ANALYTIC FORMULAS FOR ALTERNATING PROJECTION SEQUENCES FOR THE POSITIVE SEMIDEFINITE CONE AND AN APPLICATION TO CONVERGENCE ANALYSIS

HIROYUKI OCHIAI, YOSHIYUKI SEKIGUCHI, AND HAYATO WAKI

ABSTRACT. We derive analytic formulas for the alternating projection method applied to the cone $\mathbb{S}^n_+$ of positive semidefinite matrices and an affine subspace. More precisely, we find recursive relations on parameters representing a sequence constructed by the alternating projection method. By applying these formulas, we analyze the alternating projection method in detail and show that the upper bound given by the singularity degree is actually tight when the alternating projection method is applied to $\mathbb{S}^3_+$ and a 3-plane whose intersection is a singleton with singularity degree 2.

1. INTRODUCTION

1.1. **The alternating projection method.** Let $\mathbb{S}^n, \mathbb{S}^n_+$ be the sets of $n \times n$ symmetric matrices and positive semidefinite matrices, respectively. For an affine subspace E of $\mathbb{S}^n$, $E \cap \mathbb{S}^n_+$ represents the feasible region of a semidefinite programming problem. Thus, it is important to find a point in $E \cap \mathbb{S}^n_+$ in numerous applications across a wide range of areas [4, 5, 10]. The alternating projection method for E and $\mathbb{S}^n_+$ constructs a sequence $\{U_k\}$ with $U_0 \in E$ by

$$U_{k+1} = P_E \circ P_{\mathbb{S}^n_+}(U_k),$$

where P_E and $P_{\mathbb{S}^n_+}$ are the projections onto E and $\mathbb{S}^n_+$, respectively. We call $\{U_k\}$ an *AP sequence* for short. It is known that $\{U_k\}$ converges to a point in $E \cap \mathbb{S}^n_+$ if $E \cap \mathbb{S}^n_+$ is nonempty, or to a point with displacement if $E \cap \mathbb{S}^n_+$ is empty; see, e.g., [6] and the references therein.

The behavior of an AP sequence has been analyzed in most studies using inequalities related to the projections, and only upper bounds for the convergence rate are given. In particular, [6] showed that an upper bound is given by the singularity degree of $E \cap \mathbb{S}^n_+$. However, as discussed in an open question proposed in [3], known upper bounds might not be tight and thus it would be interesting to find a tight upper bound. In fact, we construct an affine subspace E in Example 3.2, where the singularity degree of $E \cap \mathbb{S}^3_+$ is 2 and the tight upper bound for the convergence rate of the AP sequence is $O(k^{-1/2})$, although the upper bound given by the singularity degree is $O(k^{-1/6})$. In examples in Section 3 and [3, Example 5.2, 5.4, 5.6], the gaps between known upper bounds and the actual convergence rates are found by directly analyzing defining equations for AP sequences, instead of the inequalities related to the projections.

2010 *Mathematics Subject Classification.* Primary 90C25, 41A25; Secondary 65K10.

Key words and phrases. Alternating projection method, Positive semidefinite cone, Nontransversal intersection, Determinantal variety, Grassmannian, Exact convergence rate.

The purpose of this paper is to shed new light on the recursive relation defining AP sequences for $\mathbb{S}^n_+$ and an affine subspace with the aim of convergence analysis. It is observed that AP sequence for E and $\mathbb{S}^n_+$ is defined with simple projections and it can be parameterized with respect to a basis for $E - U_*$ for $U_* \in E \cap \mathbb{S}^n_+$. Thus we obtain a parametric representation for the projections that is suitable for direct calculation in the convergence analysis.

1.2. **Contributions.** It is usually a hard problem to obtain an exact convergence rate for a sequence generated by an iterative method for an optimization problem. We need to find a recursive relation that is explicit and appropriate for detailed computations. For this purpose, we mainly consider the case where $E \cap \mathbb{S}^n_+$ is a singleton as assumed in examples in [3], to obtain exact convergence rates. For a non-singleton case, we present only Example 3.3 as an application of the first formula to a local analysis around a point in the intersection. It is future work to investigate the case where $E \cap \mathbb{S}^n_+$ is not a singleton in detail. Under this assumption we obtain the following formulas. After discussing the first formula, we concentrate on the case of $\mathbb{S}^3_+$. In low dimensions, we can analyze the alternating projection method in significantly more detail than in higher dimensions, thereby deepening our understanding of the method. We hope that these results also grasp the essential nature of a higher dimensional problem that is generic.

Eigenvalue formula. For a general affine subspace E and $\mathbb{S}^n_+$, we obtain the first analytic formula for the parameters for an AP sequence by using eigenvalues (Proposition 3.1). In general, the eigenvalues of a parametric matrix is not readily available, and hence this formula is not easily applied to convergence analysis. However, in some special cases, the formula has useful applications, such as constructing interesting examples (Example 3.2,3.3, 3.4), and estimating convergence rates when E is a line [12].

Analytic formula when $P_{\mathbb{S}^3_+}(U_k)$ is rank 1. We consider $\mathbb{S}^3_+$ and a 3-plane E for simplicity. By numerical experiments, we see that $P_{\mathbb{S}^3_+}(U_k)$ is often rank 1, and this case appears to be crucial for the convergence analysis in Section 7. Thus, we additionally assume that $P_{\mathbb{S}^3_+}(U_k)$ is rank 1. Then we obtain the second analytic formula (Theorem 4.1). This formula allows us to construct a curve such that an AP sequence converges most slowly if the initial point is taken from the curve. Finding such an initial point is crucial for showing the tightness of an upper bound.

Rational formula. We find a parameterization of the family of 3-planes E such that $E \cap \mathbb{S}^3_+$ is a singleton (Proposition 5.1). Then the set of such planes with singularity degree 2 is fully characterized. With this characterization, we find a rational formula (the matrix (5)) for the curve giving the slowest convergence rate. By using this formula, we obtain explicit expressions of $P_{\mathbb{S}^3_+}(U_k)$ (Theorem 6.2) and $P_E \circ P_{\mathbb{S}^3_+}(U_k)$ (Theorem 6.3), and then we show that the upper bound given by the singularity degree is actually tight (Theorem 7.1).

1.3. **Organization.** Section 2 provides the basic notation. The first analytic formula is obtained in Section 3. Section 4 contains the second analytic formula. In Section 5, we parameterize the set of 3-planes whose intersections with $\mathbb{S}^3_+$ are a singleton. The rational formula for the curve giving the slowest convergence rate is given in Section 6. Section 7 deals with the case where the upper bound given by the singularity degree is tight.

2. Preliminaries

Let $[n] = \{1, \dots, n\}$, $\langle A, B\rangle = \operatorname{tr} A^T B = \sum_{i,j=1}^n A_{ij}B_{ij}$, $\|A\|_F = \sqrt{\langle A, A\rangle}$ and $\|A\|_2$ be the spectral norm of A. If there is no confusion, we simply use $\|A\|$ to mean $\|A\|_F$. The distance $d(A, E)$ from a matrix $A \in \mathbb{S}^n$ to a set $E \subset \mathbb{S}^n$ is defined by $d(A, E) = \inf_{X\in E} \|A - X\|_F$. If E is a closed convex subset of $\mathbb{S}^n$, then there exists a unique optimal solution to $\min_{X\in E} \|A - X\|_F$, and the optimal solution is called the projection of A onto E and denoted by $P_E(A)$.

For $f, g : \mathbb{R} \to \mathbb{R}$, we write $f(x) = O(g(x))$ as $x \to \infty$ if there exist $C, M > 0$ such that $|f(x)| \le Cg(x)$ for all x with $|x| > M$. We also write $f(x) = \Theta(g(x))$ as $x \to \infty$ if there exist $C_1, C_2 > 0$ such that $C_1g(x) \le f(x) \le C_2g(x)$ for all x with $|x| > M$. The meaning of the statement $f(x) = O(g(x))$ as $x \to 0$ is defined similarly. If there is no ambiguity, we simply write $f(x) = O(g(x))$, or $f(x) = \Theta(g(x))$. For $F : \mathbb{R} \to \mathbb{R}^n$, we also write $F(x) = O(g(x))$ if $F_i(x) = O(g(x))$ for $i \in [n]$. Similarly, $F(x) = \Theta(g(x))$ is defined.

3. Analytic formula for a general case

3.1. **Eigenvalue formula.** Suppose that E is an affine subspace of $\mathbb{S}^n$ and $U_* \in E\cap\mathbb{S}^n_+$. Let $B_1, \dots, B_m$ be an orthogonal basis for $E-U_* := \{U-U_* \in \mathbb{S}^n : U \in E\}$, and define

$$\varphi_0(p) = \sum_{i=1}^m p_iB_i,\ \ \varphi(p) = U_* + \varphi_0(p).$$

The following proposition gives the first analytic formula for the alternating projections $P_E \circ P_{\mathbb{S}^n_+}$. Note that $\varphi : \mathbb{R}^m \to E$ is bijective and thus φ has the inverse map φ^{-1}.

Proposition 3.1. *Let $\tilde{p} = \varphi^{-1} \circ P_E \circ P_{\mathbb{S}^n_+} \circ \varphi(p)$ and $\lambda_1(p), \dots, \lambda_n(p)$ be eigenvalues of $\varphi(p)$. Then we have*

$$\tilde{p}_i = p_i - \frac{1}{\|B_i\|^2}\frac{\partial}{\partial p_i}\sum_{\ell\in n(p)}\frac{1}{2}\lambda_\ell^2(p), \quad i \in [m],$$

where $n(p) = \{\ell \in [n] : \lambda_\ell(p) < 0\}$.

Proof. If $\varphi(p) \in \mathbb{S}^n_+$, then $n(p) = \emptyset$ and hence the formula obviously holds. Thus we assume $\varphi(p) \notin \mathbb{S}^n_+$. Let $U = \varphi(p)$, $V = P_{\mathbb{S}^n_+}(U)$ and $\widetilde{U} = P_E(V)$. By [8, Thm 4.8], we have that $d^2(U, \mathbb{S}^n_+)$ is continuously differentiable and

$$\nabla\frac{1}{2}d^2(U, \mathbb{S}^n_+) = U - P_{\mathbb{S}^n_+}(U),$$

where ∇ corresponds to differentiation with respect to each component of U. Since P_E is the orthogonal projection onto E, we can easily show that

$$P_E(V) = U_* + \sum_i \frac{\langle B_i, V - U_*\rangle}{\|B_i\|^2}B_i = \sum_i \frac{\langle B_i, V\rangle}{\|B_i\|^2}B_i - \sum_i \frac{\langle B_i, U_*\rangle}{\|B_i\|^2}B_i + U_*.$$

Thus we obtain

$$\begin{aligned}\widetilde{U} &= P_E(V) = P_E\left(U - \nabla\frac{1}{2}d^2(U, \mathbb{S}^n_+)\right)\\ &= \sum_i \frac{\langle B_i, U - \nabla\frac{1}{2}d^2(U, \mathbb{S}^n_+)\rangle}{\|B_i\|^2}B_i - \sum_i \frac{\langle B_i, U_*\rangle}{\|B_i\|^2}B_i + U_*\\ &= P_E(U) - \sum_i \frac{\langle B_i, \nabla\frac{1}{2}d^2(U, \mathbb{S}^n_+)\rangle}{\|B_i\|^2}B_i = U - \sum_i \frac{1}{2\|B_i\|^2}\frac{\partial}{\partial p_i}d^2(\varphi(p), \mathbb{S}^n_+)B_i.\end{aligned}$$

Note that the last equality follows the chain rule

$$\frac{\partial}{\partial p_i}d^2(\varphi(p), \mathbb{S}^n_+) = \frac{\partial}{\partial p_i}d^2(U_* + \textstyle\sum_j p_k B_k, \mathbb{S}^n_+) = \langle \nabla d^2(U, \mathbb{S}^n_+), B_i\rangle.$$

Here, we see that

$$d^2(\varphi(p), \mathbb{S}^n_+) = \|\varphi(p) - P_{\mathbb{S}^n_+}(\varphi(p))\|^2 = \sum_{\ell\in n(p)} \lambda_\ell^2(p),$$

and thus we have $\widetilde{U} = U - \sum_i \frac{1}{2\|B_i\|^2}\frac{\partial}{\partial p_i}\sum_{\ell\in n(p)}\lambda_\ell^2(p)B_i$. Since $\widetilde{U} = \varphi(\tilde{p}) = U_* + \sum_i \tilde{p}_i B_i$ and $U = \varphi(p) = U_* + \sum_i p_i B_i$, we obtain the desired equality by comparing the coefficients of B_i on the both side. □

3.2. **Applications of the eigenvalue formula.** Since the computation of the eigenvalues of a parameterized matrix is usually difficult, this formula is not so useful to analyze a general AP sequence. However, this formula can be used to investigate a simpler case [12], or to construct examples with interesting properties as below.

