao was supported in part by grants HKBU12302019 and HKBU12300920 from General Research Fund (GRF) of Hong Kong.

†The work of J. Sun was supported in part by Australia Research Council under Grant DP-160102918.

‡Masao was a funding co-editor of Pacific Journal of Optimization (2005–2022). Under his leadership, the journal has become a flagship journal of the Pacific Optimization Research Activity Group and an important publication platform for optimization researchers in the world. We have known Masao since 1990s and he has been a long-term colleague and close friend of us. We wish him the best at this glorious moment of his life.

DOI: https://doi.org/10.61208/pjo-2023-042

By $X \succeq 0$ ($X \succ 0$), where $X \in \mathcal{S}^n$, we mean that X is positive semidefinite (positive definite). Consider the following convex semidefinite programming (**SDP**) problem

$$\begin{array}{ll} \min\limits_{X \in \mathcal{S}^n} & f(X) \\ s.t. & A_k \bullet X = b_k,\ k = 1, \ldots, m, \\ & X \succeq 0, \end{array} \tag{P}$$

where $f : \mathcal{S}^n \to R$ is convex, $b \in R^m$, and $A_k \in \mathcal{S}^n$, $k = 1, \ldots, m$. As a blanket assumption, we assume that the optimal value for problem (P) is finite and attainable, therefore, we use min rather than inf in problem (P).

The following notations are used in our later discussions.

$$\begin{aligned} &\mathcal{S}^n_+ = \{X \in \mathcal{S}^n | X \succeq 0\}, \quad \mathcal{S}^n_{++} = \{X \in \mathcal{S}^n | X \succ 0\}, \\ &\mathcal{P}^+ = \{X \in \mathcal{S}^n | A_k \bullet X = b_k,\ k = 1, \ldots, m,\ X \succeq 0\}, \\ &\mathcal{P}^{++} = \{X \in \mathcal{S}^n | A_k \bullet X = b_k,\ k = 1, \ldots, m,\ X \succ 0\}, \end{aligned}$$

where $\mathcal{P}^{++}$ is called the relative interior of $\mathcal{P}^+$. It is conventional to assume that $\mathcal{P}^{++}$ is nonempty in the analysis of interior point methods. A comprehensive study of semidefinite programming can be found in [39].

There are many interior point algorithms for solving problem (P), for example, [1, 20, 21, 31, 33, 40] for linear $f(X)$, and [15, 23, 34, 35] for convex and quadratic $f(X)$. There are also some continuous methods for linear SDP. For example, a recurrent neural network for real-time SDP was proposed and studied in [14]; in [19], it was shown that the primal-dual central path converges to the analytic center of the primal-dual optimal solution set under the strict complementarity assumption. Many of the interior point algorithms for SDP are primal-dual path-following algorithms that are closely related to the central path [37]. In the linear case with $f(X) = \mathrm{tr}(CX)$ where $C \in \mathcal{S}^n$, the central path is the set of the solutions of the following system with the parameter $\mu > 0$ [39]

$$\left\{ \begin{array}{l} A_k \bullet X = b_k,\ k = 1, \ldots, m, \\ \sum_{k=1}^m y_k A_k + Z = C, \\ XZ = \mu I,\ X \succeq 0,\ Z \succeq 0, \end{array} \right. \tag{1.1}$$

where I is the identity matrix. In [27], Shida and Shindoh studied the existence and convergence of the infeasible central path for the monotone semidefinite complementarity problem and showed that for the monotone semidefinite linear complementarity problem, the trajectory converges to the analytic center of the solution set provided that there exists a strictly complementary solution. Under the assumption of primal and dual strict feasibility, Goldfarb and Scheinberg [6] showed that the primal and dual central paths exist and converge to the analytic centers of the optimal faces of the primal and the dual problems, respectively. But later, Halická *et al*. [10] showed that the result is not correct in the absence of strict complementarity by a counterexample, where the central path converges to a different optimal solution, and they also gave a short proof that the central path always converges in SDP by using ideas from algebraic geometry. The dynamical system characterization of the central path and its variants in linear programming (LP) and SDP was also studied in [4]. Furthermore, the study of limiting behavior of some infeasible weighted central paths for SDP can be found in [11, 18, 24]. There is also some research work on the central path for nonlinear SDP, for instance, [7, 8, 16]. For problem (P), López and Ramírez [16] showed the convergence of the central path where the logarithm barrier function is used under the analyticity of $f(X)$ by a similar method to [10], and other central paths defined with a large class of penalty and barrier functions were also studied there.

It should be noted that there have been some studies on other continuous trajectories. Sim and Zhao [28] 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 strict complementarity condition in [29]. The relationship between the interior point methods and the underlying paths is also discussed in [28].

In this paper we are particularly interested in the interior point continuous trajectories for problem (P). In order to write down the equations explicitly, we need the following notations.

- Let svec map $\mathcal{S}^n$ to $R^{n(n+1)/2}$. If $U \in \mathcal{S}^n$, then $\mathrm{svec}(U)$ is defined by

$$\mathrm{svec}(U) := (u_{11}, \sqrt{2}u_{21}, \ldots, \sqrt{2}u_{n1}, u_{22}, \sqrt{2}u_{32}, \ldots, \sqrt{2}u_{n2}, \ldots, u_{nn})^T.$$

- The symmetrized Kronecker product $\otimes_s$ is defined by

$$(G \otimes_s K)\mathrm{svec}(H) = \frac{1}{2}\mathrm{svec}(KHG^T + GHK^T),$$

  where $G, K \in R^{n\times n}$ and $H \in \mathcal{S}^n$. The properties of operator $\otimes_s$ can be found in Appendix of [1] and [29].

- Let matrix $\mathcal{A}$ be defined as follows

$$\mathcal{A} = \begin{pmatrix} \mathrm{svec}(A_1)^T \\ \vdots \\ \mathrm{svec}(A_m)^T \end{pmatrix} \in R^{m\times n(n+1)/2}.$$

- For any $X \in \mathcal{S}^n$, let $\nabla^2 f(X)$ be the following matrix

$$\nabla^2 f(X) = \begin{pmatrix} \mathrm{svec}(\frac{\partial(\frac{\partial f}{\partial X})_{11}}{\partial X})^T \\ \sqrt{2}\mathrm{svec}(\frac{\partial(\frac{\partial f}{\partial X})_{21}}{\partial X})^T \\ \vdots \\ \sqrt{2}\mathrm{svec}(\frac{\partial(\frac{\partial f}{\partial X})_{n1}}{\partial X})^T \\ \mathrm{svec}(\frac{\partial(\frac{\partial f}{\partial X})_{22}}{\partial X})^T \\ \sqrt{2}\mathrm{svec}(\frac{\partial(\frac{\partial f}{\partial X})_{32}}{\partial X})^T \\ \vdots \\ \sqrt{2}\mathrm{svec}(\frac{\partial(\frac{\partial f}{\partial X})_{n2}}{\partial X})^T \\ \vdots \\ \mathrm{svec}(\frac{\partial(\frac{\partial f}{\partial X})_{nn}}{\partial X})^T \end{pmatrix} \in R^{n(n+1)/2\times n(n+1)/2}.$$

Now we present the following four ordinary differential equation (ODE) systems

$$\mathrm{svec}(\dot{X}) = -\big(I - (X \otimes_s X)P_{\mathcal{A}X}\big)(X \otimes_s X)\mathrm{svec}\left(\frac{\partial f}{\partial X}\right), \tag{1.2}$$

$$\mathrm{svec}(\dot{X}) = -\Big(I + t\big(I - (X \otimes_s X)P_{\mathcal{A}X}\big)\big((X \otimes_s X)\nabla^2 f(X)\big)\Big)^{-1} \big(I - (X \otimes_s X)P_{\mathcal{A}X}\big)(X \otimes_s X)\mathrm{svec}\left(\frac{\partial f}{\partial X}\right), \tag{1.3}$$

$$\mathrm{svec}(\dot{X}) = -\big(I - (X \otimes_s X^{\frac{1}{2}})P_{\mathcal{A}X^{\frac{1}{2}}}\big)(X \otimes_s X^{\frac{1}{2}})\mathrm{svec}\left(\frac{\partial f}{\partial X}\right), \tag{1.4}$$

$$\mathrm{svec}(\dot{X}) = -\Big(I + t\big(I - (X \otimes_s X^{\frac{1}{2}})P_{\mathcal{A}X^{\frac{1}{2}}}\big)\big((X \otimes_s X^{\frac{1}{2}})\nabla^2 f(X)\big)\Big)^{-1} \big(I - (X \otimes_s X^{\frac{1}{2}})P_{\mathcal{A}X^{\frac{1}{2}}}\big)(X \otimes_s X^{\frac{1}{2}})\mathrm{svec}\left(\frac{\partial f}{\partial X}\right), \tag{1.5}$$

which have the same initial condition: $X(t_0) = X^0 \in \mathcal{P}^{++}$ and $t_0 > 0$, where

$$X \in \mathcal{S}^n_{++}, \quad X^{\frac{1}{2}} \in \mathcal{S}^n_{++} \text{ is the unique square root matrix of X,}$$
$$P_{\mathcal{A}X^\gamma} = \mathcal{A}^T(\mathcal{A}(X \otimes_s X^\gamma)\mathcal{A}^T)^{-1}\mathcal{A}, \ \gamma \in \left\{\frac{1}{2},\ 1\right\},$$
$$I \text{ stands for the } \frac{n(n+1)}{2} \times \frac{n(n+1)}{2} \text{ identity matrix.}$$

For ODE systems (1.3) and (1.5), we sometimes use the following equivalent implicit forms

$$\mathrm{svec}(\dot{X}) = -\big(I - (X \otimes_s X)P_{\mathcal{A}X}\big)(X \otimes_s X)\left(t\nabla^2 f(X)\mathrm{svec}(\dot{X}) + \mathrm{svec}\left(\frac{\partial f}{\partial X}\right)\right), \tag{1.6}$$

$$\mathrm{svec}(\dot{X}) = -\big(I - (X \otimes_s X^{\frac{1}{2}})P_{\mathcal{A}X^{\frac{1}{2}}}\big)(X \otimes_s X^{\frac{1}{2}})\left(t\nabla^2 f(X)\mathrm{svec}(\dot{X}) + \mathrm{svec}\left(\frac{\partial f}{\partial X}\right)\right). \tag{1.7}$$

For ODE systems (1.2) and (1.4), we need $f(X) \in C^2$ on $\mathcal{S}^n_+$, and for ODE systems (1.3) and (1.5), we need $f(X) \in C^3$ on $\mathcal{S}^n_+$.

The right-hand side of ODE (1.2) comes from the affine scaling direction for SDP in [22]. The right-hand side of ODE (1.3) comes from the central path. In fact, in the above central path system (1.1), if we replace the matrix C by $\frac{\partial f}{\partial X}$ and take the derivative with respect to μ, we can get

$$\mathrm{svec}\left(\frac{dX}{d\mu}\right) = \frac{1}{\mu^2}\Big(I + \frac{1}{\mu}(I - (X \otimes_s X)\mathcal{A}^T(\mathcal{A}(X \otimes_s X)\mathcal{A}^T)^{-1}\mathcal{A})((X \otimes_s X)\nabla^2 f(X))\Big)^{-1} (I - (X \otimes_s X)\mathcal{A}^T(\mathcal{A}(X \otimes_s X)\mathcal{A}^T)^{-1}\mathcal{A})(X \otimes_s X)\mathrm{svec}\left(\frac{\partial f}{\partial X}\right), \tag{1.8}$$

then we use the new variable t by setting $t = \frac{1}{\mu}$, and we have

$$\mathrm{svec}\left(\frac{dX}{dt}\right) = -\Big(I + t(I - (X \otimes_s X)\mathcal{A}^T(\mathcal{A}(X \otimes_s X)\mathcal{A}^T)^{-1}\mathcal{A})((X \otimes_s X)\nabla^2 f(X))\Big)^{-1}$$

$$(I - (X \otimes_s X)\mathcal{A}^T(\mathcal{A}(X \otimes_s X)\mathcal{A}^T)^{-1}\mathcal{A})(X \otimes_s X)\text{svec}\left(\frac{\partial f}{\partial X}\right), \tag{1.9}$$

which is exactly ODE system (1.3) except the initial points, where (1.1) or (1.9) requires the initial point must be on the central path but (1.3) only requires the initial point $X^0 \in \mathcal{P}^{++}$. ODEs (1.4) and (1.5) are some variants of ODEs (1.2) and (1.3), respectively. This kind of variants also exists in the linearly constrained smooth optimization [36]. For linear programming, ODE systems (1.3) and (1.5) reduce to ODE systems (1.2) and (1.4), respectively, and actually become the search direction in (6) of [36] with $\gamma = 1$ and $\gamma = \frac{3}{4}$ respectively. Later we will show that the solutions of ODE systems (1.4) and (1.5) have stronger properties under weaker conditions than the solutions of ODE systems (1.2) and (1.3), and the reason is that the potential functions of ODE systems (1.4) and (1.5) are bounded below naturally, while those of ODE systems (1.2) and (1.3) may not. It should be noted that the solutions of the four ODE systems define four interior point (verified in Section 2) continuous trajectories for problem (P).

For simplicity, in what follows, we use $\|\cdot\|_2$ to denote either the vector 2-norm or the matrix 2-norm. C^k stands for the class of kth order continuously differentiable functions. Unless otherwise specified, x_j denotes the jth component of a vector x, e denotes the column vector of all ones, and e_i denotes the unity column vector whose ith component is 1, the dimensions of e and e_i are clear from the context. For any index subset $J \subseteq \{1, \ldots, n\}$, we denote by x_J the vector composed of those components of $x \in \mathbb{R}^n$ indexed by $j \in J$, rank(Q) denotes the rank of matrix Q. For any $Q \in \mathcal{S}^n_+$, $\lambda_{\max}(Q)$ and $\lambda_{\min}(Q)$ denote the largest and smallest eigenvalues of Q, respectively.

The rest of this paper is organized as follows. In Section 2, we (i) introduce four potential functions for the four ODE systems (1.2), (1.3), (1.4), and (1.5), respectively; (ii) verify that each ODE system has a unique solution in $[t_0, +\infty)$; and (iii) show the weak convergence for the solution of ODE system (1.2). In Section 3, we prove that every accumulation point of the solutions of the four ODE systems is an optimal solution for problem (P), and show the weak convergence for the solution of ODE system (1.3). In Section 4, we first show the strong convergence of the solutions of the last two ODE systems under certain conditions, and verify that each limiting point has the maximal rank among the optimal solution set of problem (P), then we prove the convergence for the solutions of the first two ODE systems in the linear case. Finally, some conclusions are drawn in Section 5.

2 Properties of the Continuous Trajectories

The following assumptions are made throughout this paper.

Assumption 1. *There exists a point $X^* \in \mathcal{P}^+$ such that $f(X^*)$ is the optimal value of problem (P).*

Assumption 2. *The matrix $\mathcal{A}$ has full row rank m.*

Assumption 3. *For ODE systems (1.2) and (1.4), we assume $f(X) \in C^2$ on $\mathcal{S}^n_+$, and for ODE systems (1.3) and (1.5), we assume $f(X) \in C^3$ on $\mathcal{S}^n_+$.*

Theorem 2.1. *$P_{\mathcal{A}X^\gamma} \in C^1$ on $\mathcal{S}^n_{++}$, $\gamma \in \{\frac{1}{2},\ 1\}$.*

Proof. According to Assumption 2, $\mathcal{A}$ has full row rank. If $X \in \mathcal{S}^n_{++}$, $X \otimes_s X$ and $X \otimes_s X^{\frac{1}{2}}$ are both symmetric and positive definite. So $\mathcal{A}(X \otimes_s X)\mathcal{A}^T$ and $\mathcal{A}(X \otimes_s X^{\frac{1}{2}})\mathcal{A}^T$ are also symmetric and positive definite, thus invertible.

Notice that the inverse of a matrix and $X \otimes_s X$ are both continuous differentiable, we get $P_{\mathcal{A}X} \in C^1$ on $\mathcal{S}^n_{++}$. Furthermore, according to Chapter 6 in [12] we have

$$X^{\frac{1}{2}} = \frac{2}{\pi} X \int_0^\infty (t^2 I + X)^{-1} dt,$$

which indicates that the square root of a symmetric positive definite matrix is continuous differentiable. So $P_{\mathcal{A}X^{\frac{1}{2}}} \in C^1$ on $\mathcal{S}^n_{++}$. Thus the proof is complete. □

Lemma 2.2. *If A, $B \in \mathbb{R}^{n\times n}$ are both symmetric and positive semidefinite, then all eigenvalues of AB are nonnegative.*

Proof. We give two proofs.

First proof: From Corollary 4.6.3 on page 99 in [38], the result is evident.

Second proof: From Theorem 1.3.20 in [13], for any matrices $X \in \mathbb{R}^{n\times n}$ and $Y \in \mathbb{R}^{n\times n}$, we have that XY has the same eigenvalues as YX. By setting $X = A^{\frac{1}{2}}$ and $Y = A^{\frac{1}{2}}B$, we obtain that AB has the same eigenvalues as $A^{\frac{1}{2}}BA^{\frac{1}{2}}$, which is symmetric and positive semidefinite. Hence all eigenvalues of AB are nonnegative. □

Theorems 2.3 and 2.4 below guarantee the existence, uniqueness, and feasibility for the solutions of the four ODE systems (1.2), (1.3), (1.4), and (1.5).

Theorem 2.3. *For each of the four ODE systems (1.2), (1.3), (1.4), and (1.5), there exists a unique solution $X(t)$ with a maximal existence interval $[t_0, \alpha_1)$, $[t_0, \alpha_2)$, $[t_0, \beta_1)$, and $[t_0, \beta_2)$, respectively. In addition, $X(t) \succ 0$ on the existence intervals for all four ODE systems.*

Proof. For ODE system (1.3), notice that

$$\big(I - (X \otimes_s X)P_{\mathcal{A}X}\big)(X \otimes_s X) = (X^{\frac{1}{2}} \otimes_s X^{\frac{1}{2}})(I - X^{\frac{1}{2}} \otimes_s X^{\frac{1}{2}} P_{\mathcal{A}X} X^{\frac{1}{2}} \otimes_s X^{\frac{1}{2}})(X^{\frac{1}{2}} \otimes_s X^{\frac{1}{2}}).$$

Since $I - X^{\frac{1}{2}} \otimes_s X^{\frac{1}{2}} P_{\mathcal{A}X} X^{\frac{1}{2}} \otimes_s X^{\frac{1}{2}}$ is symmetric and idempotent, we know it is positive semidefinite. So $\big(I - (X \otimes_s X)P_{\mathcal{A}X}\big)(X \otimes_s X)$ is symmetric and positive semidefinite. Since $f(X) \in C^3$ on $\mathcal{S}^n_+$ and is convex, we have $\nabla^2 f(X)$ is symmetric and positive semidefinite. From Lemma 2.2, we know that for any $t > 0$, $I + t\big(I - (X \otimes_s X)P_{\mathcal{A}X}\big)\big((X \otimes_s X)\nabla^2 f(X)\big)$ is always invertible.

For ODE system (1.5), since $X \otimes_s X^{\frac{1}{2}}$ is also symmetric and positive definite, similarly, we can get that for any $t > 0$, $I + t\big(I - (X \otimes_s X^{\frac{1}{2}})P_{\mathcal{A}X^{\frac{1}{2}}}\big)\big((X \otimes_s X^{\frac{1}{2}})\nabla^2 f(X)\big)$ is also invertible.

Now from Assumption 3 and Theorem 2.1, along with the fact that the inverse of a matrix is continuous differentiable, we know the right-hand sides of the four ODE systems are all continuous differentiable and thus locally Lipschitz continuous on $(0, +\infty) \times \mathcal{S}^n_{++}$. Since $(0, +\infty) \times \mathcal{S}^n_{++}$ is an open set, from the Cauchy-Peano theorem and Picard-Lindelöf theorem, for each of the four ODE systems (1.2), (1.3), (1.4), and (1.5), there exists a unique solution $X(t)$ with a maximal existence interval $[t_0, \alpha_1)$, $[t_0, \alpha_2)$, $[t_0, \beta_1)$, and $[t_0, \beta_2)$, respectively.

Because the right-hand sides of the four ODE systems are all defined on the open set $(0, +\infty) \times \mathcal{S}^n_{++}$ and the initial points are also symmetric and positive definite, the solutions of the four ODE systems are of course in the open set $\mathcal{S}^n_{++}$, so they are all symmetric and positive definite on the existence intervals. □

Later in this section, it will be shown that $\alpha_1 = +\infty$, $\alpha_2 = +\infty$ (Theorem 2.12) and $\beta_1 = +\infty$, $\beta_2 = +\infty$ (Theorem 2.13). To simplify presentation in the sequel, in the remaining of this paper, $X(t)$ will be replaced by X whenever no confusion would occur.

Theorem 2.4. *Each of the unique solutions $X(t)$ of the four ODE systems (1.2), (1.3), (1.4), and (1.5) satisfies $\mathcal{A}\mathrm{svec}(X(t)) = b$ on its own maximal existence interval.*

Proof. For the four ODE systems (1.2), (1.3), (1.4), and (1.5), we know that for any t belonging to their own maximal existence interval, the unique solutions $X(t)$ satisfy

$$X(t) = X^0 + \int_{t_0}^{t} (\dot{X}|_{t=\tau}) d\tau.$$

Notice

$$\mathcal{A}\big(I - (X \otimes_s X)P_{\mathcal{A}X}\big)(X \otimes_s X) = \mathcal{A}(X \otimes_s X) - \mathcal{A}(X \otimes_s X) = 0$$

and

$$\mathcal{A}\big(I - (X \otimes_s X^{\frac{1}{2}})P_{\mathcal{A}X^{\frac{1}{2}}}\big)(X \otimes_s X^{\frac{1}{2}}) = \mathcal{A}(X \otimes_s X^{\frac{1}{2}}) - \mathcal{A}(X \otimes_s X^{\frac{1}{2}}) = 0,$$

we can get (for ODE systems (1.3) and (1.5), we use the implicit forms (1.6) and (1.7) instead)

$$\mathcal{A}\mathrm{svec}(\dot{X}) = 0,$$

so

$$\mathcal{A}\mathrm{svec}(X(t)) = \mathcal{A}X^0 + \int_{t_0}^{t} \mathcal{A}\mathrm{svec}(\dot{X}|_{t=\tau}) d\tau = \mathcal{A}X^0 = b.$$

Thus the theorem is proved. □

Theorem 2.5. *Let $X(t)$ be any unique solution of the four ODE systems (1.2), (1.3), (1.4), and (1.5). Then $f(X(t))$ is a nonincreasing function of t on its own maximal existence interval.*

Proof. For ODE systems (1.2) and (1.3), we use $\mathcal{X}$ to denote $X^{\frac{1}{2}} \otimes_s X^{\frac{1}{2}}$, and $\mathcal{P}$ to denote $I - \mathcal{X}P_{\mathcal{A}X}\mathcal{X}$. For ODE systems (1.4) and (1.5), we use $\mathcal{X}$ to denote $(X \otimes_s X^{\frac{1}{2}})^{\frac{1}{2}}$, and $\mathcal{P}$ to denote $I - \mathcal{X}P_{\mathcal{A}X^{\frac{1}{2}}}\mathcal{X}$. From Theorem 2.3 it is clear that $\mathcal{X}$ and $\mathcal{P}$ are all symmetric and positive semidefinite and $\mathcal{P}^2 = \mathcal{P}$.

Now we can write ODE systems (1.2) and (1.4) in the same form as

$$\mathrm{svec}(\dot{X}) = -\mathcal{X}\mathcal{P}\mathcal{X}\mathrm{svec}\left(\frac{\partial f}{\partial X}\right), \tag{2.1}$$

and can write ODE systems (1.3) and (1.5) in the same form as

$$\mathrm{svec}(\dot{X}) = -(I + t\mathcal{X}\mathcal{P}\mathcal{X}\nabla^2 f(X))^{-1}\mathcal{X}\mathcal{P}\mathcal{X}\mathrm{svec}\left(\frac{\partial f}{\partial X}\right). \tag{2.2}$$

So for ODE systems (1.2) and (1.4), we have

$$\begin{aligned} \frac{df(X(t))}{dt} &= \mathrm{tr}\left(\frac{\partial f}{\partial X}\dot{X}\right) = \mathrm{svec}\left(\frac{\partial f}{\partial X}\right)^T \mathrm{svec}(\dot{X}) \\ &= -\mathrm{svec}\left(\frac{\partial f}{\partial X}\right)^T \mathcal{X}\mathcal{P}\mathcal{X}\mathrm{svec}\left(\frac{\partial f}{\partial X}\right) \end{aligned} \tag{2.3}$$

$$= -\left\|\mathcal{P}\mathcal{X}\mathrm{svec}\left(\frac{\partial f}{\partial X}\right)\right\|_2^2 \leq 0. \tag{2.4}$$

Thus $f(X(t))$ is a nonincreasing function of t on its own maximal existence interval for ODE systems (1.2) and (1.4).

Similarly, we can prove the same conclusion for ODE systems (1.3) and (1.5) if we can show that $\big(I + t\mathcal{X}\mathcal{P}\mathcal{X}\nabla^2 f(X)\big)^{-1}\mathcal{X}\mathcal{P}\mathcal{X}$ is a symmetric and positive semidefinite matrix. This is actually true because we have

$$\big(I + t\mathcal{X}\mathcal{P}\mathcal{X}\nabla^2 f(X)\big)^{-1}\mathcal{X}\mathcal{P}\mathcal{X} = \mathcal{X}\mathcal{P}\big(I + t\mathcal{P}\mathcal{X}\nabla^2 f(X)\mathcal{X}\mathcal{P}\big)^{-1}\mathcal{P}\mathcal{X}, \tag{2.5}$$

where the proof is similar to Lemma 12 in [26]. Thus the theorem is proved. □

The following lemma reveals an essential property for any convex function.

Lemma 2.6 (Section 3.1.3, [3])**.** *Suppose f is differentiable (i.e., its gradient ∇f exists at each point in $\mathrm{dom} f$). Then f is convex if and only if $\mathrm{dom} f$ is convex and*

$$f(y) \geq f(x) + \nabla f(x)^T (y - x) \tag{2.6}$$

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

Now we will introduce four potential functions for the four ODE systems, respectively. In 1983, Losert and Akin [17] introduced a kind of potential function for both the discrete and continuous dynamical systems in a classical model of population genetics. Their potential function can be extended for our purpose. The potential function $I_1(X, Y)$ for ODE system (1.2) can be defined as

$$I_1(X, Y) = \ln\det X + \mathrm{tr}(X^{-1}Y), \tag{2.7}$$

where $X \in \mathcal{S}^n_{++}$ is the variable and $Y \in \mathcal{S}^n_+$ is a parameter.