Example 3.2 (Known upper bounds and actual convergence rates)**.** It is well known that an upper bound of the convergence rate of alternating projections for an affine subspace and $\mathbb{S}^n_+$ is given using the singularity degree; [6]. The *singularity degree* is a nonnegative integer determined by the iterative process called *facial reduction*. For the detail; see, e.g. [2, 7, 13]. We note that the singularity degree of the intersection of an affine subspace and $\mathbb{S}^n_+$ is less than or equal to $n-1$.

Consider

$$\begin{aligned}E = \Bigg\{U \in \mathbb{S}^3 : \left\langle\begin{pmatrix}1&0&0\\0&0&0\\0&0&0\end{pmatrix}, U\right\rangle = 1,\ \left\langle\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}, U\right\rangle = 0,\ \left\langle\begin{pmatrix}0&1&0\\1&0&0\\0&0&0\end{pmatrix}, U\right\rangle = 0,\\ \left\langle\begin{pmatrix}0&0&0\\0&0&1\\0&1&0\end{pmatrix}, U\right\rangle = 0,\ \left\langle\begin{pmatrix}0&0&0\\0&0&0\\0&0&1\end{pmatrix}, U\right\rangle = 0\Bigg\}.\end{aligned}$$

Then a matrix in E can be written as

$$U(t) := \begin{pmatrix}1&0&0\\0&0&0\\0&0&0\end{pmatrix} + t\begin{pmatrix}0&0&-1\\0&2&0\\-1&0&0\end{pmatrix}$$

and hence $E \cap \mathbb{S}^3_+ = \{U(0)\}$. Using the definition of singularity degree in [7], the sigularity degree of $E \cap \mathbb{S}^3_+$ is 2. By the bound based on the singularity degree given in [6, Theorem 2.4], an upper bound for the convergence rate of an AP sequence for E and $\mathbb{S}^3_+$ is $O(k^{-\frac{1}{6}})$. However, the formula in Proposition 3.1 ensures that the tight upper bound for the convergence rate is $O(k^{-1/2})$ as follows.

The eigenvalues of $U(t)$ is $\lambda_1(t) = 2t$, $\lambda_2(t) = 1 + t^2 - t^4 + O(t^6)$, $\lambda_3(t) = -t^2 + t^4 + O(t^6)$. Let $U(t_{k+1}) = P_E \circ P_{\mathbb{S}^3_+}(U(t_k))$. If $t_0 > 0$ sufficiently small, the formula in Proposition 3.1 gives that

$$t_{k+1} = t_k - \frac{1}{6}\frac{d}{dt}\frac{1}{2}\lambda_3^2(t_k) = t_k - \frac{1}{3}t_k^3 + O(t_k^4).$$

Then [12, Lemma 5.2] implies that $t_k \to 0$ with $t_k > 0$ and $t_k = \Theta(k^{-1/2})$. More precisely, $t_k \approx (3/2)^{1/2}k^{-1/2}$; see also [11]. If $t_0 < 0$ sufficiently close to 0, then

$$t_{k+1} = t_k - \frac{1}{6}\frac{d}{dt}\frac{1}{2}(\lambda_1^2(t_k) + \lambda_3^2(t_k)) = t_k - \frac{2}{3}t_k - \frac{1}{3}t_k^3 + O(t_k^4) = \frac{1}{3}t_k + O(t_k^3).$$

Thus $\frac{2}{3}t_k < t_{k+1} < 0$. Hence $t_k \to 0$ with $t_k < 0$ and t_k converges linearly. Combining the two cases, we see that $\|U(t_k) - U_*\| = \sqrt{6}|t_k| \leq O(k^{-1/2})$ for an arbitrary initial point. Figure 1 illustrates these rates of convergence. In the case that $t_0 > 0$, we observe from Figure 2 that the plot of $1/\|U_k - U_*\|^2$ approximately coincides with the line $33.51 + 0.111k$. Hence $\|U_k - U_*\| \approx (33.51 + 0.111k)^{-1/2} \approx 3k^{-1/2}$. This is consistent with our estimate $\|U(t_k) - U_*\| = \sqrt{6}t_k \approx 3k^{-1/2}$. General cases are investigated in [12, Section 5].

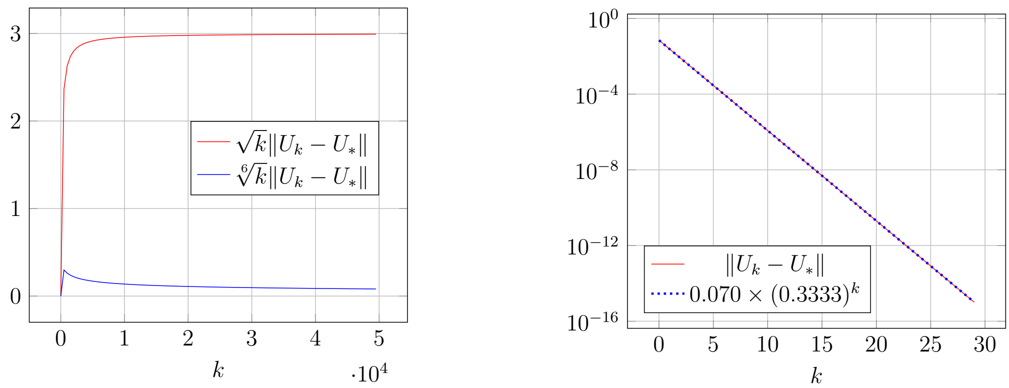

FIGURE 1. The left figure displays a plot of $\|U_k - U_*\|$ with $t_0 > 0$ and the right figure displays a plot with $t_0 < 0$ in Example 3.2, and the line fitting for the plot in the right figure.

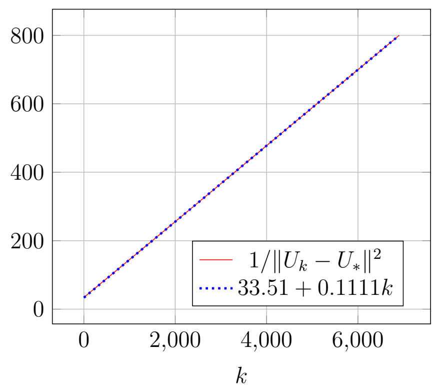

FIGURE 2. Plot of $1/\|U_k - U_*\|^2$ with $t_0 > 0$ in Example 3.2 and the line fitting.

Example 3.3 (Positive dimensional intersection)**.** Proposition 3.1 can be used in the case that an intersection has a positive dimension. Consider

$$E = \Big\{ U \in \mathbb{S}^3 : \langle \left(\begin{smallmatrix} 1&0&0\\0&1&0\\0&0&0 \end{smallmatrix}\right), U\rangle = 1,\ \langle \left(\begin{smallmatrix} 0&0&0\\0&1&0\\0&0&-1 \end{smallmatrix}\right), U\rangle = 0,\ \langle \left(\begin{smallmatrix} 0&0&0\\0&1&-1\\0&-1&1 \end{smallmatrix}\right), U\rangle = 0,$$
$$\langle \left(\begin{smallmatrix} 0&1&0\\1&0&0\\0&0&0 \end{smallmatrix}\right), U\rangle = 0,\ \langle \left(\begin{smallmatrix} 0&0&1\\0&0&0\\1&0&0 \end{smallmatrix}\right), U\rangle = 0 \Big\}.$$

Then a matrix in E can be written as

$$U(t) = \left(\begin{smallmatrix} 1&0&0\\0&0&0\\0&0&0 \end{smallmatrix}\right) + t\left(\begin{smallmatrix} -1&0&0\\0&1&1\\0&1&1 \end{smallmatrix}\right),$$

and hence $E \cap \mathbb{S}^3_+ = \{U(t) : 0 \le t \le 1\}$. The singularity degree of $E \cap \mathbb{S}^3_+$ is 2. Now the eigenvalues of $U(t)$ are $\lambda_1(t) = 1 - t,\ \lambda_2(t) = 2t,\ \lambda_3(t) = 0$. Consider an AP sequence $U(t_{k+1}) = P_E \circ P_{\mathbb{S}^3_+}(U(t_k))$.

If we take the initial point as $U(t_0)$ with $t_0 < 0$, then the AP sequence is expected to converge to $U(0)$. By using the formula in Proposition 3.1 with $U_* = U(0)$, we have

$$t_{k+1} = t_k - \frac{1}{5}\frac{d}{dt}\frac{1}{2}\lambda_2^2(t_k) = t_k - \frac{4}{5}t_k = \frac{1}{5}t_k.$$

Thus $t_k \to 0$ with $t_k < 0$, and in fact $U(t_k) \to U(0)$ linearly.

If we take the initial point as $U(t_0)$ with $t_0 > 1$, then the AP sequence is expected to converge to $U(1)$. To use the formula in Proposition 3.1 with $U_* = U(1)$, we define $\widehat{U}(s) = U(s+1)$ and $s_k = t_k + 1$. Then the eigenvalues of $\widehat{U}(s)$ are $\hat{\lambda}_1(s) = -s,\ \hat{\lambda}_2(s) = 2(s+1),\ \hat{\lambda}_3(s) = 0$ By Proposition 3.1, we have

$$s_{k+1} = s_k - \frac{1}{5}\frac{d}{ds}\frac{1}{2}\hat{\lambda}_1^2(s_k) = s_k - \frac{1}{5}s_k = \frac{4}{5}s_k.$$

Thus $s_k \to 0$ with $s_k > 0$ and hence $U(s_{k+1}+1) \to U(1)$ linearly. Therefore, the actual convergence rate of an AP sequence is linear. Figure 3 is consistent with our estimates that the convergence rates are $O((4/5)^k)$ and $O((1/5)^k)$, respectively.

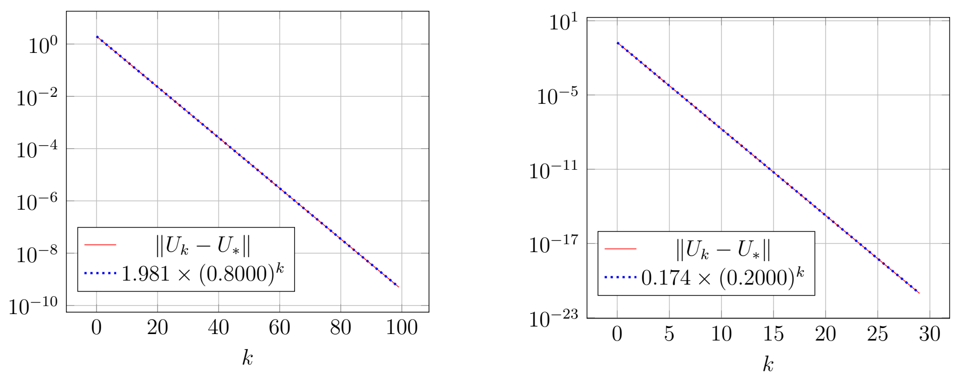

FIGURE 3. The left figure displays a log-plot of $\|U_k - U_*\|$ with $t_0 > 1$ and the right figure shows a log-plot with $t_0 < 0$ in Example 3.3, and their line fittings.

Example 3.4 (Intersection with 2-plane)**.** Consider the parametrized matrix

$$U(p) = \begin{pmatrix} 1 & 0 & p_1 & p_2 \\ 0 & 0 & p_1 & p_2 \\ p_1 & p_1 & 0 & 0 \\ p_2 & p_2 & 0 & 0 \end{pmatrix}.$$

Then $U(p)$ represents 2-plane E in $\mathbb{S}^4$. Now $E \cap \mathbb{S}^4_+ = \{U_*\}$ with $U_* = U(0,0)$, and the singularity degree of $E \cap \mathbb{S}^n_+$ is 1. By the bounds based on the singularity degree given in [6, Theorem 2.4], an upper bound for the convergence rate of AP sequence $\{U(p^{(k)})\}$ for E and $\mathbb{S}^4_+$ is $O(k^{-\frac{1}{2}})$. However, the formula in Proposition 3.1 ensures that the actual convergence rate of the AP sequence is linear as follows.