The potential function $I_2(t, X, Y)$ for ODE system (1.3) can be defined as

$$I_2(t, X, Y) = I_1(X, Y) + t\left[f(Y) - f(X) + \mathrm{tr}\Big((X - Y)\frac{\partial f}{\partial X}\Big)\right]. \tag{2.8}$$

where $X \in \mathcal{S}^n_{++}$ and $t > 0$ are variables, and $Y \in \mathcal{S}^n_+$ is a parameter.

The potential function $I_3(X, Y)$ for ODE system (1.4) can be defined as

$$I_3(X, Y) = 2\mathrm{tr}(X^{-\frac{1}{2}}Y) + 2\mathrm{tr}(X^{\frac{1}{2}}), \tag{2.9}$$

where $X \in \mathcal{S}^n_{++}$ is the variable and $Y \in \mathcal{S}^n_+$ is a parameter.

The potential function $I_4(t, X, Y)$ for ODE system (1.5) can be defined as

$$I_4(t, X, Y) = I_3(X, Y) + t\left[f(Y) - f(X) + \mathrm{tr}\Big((X - Y)\frac{\partial f}{\partial X}\Big)\right], \tag{2.10}$$

where $X \in \mathcal{S}^n_{++}$ and $t > 0$ are variables, and $Y \in \mathcal{S}^n_+$ is a parameter.

A direct application of function $I_3(X, Y)$ and $I_4(t, X, Y)$ is the boundedness of the solutions of ODE systems (1.4) and (1.5).

Theorem 2.7. *The unique solution $X(t)$ of ODE system (1.4) is contained in a bounded set in $\mathcal{S}_+^n$, and the bound only depends on X^0 and X^*, where X^* is a finite optimal solution for problem (P).*

Proof. According Theorem 2.3 and Assumption 1, we can define

$$V_1(t) = I_3(X, X^*) = 2\mathrm{tr}(X^{-\frac{1}{2}}X^*) + 2\mathrm{tr}(X^{\frac{1}{2}}) \quad \forall t \in [t_0, \beta_1). \tag{2.11}$$

Then from Theorem 2.4, ODE system (1.4), and the properties of $\otimes_s$ in [29], we have

$$\begin{aligned}
\frac{dV_1(t)}{dt} &= -\mathrm{svec}(X^*)^T(X^{-\frac{1}{2}} \otimes_s I)^{-1}(X \otimes_s X)^{-1}\mathrm{svec}(\dot{X}) \\
&\quad +\mathrm{svec}(I)^T(X^{\frac{1}{2}} \otimes_s I)^{-1}\mathrm{svec}(\dot{X}) \\
&= -\mathrm{svec}(X^*)^T(X^{-\frac{1}{2}} \otimes_s I)^{-1}(X \otimes_s X)^{-1}\mathrm{svec}(\dot{X}) \\
&\quad +\mathrm{svec}(X)^T(X^{-\frac{1}{2}} \otimes_s I)^{-1}(X \otimes_s X)^{-1}\mathrm{svec}(\dot{X}) \\
&= \mathrm{svec}(X - X^*)^T(X^{-\frac{1}{2}} \otimes_s I)^{-1}(X \otimes_s X)^{-1}\mathrm{svec}(\dot{X}) \\
&= \mathrm{svec}(X^* - X)^T\big(I - P_{\mathcal{A}X^{\frac{1}{2}}}(X \otimes_s X^{\frac{1}{2}})\big)\mathrm{svec}(\frac{\partial f}{\partial X}) \\
&= \mathrm{svec}(X^* - X)^T\mathrm{svec}(\frac{\partial f}{\partial X}),
\end{aligned} \tag{2.12}$$

where the second equality comes from that

$$\begin{aligned}
(X \otimes_s X)^{-1}(X^{-\frac{1}{2}} \otimes_s I)^{-1}\mathrm{svec}(X) &= (X^{\frac{1}{2}} \otimes_s X)^{-1}\mathrm{svec}(X) \\
&= (X \otimes_s X^{\frac{1}{2}})^{-1}\mathrm{svec}(X) \\
&= \Big((X^{\frac{1}{2}} \otimes_s X^{\frac{1}{2}}) \cdot (X^{\frac{1}{2}} \otimes_s I)\Big)^{-1}\mathrm{svec}(X) \\
&= (X^{\frac{1}{2}} \otimes_s I)^{-1}(X^{\frac{1}{2}} \otimes_s X^{\frac{1}{2}})^{-1}\mathrm{svec}(X) \\
&= (X^{\frac{1}{2}} \otimes_s I)^{-1}(X^{-\frac{1}{2}} \otimes_s X^{-\frac{1}{2}})\mathrm{svec}(X) \\
&= (X^{\frac{1}{2}} \otimes_s I)^{-1}\mathrm{svec}(I).
\end{aligned}$$

From Lemma 2.6, we know

$$\mathrm{svec}(X^* - X)^T\mathrm{svec}\left(\frac{\partial f}{\partial X}\right) \le f(X^*) - f(X) \le 0,$$

so we get

$$\frac{dV_1(t)}{dt} \le f(X^*) - f(X) \le 0. \tag{2.13}$$

Then for any $T \in [t_0, \beta_1)$, we have

$$2\mathrm{tr}(X(T)^{-\frac{1}{2}}X^*) + 2\mathrm{tr}(X(T)^{\frac{1}{2}}) = V_1(T) \le V_1(t_0) = 2\mathrm{tr}(X(t_0)^{-\frac{1}{2}}X^*) + 2\mathrm{tr}(X(t_0)^{\frac{1}{2}}).$$

Since $\|X(T)\|_2 \le \mathrm{tr}(X(T))$ and $\mathrm{tr}(X(T)^{-\frac{1}{2}}X^*) \ge 0$ (Lemma 2.2), we have

$$\|X(T)\|_2 \le \|X(T)^{\frac{1}{2}}\|_2^2 \le \big(\mathrm{tr}(X(T)^{\frac{1}{2}})\big)^2 \le \frac{V_1(t_0)^2}{4},$$

where $\frac{V_1(t_0)^2}{4}$ only depends on X^0 and X^*. ∎

Theorem 2.8. *The unique solution $X(t)$ of ODE system (1.5) is contained in a bounded set in $\mathcal{S}^n_+$, and the bound only depends on X^0 and X^*, where X^* is a finite optimal solution for problem (P).*

Proof. According Theorem 2.3 and Assumption 1, we can define

$$V_2(t) = I_4(t, X, X^*) = I_3(X, X^*) + t\left[f(X^*) - f(X) + \mathrm{tr}\Big((X - X^*)\frac{\partial f}{\partial X}\Big)\right], \qquad (2.14)$$

where $t \in [t_0, \beta_2)$.

Then from Theorem 2.4 and similar to the calculation of $\frac{dV_1(t)}{dt}$ in (2.12) with the implicit form (1.7), we have

$$\begin{aligned}\frac{dV_2(t)}{dt} &= \mathrm{svec}(X^* - X)^T\mathrm{svec}\big(t\nabla^2 f(X)\mathrm{svec}(\dot{X}) \\ &\quad + \frac{\partial f}{\partial X}) + t\mathrm{svec}(\dot{X})^T\nabla^2 f(X)\mathrm{svec}(X - X^*) \\ &\quad + \left[f(X^*) - f(X) + \mathrm{tr}\Big((X - X^*)\frac{\partial f}{\partial X}\Big)\right] \\ &= f(X^*) - f(X) \le 0, \end{aligned} \qquad (2.15)$$

so for any $T \in [t_0, \beta_2)$, from the definition (2.14), we have

$$V_2(T) \le V_2(t_0) = I_3(X^0, X^*) + t_0\left[f(X^*) - f(X^0) + \mathrm{tr}\Big((X^0 - X^*)\frac{\partial f}{\partial X}|_{X=X^0}\Big)\right].$$

From Lemma 2.6, we know

$$f(X^*) - f(X(T)) + \mathrm{tr}\Big((X(T) - X^*)\frac{\partial f}{\partial X}|_{X=X(T)}\Big) \ge 0,$$

this along with $\|X(T)\|_2 \le \mathrm{tr}(X(T))$ and $\mathrm{tr}(X(T)^{-\frac{1}{2}}X^*) \ge 0$ implies

$$\|X(T)\|_2 \le \|X(T)^{\frac{1}{2}}\|_2^2 \le \big(\mathrm{tr}(X(T)^{\frac{1}{2}})\big)^2 \le \frac{V_2(t_0)^2}{4},$$

where $\frac{V_2(t_0)^2}{4}$ only depends on X^0 and X^*. □

For ODE systems (1.2) and (1.3), we need additional conditions to guarantee the boundedness of the solutions for the general convex function $f(X)$.

Theorem 2.9. *For ODE systems (1.2) and (1.3), if the level set $\{X \in \mathcal{P}^+ | f(X) \le f(X^0)\}$ is bounded, then the unique solutions $X(t)$ of ODE systems (1.2) and (1.3) are contained in a bounded set in $\mathcal{S}^n_+$.*

Proof. From Theorem 2.5, for the unique solutions $X(t)$ of ODE systems (1.2) and (1.3) we have $f(X(t)) \le f(X^0)$, then $X(t)$ will be contained in the level set $\{X \in \mathcal{P}^+ | f(X) \le f(X^0)\}$ which is bounded according to the assumption. □

However for ODE systems (1.2) and (1.3), if $f(X)$ is linear, then we do not need the boundedness of the level set to guarantee the boundedness of the solutions. First, we state a lemma which will be used later.

Lemma 2.10 (Theorem 4.3.26, [13])**.** *Let A be Hermitian. The vector of diagonal entries of A majorizes the vector of eigenvalues of A.*

According to [13], a vector β is said to majorize a vector α if

$$\min\left\{\sum_{j=1}^{k}\beta_{i_j} : 1 \le i_1 < \cdots < i_k \le n\right\} \ge \min\left\{\sum_{j=1}^{k}\alpha_{i_j} : 1 \le i_1 < \cdots < i_k \le n\right\},$$

for any $k = 1, 2, \ldots, n$ with equality for $k = n$.

Theorem 2.11. *If $f(X) = \mathrm{tr}(CX)$, where $C \in R^{n\times n}$ is a symmetric matrix, then the unique solutions $X(t)$ of ODE systems (1.2) and (1.3) are contained in a bounded set in $\mathcal{S}^n_+$.*

Proof. For ODE system (1.2), from Theorem 2.3, for any $T \in [t_0, \alpha_1)$, $X(T) \succ 0$, so we define

$$\tilde{V}(t) = \mathrm{tr}\big(X^{-1}(X(T) - X^*)\big),$$

where $t \in [t_0, \alpha_1)$. From Theorem 2.4,

$$\begin{aligned}
\frac{d\tilde{V}(t)}{dt} &= -\mathrm{tr}\big(X^{-1}(X(T) - X^*)X^{-1}\dot{X}\big) \\
&= -\mathrm{svec}(X(T) - X^*)^T(X^{-1} \otimes_s X^{-1})\mathrm{svec}(\dot{X}) \\
&= \mathrm{svec}(X(T) - X^*)^T(X^{-1} \otimes_s X^{-1})\big(I - (X \otimes_s X)P_{\mathcal{A}X}\big)(X \otimes_s X)\mathrm{svec}\left(\frac{\partial f}{\partial X}\right) \\
&= \mathrm{svec}(X(T) - X^*)^T\big(I - P_{\mathcal{A}X}(X \otimes_s X)\big)\mathrm{svec}(C) \\
&= \mathrm{svec}(X(T) - X^*)^T\mathrm{svec}(C) = f(X(T)) - f(X^*) \ge 0,
\end{aligned}$$

then

$$\tilde{V}(t_0) \le \tilde{V}(T) = n - \mathrm{tr}\big(X(T)^{-1}X^*\big) \le n. \tag{2.16}$$

From the eigenvalue decomposition, we have $X(T) = Q(T)\Lambda(T)Q(T)^T$, then

$$\begin{aligned}
\tilde{V}(t_0) &= \mathrm{tr}((X^0)^{-1}X(T)) - \mathrm{tr}((X^0)^{-1}X^*) \\
&= \mathrm{tr}\big(Q(T)^T(X^0)^{-1}Q(T)\Lambda(T)\big) - \mathrm{tr}((X^0)^{-1}X^*).
\end{aligned}$$

From Lemma 2.10, the diagonal entries of $Q(T)^T(X^0)^{-1}Q(T)$ are all greater than

$$\lambda_{\min}(Q(T)^T(X^0)^{-1}Q(T)) = \lambda_{\min}\left((X^0)^{-1}\right),$$

thus

$$\mathrm{tr}(Q(T)^T(X^0)^{-1}Q(T)\Lambda(T)) \ge \lambda_{\min}\left((X^0)^{-1}\right)\mathrm{tr}(\Lambda(T)).$$

From the above inequality and (2.16) we have

$$\begin{aligned}
\|X(T)\|_2 &\le \mathrm{tr}(X(T)) = \mathrm{tr}(\Lambda(T)) \\
&\le \frac{1}{\lambda_{\min}\left((X^0)^{-1}\right)}\mathrm{tr}(Q(T)^T(X^0)^{-1}Q(T)\Lambda(T)) \\
&\le \frac{1}{\lambda_{\min}\left((X^0)^{-1}\right)}\big(n + \mathrm{tr}((X^0)^{-1}X^*)\big).
\end{aligned}$$

So $X(T)$ is bounded, and the bound depends only on X^0 and X^*. Notice that if $f(X)$ is linear, ODE systems (1.2) and (1.3) are the same. □

Theorem 2.12. *If the unique solutions $X(t)$ of ODE systems (1.2) and (1.3) are contained in a bounded set in $\mathcal{S}^n_+$, then the maximal existence interval for $X(t)$ can be extended to infinity, that is, $\alpha_1 = \alpha_2 = +\infty$.*

Proof. First, we prove this for ODE system (1.2) by contradiction. According to the Extension Theorem in §2.5, [2], we know that the solution $X(t)$ will go to the boundary of the open set $(0, +\infty) \times \mathcal{S}^n_{++}$. If $\alpha_1 \neq +\infty$, $X(t)$ will go to the boundary of $\mathcal{S}^n_{++}$, but from the condition, $X(t)$ is contained in a bounded set in $\mathcal{S}^n_+$. Then $\lambda_{\min}(X(t)) \to 0$ as $t \to \alpha_1$. So $\ln\det(X(t)) \to -\infty$ as $t \to \alpha_1$.

Let us define

$$V_3(t) = \ln\det(X),$$

where $t \in [t_0, \alpha_1)$ and X (or $X(t)$) is the unique solution of ODE system (1.2). Then from (2.1) and using the same notations $\mathcal{X}$ and $\mathcal{P}$ as in the proof of Theorem 2.5, we have

$$\frac{dV_3(t)}{dt} = \mathrm{tr}(X^{-1}\dot{X}) = \mathrm{svec}(I)^T(X^{-\frac{1}{2}} \otimes_s X^{-\frac{1}{2}})\mathrm{svec}(\dot{X}) = -\mathrm{svec}(I)^T\mathcal{P}\mathcal{X}\mathrm{svec}\left(\frac{\partial f}{\partial X}\right).$$

Since $\|\mathcal{P}\|_2 \leq 1$, along with Assumption 3 and the assumption that $X(t)$ is contained in a bounded set in $\mathcal{S}^n_+$, there exists a bound $M_1 > 0$ such that

$$\left|\frac{dV_3(t)}{dt}\right| \leq M_1.$$

Then for any $t \in [t_0, \alpha_1)$, we have

$$V_3(t) \geq V_3(t_0) - M_1(\alpha_1 - t_0),$$

which is contrary with $\ln\det(X) \to -\infty$ as $t \to \alpha_1$. So the hypothesis is not true, thus $\alpha_1 = +\infty$.

For ODE system (1.3), from (2.2) and (2.5),

$$\mathrm{svec}(\dot{X}) = -\mathcal{X}\mathcal{P}(I + t\mathcal{P}\mathcal{X}\nabla^2 f(X)\mathcal{X}\mathcal{P})^{-1}\mathcal{P}\mathcal{X}\mathrm{svec}\left(\frac{\partial f}{\partial X}\right),$$

since

$$\|(I + t\mathcal{P}\mathcal{X}\nabla^2 f(X)\mathcal{X}\mathcal{P})^{-1}\|_2 = \lambda_{\max}((I + t\mathcal{P}\mathcal{X}\nabla^2 f(X)\mathcal{X}\mathcal{P})^{-1}) \leq 1,$$

we know

$$\begin{aligned}\frac{dV_3(t)}{dt} &= \mathrm{tr}(X^{-1}\dot{X}) = \mathrm{svec}(I)^T(X^{-\frac{1}{2}} \otimes_s X^{-\frac{1}{2}})\mathrm{svec}(\dot{X}) \\ &= -\mathrm{svec}(I)^T\mathcal{P}(I + t\mathcal{P}\mathcal{X}\nabla^2 f(X)\mathcal{X}\mathcal{P})^{-1}\mathcal{P}\mathcal{X}\mathrm{svec}\left(\frac{\partial f}{\partial X}\right)\end{aligned}$$

is also bounded under the Assumption 3 and the assumption that $X(t)$ is contained in a bounded set in $\mathcal{S}^n_+$.

Then we can prove $\alpha_2 = +\infty$ in the same way. □

For ODE systems (1.4) and (1.5), if we can prove that the matrix

$$(\mathcal{A}(X \otimes_s X^{\frac{1}{2}})\mathcal{A}^T)^{-1}\mathcal{A}(X \otimes_s X^{\frac{1}{2}}),$$

is bounded for any bounded subset of $\mathcal{S}^n_{++}$, then their solutions can be extended to infinity by the same way as in the proof of Theorem 2.12. This is true if X is a diagonal matrix [32], but is not correct in the general case. We show this by the following example.

Example. $m = 1, n = 2, \mathcal{A} = (1, 0, 0), h = (0, 0, 1)^T$. For $\epsilon \in (0, 1)$, $X_\epsilon = Q_\epsilon \begin{pmatrix} \epsilon & \\ & 1-\epsilon \end{pmatrix} Q_\epsilon^T$, where $Q_\epsilon = \begin{pmatrix} \sqrt{1-\epsilon} & \sqrt{\epsilon} \\ -\sqrt{\epsilon} & \sqrt{1-\epsilon} \end{pmatrix}$ is an orthogonal matrix. Then

$$\begin{aligned} &\lim_{\epsilon \to 0^+} (\mathcal{A}(X_\epsilon \otimes_s X_\epsilon^{\frac{1}{2}})\mathcal{A}^T)^{-1}\mathcal{A}(X_\epsilon \otimes_s X_\epsilon^{\frac{1}{2}})h \\ = &\lim_{\epsilon \to 0^+} \frac{(1-\epsilon)\epsilon^4 - \frac{1}{2}\epsilon^2(1-\epsilon)^2(\sqrt{\epsilon}+\sqrt{1-\epsilon})^2 + \epsilon(1-\epsilon)^4}{(1-\epsilon)^2\epsilon^3 + \frac{1}{2}\epsilon^2(1-\epsilon)^2(\sqrt{\epsilon}+\sqrt{1-\epsilon})^2 + \epsilon^2(1-\epsilon)^3} = +\infty. \end{aligned}$$

By this example, we also show that $(\mathcal{A}(X \otimes_s X)\mathcal{A}^T)^{-1}\mathcal{A}(X \otimes_s X)$ can be unbounded on certain bounded subset of $\mathcal{S}^n_{++}$. However, by using some potential functions we can still extend the solutions of ODE systems (1.4) and (1.5) to infinity.

Theorem 2.13. *The maximal existence interval for the unique solutions $X(t)$ of ODE systems (1.4) and (1.5) can be extended to infinity, that is, $\beta_1 = \beta_2 = +\infty$.*

Proof. We first prove this for ODE system (1.4) by contradiction. According to the Extension Theorem in §2.5, [2], we know that the solution $X(t)$ will go to the boundary of the open set $(0, +\infty) \times \mathcal{S}^n_{++}$. If $\beta_1 \neq +\infty$, $X(t)$ will go to the boundary of $\mathcal{S}^n_{++}$, but from Theorem 2.7, $X(t)$ is contained in a bounded set in $\mathcal{S}^n_{+}$. So $\lambda_{\min}(X(t)) \to 0$ as $t \to \beta_1$.

Let us define

$$V_4(t) = I_3(X, X^0) = 2\mathrm{tr}(X^{-\frac{1}{2}}X^0) + 2\mathrm{tr}(X^{\frac{1}{2}}),$$

where $t \in [t_0, \beta_1)$ and X (or $X(t)$) is the unique solution of ODE system (1.4).

Then we have

$$\frac{dV_4(t)}{dt} = \mathrm{svec}(X^0 - X)^T \mathrm{svec}\left(\frac{\partial f}{\partial X}\right),$$

from Theorem 2.7 and Assumption 3, we know there exists a bound $M > 0$ which depends only on X^0, X^* and $f(X)$ such that for every $t \in [t_0, \beta_1)$,

$$\left|\frac{dV_4(t)}{dt}\right| = \left|\mathrm{svec}(X^0 - X)^T \mathrm{svec}\Big(\frac{\partial f}{\partial X}\Big)\right| \le M.$$

Hence for any $t \in [t_0, \beta_1)$,

$$V_4(t) \le V_4(t_0) + M(\beta_1 - t_0) < +\infty. \tag{2.17}$$

But $\lambda_{\min}(X(t)) \to 0$ as $t \to \beta_1$. Let $X(t) = Q(t)\Lambda(t)Q(t)^T$ be an eigenvalue decomposition of $X(t)$, then $\mathrm{tr}(X(t)^{-\frac{1}{2}}X^0) = \mathrm{tr}(\Lambda(t)^{-\frac{1}{2}}Q(t)^T X^0 Q(t))$. According to Lemma 2.10, the diagonal entries of $Q(t)^T X^0 Q(t)$ are all greater than

$$\lambda_{\min}(Q(t)^T X^0 Q(t)) = \lambda_{\min}(X^0) > 0.$$

So $\mathrm{tr}(X(t)^{-\frac{1}{2}}X^0) \to +\infty$ as $t \to \beta_1$ which is contrary with (2.17). So the hypothesis is not true, thus $\beta_1 = +\infty$.

For ODE system (1.5), from Theorem 2.8, Lemmas 2.6 and 2.10, we can prove $\beta_2 = +\infty$ in the similar way. ☐

From Theorems 2.12 and 2.13, we can define the limit set for the unique solutions $X(t)$ of the four ODE systems (1.2), (1.3), (1.4), and (1.5). For $i = 1, 2, 3, 4$, the limit set $\Omega^i(X^0)$ of $\{X(t)\}$ of the ODE system $(1, i+1)$ can be defined as follows

$$\Omega^i(X^0) = \left\{ X \in \mathcal{S}^n \mid \exists \ \{t_k\}_{k=0}^{+\infty} \text{ with } \lim_{k\to+\infty} t_k = +\infty \text{ such that } \lim_{k\to+\infty} X(t_k) = X \right\}. \tag{2.18}$$

Theorem 2.14. *If the unique solutions $X(t)$ of ODE systems (1.2) and (1.3) are contained in a bounded set in $\mathcal{S}^n_+$, then for each $i = 1, 2, 3, 4$, the limit set $\Omega^i(X^0)$ is nonempty, compact, and connected. Furthermore $\Omega^i(X^0)$ is contained in $\mathcal{P}^+$, $i = 1, 2, 3, 4$.*

Proof. From Theorems 2.3, 2.4, 2.13, and 2.12, we know that the limit set $\Omega^i(X^0)$ is contained in $\mathcal{P}^+$, $i = 1, 2, 3, 4$. From the proof of Theorem 2.13 and the assumption, we know that the unique solutions $X(t)$ of the four ODE systems are contained in a bounded closed set. So similar to the proof of Theorem 1.1 on page 390 in [5] (the proof in [5] is for $n = 2$, but it can be easily extended to the general case), it can be verified that for each $i = 1, 2, 3, 4$, $\Omega^i(X^0)$ is nonempty, compact, and connected. □

At the end of this section, we prove the weak convergence of ODE system (1.2), i.e., $\text{svec}(\dot{X}) \to 0$ as $t \to +\infty$. First, we need the following Lemma.

Lemma 2.15 (Barbalat's Lemma [30])**.** *If the differentiable function $f(t)$ has a finite limit as $t \to +\infty$, and $\dot{f}$ is uniformly continuous, then $\dot{f} \to 0$ as $t \to +\infty$.*

Theorem 2.16. *If the unique solution $X(t)$ of ODE system (1.2) is contained in a bounded set in $\mathcal{S}^n_+$, then*

$$\lim_{t\to+\infty} \big(I - (X \otimes_s X)P_{\mathcal{A}X}\big)(X \otimes_s X)\text{svec}\left(\frac{\partial f}{\partial X}\right) = 0.$$

Proof. From (2.3) and using the same notations $\mathcal{X}$ and $\mathcal{P}$ as in the proof of Theorem 2.5, we know

$$\begin{aligned} \frac{df(X(t))}{dt} &= -\text{svec}\left(\frac{\partial f}{\partial X}\right)^T \mathcal{X}\mathcal{P}\mathcal{X}\text{svec}\left(\frac{\partial f}{\partial X}\right) \\ &= -\text{svec}\left(\frac{\partial f}{\partial X}\right)^T ((X \otimes_s X) - (X \otimes_s X)P_{\mathcal{A}X}(X \otimes_s X))\text{svec}\left(\frac{\partial f}{\partial X}\right). \end{aligned}$$