For $r = \sqrt{p_1^2 + p_2^2}$, the characteristic equation of $U(p)$ is written as

$$\lambda\left(\lambda^3 - \lambda^2 - 2r^2\lambda + r^2\right) = 0.$$

We consider the parametric equation $\lambda^3 - \lambda^2 - 2r^2\lambda + r^2 = 0$ with the parameter r. When $r = 0$, the polynomial $\lambda^3 - \lambda^2$ has 1 as a simple zero and 0 as a double zero. Since the constant term is positive for $r > 0$, by considering the graph of $\lambda^3 - \lambda^2$, we see that the solutions to the equation are $1 + O(r)$, a positive and a negative solution for sufficiently small $r > 0$. To apply Proposition 3.1, we will find the negative solution. By putting $\lambda = -ru$, we obtain $-r^3u^3 - r^2u^2 + 2r^3u + r^2 = 0$ and thus $r = \dfrac{(u+1)(u-1)}{u(2-u^2)}$. Additionally, if we put $u = 1 + v$, then $r = \dfrac{v(v+2)}{(v+1)(1-2v-v^2)}$. This means that r is a rational function in v. By computation, we have $dr/dv|_{v=0} = 2 \neq 0$, and then the inverse function $v = v(r)$ near $r = 0$ is also analytic. Since $r = 2v + 3v^2 + O(v^3)$, we have $v = r/2 - 3r^2/8 + O(r^3)$. Thus the negative eigenvalue can be written as $\lambda = -r + h(r) = -\sqrt{p_1^2 + p_2^2} + h(\sqrt{p_1^2 + p_2^2})$, where $h(r) = O(r^2)$ is a convergent power series around $r = 0$. Then

$$\frac{\partial}{\partial p_1}\lambda^2 = 2p_1 - \frac{2p_1}{r}h(r) - 2p_1 h'(r) + \frac{2p_1}{r}h(r)h'(r) = 2p_1\left(1 + O(r)\right),$$

and hence

$$\tilde{p}_1 = p_1 - \frac{1}{8}\frac{\partial}{\partial p_1}\lambda^2 = \frac{3}{4}p_1 + p_1 O(r).$$

Similarly, $\tilde{p}_2 = \frac{3}{4}p_2 + p_2 O(r)$. Thus we have

$$\tilde{r} := \sqrt{\tilde{p}_1^2 + \tilde{p}_2^2} = \sqrt{\frac{9}{16}r^2 + O(r^3)} = \frac{3}{4}r + O(r^2).$$

Since $\|U(p) - U_*\| = 2r$, we see that for small $\delta > 0$,

$$\|U(\tilde{p}) - U_*\| = 2\tilde{r} = 2 \cdot \frac{3}{4}r(1 + O(r)) < 2 \cdot \left(\frac{3}{4} + \delta\right) r = \left(\frac{3}{4} + \delta\right)\|U(p) - U_*\|$$

for p sufficiently close to $(0,0)$. Therefore, the actual convergence rate of an AP sequence is linear. Figure 4 is consistent with our estimate that the convergence rate of $\|U_k - U_*\|$ is approximately $O((3/4)^k)$.

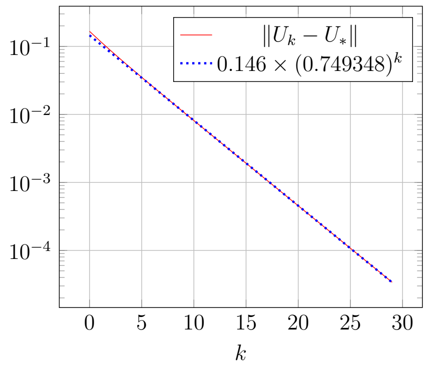

FIGURE 4. Log plot of $\|U_k - U_*\|$ in Example 3.4

4. Analytic formula when $P_{\mathbb{S}^3_+}(U)$ is rank 1

In the rest of the paper, we consider $\mathbb{S}^3_+$ and an affine subspace E whose intersection with $\mathbb{S}^3_+$ is $U_* = \left(\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right)$. Let $D_k = \{V \in \mathbb{S}^3 : \operatorname{rank} V \leq k\}$; i.e., the determinantal variety of rank at most k. Then D_2 contains the boundary of $\mathbb{S}^3_+$ and its singular locus is D_1. Since $\mathbb{S}^3_+$ is convex, $D_1 \cap \mathbb{S}^3_+$ is geometrically interpreted as the ridge of the boundary of $\mathbb{S}^3_+$. Hence, $P_{\mathbb{S}^3_+}(U)$ is expected to be frequently included in D_1 for $U \in E$. In fact, this often happens in numerical experiments. Thus we consider the case that $P_{\mathbb{S}^3_+}(U) \in D_1$ for $U \in E$ sufficiently close to U_*. In Section 7, this case will appear to be crucial for the convergence analysis. We note that $\dim D_1 = 3$ and $\dim \mathbb{S}^3 = 6$ and then the complementary dimension of D_1 is 3. Since $E \cap \mathbb{S}^3_+$ is a singleton and contained in D_1, we assume that the dimension of E is 3 so that the intersection of E and D_1 is zero-dimensional in general.

4.1. Analytic formula with a distance function.

Let B_1, B_2, B_3 be an orthogonal basis for $E - U_*$ and

$$\varphi(p) = U_* + p_1 B_1 + p_2 B_2 + p_3 B_3, \quad \psi(x) = \frac{1}{x_1}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix}$$

and $x_* = (1, 0, 0)$. Then $\varphi : \mathbb{R}^3 \to E$ and $\psi : \{x \in \mathbb{R}^3 : x_1 \neq 0\} \to D_1$. It is easily verified that the image of ψ contains an open neighborhood of U_* in D_1.

For $f : \mathbb{R}^3 \to \mathbb{R}$, define $\partial_k f(x) = \frac{\partial}{\partial x_k} f(x)$ for $k = 1, 2, 3$.

Theorem 4.1. *Suppose that E is a 3-plane in $\mathbb{S}^3$, $E \cap \mathbb{S}^3_+ = \{U_*\}$ and $P_{\mathbb{S}^3_+} \circ \varphi(p)$ has rank 1. Let $\tilde{p}, x \in \mathbb{R}^3$ satisfy the relation $\varphi(p) \overset{P_{\mathbb{S}^3_+}}{\longmapsto} \psi(x) \overset{P_E}{\longmapsto} \varphi(\tilde{p})$. Then we have*

$$M(x)(\tilde{p} - p) + \nabla_x \frac{1}{2} d^2(\psi(x), E) = 0,$$

where $M(x) = \Big(\langle \partial_k \psi(x), B_i \rangle\Big)_{k,i} \in \mathbb{R}^{3\times 3}$. Moreover, if $\det M(x_*) \neq 0$ and p is sufficiently close to $(0, 0, 0)$, then

$$\tilde{p} = p - M(x)^{-1} \nabla_x \frac{1}{2} d^2(\psi(x), E).$$

Proof. Since $P_{S^3_+}(\varphi(p)) \in D_1 \subset \mathbb{S}^3_+$, we see that $P_{S^3_+}(\varphi(p)) = \operatorname{argmin}_{x' \in \mathbb{R}^3, x'_1 \neq 0} \|\psi(x') - \varphi(p)\|_F^2$. In addition, $\varphi(\tilde{p}) = P_E(\psi(x))$ is equivalent to $(\psi(x) - \varphi(\tilde{p})) \perp E$. Thus, we have

$$\psi(x) = P_{\mathbb{S}^3_+}(\varphi(p)) \quad \Longleftrightarrow (*) \left\{ \begin{array}{l} \langle \partial_k \psi(x), \psi(x) - \varphi(p) \rangle = 0, \;\; k = 1, 2, 3 \\ \tilde{p} = \left(\frac{\langle B_1, \psi(x) - U_* \rangle}{\|B_1\|^2}, \frac{\langle B_2, \psi(x) - U_* \rangle}{\|B_2\|^2}, \frac{\langle B_3, \psi(x) - U_* \rangle}{\|B_3\|^2} \right)^T. \end{array} \right.$$

By extending the basis for $E - U_*$, we obtain an orthogonal basis $B_1, \ldots, B_6$ for $\mathbb{S}^3$. Then the distance function can be written as

$$\begin{aligned} d^2(\psi(x), E) &= \|\psi(x) - P_E(\psi(x))\|^2 \\ &= \left\| U_* + \sum_{j=1}^{6} \frac{\langle B_j, \psi(x) - U_* \rangle}{\|B_j\|^2} B_j - \left(U_* + \sum_{i=1}^{3} \frac{\langle B_i, \psi(x) - U_* \rangle}{\|B_i\|^2} B_i \right) \right\|^2 \\ &= \left\| \sum_{j=4}^{6} \frac{\langle B_j, \psi(x) - U_* \rangle}{\|B_j\|^2} B_j \right\|^2 = \sum_{j=4}^{6} \frac{\langle B_j, \psi(x) - U_* \rangle^2}{\|B_j\|^2}. \end{aligned}$$

By rewriting the condition $(*)$ using the basis, we have, for $k = 1, 2, 3$,

$$\begin{aligned} 0 &= \langle \partial_k \psi(x), \;\psi(x) - \varphi(p) \rangle \\ &= \left\langle \sum_{j=1}^{6} \partial_k \psi_i(x), \; U_* + \sum_{j=1}^{6} \psi_j(x) - \left(U_* + \sum_{i=1}^{3} \varphi_i(p) \right) \right\rangle \\ &= \left\langle \sum_{j=1}^{6} \frac{\langle B_j, \partial_k \psi(x) \rangle}{\|B_j\|^2} B_j, \; \sum_{j=1}^{6} \frac{\langle B_j, \psi(x) - U_* \rangle}{\|B_j\|^2} B_j - \sum_{i=1}^{3} p_i B_i \right\rangle \\ &= \left\langle \sum_{j=1}^{6} \frac{\langle B_j, \partial_k \psi(x) \rangle}{\|B_j\|^2} B_j, \; \sum_{i=1}^{3} \left(\frac{\langle B_i, \psi(x) - U_* \rangle}{\|B_i\|^2} - p_i \right) B_i + \sum_{j=4}^{6} \frac{\langle B_j, \psi(x) - U_* \rangle}{\|B_j\|^2} B_j \right\rangle \\ &= \sum_{i=1}^{3} \langle B_i, \partial_k \psi(x) \rangle \left(\frac{\langle B_i, \psi(x) - U_* \rangle}{\|B_i\|^2} - p_i \right) + \sum_{j=4}^{6} \langle B_j, \partial_k \psi(x) \rangle \frac{\langle B_j, \psi(x) - U_* \rangle}{\|B_j\|^2} \\ &= \sum_{i=1}^{3} \langle B_i, \partial_k \psi(x) \rangle (\tilde{p}_i - p_i) + \partial_k \cdot \frac{1}{2} \sum_{j=4}^{6} \left(\frac{\langle B_j, \psi(x) - U_* \rangle}{\|B_j\|} \right)^2 \\ &= \sum_{i=1}^{3} \langle B_i, \partial_k \psi(x) \rangle (\tilde{p}_i - p_i) + \partial_k \cdot \frac{1}{2} d^2(\psi(x), E). \end{aligned}$$

□

Remark 4.2. Since both of D_1 and E have 3 parameters, the matrix $M(x)$ defined in Theorem 4.1 is a square matrix. If we consider general $D_k \subset \mathbb{S}^n_+$ and E, then $M(x)$ may become a rectangle matrix.

Remark 4.3. We can connect the formula in Proposition 3.1 with that in Theorem 4.1 in the following way. Let $\tilde{p}, x \in \mathbb{R}^3$ be such that $\varphi(p) \xmapsto{P_{\mathbb{S}^3_+}} \psi(x) \xmapsto{P_E} \varphi(\tilde{p})$. Since

$\tilde{p}_i - p_i = -\frac{\partial}{\partial p_i}\frac{1}{2\|B_i\|^2}d^2(\varphi(p), \mathbb{S}^3_+)$, we have

$$\nabla_x \frac{1}{2} d^2(\psi(x), E) = \widehat{M}(x)\nabla_p \frac{1}{2} d^2(\varphi(p), \mathbb{S}^3_+), \text{ where } \widehat{M}(x) = \left(\frac{\langle \partial_k \psi(x), B_i\rangle}{\|B_i\|^2}\right)_{k,i}.$$

Let $U = \varphi(p)$. Then we can also write

$$\nabla_x \frac{1}{2} d^2(P_{\mathbb{S}^3_+}(U), E) = \widehat{M}(x)\nabla_p \frac{1}{2} d^2(U, \mathbb{S}^3_+).$$

4.2. **Equations for the slowest curve.** The known results give upper bounds for the convergence rate of an AP sequence. However, it is hard to show that a given upper bound is actually tight. A key to show the tightness is to obtain a candidate for the initial point with which the AP sequence converges most slowly.

If $\det M(x_*) \neq 0$, then we have from Theorem 4.1 that

$$\tilde{p} = p - M(x)^{-1}\nabla_x \frac{1}{2} d^2(\psi(x), E).$$

Thus if we find a point that is a minimizer of $\min_{\|x\|=\delta} \|M(x)^{-1}\nabla \frac{1}{2} d^2(\psi(x), E)\|$ for $\delta \neq 0$, then the point gives the shortest step size with respect to the parameters of the alternating projection method.