From Assumption 3, $f(X) \in C^2$ on $\mathcal{S}^n_+$. Furthermore, $\text{svec}(\dot{X})$ is bounded because $X(t)$ is contained in a bounded set in $\mathcal{S}^n_+$. So if we want to show $\frac{d^2 f(X(t))}{dt^2}$ is bounded, we only need to show that $\frac{d(X\otimes_s X)}{dt}$ and $\frac{d((X\otimes_s X)P_{\mathcal{A}X}(X\otimes_s X))}{dt}$ are both bounded. Notice

$$\frac{d(X \otimes_s X)}{dt} = 2X \otimes_s \dot{X},$$

since $X(t)$ and $\dot{X}$ are both bounded, thus $\frac{d(X\otimes_s X)}{dt}$ is bounded. Notice

$$\begin{aligned} & \frac{d((X \otimes_s X)P_{\mathcal{A}X}(X \otimes_s X))}{dt} \\ = \quad & 2(X \otimes_s \dot{X})P_{\mathcal{A}X}(X \otimes_s X) + 2(X \otimes_s X)P_{\mathcal{A}X}(X \otimes_s \dot{X}) \end{aligned}$$

$$
\begin{aligned}
&-2(X \otimes_s X)P_{\mathcal{A}X}(X \otimes_s \dot{X})P_{\mathcal{A}X}(X \otimes_s X) \\
= \;& 2(X \otimes_s \dot{X})\mathcal{X}^{-1}\mathcal{X}P_{\mathcal{A}X}\mathcal{X}^2 + 2\mathcal{X}^2 P_{\mathcal{A}X}\mathcal{X}\mathcal{X}^{-1}(X \otimes_s \dot{X}) \\
&-2\mathcal{X}^2 P_{\mathcal{A}X}\mathcal{X}\mathcal{X}^{-1}(X \otimes_s \dot{X})\mathcal{X}^{-1}\mathcal{X}P_{\mathcal{A}X}\mathcal{X}^2,
\end{aligned}
$$

where $\mathcal{X}^2 = \mathcal{X} \cdot \mathcal{X} = X \otimes_s X$. Since $\mathcal{X}P_{\mathcal{A}X}\mathcal{X}$ is symmetric and idempotent, it's always bounded. So if we can show $\mathcal{X}^{-1}(X \otimes_s \dot{X})\mathcal{X}^{-1}$ is bounded, then $\frac{d((X\otimes_s X)P_{\mathcal{A}X}(X\otimes_s X))}{dt}$ will be bounded. Let smat be the inverse map of svec. From (2.1), if we denote $B(t) = -\text{smat}(\mathcal{P}\mathcal{X}\text{svec}(\frac{\partial f}{\partial X}))$, then $B(t)$ is also bounded, and $\dot{X} = X^{\frac{1}{2}}B(t)X^{\frac{1}{2}}$. Therefore we get

$$\mathcal{X}^{-1}(X \otimes_s \dot{X})\mathcal{X}^{-1} = (I \otimes_s B(t)),$$

which is bounded. Thus $\frac{d^2 f(X(t))}{dt^2}$ is bounded, and as a consequence, $\frac{df(X(t))}{dt}$ is uniformly continuous. Furthermore from Theorem 2.5 and Assumption 1, $f(X(t))$ has a finite limit as $t \to +\infty$. So from Barbalat's Lemma, we have

$$\lim_{t\to+\infty} \frac{df(X(t))}{dt} = \lim_{t\to+\infty} -\left\| \mathcal{P}\mathcal{X}\text{svec}\left(\frac{\partial f}{\partial X}\right)\right\|_2^2 = 0.$$

□

3 Optimality of the Cluster Point(s)

In this section, we will show that every accumulation point of the solutions of the four ODE systems (1.2), (1.3), (1.4), and (1.5) is an optimal solution for problem (P).

Theorem 3.1. *If the unique solutions $X(t)$ of ODE systems (1.2) and (1.3) are contained in a bounded set in $\mathcal{S}^n_+$, then for any $X^{(1)} \in \Omega^1(X^0)$ (defined in (2.18)) and $X^{(2)} \in \Omega^2(X^0)$, $X^{(1)}$ and $X^{(2)}$ are both optimal solutions for problem (P).*

Proof. We prove this by contradiction. From Theorems 2.5 and 2.12, we know $\lim_{t\to+\infty} f(X(t))$ exists since $f(X)$ is bounded below in $\mathcal{P}^+$. Then if $X^{(1)} \in \Omega^1(X^0)$ is not an optimal solution for problem (P), we have

$$f(X^0) \geq f(X^{(1)}) = \lim_{k\to+\infty} f(X(t_k)) > f(X^*).$$

Let us define

$$Y^{(1)} = \frac{f(X^{(1)}) - f(X^*)}{2(f(X^0) - f(X^*))}X^0 + \left(1 - \frac{f(X^{(1)}) - f(X^*)}{2(f(X^0) - f(X^*))}\right)X^*,$$

then $Y^{(1)} \in \mathcal{P}^{++}$. Since $f(X)$ is convex, we have

$$
\begin{aligned}
f(Y^{(1)}) &\leq \frac{f(X^{(1)}) - f(X^*)}{2(f(X^0) - f(X^*))}f(X^0) + \left(1 - \frac{f(X^{(1)}) - f(X^*)}{2(f(X^0) - f(X^*))}\right)f(X^*) \\
&= \frac{f(X^{(1)}) + f(X^*)}{2}.
\end{aligned}
$$

From (2.7), we can define

$$V_5(t) = I_1(X, Y^{(1)}) = \ln\det X + \text{tr}(X^{-1}Y^{(1)}),$$

where $t \in [t_0, +\infty)$ and X (or $X(t)$) is the unique solution of ODE system (1.2). Then from Theorem 2.4, Lemma 2.6, and the properties of $\otimes_s$ [29], we have

$$\begin{aligned}
\frac{dV_5(t)}{dt} &= \mathrm{tr}(X^{-1}\dot{X}) - \mathrm{tr}(X^{-1}Y^{(1)}X^{-1}\dot{X}) \\
&= \mathrm{svec}(X)^T(X^{-1} \otimes_s X^{-1})\mathrm{svec}(\dot{X}) - \mathrm{svec}(Y^{(1)})^T(X^{-1} \otimes_s X^{-1})\mathrm{svec}(\dot{X}) \\
&= -\mathrm{svec}(X - Y^{(1)})^T(I - P_{\mathcal{A}X})\mathrm{svec}\left(\frac{\partial f}{\partial X}\right) = \mathrm{svec}(Y^{(1)} - X)^T\mathrm{svec}\left(\frac{\partial f}{\partial X}\right) \\
&\le f(Y^{(1)}) - f(X) \le f(Y^{(1)}) - f(X^{(1)}) \le \frac{f(X^{(1)}) + f(X^*)}{2} - f(X^{(1)}) \\
&= \frac{f(X^*) - f(X^{(1)})}{2} < 0,
\end{aligned}$$

where the second inequality comes from the fact that $f(X(t))$ is a nonincreasing function with respect to t. So $V_5(t) \to -\infty$ as $t \to +\infty$. We next show $V_5(t)$ is bounded below.

For any $t \in [t_0, +\infty)$, let $X(t) = Q(t)\Lambda(t)Q(t)^T$ be an eigenvalue decomposition of $X(t)$, and $\{\lambda_i(t)\}_{i=1}^n$ be the eigenvalues of $X(t)$. Then

$$\begin{aligned}
V_5(t) &= \ln\det X(t) + \mathrm{tr}\big(Q(t)\Lambda(t)^{-1}Q(t)^TY^{(1)}\big) \\
&= \sum_{i=1}^n \ln\lambda_i(t) + \mathrm{tr}\big(\Lambda(t)^{-1}Q(t)^TY^{(1)}Q(t)\big),
\end{aligned}$$

since $Y^{(1)} \in \mathcal{P}^{++}$, we have

$$\lambda_{\min}(Q(t)^TY^{(1)}Q(t)) = \lambda_{\min}(Y^{(1)}) > 0.$$

Hence by applying Lemma 2.10 to $Q(t)^TY^{(1)}Q(t)$, we have

$$\begin{aligned}
V_5(t) &= \sum_{i=1}^n \ln\lambda_i(t) + \mathrm{tr}(\Lambda(t)^{-1}Q(t)^TY^{(1)}Q(t)) \\
&\ge \sum_{i=1}^n \ln\lambda_i(t) + \sum_{i=1}^n \lambda_i(t)^{-1}\lambda_{\min}(Y^{(1)}) = \sum_{i=1}^n(\ln\lambda_i(t) + \lambda_i(t)^{-1}\lambda_{\min}(Y^{(1)})) \\
&\ge \sum_{i=1}^n(\ln\lambda_{\min}(Y^{(1)}) + 1) = n(\ln\lambda_{\min}(Y^{(1)}) + 1) > -\infty,
\end{aligned}$$

where the second inequality comes from $\ln\lambda + \lambda^{-1}\lambda_{\min}(Y^{(1)}) \ge \ln\lambda_{\min}(Y^{(1)}) + 1$ for any $\lambda > 0$. It is contrary with $V_5(t) \to -\infty$ as $t \to +\infty$. So the hypothesis is not true, thus $X^{(1)}$ is an optimal solution for problem (P).

As for $X^{(2)} \in \Omega^2(X^0)$, let X (or $X(t)$) be the unique solution of ODE system (1.3). If $X^{(2)} \in \Omega^2(X^0)$ is not an optimal solution for problem (P), we have

$$f(X^0) \ge f(X^{(2)}) = \lim_{k\to+\infty} f(X(t_k)) > f(X^*).$$

We define

$$Y^{(2)} = \frac{f(X^{(2)}) - f(X^*)}{2(f(X^0) - f(X^*))}X^0 + \left(1 - \frac{f(X^{(2)}) - f(X^*)}{2(f(X^0) - f(X^*))}\right)X^*,$$

then $Y^{(2)} \in \mathcal{P}^{++}$, and similar to $f(Y^{(1)})$, we have

$$f(Y^{(2)}) \leq \frac{f(X^{(2)}) + f(X^*)}{2}.$$

Notice if $Y \in \mathcal{P}^{+}$, then

$$\begin{aligned}\frac{dI_2(t,X,Y)}{dt} &= \mathrm{svec}(Y-X)^T \mathrm{svec}\left(t\nabla^2 f(X)\mathrm{svec}(\dot{X}) + \frac{\partial f}{\partial X}\right) \\ &\quad + t\mathrm{svec}(\dot{X})^T \nabla^2 f(X)\mathrm{svec}(X-Y) + \left[f(Y) - f(X) + \mathrm{tr}(X-Y)\frac{\partial f}{\partial X}\right] \\ &= f(Y) - f(X).\end{aligned}$$

Thus we can show that $\frac{dI_2(t,X,Y^{(2)})}{dt} = f(Y^{(2)}) - f(X) \leq \frac{f(X^*) - f(X^{(1)})}{2} < 0$. Then noticing that $I_2(t,X,Y^{(2)})$ is bounded below, we can prove that $X^{(2)}$ is also an optimal solution for problem (P) in the same way as $X^{(1)}$. □

Now we are ready to prove the weak convergence of ODE system (1.3).

Theorem 3.2. *For ODE system (1.3), if the level set* $\{X \in \mathcal{P}^{+} | f(X) \leq f(X^0)\}$ *is bounded, then the unique solution* $X(t)$ *of ODE system (1.3) satisfies*

$$\lim_{t\to+\infty} \dot{X} = 0.$$

Proof. From (2.5) and using the same notations $\mathcal{X}$ and $\mathcal{P}$ as in the proof of Theorem 2.5, we know that if we can prove

$$\lim_{t\to+\infty} - \left\| \mathcal{P}\mathcal{X}\mathrm{svec}\left(\frac{\partial f}{\partial X}\right)\right\|_2 = 0, \tag{3.1}$$

then the theorem holds.

We prove this by contradiction. If (3.1) is not true, there must exist a constant $c_0 > 0$ such that for any $T > t_0$, there always exists a $t > T$ such that $\|\mathcal{P}\mathcal{X}\mathrm{svec}(\frac{\partial f}{\partial X})\|_2 > c_0$.

Let us consider the following cluster of trajectories: each trajectory is defined by the solution of ODE system (1.2) with initial point $X(t)$ at initial time t_0, where $X(t)$ denotes the solution of ODE system (1.3) at time t. We use $\tilde{X}(\tau,t)$ to denote this trajectory. From Theorem 2.5, each trajectory $\tilde{X}(\tau,t)$ is contained in the bounded level set. Then with the same analysis as in the proof of Theorem 2.16, we know there exists a constant $L_0 > 0$ which is independent of t such that

$$\left|\frac{d^2 f(\tilde{X}(\tau,t))}{d\tau^2}\right| \overset{(2.4)}{=} \frac{d\|\mathcal{P}\mathcal{X}\mathrm{svec}(\frac{\partial f}{\partial X})\|_2^2|_{X=\tilde{X}(\tau,T_1)}}{d\tau} \leq L_0. \tag{3.2}$$

From the hypothesis and Theorem 3.1, there exists a $T_1 > t_0$, such that

$$f(\tilde{X}(t_0,T_1)) = f(X(T_1)) < f(X^*) + \frac{c_0^4}{4L_0}, \tag{3.3}$$

and

$$\left\|\mathcal{P}\mathcal{X}\mathrm{svec}\left(\frac{\partial f}{\partial X}\right)\right\|_2 |_{X=\tilde{X}(t_0,T_1)} = \left\|\mathcal{P}\mathcal{X}\mathrm{svec}\left(\frac{\partial f}{\partial X}\right)\right\|_2 |_{X=X(T_1)} > c_0. \tag{3.4}$$

From (3.2) and (3.4), we have

$$\left\|\mathcal{P}\mathcal{X}\text{svec}\left(\frac{\partial f}{\partial X}\right)\right\|_2^2 |_{X=\tilde{X}(\tau,T_1)} \geq \max(c_0^2 - L_0(\tau - t_0), 0),$$

then from (2.4), we can obtain

$$\begin{aligned}\int_{t_0}^{+\infty} -\frac{df(\tilde{X}(\tau,T_1))}{d\tau} d\tau &= \int_{t_0}^{+\infty} \left\|\mathcal{P}\mathcal{X}\text{svec}\left(\frac{\partial f}{\partial X}\right)\right\|_2^2 |_{X=\tilde{X}(\tau,T_1)} d\tau \\ &\geq \int_{t_0}^{+\infty} \max(c_0^2 - L_0(\tau - t_0), 0) d\tau = \frac{c_0^4}{2L_0},\end{aligned}$$

however, from Theorem 3.1 and (3.3), we have

$$\int_{t_0}^{+\infty} -\frac{df(\tilde{X}(\tau,T_1))}{d\tau} d\tau = f(X(T_1)) - f(X^*) < \frac{c_0^4}{4L_0},$$

which contradicts with the previous inequality. Thus the theorem is proved. □

Theorem 3.3. *For any $X^{(3)} \in \Omega^3(X^0)$ and $X^{(4)} \in \Omega^4(X^0)$, $X^{(3)}$ and $X^{(4)}$ are both optimal solutions for problem (P).*

Proof. We prove this by contradiction. Similar to the proof of Theorem 3.1, if $X^{(3)} \in \Omega^3(X^0)$ is not an optimal solution for problem (P), then

$$f(X^{(3)}) = \lim_{k\to+\infty} f(X(t_k)) > f(X^*).$$

From (2.13), we can see $V_1(t)$ defined by (2.11) will go to $-\infty$ as $t \to +\infty$. However,

$$V_1(t) = I_3(X, X^*) = 2\text{tr}(X^{-\frac{1}{2}}X^*) + 2\text{tr}(X^{\frac{1}{2}}) \geq 0 \quad \forall t \in [t_0, +\infty),$$

where X (or $X(t)$) is the unique solution for ODE system (1.4). So the hypothesis is not true, thus $X^{(3)}$ must be an optimal solution for problem (P).

As for $X^{(4)} \in \Omega^4(X^0)$, by using $V_2(t)$ defined in (2.14), the property in (2.15), and the fact that $V_2(t)$ is also bounded below by zero, we can prove that $X^{(4)}$ is also an optimal solution for problem (P) by contradiction in the same way as $X^{(3)}$. □

4 Convergence of the Continuous Trajectories

Now, it comes to the key results of the paper. Theorem 4.1 below shows that if the maximal rank among the optimal solution set of problem (P) is equal to one, then the solution of ODE system (1.4) converges as $t \to +\infty$. Theorem 4.2 shows that the solution of ODE system (1.5) always converges as $t \to +\infty$. Theorem 4.6 shows that in the linear case of $f(X)$, the solutions of ODE systems (1.2) and (1.3) also converge as $t \to +\infty$.

Theorem 4.1. *Every point in the limit set $\Omega^3(X^0)$ has the maximal rank among the optimal solution set of problem (P). Furthermore, if the maximal rank among the optimal solution set of problem (P) is equal to one, then the limit set $\Omega^3(X^0)$ only contains a single point.*

Proof. From Theorem 2.14, we know that $\Omega^3(X^0)$ is not empty. So we can choose a point $\bar{X} \in \Omega^3(X^0)$, and evidently $\bar{X} \in \mathcal{P}^+$. Without loss of generality, we assume the optimal solution X^* has the maximal rank among the optimal solution set of problem (P), and $rank(X^*) = r$. Let $X^* = Q\Lambda Q^T$ be an eigenvalue decomposition of X^* and

$$\Lambda = \begin{pmatrix} \Lambda_1 & \\ & 0 \end{pmatrix},$$

where Λ_1 is a $r \times r$ diagonal matrix and Λ_1 is invertible. Since X^* has the maximal rank among the optimal solution set and $\bar{X}$ is an optimal solution (Theorem 3.1), $rank(\bar{X}) \leq rank(X^*)$. Following the same claim as Lemma 4.1 in [6], there exists an eigenvalue decomposition $\bar{X} = \bar{Q}\bar{\Lambda}\bar{Q}^T$ with

$$\bar{\Lambda} = \begin{pmatrix} \bar{\Lambda}_1 & \\ & 0 \end{pmatrix},$$

where $\bar{\Lambda}_1$ is a $r \times r$ diagonal matrix, and a sequence $\{\bar{t}_k\}_{k=1}^{+\infty}$ with $\lim_{k \to +\infty} \bar{t}_k = +\infty$ such that $X(\bar{t}_k) \to \bar{X}$, $Q(\bar{t}_k) \to \bar{Q}$, and $\Lambda(\bar{t}_k) \to \bar{\Lambda}$, where $Q(\bar{t}_k)\Lambda(\bar{t}_k)Q(\bar{t}_k)^T$ is an eigenvalue decomposition of $X(\bar{t}_k)$ with $\Lambda(\bar{t}_k) = \begin{pmatrix} \Lambda_1(\bar{t}_k) & \\ & \Lambda_2(\bar{t}_k) \end{pmatrix}$, $\Lambda_1(\bar{t}_k) \in R^{r \times r}$. Notice $V_1(t)$ defined by (2.11) is a nonincreasing function in $[t_0, +\infty)$ and bounded below, we know $V_1(t)$ has a finite limit as $t \to +\infty$. Therefore,

$$\begin{aligned}
\lim_{t \to +\infty} V_1(t) = \lim_{k \to +\infty} V_1(\bar{t}_k) &= \lim_{k \to +\infty} \left[2\mathrm{tr}(X(\bar{t}_k)^{-\frac{1}{2}} X^*) + 2\mathrm{tr}(X(\bar{t}_k)^{\frac{1}{2}}) \right] \\
&= \lim_{k \to +\infty} \left[2\mathrm{tr}(Q(\bar{t}_k)\Lambda(\bar{t}_k)^{-\frac{1}{2}} Q(\bar{t}_k)^T Q\Lambda Q^T) + 2\mathrm{tr}(Q(\bar{t}_k)\Lambda(\bar{t}_k)^{\frac{1}{2}} Q(\bar{t}_k)^T) \right] \\
&= \lim_{k \to +\infty} \left[2\mathrm{tr}(\Lambda(\bar{t}_k)^{-\frac{1}{2}} Q(\bar{t}_k)^T Q\Lambda Q^T Q(\bar{t}_k)) + 2\mathrm{tr}(\Lambda(\bar{t}_k)^{\frac{1}{2}}) \right].
\end{aligned}$$

Let $Q(\bar{t}_k)^T Q = \begin{pmatrix} (Q(\bar{t}_k)^T Q)_{11} & (Q(\bar{t}_k)^T Q)_{12} \\ (Q(\bar{t}_k)^T Q)_{21} & (Q(\bar{t}_k)^T Q)_{22} \end{pmatrix}$, where $(Q(\bar{t}_k)^T Q)_{11} \in R^{r \times r}$, then

$$\begin{aligned}
\lim_{t \to +\infty} V_1(t) = \lim_{k \to +\infty} &\left[2\mathrm{tr}(\Lambda_1(\bar{t}_k)^{-\frac{1}{2}} (Q(\bar{t}_k)^T Q)_{11} \Lambda_1 (Q(\bar{t}_k)^T Q)_{11}^T) \right. \\
&\left. + 2\mathrm{tr}(\Lambda_2(\bar{t}_k)^{-\frac{1}{2}} (Q(\bar{t}_k)^T Q)_{21} \Lambda_1 (Q(\bar{t}_k)^T Q)_{21}^T) + 2\mathrm{tr}(\Lambda(\bar{t}_k)^{\frac{1}{2}}) \right].
\end{aligned}$$

Since $\Lambda_1(\bar{t}_k) \to \bar{\Lambda}_1$, $\Lambda_2(\bar{t}_k) \to 0$, and $Q(\bar{t}_k) \to \bar{Q}$, we know the diagonal entries of $(\bar{Q}^T Q)_{21}\Lambda_1(\bar{Q}^T Q)_{21}^T$ are all zero which leads to $(\bar{Q}^T Q)_{21} = 0$. Since $\bar{Q}^T Q$ is an orthogonal matrix, $(\bar{Q}^T Q)_{21}(\bar{Q}^T Q)_{21}^T + (\bar{Q}^T Q)_{22}(\bar{Q}^T Q)_{22}^T = I$. But $(\bar{Q}^T Q)_{21} = 0$, we have $(\bar{Q}^T Q)_{22}(\bar{Q}^T Q)_{22}^T = I$, so $(\bar{Q}^T Q)_{22}$ is an orthogonal matrix. Then from $(\bar{Q}^T Q)_{11}(\bar{Q}^T Q)_{21}^T + (\bar{Q}^T Q)_{12}(\bar{Q}^T Q)_{22}^T = 0$ and $(\bar{Q}^T Q)_{21} = 0$, we have $(\bar{Q}^T Q)_{12} = 0$. From $(\bar{Q}^T Q)_{11}(\bar{Q}^T Q)_{11}^T + (\bar{Q}^T Q)_{12}(\bar{Q}^T Q)_{12}^T = I$, we know $(\bar{Q}^T Q)_{11}(\bar{Q}^T Q)_{11}^T = I$ and so $(\bar{Q}^T Q)_{11}$ is also an orthogonal matrix. From Lemma 2.10, we know the diagonal entries of $(\bar{Q}^T Q)_{11}\Lambda_1(\bar{Q}^T Q)_{11}^T$ are all positive. So $\bar{\Lambda}_1$ must be invertible which indicates that $\bar{X}$ has the maximal rank among the optimal solution set of problem (P), and

$$\begin{aligned}
\lim_{t \to +\infty} V_1(t) = &\, 2\mathrm{tr}(\bar{\Lambda}_1^{-\frac{1}{2}} (\bar{Q}^T Q)_{11} \Lambda_1 (\bar{Q}^T Q)_{11}^T) + 2\mathrm{tr}(\bar{\Lambda}_1^{\frac{1}{2}}) \\
&+ \lim_{k \to +\infty} 2\mathrm{tr}((Q(\bar{t}_k)^T Q)_{21}^T \Lambda_2(\bar{t}_k)^{-\frac{1}{2}} (Q(\bar{t}_k)^T Q)_{21} \Lambda_1).
\end{aligned}$$

If the maximal rank among the optimal solution set of problem (P) is equal to one, then the optimal solutions have the form $\lambda q_1 q_1^T$, where $\lambda \geq 0$ and $\|q\| = 1$. Hence in this case, the orthogonal matrix in the eigenvalue decomposition of every optimal solution can have the same Q as X^*. If $\bar{X}$ is not the only point of $\Omega^3(X^0)$, there must exist another point $\tilde{X} \in \Omega^3(X^0)$. Let us define

$$V_6(t) = I_3(X, \bar{X}) = 2\text{tr}(X^{-\frac{1}{2}}\bar{X}) + 2\text{tr}(X^{\frac{1}{2}}),$$

where $t \in [t_0, +\infty)$ and X (or $X(t)$) is the unique solution of ODE system (1.4). Since $\tilde{X} \in \Omega^3(X^0)$, for the same reason, $\tilde{X}$ has the maximal rank among the optimal solution set of problem (P) and there exists an eigenvalue decomposition $\tilde{X} = Q\tilde{\Lambda}Q^T = Q\begin{pmatrix}\tilde{\Lambda}_1 & \\ & 0\end{pmatrix}Q^T$, where $\tilde{\Lambda}_1 \in R$ is positive (for simiplicity, when $r = 1$, we view the 1×1 matrix as a scalar), and a sequence $\{\tilde{t}_k\}_{k=1}^{+\infty}$ with $\lim\limits_{k\to+\infty} \tilde{t}_k = +\infty$ such that $X(\tilde{t}_k) \to \tilde{X}$, $Q(\tilde{t}_k) \to Q$, and $\Lambda(\tilde{t}_k) \to \tilde{\Lambda}$, where $Q(\tilde{t}_k)\Lambda(\tilde{t}_k)Q(\tilde{t}_k)^T$ is an eigenvalue decomposition of $X(\tilde{t}_k)$ with $\Lambda(\tilde{t}_k) = \begin{pmatrix}\Lambda_1(\tilde{t}_k) & \\ & \Lambda_2(\tilde{t}_k)\end{pmatrix}$, $\Lambda_1(\tilde{t}_k) \in R$, such that

$$\lim_{t\to+\infty} V_6(t) = 2\tilde{\Lambda}_1^{-\frac{1}{2}}\bar{\Lambda}_1 + 2\tilde{\Lambda}_1^{\frac{1}{2}} + \lim_{k\to+\infty} 2(Q(\tilde{t}_k)^TQ)_{21}^T\Lambda_2(\tilde{t}_k)^{-\frac{1}{2}}(Q(\tilde{t}_k)^TQ)_{21}\bar{\Lambda}_1. \tag{4.1}$$

Hence, $\lim_{k\to+\infty} 2(Q(\tilde{t}_k)^TQ)_{21}^T\Lambda_2(\tilde{t}_k)^{-\frac{1}{2}}(Q(\tilde{t}_k)^TQ)_{21}$ exists and we denote it by $\tilde{\epsilon}$. However, $\bar{X}$ is also an accumulation point, therefore

$$\lim_{t\to+\infty} V_6(t) = 4\bar{\Lambda}_1^{\frac{1}{2}} + \bar{\epsilon}\bar{\Lambda}_1, \tag{4.2}$$

where $\bar{\epsilon} = \lim_{k\to+\infty} 2\text{tr}((Q(\bar{t}_k)^TQ)_{21}^T\Lambda_2(\bar{t}_k)^{-\frac{1}{2}}(Q(\bar{t}_k)^TQ)_{21}$.