Example 4.4. Let $E = \{U \in \mathbb{S}^3 : \langle A_1, U\rangle = 1, \langle A_2, U\rangle = 0, \langle A_3, U\rangle = 0\}$, where

$$A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},\ A_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix},\ A_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Then E is parameterized by

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + p_1 \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + p_2 \begin{pmatrix} 0 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 0 \end{pmatrix} + p_3 \begin{pmatrix} -2 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Now we have for $x_* = (1, 0, 0)$,

$$M(x_*)^{-1}\nabla_x \frac{1}{2} d^2(\psi(x), E) = \frac{1}{6x_1^3}\begin{pmatrix} 2\,x_1\,\left(2\,x_2\,x_3^2 + 2\,x_1\,x_2\,x_3 + x_1^2\,x_3 - x_1\,x_3 + x_2^3\right) \\ -2\,x_1\,\left(3\,x_3^3 + 2\,x_2^2\,x_3 + 2\,x_1^2\,x_3 + x_1\,x_2^2 + x_1^2\,x_2 - x_1\,x_2\right) \\ 3\,x_3^4 + 4\,x_2^2\,x_3^2 + 2\,x_1\,x_2^2\,x_3 - 2\,x_1\,x_2\,x_3 + x_2^4 - x_1^4 + x_1^3 \end{pmatrix}.$$

We consider the system $M(x_*)^{-1}\nabla\frac{1}{2}d^2(\psi(x), E) = \mathbf{0}$ around $x_* = (1, 0, 0)$. By the rational transformation $(x_1, x_2, x_3) = (1/(1+z), u/(1+z), v/(1+z))$, we obtain

$$\begin{cases} f_1 := 2\,u\,v - v\,z + u^3 + 2\,u\,v^2 = 0, \\ f_2 := -2\,v - u^2 + u\,z - 2\,u^2 v - 3\,v^3 = 0, \\ f_3 := z - 2\,u\,v + 2\,u^2 v - 2\,u\,v\,z + u^4 + 4\,u^2 v^2 + 3\,v^4 = 0 \end{cases}$$

Since x_* corresponds to $(u, v, z) = (0, 0, 0)$ and the Jacobian matrix of $(f_1, f_2, f_3)^T$ at $(0, 0, 0)$ is $\left(\begin{smallmatrix} 0 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right)$, the equations $f_2 = f_3 = 0$ are solved with respect to v, z by convergent power series $v(u), z(u)$, respectively. By applying the inverse of the rational transformation $(x_1, x_2, x_3) = (1/(1+z), u/(1+z), v/(1+z))$ to $v(u), z(u)$ and using $x_2 = t$ as the new parameter, we obtain

$$x_1(t) = 1 + t^3 - 2t^6 - \frac{3t^7}{8} + O(t^8),\ x_2(t) = t,\ x_3(t) = -\frac{t^2}{2} + \frac{t^5}{2} + \frac{3t^6}{16} + O(t^8).$$

Then the curve $x(t) = (x_1(t), x_2(t), x_3(t))$ satisfies

$$M(x(t))^{-1}\nabla_x \frac{1}{2} d^2(\psi(x(t)), E) = \begin{pmatrix} \frac{t^7}{8} \\ 0 \\ 0 \end{pmatrix} + O(t^8). \tag{1}$$

Thus $x(t)$ is the minimizer of $\min_{\|x\|=\delta} \|M(x)^{-1}\nabla \frac{1}{2} d^2(\psi(x), E)\|$ for some $\delta > 0$. We can also determine the degree of the leading term of the first component of (1) without actually calculating $x(t)$. We use `Singular` and calculate a weak normal form with the Mora's division algorithm and get

$$qf_1 = a_2 f_2 + a_3 f_3 - \frac{9}{512} u^7 + \text{ higher order terms},$$

where $q, a_2, a_3 \in \mathbb{R}[u, v, z]$ and $q(0,0,0) \neq 0$. Here we use the negative degree reverse lexicographical ordering for $\mathbb{R}[z, v, u]$. This means that we can find convergent power series in u that solve $(f_1, f_2, f_3) = (\Theta(u^7), 0, 0)$, and the degree 7 is the highest such degree.

With this $x(t)$, we define

$$p(t) := \left(\frac{\langle B_i, \psi(x(t)) - U_*\rangle}{\|B_i\|^2}\right)_{i=1}^3 + M(x(t))^{-1}\nabla_x \frac{1}{2} d^2(\psi(x(t)), E) = \begin{pmatrix} t + \frac{t^7}{8} \\ \frac{t^2}{2} - \frac{t^5}{2} - \frac{t^6}{16} \\ -\frac{t^3}{2} + t^6 + \frac{3t^7}{16} \end{pmatrix} + O(t^8)$$

Then we have

$$\varphi(p(t)) = \begin{pmatrix} 1 + t^3 - 2t^6 - \frac{3t^7}{8} & t + \frac{t^7}{8} & -\frac{t^2}{2} + \frac{t^5}{2} + \frac{t^6}{16} \\ t + \frac{t^7}{8} & t^2 - t^5 - \frac{t^6}{8} & -\frac{t^3}{2} + t^6 + \frac{3t^7}{16} \\ -\frac{t^2}{2} + \frac{t^5}{2} + \frac{t^6}{16} & -\frac{t^3}{2} + t^6 + \frac{3t^7}{16} & 0 \end{pmatrix} + O(t^8).$$

For sufficiently small $t > 0$, since $\det \varphi(p(t)) = \frac{t^{10}}{32} + O(t^{11}) > 0$ and the first eigenvalue of $\varphi(p(t))$ is close to 1, we see that $P_{\mathbb{S}^3_+} \circ \varphi(p(t))$ has rank 1. By reversing the modifications of the equations in the proof of Theorem 4.1, we obtain the relation $\varphi(p(t)) \overset{P_{\mathbb{S}^3_+}}{\longmapsto} \psi(x(t))$. Then we have

$$\begin{aligned} \tilde{p} &:= \varphi^{-1} \circ P_E \circ P_{\mathbb{S}^n_+} \circ \varphi(p(t)) \\ &= p(t) - M(x(t))^{-1}\nabla_x \frac{1}{2} d^2(\psi(x(t)), E) = p(t) + \begin{pmatrix} \frac{t^7}{8} \\ 0 \\ 0 \end{pmatrix} + O(t^8). \end{aligned}$$

Moreover, since the leading terms of the second and the third coordinates of $p(t)$ have degree 2 and 3 respectively, we have

$$\tilde{p} = p\left(t - \frac{t^7}{24}\right) + O(t^8),$$

and thus

$$P_E \circ P_{\mathbb{S}^3_+}(\varphi(p(t))) = \varphi\left(p\left(t - \frac{t^7}{24}\right)\right) + O(t^8).$$

Therefore, if we choose a matrix on the curve $\varphi(p(t))$ that is sufficiently close to U_*, then the matrix mapped by $P_E \circ P_{\mathbb{S}^3_+}$ can be written as $\varphi(p(t - t^7/24)) + O(t^8)$. This means that the one-step alternating projection moves the matrix in the slowest way in a sense. In Section 7, we can actually prove that an AP sequence gives the slowest convergence rate if the initial point is taken from a neighborhood of the curve $\varphi(p(t))$.

5. Family of 3-planes intersecting with $\mathbb{S}^3_+$ at a single point

5.1. **Parametrization.** In Example 4.4, we constructed a candidate for a curve that gives the slowest convergence rate for given numeric matrices A_1, A_2, A_3. To construct such a curve for a general case, we will obtain a parameterization of the family of 3-planes that intersect with $\mathbb{S}^3_+$ at a single point.

Proposition 5.1. *Let* $U_* = \begin{pmatrix} 1&0&0\\0&0&0\\0&0&0 \end{pmatrix}$. *A plane* $E \subset \mathbb{S}^3$ *satisfies* $\mathbb{S}^3_+ \cap E = \{U_*\}$ *and* $\dim E = 3$ *if and only if there exist* $c_1, \dots, c_8 \in \mathbb{R}$ *and an orthogonal matrix* $\widetilde{P} \in \mathbb{R}^{2\times 2}$ *such that E is written by*

$$E = \{X \in \mathbb{S}^3 : \langle A_1, X\rangle = 1,\ \langle A_2, X\rangle = 0,\ \langle A_3, X\rangle = 0\},$$

where A_i are given as follows:

Type 1*:*

$$A_1 = P\begin{pmatrix} 1 & c_1 & c_2\\ c_1 & c_3 & c_4\\ c_2 & c_4 & 0 \end{pmatrix}P^T,\ A_2 = P\begin{pmatrix} 0 & c_5 & c_6\\ c_5 & c_7 & c_8\\ c_6 & c_8 & 0 \end{pmatrix}P^T,\ A_3 = P\begin{pmatrix} 0 & 0 & 0\\ 0 & \mu & 0\\ 0 & 0 & 1 \end{pmatrix}P^T,$$

$\mu > 0,\ A_2 \neq O,$

or, Type 2*:*

$$A_1 = P\begin{pmatrix} 1 & c_1 & c_2\\ c_1 & 0 & c_3\\ c_2 & c_3 & 0 \end{pmatrix}P^T,\ A_2 = P\begin{pmatrix} 0 & 0 & c_4\\ 0 & 1 & c_5\\ c_4 & c_5 & 0 \end{pmatrix}P^T,\ A_3 = P\begin{pmatrix} 0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 1 \end{pmatrix}P^T,$$

where $P = \left(\begin{array}{c|cc} 1 & 0 & 0\\ \hline 0 & \multicolumn{2}{c}{\multirow{2}{*}{$\widetilde{P}$}}\\ 0 & & \end{array}\right)$ *in both types.*

Proof. If E is a Type 1 plane, then we can easily show $\mathbb{S}^3_+ \cap E = \{U_*\}$ and $\dim E = 3$. If E is a Type 2 plane, then $X \in E$ is written as

$$X = P\begin{pmatrix} 1 - 2c_1 s + 2c_2 t - 2c_3 u & s & t\\ s & -2c_4 t - 2c_5 u & u\\ t & u & 0 \end{pmatrix}P^T$$

for $s, t, u \in \mathbb{R}$. Thus $\mathbb{S}^3_+ \cap E = \{U_*\}$ and $\dim E = 3$.

Next, we will show the converse. Suppose $\mathbb{S}^3_+ \cap E = \{U_*\}$ and $\dim E = 3$. Then E is contained in a supporting hyperplane to $\mathbb{S}^3_+$ at U_*. The set of normal vectors to a supporting hyperplane of $\mathbb{S}^3_+$ at U_* is given by

$$N_{\mathbb{S}^3_+}(U_*) = \{A \in \mathbb{S}^3_+ : \langle A, U_* \rangle = 0\} = \left\{ \left(\begin{array}{c|cc} 0 & 0 & 0 \\ \hline 0 & \multicolumn{2}{c}{\multirow{2}{*}{$\widetilde{A}$}} \\ 0 & & \end{array}\right) : \widetilde{A} \in \mathbb{S}^2_+ \right\},$$

see, e.g. [9, Section 4.2.4]. Then there exists $A_3 \in N_{\mathbb{S}^3_+}(U_*)$ such that $E \subset E_3$ for

$$E_3 := \{X \in \mathbb{S}^3 : \langle A_3, X - U_* \rangle = 0\} = \{X \in \mathbb{S}^3 : \langle A_3, X \rangle = 0\}.$$

Here, for a matrix X, we partition X as in the expression of $N_{\mathbb{S}^3_+}(U_*)$ and denote the right-lower part of X as $\widetilde{X}$.

If $\operatorname{rank} \widetilde{A}_3 = 2$, then there exists an orthogonal matrix $\widetilde{P} \in \mathbb{R}^{2\times 2}$ such that $\Lambda_3 := \left(\begin{smallmatrix} 0 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{smallmatrix}\right) = P^T A_3 P$ for $P = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & \multicolumn{2}{c}{\widetilde{P}} \\ 0 & & \end{array}\right)$. We may assume $\lambda_2 = 1$. Then we have

$$X \in \mathbb{S}^3_+ \cap E_3 \iff X \in \mathbb{S}^3_+,\ \langle \widetilde{A}_3, \widetilde{X} \rangle = 0 \iff X = \left(\begin{array}{c|cc} x_{11} & 0 & 0 \\ \hline 0 & \multicolumn{2}{c}{\multirow{2}{*}{O}} \\ 0 & & \end{array}\right),\ x_{11} \geq 0.$$

Thus $\mathbb{S}^3_+ \cap E_3$ is a half line. Since $\mathbb{S}^3_+ \cap E = \{U_*\}$ and $\dim E = 3$, there exist hyperplanes $E_1 = \{X \in \mathbb{S}^3 : \langle A_1, X \rangle = 1\}$ and $E_2 = \{A_2\}^\perp$ such that $E = E_1 \cap E_2 \cap E_3$. Since $PU_*P^T = U_* \in E$, we see that A_1, A_2, A_3 satisfy

$$A_1 = P^T \begin{pmatrix} 1 & c_1 & c_2 \\ c_1 & c_3 & c_4 \\ c_2 & c_4 & c_5 \end{pmatrix} P,\ A_2 = P^T \begin{pmatrix} 0 & c_6 & c_7 \\ c_6 & c_8 & c_9 \\ c_7 & c_9 & c_{10} \end{pmatrix} P,\ A_3 = P^T \begin{pmatrix} 0 & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & 1 \end{pmatrix} P,$$
$$\mu > 0,\ \{A_1, A_2, A_3\} : \text{linearly independent},$$

for some $\mu, c_i \in \mathbb{R}$. By deleting redundant parameters, we obtain the matrices in Type 1.