Combining (4.1) and (4.2), we get

$$(\bar{\epsilon} - \tilde{\epsilon})\bar{\Lambda}_1 = 2\tilde{\Lambda}_1^{-\frac{1}{2}}\bar{\Lambda}_1 + 2\tilde{\Lambda}_1^{\frac{1}{2}} - 4\bar{\Lambda}_1^{\frac{1}{2}} = 2\bar{\Lambda}_1^{\frac{1}{2}}(\tilde{\Lambda}_1^{-\frac{1}{2}}\bar{\Lambda}_1^{\frac{1}{2}} + \tilde{\Lambda}_1^{\frac{1}{2}}\bar{\Lambda}_1^{-\frac{1}{2}} - 2) \geq 0,$$

which implies $\bar{\epsilon} \geq \tilde{\epsilon}$. If we replace $\bar{X}$ in $V_6(t)$ by $\tilde{X}$, from the similar claim, we can get

$$(\tilde{\epsilon} - \bar{\epsilon})\tilde{\Lambda}_1 \geq 0,$$

which indicates $\tilde{\epsilon} \geq \bar{\epsilon}$. Therefore $\tilde{\epsilon} = \bar{\epsilon}$, and then $\tilde{\Lambda}_1 = \bar{\Lambda}_1$, hence $\bar{X} = \tilde{X}$ and the limit set $\Omega^3(X^0)$ is a singleton. □

Theorem 4.2. *The limit set $\Omega^4(X^0)$ only contains a single point, and the limit point has the maximal rank among the optimal solution set of problem (P).*

Proof. From Theorem 2.14, we know that $\Omega^4(X^0)$ is not empty. So we can choose a point $\bar{X} \in \Omega^4(X^0)$, and evidently $\bar{X} \in \mathcal{P}^+$. Similar to the proof of Theorem 4.1, by using $V_2(t)$ defined by (2.14), we can show every accumulation point in $\Omega^4(X^0)$ has the maximal rank among the optimal solution set of problem (P). From (2.10), we can define $V_7(t)$ as follows

$$V_7(t) = I_4(t, X, \bar{X}) - 4\text{tr}(\bar{X}^{\frac{1}{2}}),$$

where $t \in [t_0, +\infty)$ and X (or $X(t)$) is the unique solution of ODE system (1.5). Since $\frac{dV_7(t)}{dt} = f(\bar{X}) - f(X(t)) \leq 0$, $V_7(t)$ is a nonincreasing function, furthermore $V_7(t)$ is bounded below, so $\lim_{t\to+\infty} V_7(t)$ exists. Similar to the claim in the proof of Theorem 4.1, if $\tilde{X} = \tilde{Q}\tilde{\Lambda}\tilde{Q}^T \in \Omega^4(X^0)$ is another point, we can have $(\tilde{Q}^T\bar{Q})_{12} = 0$, $(\tilde{Q}^T\bar{Q})_{21} = 0$, $(\tilde{Q}^T\bar{Q})_{11}$ and $(\tilde{Q}^T\bar{Q})_{22}$ are both orthogonal matrices, and

$$\begin{aligned}\lim_{t\to+\infty} V_7(t) &= 2\text{tr}\left(\tilde{\Lambda}_1^{-\frac{1}{2}}(\tilde{Q}^T\bar{Q})_{11}\bar{\Lambda}_1(\tilde{Q}^T\bar{Q})_{11}^T\right) + 2\text{tr}\left(\tilde{\Lambda}_1^{\frac{1}{2}}\right) - 4\text{tr}\left(\bar{\Lambda}_1^{\frac{1}{2}}\right)\\ &+ \lim_{k\to+\infty}\left\{2\text{tr}\left(\left(Q(\tilde{t}_k)^T\bar{Q}\right)_{21}^T\Lambda_2\left(\tilde{t}_k\right)^{-\frac{1}{2}}\left(Q(\tilde{t}_k)^T\bar{Q}\right)_{21}\bar{\Lambda}_1\right)\right.\\ &\left.+\ \tilde{t}_k\left[f(\bar{X}) - f(X(\tilde{t}_k)) + \text{tr}\left((X(\tilde{t}_k) - \bar{X})\frac{\partial f}{\partial X}|_{X=X(\tilde{t}_k)}\right)\right]\right\},\end{aligned}$$

where $\tilde{\Lambda}_1$, $\tilde{t}_k$, $(\tilde{Q}^T\bar{Q})_{11}$, $\bar{\Lambda}_1$, $Q(\tilde{t}_k)$, $(Q(\tilde{t}_k)^T\bar{Q})_{21}$, and $\Lambda_2(\tilde{t}_k)$ have the same meanings as that in the proof of Theorem 4.1. From Lemma 2.6, we know for any k,

$$\begin{aligned}&2\text{tr}\left(\left(Q(\tilde{t}_k)^T\bar{Q}\right)_{21}^T\Lambda_2\left(\tilde{t}_k\right)^{-\frac{1}{2}}\left(Q(\tilde{t}_k)^T\bar{Q}\right)_{21}\bar{\Lambda}_1\right)\\ &+\tilde{t}_k\left[f(\bar{X}) - f(X(\tilde{t}_k)) + \text{tr}\left((X(\tilde{t}_k) - \bar{X})\frac{\partial f}{\partial X}|_{X=X(\tilde{t}_k)}\right)\right] \geq 0,\end{aligned}$$

hence its limit must be nonnegative. First we assume $\lim_{t\to+\infty} V_7(t) = 0$, then we can get

$$\begin{aligned}0 &\geq \text{tr}\left(\tilde{\Lambda}_1^{-\frac{1}{2}}(\tilde{Q}^T\bar{Q})_{11}\bar{\Lambda}_1(\tilde{Q}^T\bar{Q})_{11}^T\right) + \text{tr}\left(\tilde{\Lambda}_1^{\frac{1}{2}}\right) - 2\text{tr}\left(\bar{\Lambda}_1^{\frac{1}{2}}\right)\\ &= \text{tr}\left(\bar{\Lambda}_1^{\frac{1}{2}}(\tilde{Q}^T\bar{Q})_{11}^T\tilde{\Lambda}_1^{-\frac{1}{2}}(\tilde{Q}^T\bar{Q})_{11}\bar{\Lambda}_1^{\frac{1}{2}}\right) + \text{tr}\left((\tilde{Q}^T\bar{Q})_{11}^T\tilde{\Lambda}_1^{\frac{1}{2}}(\tilde{Q}^T\bar{Q})_{11}\right) - 2\text{tr}\left(\bar{\Lambda}_1^{\frac{1}{2}}\right)\\ &= \text{tr}\left(\bar{\Lambda}_1^{\frac{1}{4}}\left[\bar{\Lambda}_1^{\frac{1}{4}}(\tilde{Q}^T\bar{Q})_{11}^T\tilde{\Lambda}_1^{-\frac{1}{2}}(\tilde{Q}^T\bar{Q})_{11}\bar{\Lambda}_1^{\frac{1}{4}} + \bar{\Lambda}_1^{-\frac{1}{4}}(\tilde{Q}^T\bar{Q})_{11}^T\tilde{\Lambda}_1^{\frac{1}{2}}(\tilde{Q}^T\bar{Q})_{11}\bar{\Lambda}_1^{-\frac{1}{4}} - 2I\right]\bar{\Lambda}_1^{\frac{1}{4}}\right),\end{aligned}$$

but for any symmetric positive definite matrix A, $A + A^{-1} - 2I \succeq 0$ and $A + A^{-1} = 2I$ if and only if $A = I$. Therefore $\bar{\Lambda}_1^{\frac{1}{4}}(\tilde{Q}^T\bar{Q})_{11}^T\tilde{\Lambda}_1^{-\frac{1}{2}}(\tilde{Q}^T\bar{Q})_{11}\bar{\Lambda}_1^{\frac{1}{4}} = I$ which leads to $(\tilde{Q}^T\bar{Q})_{11}^T\tilde{\Lambda}_1(\tilde{Q}^T\bar{Q})_{11} = \bar{\Lambda}_1$. Notice $(\tilde{Q}^T\bar{Q})_{12} = 0$, $(\tilde{Q}^T\bar{Q})_{21} = 0$, $(\tilde{Q}^T\bar{Q})_{11}$ and $(\tilde{Q}^T\bar{Q})_{22}$ are both orthogonal matrices, we have $(\tilde{Q}^T\bar{Q})_{11}^T\tilde{\Lambda}_1(\tilde{Q}^T\bar{Q})_{11} = \bar{\Lambda}_1 \iff \bar{X} = \tilde{X}$.

Now let us prove that $\lim_{t\to+\infty} V_7(t) = 0$. For any $T > t_0$ and $X(T) \succ 0$ (guaranteed by Theorems 2.3), we can define

$$V_8(t) = I_4(t, X(t), X(T)) - 4\text{tr}(X(T)^{\frac{1}{2}}),$$

where $t \in [t_0, +\infty)$. Then we have

$$\frac{V_8(t)}{dt} = f(X(T)) - f(X(t)),$$

and

$$\frac{d(V_7(t) - V_8(t))}{dt} = f(\bar{X}) - f(X(T)) \leq 0.$$

But $V_8(T) = I_4(T, X(T), X(T)) - 4\text{tr}(X(T)^{\frac{1}{2}}) = 0$, so we have

$$V_7(T) - V_8(T) = V_7(T) \leq V_7(t_0) - V_8(t_0).$$

Notice $I_4(t_0, X^0, Y) - 4\text{tr}(Y^{\frac{1}{2}})$ is continuous with respect to Y at $\bar{X}$, and $\bar{X} \in \Omega^4(X^0)$ is an accumulation point, so for any $\epsilon > 0$, we can choose $T > t_0$ such that $V_7(t_0) - V_8(t_0) < \epsilon$.

Then we get $V_7(T) < \epsilon$, furthermore $V_7(t)$ is an nonincreasing function in $[t_0, +\infty)$, therefore we have

$$\lim_{t\to+\infty} V_7(t) = 0.$$

Thus the proof is completed. □

For ODE systems (1.2) and (1.3), we cannot prove the convergence for the general convex $f(X)$, however we can prove the convergence in the linear case where $f(X) = \mathrm{tr}(CX)$ and $C \in \mathcal{S}^n$. For linear SDP, ODE systems (1.2) and (1.3) are actually the same, hence we only discuss ODE system (1.2) below. In [6], Goldfarb and Scheinberg used some auxiliary optimization problems and the auxiliary continuous trajectories $y(\mu)$ and $Z(\mu)$ to study the limiting behavior of the infeasible central paths for linear SDP. Here we adopt the same strategy. In order to propose the auxiliary optimization problems, we choose a $y^0 \in \mathbb{R}^m$, and let $Z^0 = C - \sum_{k=1}^m y_k^0 A_k$, $P = t_0 Z^0 - X(t_0)^{-1}$. Then we get the following lemma.

Lemma 4.3. *For any $t \geq t_0$, the following optimization problem*

$$\begin{array}{ll} \min\limits_{X\in\mathcal{S}^n} & C \bullet X - \frac{1}{t}(P \bullet X + \ln\det X) \\ s.t. & A_k \bullet X = b_k,\ k = 1, \ldots, m, \\ & X \succ 0, \end{array} \tag{Pt}$$

has a unique optimal solution.

Proof. For $t = t_0$, this is evident since (X^0, y^0, Z^0) satisfies the following KKT system

$$\begin{cases} \sum_{k=1}^m y_k^0 A_k + Z^0 = C, \\ A_k \bullet X^0 = b_k,\ k = 1, \ldots, m,\ X^0 \succ 0, \\ t_0 Z^0 = X(t_0)^{-1} + P. \end{cases} \tag{4.3}$$

Since for $t = t_0$, the objective function is strictly convex, we know the optimal solution set is a single point which must be bounded. Hence from Theorem 24 on page 93 in [9], the level set is bounded as well. For any $t > t_0$, $\alpha > 0$, and $X \in \mathcal{P}^{++}$, if

$$C \bullet X - \frac{1}{t}(P \bullet X + \ln\det X) \leq \alpha,$$

then for any optimal solution X^* to problem (P),

$$\begin{aligned} C \bullet X - \frac{1}{t_0}(P \bullet X + \ln\det X) &\leq \frac{t}{t_0}\alpha - \left(\frac{t}{t_0} - 1\right) C \bullet X \\ &\leq \frac{t}{t_0}\alpha - \left(\frac{t}{t_0} - 1\right) C \bullet X^* \\ &\leq \frac{t}{t_0}(\alpha + |C \bullet X^*|), \end{aligned}$$

which implies X is bounded, hence the level set for problem (Pt) for any given $t > t_0$ is bounded. Since the objective function in problem (Pt) is strictly convex for any $t \geq t_0$, it has a unique optimal solution. □

From the above lemma, we can obtain the auxiliary continuous trajectories $y(t)$ and $Z(t)$, and have the following results.

Theorem 4.4. *There exists two auxiliary continuous trajectories $y(t)$ and $Z(t)$ for $t \geq t_0$ such that $(X(t), y(t), Z(t))$ satisfies the following system*

$$\begin{cases} \sum_{k=1}^{m} y(t)_k A_k + Z(t) = C, \\ A_k \bullet X(t) = b_k,\ k = 1, \dots, m,\ X(t) \succ 0, \\ tZ(t) = X(t)^{-1} + P, \end{cases} \tag{4.4}$$

where $X(t)$ is the unique solution of ODE system (1.2).

Proof. From Lemma 4.3, for any $t \geq t_0$, there exists a unique solution for system (4.4). Then we can take derivative with respect to t and get the svec$(\frac{dX}{dt})$ which is actually the right-hand side of ODE system (1.2). Hence the unique solution $X(t)$ of system (4.4) is actually the unique solution of ODE system (1.2). □

From Theorem 2.11, we know $X(t)$ is bounded. Next we show that $y(t)$ and $Z(t)$ are also bounded.

Theorem 4.5. *The auxiliary continuous trajectories $y(t)$ and $Z(t)$ are bounded.*

Proof. From system (4.4), we have

$$\operatorname{tr}(X(t)Z(t)) = \frac{1}{t}\left[n + \operatorname{tr}(X(t)P)\right], \tag{4.5}$$

from Theorem 2.11, $X(t)$ is bounded, we know $\operatorname{tr}(X(t)Z(t)) \to 0$ as $t \to +\infty$. From system (4.4) and Theorem 2.4, we can have

$$\operatorname{tr}(X(t) - X^0)(Z(t) - Z^0) = 0,$$

hence

$$\operatorname{tr}(X^0 Z(t)) = \operatorname{tr}(X(t)Z(t)) + \operatorname{tr}(X^0 Z^0) - \operatorname{tr}(X(t)Z^0),$$

which implies that $\operatorname{tr}(X^0 Z(t))$ is bounded for $t \geq t_0$. Let $X^0 = Q_0 \Lambda_0 Q_0^T$ be the eigenvalue decomposition of X^0, then

$$\operatorname{tr}(X^0 Z(t)) = \operatorname{tr}(\Lambda_0 Q_0^T Z(t) Q_0).$$

From $Z(t) = \frac{1}{t} X(t)^{-1} + \frac{1}{t} P$ in system (4.4) and the Weyl theorem, we have

$$\lambda_{\min}(Q_0^T Z(t) Q_0) = \lambda_{\min}(Z(t)) \geq \frac{1}{t} \lambda_{\min}(P),$$

then from Lemma 2.10, the diagonal entries of $Q_0^T Z(t) Q_0$ must be bounded below by $\frac{1}{t}\lambda_{\min}(P)$. If $\|Z(t)\|_2$ is unbounded, consider $\lambda_{\min}(Z(t)) \geq \frac{1}{t}\lambda_{\min}(P)$, $\lambda_{\max}(Z(t))$ will go to $+\infty$ as $t \to +\infty$, then

$$\operatorname{tr}(Q_0^T Z(t) Q_0) = \operatorname{tr}(Z(t)) \geq \lambda_{\max}(Z(t)) + \frac{n-1}{t} \lambda_{\min}(P) \to +\infty,$$

as $t \to +\infty$, which indicates at least one diagonal entry of $Q_0^T Z(t) Q_0$ will go to $+\infty$ as $t \to +\infty$. But $\Lambda_0 \succ 0$, hence $\operatorname{tr}(X^0 Z(t)) = \operatorname{tr}(\Lambda_0 Q_0^T Z(t) Q_0)$ is unbounded. This is a contradiction, so $Z(t)$ is bounded. From $\sum_{k=1}^{m} y(t)_k A_k + Z(t) = C$ in system (4.4) and Assumption 2, $y(t)$ can be determined by $Z(t)$, so $y(t)$ is also bounded. □

Now we prove the convergence for ODE systems (1.2) and (1.3) in the linear case. In the proof, we use the similar method as Theorem $A.3$ in [10] where the curve selection lemma will be used.

Theorem 4.6. *If $f(X) = \mathrm{tr}(CX)$ is linear, where $C \in \mathcal{S}^n$, then each of $\Omega^1(X^0)$ and $\Omega^2(X^0)$ contains a single point, and the two limit points are on the optimal face of problem (P).*

Proof. Since in the linear SDP, ODE systems (1.2) and (1.3) are the same, we only need to prove ODE system (1.2). From Theorem 2.11 and Theorem 4.5, let $(\tilde{X}, \tilde{y}, \tilde{Z})$ be an accumulation point of the continuous trajectory $(X(t), y(t), Z(t))$. Let $\mu = \frac{1}{t}$ for $t \geq t_0$, and $X(\mu) = X(\frac{1}{t})$, $y(\mu) = y(\frac{1}{t})$, $Z(\mu) = Z(\frac{1}{t})$. Let the real algebraic set V be defined via

$$V = \left\{ (\bar{X}, \bar{Z}, \bar{y}, \mu) \;\middle|\; \begin{array}{l} A_k \bullet \bar{X} = 0 \ (k = 1, \ldots, m), \\ \sum_{k=1}^{m} \bar{y}_k A_k + \bar{Z} = 0, \\ (\bar{X} + \tilde{X})(\bar{Z} + \tilde{Z}) - \mu I - \mu(\bar{X} + \tilde{X})P = 0, \end{array} \right\}$$

and let the open set U be defined as the set of all $(\bar{X}, \bar{Z}, \bar{y}, \mu)$ such that all principal minors of $(\bar{X} + \tilde{X})$ are positive and $\mu > 0$.

Now from Lemma 4.3 and Theorem 4.4, we can see that $V \cap U$ corresponds to the continuous trajectory $(X(\mu), y(\mu), Z(\mu))$ excluding its limit points, in the sense that if $(\bar{X}, \bar{Z}, \bar{y}, \mu) \in V \cap U$ then $X(\mu) = \bar{X} + \tilde{X}$ and $Z(\mu) = \bar{Z} + \tilde{Z}$. Moreover, the zero element is in the closure of $V \cap U$, by construction. Then similar to the proof of Theorem $A.3$ in [10], we can prove that $(\tilde{X}, \tilde{y}, \tilde{Z})$ is the only limit point of the continuous trajectory $(X(t), y(t), Z(t))$.

Without loss of generality, we assume the optimal solution X^* is on the optimal face of problem (P), from system (4.4) and Theorem 2.4, we can get

$$\begin{aligned} \mathrm{tr}(X(t)^{-1}X^*) &= n + \mathrm{tr}\left[(X(t) - X^*)P\right] - t \cdot \mathrm{tr}\left[C(X(t) - X^*)\right] \\ &\leq n + \mathrm{tr}\left[(X(t) - X^*)P\right], \end{aligned}$$

which is bounded above, hence similar to the claim in the proof of Theorem 4.1, we can show that the limit point $\tilde{X}$ is on the optimal face of problem (P). □

5 Concluding Remarks

In this paper, four interior point continuous trajectories for convex semidefinite programming are studied. The ODE systems (1.4) and (1.5) are the variants of ODE systems (1.2) and (1.3). Compared to ODE systems (1.2) and (1.3), the solutions of ODE systems (1.4) and (1.5) can converge to the optimal solution(s) of problem (P) under a weaker condition. In this sense, ODE systems (1.4) and (1.5) seem to be more attractive. For the affine scaling direction in ODE system (1.2), Muramatsu [22] showed that the affine scaling algorithm for semidefinite programming may fail, but in his counterexample, the stepsize will go to infinity. Our results here indicate that if the stepsize of affine scaling algorithm is small enough, the accumulation point may be an optimal solution, and a strategy of stepsize in affine scaling method for problem (P) in [25] had actually been proposed to guarantee the optimality.

Acknowledgement

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, we thank one anonymous referee for providing the second proof of Lemma 2.2.

References

[1] F. Alizadeh, J.P.A. Haeberly and M.L. Overton, Primal-dual interior-point methods for semidefinite programming: convergence rates, stability and numerical results, *SIAM J. Optim.* 8 (1998) 746–768.

[2] D.V. Anosov, S.h. Aranson, V.I. Arnold, I.U. Bronshtein, V.Z. Grines and Y.S. Il'yashenko, *Ordinary Differential Equations and Smooth Dynamical Systems*, Springer, 1988.

[3] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, 2004.

[4] M. Chu and M. Lin, Dynamical system characterization of the central path and its variants– a revisit, *SIAM J. Appl. Dyn. Syst.* 10 (2011) 887–905.

[5] E.A. Coddington and N. Levinson, *Theory of Ordinary Diffenential Equations*, McGraw-Hill Book Co., New York, 1955.

[6] D. Goldfarb and K. Scheinberg, Interior point trajectories in semidefinite programming, SIAM J. Optim. 8 (1998) 871–886.

[7] L.M. Graña Drummond, A.N. Iusem and B. F. Svaiter, On the central path for nonlinear semidefinite programming, *RAIRO Oper. Res.* 34 (2000) 331–345.

[8] L.M. Graña Drummond and Y. Peterzil, The central path in smooth convex semidefinite programs, *Optimization* 51 (2002) 207–233.

[9] A.V. Fiacco and G.P. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, SIAM, Philadelphia, USA, 1990.

[10] M. Halická, E. de Klerk and C. Roos, On the convergence of the central path in semidefinite optimization, *SIAM J. Optim.* 12 (2002) 1090–1099.

[11] M. Halická and M. Trnovská, Limiting behavior and analyticity of weighted central paths in semidefinite programming, *Optim. Methods Softw.* 25 (2010) 247–262.

[12] N. J. Higham, *Functions of Matrices: Theory and Computation.* SIAM, 2008.

[13] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge university press, 2012.

[14] D. Jiang and J. Wang, A recurrent neural network for real-time semidefinite programming, *IEEE T. Neural Networ.* 10 (1999) 81–93.

[15] L. Li and K.C. Toh, A polynomial-time inexact primal-dual infeasible path-following algorithm for convex quadratic SDP, *Pac. J. Optim.* 7 (2011) 43–61.

[16] J. López and H. Ramírez C, On the central paths and cauchy trajectories in semidefinite programming, *Kybernetika* 46 (2010) 524–535.

[17] V. Losert and E. Akin, Dynamics of games and genes: disrete versus contibuous time, *J. Math. Biology* 17 (1983) 241–251.

[18] Z. Lu and R.D.C. Monteiro, Limiting behavior of the Alizadeh-Haeberly-Overton weighted paths in semidefinite programming, *Optim. Methods Softw.* 22 (2007) 849–870.

[19] Z.Q. Luo, J.F. Sturm and S. Zhang, Superlinear convergence of a symmetric primal-dual path following algorithm for semidefinite programming, *SIAM J. Optim.* 8 (1998) 59–81.

[20] R.D.C. Monteiro, Primal-dual path-following algorithms for semidefinite programming, *SIAM J. Optim.* 7 (1997) 663–678.

[21] R.D.C. Monteiro and Y. Zhang, A unified analysis for a class of long-step primal-dual path-following interior-point algorithms for semidefinite programming, *Math. Program.* 81 (1998) 281–299.

[22] M. Muramatsu, Affine scaling algorithm fails for semidefinite programming, *Math. Program.* 83 (1998) 393–406.

[23] J.W. Nie and Y.X. Yuan, A predictor–corrector algorithm for QSDP combining Dikin-type and Newton centering steps, *Ann. Oper. Res.* 103 (2001) 115–133.

[24] M. Preiß and J. Stoer, Analysis of infeasible-interior-point paths arising with semidefinite linear complenmentary problems, *Math. Program.* 99 (2004) 499–520.

[25] X. Qian, L.-Z. Liao, and J. Sun, A strategy of global convergence for the affine scaling algorithm for convex semidefinite programming, *Math. Program.* 179 (2020) 1–19.

[26] X. Qian, L.-Z. Liao, J. Sun and H. Zhu, The convergent generalized central paths for linearly constrained convex programming, *SIAM J. Optim.* 28 (2018) 1183–1204.

[27] M. Shida and S. Shindoh, Monotone semidefinite complementarity problems, Working Paper B-312, Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology, Japan, 1996

[28] C.-K. Sim and G. Zhao, Underlying paths in interior point methods for the monotone semidefinite linear complementarity problem, *Math. Program.* 110 (2007) 475–499.

[29] 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) 11–25.

[30] J.J.E. Slotine and W. Li, *Applied Nonlinear Control*, Prentice Hall, New Jersey, 1991.

[31] J. F. Sturm and S. Zhang, Symmetric primal-dual path-following algorithms for semidefinite programming, *Appl. Numer. Math.* 29 (1999) 301–315.

[32] J. Sun, A convergence proof for an affine scaling algorithm for convex quadratic programming without nondegeneracy assumptions, *Math. Program.* 60 (1993) 69–79.

[33] M.J. Todd, K.C. Toh and R.H. Tütüncü, On the Nesterov-Todd direction in semidefinite programming, *SIAM J. Optim.* 8 (1998) 769–796.

[34] K.C. Toh, An inexact primal–dual path following algorithm for convex quadratic SDP, *Math. Program.* 112 (2008) 221–254.

[35] K.C. Toh, R.H. Tütüncü and M.J. Todd, Inexact primal-dual path-following algorithms for a special class of convex quadratic SDP and related problems, *Pac. J. Optim.* 3 (2007) 135–164.

[36] P. Tseng, I.M. Bomze and W. Schachinger, A first-order interior point method for linearly constrained smooth optimization, *Math. Program.* 127 (2011) 399–424.

[37] L. Vandenberghe and S. Boyd, Semidefinite Programming, *SIAM Rev.* 38 (1996) 49–95.

[38] S.G. Wang, M.X. Wu and Z.Z. Jia, *Matrix Inequalities*, second ed., Science Press, Beijing, 2006 (in Chinese).

[39] H. Wolkowicz, R. Saigal and L. Vandenberghe, *Handbook of semidefinite programming: theory, algorithms, and applications*, vol. 27, Springer Science and Business Media, 2012.

[40] Y. Zhang, On extending some primal-dual interior-point algorithms from linear programming to semidefinite programming, *SIAM J. Optim.* 8 (1998) 365–386.

Manuscript received 25 June 2023
revised 31 October 2023
accepted for publication 7 December 2023

Xun Qian
Shanghai Artificial Intelligence Laboratory
Shanghai, China
E-mail address: qianxun@pjlab.org.cn

Li-Zhi Liao
Department of Mathematics, Hong Kong Baptist University
Hong Kong, P. R. China
E-mail address: liliao@hkbu.edu.hk

Jie Sun
Department of Mathematics and Statistics
Curtin University, Australia
E-mail address: jie.sun@curtin.edu.au