If $\operatorname{rank} \widetilde{A}_3 = 1$, then there exist $\lambda > 0$ and an orthogonal matrix $\widetilde{P} \in \mathbb{R}^{2\times 2}$ such that $A_3 = P \left(\begin{smallmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \lambda \end{smallmatrix}\right) P^T$ for $P = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & \multicolumn{2}{c}{\widetilde{P}} \\ 0 & & \end{array}\right)$. We may assume $\lambda = 1$. By a similar argument above, we see that A_1, A_2, A_3 are given by

$$A_1 = P \begin{pmatrix} 1 & c_1 & c_2 \\ c_1 & d_1 & c_3 \\ c_2 & c_3 & 0 \end{pmatrix} P^T,\ A_2 = P \begin{pmatrix} 0 & d_2 & c_4 \\ d_2 & d_3 & c_5 \\ c_4 & c_5 & 0 \end{pmatrix} P^T,\ A_3 = P \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} P^T$$

for some $c_i, d_j \in \mathbb{R}$. In addition, for $X = P \left(\begin{smallmatrix} x_{11} & x_{21} & x_{31} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{23} & x_{33} \end{smallmatrix}\right) P^T \in E$, we have

$$\begin{cases} x_{11} + 2c_1x_{21} + 2c_2x_{31} + d_1x_{22} + 2c_3x_{23} = 1, \\ 2d_2x_{21} + 2c_4x_{31} + d_3x_{22} + 2c_5x_{23} = 0, \quad x_{33} = 0. \end{cases}.$$

If $d_3 = 0$, then $P \begin{pmatrix} 1-d_1t & 0 & 0 \\ 0 & t & 0 \\ 0 & 0 & 0 \end{pmatrix} P^T \in \mathbb{S}^3_+ \cap E$ for $t \in \mathbb{R}$ and hence it contradicts to $\mathbb{S}^3_+ \cap E = \{U_*\}$. Thus $d_3 \neq 0$. In addition, if $d_2 \neq 0$, then $P \begin{pmatrix} 1-d_1t+(c_1d_3/d_2)t & -(d_3/2d_2)t & 0 \\ -(d_3/2d_2)t & t & 0 \\ 0 & 0 & 0 \end{pmatrix} P^T \in E$ for $t \in \mathbb{R}$. Since this matrix is positive semidefinite for sufficiently small $t > 0$, it is a contradiction and hence $d_2 = 0$. Therefore we can write $A_2 = P \begin{pmatrix} 0 & 0 & c_4 \\ 0 & 1 & c_5 \\ c_4 & c_5 & 0 \end{pmatrix} P^T$ and hence $A_1 = P \begin{pmatrix} 1 & c_1 & c_2 \\ c_1 & 0 & c_3 \\ c_2 & c_3 & 0 \end{pmatrix} P^T$ by deleting $(2,2)$ element of A_1 using A_2.

□

Remark 5.2 (relations to singularity degrees)**.** It is well known that an upper bound of the convergence rate of alternating projections of

$$E = \{X \in \mathbb{S}^3 : \langle A_1, X\rangle = 1,\ \langle A_2, X\rangle = 0,\ \langle A_3, X\rangle = 0\}$$

and $\mathbb{S}^3_+$ is given using the singularity degree; [6]. As explained in Example 3.4, the singularity degree of $E \cap \mathbb{S}^3_+$ is either $0, 1$ or 2. A Type 1 plane in Proposition 5.1 has the singularity degree 1 since A_3 itself is positive semidefinite. For a Type 2 plane, a linear combination of A_2 and A_3 can be positive semidefinite if and only if $c_4 = 0$, in case the singularity degree is 1. Therefore the singularity degree is 2 if and only if $c_4 \neq 0$.

5.2. **Plücker embedding.** Let $\mathcal{D}$ be the family of 3-planes that intersect with $\mathbb{S}^3_+$ at $U_* = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. Let B_1, B_2, B_3 be a basis of the linear space

$$E' = \{X \in \mathbb{S}^3 : \langle A_1, X\rangle = 0,\ \langle A_2, X\rangle = 0,\ \langle A_3, X\rangle = 0\}.$$

For the standard basis $e_1, \dots, e_6 \in \mathbb{S}^3$, the Plücker coordinate $\{C_{i_1,i_2,i_3}\}$ is given by

$$B_1 \wedge B_2 \wedge B_3 = \sum \{C_{i_1,i_2,i_3} e_{i_1} \wedge e_{i_2} \wedge e_{i_3} : 1 \le i_1 < i_2 < i_3 \le 6\} \in \bigwedge^3 \mathbb{S}^3.$$

With these coordinates, $\mathcal{D}$ is considered as a subset of the Grassmannian of 3-planes in $\mathbb{S}^3$. By counting the number of parameters and noting that P is a one-parameter matrix, we see that the dimensions of the family of planes of Type 2 is 6. Consider the dimensions of the family of planes of Type 1. Since $A_2 \neq O$, one of c_5, c_6, c_7, c_8 can be replaced with 1 and then the corresponding parameter in A_1 is redundant. Thus we conclude that the dimension of the family of planes of Type 1 is 8.

We can explicitly calculate the defining ideals of these family of planes with `Macaulay 2` as follows. First, define the ring with `SkewCommutative` elements $e_1, \dots, e_6$. We write B_1, B_2, B_3 as linear combinations of $e_1, \dots, e_6$, and then the coefficients of the product of these linear combinations are the coordinates C_{i_1,i_2,i_3}. Let

$$\begin{aligned} I = \langle t \cdot C_{i_1,i_2,i_3} - x_{i_1,i_2,i_3} : 1 \le i_1 < i_2 < i_3 \le 6\rangle + \langle u_1^T u_2, \|u_1\|^2 - 1, \|u_2\|^2 - 1\rangle \\ \subset \mathbb{R}[c_1, \dots, c_5, t, x_{1,2,3}, \dots, x_{4,5,6}]. \end{aligned}$$

Then the ideal

$$J = I \cap \mathbb{R}[x_{1,2,3}, \dots, x_{4,5,6}]$$

is the defining ideal of the Plücker embedding of the 3-planes of Type 2. We can calculate J with the command `elimination` in `Macaulay 2`. Similarly, we obtain the defining ideal of the 3-planes of Type 1.

Then we have the following relation:

$$\begin{array}{llllll} \mathrm{Gr}(3,6) & \supset & \mathcal{D} & = & \{\text{Type 1}\} & \cup\ \{\text{Type 2}\} \\ \dim 9 & & & & \mathrm{SD}=1,\ \dim 8,\ \text{semialg.} & \mathrm{SD}=2,\ \dim 6 \\ & & & & & \mathrm{SD}=1,\ \dim 5 \end{array}$$

The Grassmannian of 3-planes in $\mathbb{S}^3$ has dimension 9. $\mathcal{D}$ is a semialgebraic subset with dimension 8, in which a generic plane intersects $\mathbb{S}^3_+$ with singularity degree 1. The planes with singularity degree 2 form a 6 dimensional subvariety. Within the subvariety, there is a 5 dimensional subvariety whose point corresponds to a plane with singularity degree 1.

6. Rational formulas for a special curve

In this section, we consider a Type 2 plane with $c_4 \neq 0$. As explained in Remark 5.2, the set of such planes is exactly the set of 3-planes whose intersections with $\mathbb{S}^3_+$ are $\{U_*\}$ and have singularity degree 2.

A matrix in a Type 2 plane is written as

$$U_* + p_1 B_1 + p_2 B_2 + p_3 B_3,$$

where $U_* = \left(\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right)$ and

$$B_1 = \begin{pmatrix} -2c_1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 2c_2 & 0 & -1 \\ 0 & 2c_4 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad B_3 = \begin{pmatrix} -2c_3 & 0 & 0 \\ 0 & -2c_5 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \tag{2}$$

and $c_4 \neq 0$. Note that B_1, B_2, B_3 are not necessarily orthogonal to each other.

6.1. **Special curve.** First, we present the special curve $G(t)$ for a Type 2 plane with $c_4 \neq 0$, which is, in fact, the slowest curve for such a plane, as obtained in Example 4.4 and will be used to show the tightness of a known upper bound for the convergence rate of alternating projections in Section 7. Then for a matrix $G(t)$ on the curve, we give rational formulas for $P_{\mathbb{S}^3_+}(G(t))$ and $P_E \circ P_{\mathbb{S}^3_+}(G(t))$ up to degree 7.

Let

$$g_{13} = \tilde{g}_{13} + r_{13}, \quad g_{23} = -\frac{\tilde{g}_{13}}{w} t + r_{23}, \tag{3}$$

where

$$\tilde{g}_{13} = \frac{1}{2c_4 w + 2c_5 t} t^2 - \frac{2(2c_5^2+1)}{(2c_4 w + 2c_5 t)^5} t^6,$$

$$r_{13} = \frac{c_5}{c_4} r_0 t^7 + \frac{\langle B_2, B_1\rangle}{16\, c_4^6\, \|B_1\|^2} t^7, \quad r_{23} = -\frac{2c_5}{(2c_4 w + 2c_5 t)^4 w} t^6 + r_0 t^7,$$

$$r_0 = \left(\frac{c_4\langle B_3, B_1\rangle + c_5\langle B_2, B_1\rangle}{8c_4^5\|B_1\|^2} + \frac{1}{8c_4^3}\right),$$

$$w = 1 - 2\,c_1\,t + \frac{c_2\,t^2}{c_4\left(1 - 2\,c_1\,t + \frac{c_2\,t^2}{c_4\,(1-2\,c_1\,t) + c_5\,t} + \frac{c_3\,t^3}{c_4}\right) + c_5\,t} \tag{4}$$

$$+\frac{c_3\,t^3}{\left(c_4\left(1-2\,c_1\,t+\frac{c_2\,t^2}{c_4}\right)+c_5\,t\right)\left(1-2\,c_1\,t+\frac{c_2\,t^2}{c_4}\right)},$$

Define $g(t)=(t,g_{13},g_{23})$, $h=\dfrac{2t^6}{(2c_4w+2c_5t)^4w}$ and

$$G(t):=\varphi(g(t))=\begin{pmatrix}1-2c_1t+2c_2g_{13}-2c_3g_{23} & t & -g_{13}\\ t & 2c_4g_{13}-2c_5g_{23} & g_{23}\\ -g_{13} & g_{23} & 0\end{pmatrix}. \tag{5}$$

Example 6.1. For $c_1=c_4=1,\ c_2=c_3=c_5=0$, we have $w=1-2t$ and

$$g_{13}=\frac{t^2}{2\,(1-2\,t)}+\frac{t^6}{16\,(1-2\,t)^5},\quad g_{23}=-\frac{t^3}{2\,(1-2\,t)^2}+\frac{3\,t^7}{16\,(1-2\,t)^6},$$
$$G(t)=\begin{pmatrix}1-2t & * & *\\ t & 2g_{13} & *\\ -g_{13} & g_{23} & 0\end{pmatrix},$$

where $*$ in the matrix stands for the corresponding element of the transpose of the lower triangular block part. By changing the variable as $s=t/(1-t)$, we see that

$$\frac{G(t)}{1-2t}=\begin{pmatrix}1 & * & *\\ s & s^2 & *\\ -\frac{1}{2}s^2 & -\frac{1}{2}s^3 & 0\end{pmatrix}+\begin{pmatrix}0 & * & *\\ 0 & -\frac{1}{8}s^6 & *\\ \frac{1}{16}s^6 & \frac{3}{16}s^7 & 0\end{pmatrix}+O(s^8).$$

This is the curve obtained by perturbing a moment curve with the higher order terms.

6.2. **Rational formulas.** The following two formulas give rational expressions for projections $P_{\mathbb{S}^3_+}$ and $P_E\circ P_{\mathbb{S}^3_+}$ along the curve $G(t)$ up to degree 7. The proofs are given in the next section.

Theorem 6.2. *For sufficiently small $t>0$, we have*

$$P_{\mathbb{S}^3_+}(G(t))=G(t)+\begin{pmatrix}0 & * & *\\ -\frac{t^7}{8c_4^4} & h-\frac{\langle B_2,B_1\rangle}{8\,c_4^5\,\|B_1\|^2}t^7 & *\\ c_4h & c_5h-\frac{c_5\langle B_2,B_1\rangle}{8\,c_4^5\,\|B_1\|^2}t^7-\frac{\langle B_3,B_1\rangle}{8c_4^4\|B_1\|^2}t^7 & \frac{g_{13}^2}{w}\end{pmatrix}+O(t^8). \tag{6}$$

Theorem 6.3. *For sufficiently small $t>0$, we have*

$$P_E\circ P_{\mathbb{S}^3_+}(G(t))=G\left(t-\frac{t^7}{4c_4^4\|B_1\|^2}\right)+O(t^8). \tag{7}$$

6.3. **Proof of the formulas.** We use four lemmas to prove the formula. Recall that the rational function w in (4) is used to define g_{13} and g_{23} as in (3). We start with the lemma that shows that w has the following recursive property.

Lemma 6.4. *Let w is the rational function in* (4)*. Then*

$$w=1-2c_1t+2c_2g_{13}-2c_3g_{23}+O(t^6).$$

Proof. Let w_i be the sum of the terms of w with the degree less than or equal to i. Then

$$\begin{aligned} w_2 &= 1 - 2c_1 t + \frac{c_2}{c_4} t^2, \\ w_3 &= 1 - 2c_1 t + \frac{c_2}{c_4(1-2c_1 t) + c_5 t} t^2 + \frac{c_3}{c_4} t^3. \end{aligned}$$

By considering the Taylor expansion, we see that

$$\begin{aligned} 1 - 2c_1 t + 2c_2 g_{13} - 2c_3 g_{23} &= 1 - 2c_1 t + \frac{2c_2}{2c_4 w_3 + 2c_5 t} t^2 - \frac{2c_3}{(2c_4 w_2 + 2c_5 t) w_2} t^3 + O(t^6) \\ &= w + O(t^6). \end{aligned}$$

□

Here, we write $g_{13} = g_{13}(w)$ and $g_{23} = g_{23}(w)$ to specify w in their definitions. Let

$$\widehat{w} := 1 - 2c_1 t + 2c_2 g_{13}(w) - 2c_3 g_{23}(w).$$

By Lemma 6.4, $\widehat{w}$ equals to w up to degree 5. By the similar argument to the proof of Lemma 6.4, we have $g_{13}(\widehat{w}) = g_{13}(w) + O(t^8)$, $g_{23}(\widehat{w}) = g_{23}(w) + O(t^9)$ and hence

$$\widehat{w} = 1 - 2c_1 t + 2c_2 g_{13}(\widehat{w}) - 2c_3 g_{23}(\widehat{w}) + O(t^8).$$

Thus, if $\widehat{w}$ is used in the definition (5) of $G(t)$ instead of w, then $G(t)$ does not change up to degree 7. Since the following arguments only treat equations up to degree 7 except for Lemma 6.7, we can replace w with $\widehat{w}$. To be precise, we keep using w and $\widehat{w}$ separately, but the reader may consider w as $\widehat{w}$.

Decomposition. We will investigate the basic structure of the matrix $G(t)$, such as a decomposition of the low degree terms, the first eigenvalue, and the determinant. First, we will show that $G(t)$ is obtained by perturbing a rank 1 matrix with higher order terms. Recall that $h = \dfrac{2t^6}{(2c_4 w + 2c_5 t)^4 w}$.

Lemma 6.5.

$$G(t) = \widehat{w} \begin{pmatrix} 1 \\ \frac{t}{w} \\ \frac{-g_{13}}{w} \end{pmatrix} \begin{pmatrix} 1 & \frac{t}{w} & \frac{-g_{13}}{w} \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & -h + \dfrac{\langle B_2, B_1 \rangle}{8\, c_4^5 \, \|B_1\|^2} t^7 & g_{23} + \dfrac{t g_{13}}{w} \\ 0 & g_{23} + \dfrac{t g_{13}}{w} & -\dfrac{g_{13}^2}{w} \end{pmatrix} + O(t^8).$$

In addition,

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & -h + \frac{\langle B_2, B_1 \rangle}{8\, c_4^5 \|B_1\|^2} t^7 & g_{23} + \frac{t g_{13}}{w} \\ 0 & g_{23} + \frac{t g_{13}}{w} & -\frac{g_{13}^2}{w} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & O(t^6) & O(t^6) \\ 0 & O(t^6) & O(t^4) \end{pmatrix}.$$

Proof. Recall $\tilde{g}_{13} = \dfrac{1}{2c_4 w + 2c_5 t} t^2 - \dfrac{2(2c_5^2 + 1)}{(2c_4 w + 2c_5 t)^5} t^6$. Let $\tilde{g}_{23} = -\frac{t \tilde{g}_{13}}{w}$. Then we have

$$2c_4 \tilde{g}_{13} - 2c_5 \tilde{g}_{23} = 2c_4 \tilde{g}_{13} + 2c_5 \frac{t \tilde{g}_{13}}{w} = \frac{2c_4 w + 2c_5 t}{w} \tilde{g}_{13} = \frac{t^2}{w} - \frac{2(2c_5^2 + 1) t^6}{(2c_4 w + 2c_5 t)^4 w}$$

$$= \frac{t^2}{w} - (2c_5^2+1)h.$$

In addition,

$$2c_4 r_{13} - 2c_5 r_{23} = \frac{2c_4\langle B_2, B_1\rangle}{16\, c_4^6\, \|B_1\|^2} t^7 + \frac{4c_5^2 t^6}{(2c_4 w + 2c_5 t)^4 w} = \frac{\langle B_2, B_1\rangle}{8\, c_4^5\, \|B_1\|^2} t^7 + 2c_5^2 h.$$

Since $g_{23} = \tilde{g}_{23} + r_{23}$, we obtain

$$2c_4 g_{13} - 2c_5 g_{23} = \frac{t^2}{w} + \frac{\langle B_2, B_1\rangle}{8\, c_4^5\, \|B_1\|^2} t^7 - h. \tag{8}$$

Since $\widehat{w}\frac{t^2}{w^2} = \frac{t^2}{w} + O(t^8)$, we have

$$G(t) - \widehat{w}\begin{pmatrix} 1 \\ \frac{t}{w} \\ \frac{-g_{13}}{w}\end{pmatrix}\begin{pmatrix} 1 & \frac{t}{w} & \frac{-g_{13}}{w}\end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -h + \frac{\langle B_2,B_1\rangle}{8\, c_4^5\, \|B_1\|^2} t^7 & g_{23} + \dfrac{tg_{13}}{w} \\ 0 & g_{23} + \dfrac{tg_{13}}{w} & -\dfrac{g_{13}^2}{w}\end{pmatrix} + O(t^8).$$

In addition,

$$\begin{aligned} g_{23} &= -\frac{tg_{13}}{w} + \left(\frac{c_4\langle B_3, B_1\rangle + c_5\langle B_2, B_1\rangle}{8c_4^5\|B_1\|^2} + \frac{1}{8c_4^3}\right)t^7 - c_5 h \\ &= -\frac{tg_{13}}{w} + \frac{c_4\langle B_3, B_1\rangle + c_5\langle B_2, B_1\rangle}{8c_4^5\|B_1\|^2} t^7 + (-c_5 + c_4 t)h + O(t^8). \end{aligned} \tag{9}$$

In the last equality, we use $\frac{1}{8c_4^3}t^7 = c_4 t h = O(t^8)$. Thus we have

$$g_{23} + \frac{tg_{13}}{w} = -c_5 h + O(t^7) = O(t^6). \tag{10}$$

□

The first eigenvalue and eigenvector. Using the decomposition of the low degree term, we obtain the first eigenvalue and the associated eigenvector of $G(t)$ up to degree 7.

Lemma 6.6. *Let*

$$v = \begin{pmatrix} 1 \\ \dfrac{t}{\widehat{w}} \\ \dfrac{-g_{13}}{w}\end{pmatrix}, \; \delta = \frac{1}{w\|v\|^2}\begin{pmatrix} 0 \\ -\dfrac{th}{w} \\ c_4 h\end{pmatrix}.$$

Then the first eigenvalue $\tilde{\lambda}$ and the associated eigenvector $\tilde{v}$ of $G(t)$ can be written as

$$\tilde{\lambda} = \widehat{w}\|v+\delta\|^2 + O(t^8), \;\; \tilde{v} = v + \delta + O(t^8).$$

Proof. Since $v = (1, O(t), O(t^2))^T$, Lemma 6.5 implies that

$$G(t)v = \widehat{w}\|v\|^2 v + \begin{pmatrix} 0 \\ -\dfrac{th}{w} \\ \dfrac{tg_{23}}{w} + \dfrac{t^2 g_{13}}{w^2} + \dfrac{g_{13}^3}{w^2}\end{pmatrix} + O(t^8).$$

Here, by $h = \frac{2t^6}{(2c_4w+2c_5t)^4w}$, we see that

$$\frac{g_{13}^3}{w^2} = \frac{t^6}{(2c_4w+2c_5t)^3w^2} = \frac{c_4w+c_5t}{w}h.$$

In addition, since $g_{23} + \frac{tg_{13}}{w} = -c_5h + O(t^7)$ as in (10), we have

$$\begin{aligned}\frac{tg_{23}}{w} + \frac{t^2g_{13}}{w^2} + \frac{g_{13}^3}{w^2} = \frac{t}{w}\left(g_{23} + \frac{tg_{13}}{w}\right) + \frac{g_{13}^3}{w^2} &= \frac{-c_5th}{w} + \frac{c_4w+c_5t}{w}h + O(t^8)\\ &= c_4h + O(t^8).\end{aligned}$$

Thus we obtain

$$G(t)v = \widehat{w}\|v\|^2v + (0, -\frac{th}{w}, c_4h)^T + O(t^8) = \widehat{w}\|v\|^2v + \widehat{w}\|v\|^2\delta + O(t^8).$$

Since $v = (1, O(t), O(t^2))$, $\delta = (0, O(t^7), O(t^6))^T$ and $G(t) = \begin{pmatrix} O(1) & * & * \\ O(t) & O(t^2) & * \\ O(t^2) & O(t^3) & 0 \end{pmatrix}$, we have

$$G(t)(v+\delta) = G(t)v + O(t^8) = \widehat{w}\|v+\delta\|^2(v+\delta) + O(t^8). \tag{11}$$

Let $\tilde{\lambda}(t)$ and $\tilde{v}(t)$ be the first eigenvalue and the associated eigenvector of $G(t)$. Then $\tilde{\lambda}(0) = 1$, and we may assume $\tilde{v}(0) = (1,0,0)^T$. Let $\varepsilon = \tilde{\lambda} - \widehat{w}\|v+\delta\|^2$ and $(0, r_1, r_2)^T = \tilde{v} - (v+\delta)$. Then we have

$$G(t)(v+\delta+(0,r_1,r_2)^T) = (\widehat{w}\|v+\delta\|^2 + \varepsilon)(v+\delta+(0,r_1,r_2)^T).$$

Thus the equation (11) implies that

$$\widehat{w}\|v+\delta\|^2\begin{pmatrix}0\\r_2\\r_3\end{pmatrix} + \varepsilon(v+\delta) + \varepsilon\begin{pmatrix}0\\r_2\\r_3\end{pmatrix} + O(t^8) = G(t)\begin{pmatrix}0\\r_2\\r_3\end{pmatrix} = \begin{pmatrix} tr_2 + O(t^2)r_3 \\ O(t^2)r_2 + O(t^3)r_3 \\ O(t^3)r_2 \end{pmatrix}.$$

Since $v+\delta = (1, O(t), O(t^2))$, we have $\varepsilon = tr_2 + O(t^2)r_3 + O(t^8)$. By $\widehat{w}\|v+\delta\|^2 = 1 + O(t)$, we obtain

$$\begin{pmatrix}r_2\\r_3\end{pmatrix} + \begin{pmatrix}O(t^2)r_2 + O(t^3)r_3\\O(t^3)r_2 + O(t^4)r_3\end{pmatrix} + O(t)\begin{pmatrix}r_2\\r_3\end{pmatrix} + O(t^8) = \begin{pmatrix}O(t^2)r_2 + O(t^3)r_3\\O(t^3)r_2\end{pmatrix}.$$

By the second equality, we see that $r_3 = O(t^3)r_2 + O(t^8)$. Thus the first equality gives $r_2 = O(t^8)$. Then we obtain $r_3 = O(t^8)$ and $\varepsilon = O(t^8)$. Therefore, we complete the proof. □

Determinant. When we calculate $P_{\mathbb{S}^3_+}(U)$ for U sufficiently close to U_*, we have to consider two cases; $P_{\mathbb{S}^3_+}(U)$ is rank 2 or rank 1. Suppose that $\lambda_1 \geq \lambda_2 \geq \lambda_3$ are the eigenvalues of U and v_1, v_2, v_3 are the associated eigenvectors of U respectively. Since U is sufficiently close to U_*, we see that λ_1 is close to 1 and $\lambda_3 < 0$. If $\lambda_1, \lambda_2 > 0$, then $P_{\mathbb{S}^3_+}(U) = \lambda_1\frac{v_1v_1^T}{\|v_1\|^2} + \lambda_2\frac{v_2v_2^T}{\|v_2\|^2}$. If $\lambda_1 > 0$ and $\lambda_2 \leq 0$, then $P_{\mathbb{S}^3_+}(U) = \lambda_1\frac{v_1v_1^T}{\|v_1\|^2}$. Thus the rank of $P_{\mathbb{S}^3_+}(U)$ is determined by the sign of $\det U$. We show that the determinant of $G(t)$ is positive for t sufficiently close 0 under higher order perturbations. Then we have that $P_{\mathbb{S}^3_+}(G(t))$ has rank 1.

Lemma 6.7.

$$\det\left(G(t)+\begin{pmatrix}O(t^7) & O(t^7) & O(t^7)\\ O(t^7) & O(t^7) & O(t^7)\\ O(t^7) & O(t^7) & 0\end{pmatrix}\right)=\frac{t^{10}}{32c_4^6}+O(t^{11}).$$

In particular, $P_{\mathbb{S}^3_+}(G(t))$ *has rank* 1 *for* t *sufficiently close to* 0.

Proof. Recall that $g_{13}=O(t^2)$, $g_{23}=O(t^3)$, and that $2c_4g_{13}-2c_5g_{23}=\frac{t^2}{w}-h+O(t^7)$ and $\frac{tg_{13}}{w}+g_{23}=O(t^6)$ by (8) and (10). Then we have

$$\begin{aligned}
&\det(G(t)+O(t^7))\\
&=\begin{vmatrix}\widehat{w}+O(t^7) & t+O(t^7) & -g_{13}+O(t^7)\\ t+O(t^7) & 2c_4g_{13}-2c_5g_{23}+O(t^7) & g_{23}+O(t^7)\\ -g_{13}+O(t^7) & g_{23}+O(t^7) & 0\end{vmatrix}\\
&=(-g_{13}+O(t^7))\begin{vmatrix}t+O(t^7) & -g_{13}+O(t^7)\\ \frac{t^2}{w}-h+O(t^7) & g_{23}+O(t^7)\end{vmatrix}-(g_{23}+O(t^7))\begin{vmatrix}w+O(t^7) & -g_{13}+O(t^7)\\ t+O(t^7) & g_{23}+O(t^7)\end{vmatrix}\\
&=\begin{vmatrix}-tg_{13}-wg_{23}+O(t^8) & -g_{13}+O(t^7)\\ -\frac{t^2g_{13}}{w}+hg_{13}-tg_{23}+O(t^9) & g_{23}+O(t^7)\end{vmatrix}\\
&=\begin{vmatrix}-tg_{13}+wg_{23} & -g_{13}+O(t^7)\\ -\frac{t^2g_{13}}{w}-tg_{23} & g_{23}+O(t^7)\end{vmatrix}+\begin{vmatrix}O(t^8) & -g_{13}+O(t^7)\\ hg_{13}+O(t^9) & g_{23}+O(t^7)\end{vmatrix}\\
&=\begin{vmatrix}-w\left(\frac{tg_{13}}{w}+g_{23}\right) & -g_{13}\\ -t\left(\frac{tg_{13}}{w}+g_{23}\right) & g_{23}\end{vmatrix}+\begin{vmatrix}0 & -g_{13}\\ hg_{13} & g_{23}\end{vmatrix}+O(t^{11})\\
&=\left(g_{23}+\frac{t}{w}g_{13}\right)(-tg_{13}-wg_{23})+hg_{13}^2+O(t^{11})\\
&=hg_{13}^2+O(t^{11})=\frac{2t^{10}}{(2c_4w+2c_5t)^6w}+O(t^{11})=\frac{t^{10}}{32c_4^6}+O(t^{11}).
\end{aligned}$$

□

Finally, we show the Theorem 6.2 and 6.3.

Proof of Theorem 6.2. By Lemma 6.7, we have $\det G(t)>0$. Then $P_{\mathbb{S}^3_+}(G(t))$ has rank 1 and Lemma 6.6 implies that

$$P_{\mathbb{S}^3_+}(G(t))=\tilde{\lambda}\frac{\tilde{v}\tilde{v}^T}{\|\tilde{v}\|^2}=\widehat{w}\|v+\delta\|^2\frac{(v+\delta)(v+\delta)^T}{\|v+\delta\|^2}+O(t^8)=\widehat{w}(v+\delta)(v+\delta)^T+O(t^8).$$

Here we have, for $\delta=(0,\delta_2,\delta_3)$,

$$\begin{aligned}
&(v+\delta)(v+\delta)^T\\
&=vv^T+\delta v^T+v\delta v^T+\delta\delta^T\\
&=vv^T+\begin{pmatrix}0&0&0\\ \delta_2&0&0\\ \delta_3&t\delta_3&0\end{pmatrix}+\begin{pmatrix}0&\delta_2&\delta_3\\ 0&0&t\delta_3\\ 0&0&0\end{pmatrix}+O(t^8)=vv^T+\begin{pmatrix}0&\delta_2&\delta_3\\ \delta_2&0&t\delta_3\\ \delta_3&t\delta_3&0\end{pmatrix}+O(t^8).
\end{aligned}$$

Since $\|v\|^2=1+O(t^2)$, we obtain

$$P_{\mathbb{S}^3_+}(G(t))$$

$$= \widehat{w}vv^T + w\begin{pmatrix} 0 & \delta_2 & \delta_3 \\ \delta_2 & 0 & t\delta_3 \\ \delta_3 & t\delta_3 & 0 \end{pmatrix} + O(t^8) = \widehat{w}vv^T + \begin{pmatrix} 0 & -\frac{th}{w} & c_4 h \\ -\frac{th}{w} & 0 & c_4 th \\ c_4 h & c_4 th & 0 \end{pmatrix} + O(t^8)$$

By Lemma 6.5, we have $\widehat{w}vv^T = G(t) - \begin{pmatrix} 0 & 0 & 0 \\ 0 & -h + \frac{\langle B_2, B_1\rangle}{8\,c_4^5\,\|B_1\|^2}t^7 & g_{23} + \frac{tg_{13}}{w} \\ 0 & g_{23} + \frac{tg_{13}}{w} & -\frac{g_{13}^2}{w} \end{pmatrix}$ and hence

$$P_{\mathbb{S}^3_+}(G(t)) = G(t) + \begin{pmatrix} 0 & -\dfrac{th}{w} & c_4 h \\ -\dfrac{th}{w} & h - \dfrac{\langle B_2, B_1\rangle}{8\,c_4^5\,\|B_1\|^2}t^7 & -g_{23} - \dfrac{tg_{13}}{w} + c_4 th \\ c_4 h & -g_{23} - \dfrac{tg_{13}}{w} + c_4 th & \dfrac{g_{13}^2}{w} \end{pmatrix} + O(t^8)$$

Here the equation (9) implies

$$-g_{23} - \frac{tg_{13}}{w} + c_4 h = -\frac{c_4\langle B_3, B_1\rangle + c_5\langle B_2, B_1\rangle}{8c_4^5\|B_1\|^2}t^7 + c_5 h + O(t^8).$$

This gives the equation (6). □

Proof of Theorem 6.3. Let $P_E(X) = U_* + s_1B_1 + s_2B_2 + s_3B_3$. Then (s_1, s_2, s_3) satisfies

$$\begin{pmatrix} \|B_1\|^2 & \langle B_1, B_2\rangle & \langle B_1, B_3\rangle \\ \langle B_2, B_1\rangle & \|B_2\|^2 & \langle B_2, B_3\rangle \\ \langle B_3, B_1\rangle & \langle B_3, B_2\rangle & \|B_3\|^2 \end{pmatrix}\begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix} = \begin{pmatrix} \langle B_1, X - U_*\rangle \\ \langle B_2, X - U_*\rangle \\ \langle B_3, X - U_*\rangle \end{pmatrix}.$$

Let H be the coefficient matrix of the left hand side and $D(t) = P_{\mathbb{S}^3_+}(G(t)) - G(t)$. Then we have

$$P_E(X) = \begin{pmatrix} B_1 & B_2 & B_3 \end{pmatrix} H^{-1}\begin{pmatrix} \langle B_1, X\rangle \\ \langle B_2, X\rangle \\ \langle B_3, X\rangle \end{pmatrix} + U_* - \begin{pmatrix} B_1 & B_2 & B_3 \end{pmatrix} H^{-1}\begin{pmatrix} \langle B_1, U_*\rangle \\ \langle B_2, U_*\rangle \\ \langle B_3, U_*\rangle \end{pmatrix},$$

and hence

$$\begin{aligned} P_E \circ P_{\mathbb{S}^3_+}(G(t)) = P_E((G+D)(t)) &= P_E(G(t)) + \begin{pmatrix} B_1 & B_2 & B_3 \end{pmatrix} H^{-1}\begin{pmatrix} \langle B_1, D(t)\rangle \\ \langle B_2, D(t)\rangle \\ \langle B_3, D(t)\rangle \end{pmatrix} \\ &= G(t) + \begin{pmatrix} B_1 & B_2 & B_3 \end{pmatrix} H^{-1}\begin{pmatrix} \langle B_1, D(t)\rangle \\ \langle B_2, D(t)\rangle \\ \langle B_3, D(t)\rangle \end{pmatrix}. \end{aligned}$$

By Theorem 6.2, we have

$$\begin{pmatrix} \langle B_1, D(t)\rangle \\ \langle B_2, D(t)\rangle \\ \langle B_3, D(t)\rangle \end{pmatrix} = \begin{pmatrix} -\dfrac{2th}{w} \\ -2c_4\dfrac{\langle B_2, B_1\rangle}{8\,c_4^5\,\|B_1\|^2}t^7 \\ -\dfrac{\langle B_3, B_1\rangle}{4c_4^4\|B_1\|^2}t^7 \end{pmatrix} + O(t^8) = \begin{pmatrix} -\dfrac{t^7}{4c_4^4} \\ -\dfrac{\langle B_2, B_1\rangle}{4c_4^4\|B_1\|^2}t^7 \\ -\dfrac{\langle B_3, B_1\rangle}{4c_4^4\|B_1\|^2}t^7 \end{pmatrix} + O(t^8)$$

$$= -\frac{t^7}{4c_4^4\|B_1\|^2}\begin{pmatrix}\|B_1\|^2\\ \langle B_2, B_1\rangle\\ \langle B_3, B_1\rangle\end{pmatrix} + O(t^8).$$

By the Cramel's rule, we obtain

$$P_E \circ P_{\mathbb{S}^3_+}((G(t))) = G(t) - \frac{t^7}{4c_4^4\|B_1\|^2}B_1 + O(t^8) = G(t) - \frac{t^7}{8c_4^4(2c_1^2+1)}B_1 + O(t^8).$$

□

7. Application to convergence analysis

If the singularity degree of the intersection of $\mathbb{S}^n_+$ and a plane is 2, then an upper bound for the convergence rate is given as $O(k^{-1/6})$. However, as shown in Example 3.4, an upper bound based on the singularity degree is far from tight in general.

Throughout this section, we consider $\mathbb{S}^3_+$ and a Type 2 plane E with $c_4 \neq 0$. Using rational formulas, we have the following theorem, which shows that the upper bound based on the singularity degree is actually tight for alternating projections for E and $\mathbb{S}^3_+$.

Theorem 7.1. *For sufficiently small $t > 0$, let $U_0 = G(t)$ in* (5) *be the initial point and construct the AP sequence $U_1, U_2, \dots$ for E and $\mathbb{S}^3_+$. Then $\|U_k - U_*\| = \Theta\left(k^{-\frac{1}{6}}\right)$. Moreover,*

$$\lim_{k\to\infty}\left(\frac{3}{32c_4^4(2c_1^2+1)^4}\right)^{\frac{1}{6}} k^{\frac{1}{6}}\|U_k - U_*\| = 1.$$

To prove this theorem, we use the following lemmas, which deal with the first eigenvalue and rational expressions for projections of a matrix obtained by perturbing $G(t)$ with terms of degree 7.

By applying the Gram-Schmidt process to B_1, B_2, B_3 in this order, we obtain an orthogonal basis of E and denote it by C_1, C_2, C_3. Note that $C_1 = B_1$. Let λ be the first eigenvalue of $G(t)$ and $v(t)$ be the the associated eigenvector with $v_1(t) = 1$. By Lemma 6.6, we see that

$$\lambda(t) = 1 + O(t), \quad v(t) = \left(1, O(t), O(t^2)\right)^T.$$

Let $H(t) = \beta t^7 C_2 + \gamma t^7 C_3$ and $\begin{pmatrix}\eta_1(t) & \eta_2(t) & \eta_3(t)\\ \eta_2(t) & \eta_4(t) & \eta_5(t)\\ \eta_3(t) & \eta_5(t) & \eta_6(t)\end{pmatrix} = H(t).$

Lemma 7.2. *Let $\tilde\lambda$ be the first eigenvalue of $(G+H)(t)$ and $\tilde v(t) = (\tilde v_1(t), \tilde v_2(t), \tilde v_3(t))$ be the associated eigenvector with $\tilde v_1(t) = 1$. Then*

$$\tilde\lambda(t) = \lambda + \eta_1(t) + O(t^8), \quad \tilde v(t) = v(t) + \begin{pmatrix}0\\ \eta_2(t) + O(t^8)\\ \eta_3(t) + O(t^8)\end{pmatrix}$$

Proof. Since λ and $\tilde\lambda$ are simple eigenvalues of $G(t)$ and $(G+H)(t)$ respectively, both eigenvalues and associated eigenvectors are analytic functions in t. Let ith homogeneous parts of G, λ, v be G_i, λ_i, v_i respectively. Then these are decomposed as

$$G = G_0 + G_1 + \cdots, \quad \lambda = \lambda_0 + \lambda_1 + \cdots, \quad \tilde v = v_0 + v_1 + \cdots.$$

Note that $G_0 = \begin{pmatrix} 1&0&0\\0&0&0\\0&0&0 \end{pmatrix}$, $\lambda_0 = 1$, $v_{0,1} = 1$, $v_{i,1} = 0$ $(i = 1, 2, \ldots)$. Then we have

$$(G_0 + G_1 + \cdots)(v_0 + v_1 + \cdots) = (\lambda_0 + \lambda_1 + \cdots)(v_0 + v_1 + \cdots).$$

Let I be the identity matrix. By comparing the terms of degree n, we obtain

$$G_0 v_n + G_1 v_{n-1} + \cdots + G_n v_0 = \lambda_0 v_n + \lambda_1 v_{n-1} + \cdots + \lambda_n v_0,$$

$$\lambda_n v_0 + (\lambda_0 I - G_0) v_n = G_1 v_{n-1} + \cdots + G_n v_0 - (\lambda_1 v_{n-1} + \cdots \lambda_{n-1} v_1),$$

$$\begin{pmatrix} \lambda_n \\ v_{n,2} \\ v_{n,3} \end{pmatrix} = G_1 v_{n-1} + \cdots + G_n v_0 - (\lambda_1 v_{n-1} + \cdots \lambda_{n-1} v_1). \tag{12}$$

Since the right hand side of (12) has only the terms of the eigenvalues and eigenvectors of degree less than or equal to $n - 1$, λ_n and v_n are determined iteratively by the lower degree parts and the parts of $G(t)$ of degree less than or equal to n.

Next, we consider the eigenpair $\tilde{\lambda}$ and $\tilde{v}$ of $G+H$. Since H consists of homogeneous polynomials of degree 7, we see that $G + H$ coincides with G up to degree 6. Thus $\tilde{\lambda}$ and $\tilde{v}$ satisfy the equation (12) for $n \le 6$ and

$$\begin{aligned} \begin{pmatrix} \tilde{\lambda}_7 \\ \tilde{v}_{7,2} \\ \tilde{v}_{7,3} \end{pmatrix} &= G_1 v_6 + \cdots + (G_7 + H) v_0 - (\lambda_1 v_6 + \cdots \lambda_6 v_1) \\ &= \begin{pmatrix} \lambda_7 \\ v_{7,2} \\ v_{7,3} \end{pmatrix} + H v_0 = \begin{pmatrix} \lambda_7 + \eta_1 \\ v_{7,2} + \eta_2 \\ v_{7,3} + \eta_3 \end{pmatrix}. \end{aligned}$$

□

By using the expressions of the eigenpair of $G + H$, the following lemma shows that the projections of G and $G + H$ onto $\mathbb{S}^3_+$ coincide up to degree 6.

Lemma 7.3. *If $G(t)$, $(G + H)(t)$ are mapped to rank 1 matrices by $P_{\mathbb{S}^3_+}$, then*

$$P_{\mathbb{S}^3_+}((G + H)(t)) = P_{\mathbb{S}^3_+}(G(t)) + \begin{pmatrix} \eta_1 & \eta_2 & \eta_3 \\ \eta_2 & 0 & 0 \\ \eta_3 & 0 & 0 \end{pmatrix} + O(t^8)$$

Proof. Let $\tilde{\lambda}$ and $\tilde{v}$ be the first eigenvalue and the associated eigenvector of $G + H$ with $\tilde{v}_1 = 1$, and $\tilde{\eta} = (0, \eta_2, \eta_3)^T$. Since $\lambda = 1 + O(t)$, $v = (1, O(t), O(t^2))^T$ and $\eta_1, \eta_2, \eta_3 = O(t^7)$, Lemma 7.2 implies that

$$\begin{aligned} &P_{\mathbb{S}^3_+}((G + H)(t)) = \tilde{\lambda} \frac{\tilde{v}\tilde{v}^T}{\|\tilde{v}\|^2} \\ &= (\lambda + \eta_1) \frac{(v + \tilde{\eta})(v + \tilde{\eta})^T}{\|v + \tilde{\eta}\|^2} + O(t^8) = (\lambda + \eta_1) \frac{v v^T + \tilde{\eta} v^T + v \tilde{\eta}^T + \tilde{\eta}\tilde{\eta}^T}{\|v + \tilde{\eta}\|^2} + O(t^8) \\ &= \frac{(\lambda + \eta_1)}{\|v + \tilde{\eta}\|^2} \left(v v^T + \begin{pmatrix} 0 & \eta_2 & \eta_3 \\ \eta_2 & 0 & 0 \\ \eta_3 & 0 & 0 \end{pmatrix} \right) + O(t^8) \end{aligned}$$

$$= \frac{\lambda}{\|v+\tilde{\eta}\|^2} vv^T + +\eta_1 vv^T + \begin{pmatrix} 0 & \eta_2 & \eta_3 \\ \eta_2 & 0 & 0 \\ \eta_3 & 0 & 0 \end{pmatrix} + O(t^8)$$

$$= \frac{\lambda}{\|v\|^2} vv^T + \begin{pmatrix} \eta_1 & \eta_2 & \eta_3 \\ \eta_2 & 0 & 0 \\ \eta_3 & 0 & 0 \end{pmatrix} + O(t^8).$$

□

Let $\widetilde{C}_2,\ \widetilde{C}_3$ be the matrices that are equal to $C_2,\ C_3$ except for the first low and column being zero vectors, respectively. Let $\widetilde{P}_E(X) = P_E(X) - P_E(O)$. Then $\widetilde{P}_E$ is the linear part of P_E and we have

$$P_E \circ P_{\mathbb{S}^3_+}((G+H)(t)) = P_E \left(P_{\mathbb{S}^3_+}(G(t)) + \begin{pmatrix} \eta_1 & \eta_2 & \eta_3 \\ \eta_2 & 0 & 0 \\ \eta_3 & 0 & 0 \end{pmatrix} + O(t^8) \right)$$

$$= P_E \circ P_{\mathbb{S}^3_+}(G(t)) + \widetilde{P}_E \left(\begin{pmatrix} \eta_1 & \eta_2 & \eta_3 \\ \eta_2 & 0 & 0 \\ \eta_3 & 0 & 0 \end{pmatrix} \right) + O(t^8).$$

A representing matrix for $\widetilde{P}_E$ is given as follows.

Lemma 7.4.

$$\widetilde{P}_E \left(\begin{pmatrix} \eta_1 & \eta_2 & \eta_3 \\ \eta_2 & 0 & 0 \\ \eta_3 & 0 & 0 \end{pmatrix} \right) = \begin{pmatrix} C_2 & C_3 \end{pmatrix} \begin{pmatrix} 1 - \frac{\|\widetilde{C}_2\|^2}{\|C_2\|^2} & -\frac{\langle \widetilde{C}_2, \widetilde{C}_3 \rangle}{\|C_2\|^2} \\ -\frac{\langle \widetilde{C}_2, \widetilde{C}_3 \rangle}{\|C_3\|^2} & 1 - \frac{\|\widetilde{C}_3\|^2}{\|C_3\|^2} \end{pmatrix} \begin{pmatrix} \beta t^7 \\ \gamma t^7 \end{pmatrix}. \tag{13}$$

Proof. Recall that $H(t) = \beta t^7 C_2 + \gamma t^7 C_3 = \begin{pmatrix} \eta_1 & \eta_2 & \eta_3 \\ \eta_2 & \eta_4 & \eta_5 \\ \eta_3 & \eta_5 & \eta_6 \end{pmatrix}$. Since we can write $\begin{pmatrix} \eta_1 & \eta_2 & \eta_3 \\ \eta_2 & 0 & 0 \\ \eta_3 & 0 & 0 \end{pmatrix} = \beta t^7 (C_2 - \widetilde{C}_2) + \gamma t^7 (C_3 - \widetilde{C}_3)$, the orthogonality of C_1, C_2, C_3 implies that

$$\widetilde{P}_E \left(\begin{pmatrix} \eta_1 & \eta_2 & \eta_3 \\ \eta_2 & 0 & 0 \\ \eta_3 & 0 & 0 \end{pmatrix} \right) = \sum_{i=1}^{3} \frac{\langle C_i, \beta t^7 (C_2 - \widetilde{C}_2) + \gamma t^7 (C_3 - \widetilde{C}_3) \rangle}{\|C_i\|^2} C_i$$

$$= \frac{-\beta t^7 \langle C_1, \widetilde{C}_2 \rangle - \gamma t^7 \langle C_1, \widetilde{C}_3 \rangle}{\|C_1\|^2} C_1 + \frac{\beta t^7 (\|C_2\|^2 - \langle C_2, \widetilde{C}_2 \rangle) - \gamma t^7 \langle C_2, \widetilde{C}_3 \rangle}{\|C_2\|^2} C_2$$

$$+ \frac{-\beta t^7 \langle C_3, \widetilde{C}_2 \rangle + \gamma t^7 (\|C_3\|^2 - \langle C_3, \widetilde{C}_3 \rangle)}{\|C_3\|^2} C_3.$$

Since $C_1 = B_1 = \begin{pmatrix} -2c_1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, we have $\langle C_1, \widetilde{C}_2 \rangle = \langle C_1, \widetilde{C}_3 \rangle = 0$ and thus

$$\widetilde{P}_E \left(\begin{pmatrix} \eta_1 & \eta_2 & \eta_3 \\ \eta_2 & 0 & 0 \\ \eta_3 & 0 & 0 \end{pmatrix} \right) = \left(\left(1 - \frac{\langle C_2, \widetilde{C}_2 \rangle}{\|C_2\|^2} \right) \beta t^7 - \frac{\langle C_2, \widetilde{C}_3 \rangle}{\|C_2\|^2} \gamma t^7 \right) C_2$$

$$+\left(-\frac{\langle C_3, \widetilde{C}_2\rangle}{\|C_3\|^2}\beta t^7 + \left(1-\frac{\langle C_3, \widetilde{C}_3\rangle}{\|C_3\|^2}\right)\gamma t^7\right) C_3.$$

By the definition of $\widetilde{C}_2$ and $\widetilde{C}_3$, we have the desired expression. □

Let $c = \frac{1}{4c_4^4\|C_1\|^2}$. By Theorem 6.3, we have $P_E \circ P_{\mathbb{S}^3_+}(G(t)) = G(t-ct^7) + O(t^8)$ and hence

$$P_E \circ P_{\mathbb{S}^3_+}((G+H)(t)) = G(t-ct^7) + \widetilde{P}_E\left(\begin{pmatrix}\eta_1 & \eta_2 & \eta_3\\ \eta_2 & 0 & 0\\ \eta_3 & 0 & 0\end{pmatrix}\right) + O(t^8).$$

Thus the distance between an AP sequence and the slowest curve is $N + O(t^8)$, where

$$N := \left\|\widetilde{P}_E\left(\begin{pmatrix}\eta_1 & \eta_2 & \eta_3\\ \eta_2 & 0 & 0\\ \eta_3 & 0 & 0\end{pmatrix}\right)\right\|_F.$$

We will show that N is strictly less than

$$\|H(t)\|_F = \left\|\beta t^7 C_2 + \gamma t^7 C_3\right\|_F = \sqrt{(\|C_2\|_F\beta t^7)^2 + (\|C_3\|_F\gamma t^7)^2}.$$

Lemma 7.5. *Let*

$$R = \begin{pmatrix} 1-\dfrac{\|\widetilde{C}_2\|^2}{\|C_2\|^2} & -\dfrac{\langle \widetilde{C}_2, \widetilde{C}_3\rangle}{\|C_2\|\|C_3\|}\\ -\dfrac{\langle \widetilde{C}_2, \widetilde{C}_3\rangle}{\|C_2\|\|C_3\|} & 1-\dfrac{\|\widetilde{C}_3\|^2}{\|C_3\|^2}\end{pmatrix}.$$

Then we have $\|R\|_2 < 1$ *and*

$$N \le \|R\|_2 \cdot \|H(t)\|_F.$$

Proof. Let $Q = \left(\begin{smallmatrix} Q_{11} & Q_{21}\\ Q_{21} & Q_{22}\end{smallmatrix}\right)$ be the matrix that appears in the RHS of (13). By Lemma 7.4, we have

$$\begin{aligned} N^2 &= \|(Q_{11}\beta t^7 + Q_{21}\gamma t^7)C_2 + (Q_{21}\beta t^7 + Q_{22}\gamma t^7)C_3\|_F^2\\ &= (Q_{11}\beta t^7 + Q_{21}\gamma t^7)^2\|C_2\|_F^2 + (Q_{21}\beta t^7 + Q_{22}\gamma t^7)^2\|C_3\|_F^2\\ &= \left(\left(1-\frac{\|\widetilde{C}_2\|^2}{\|C_2\|^2}\right)\|C_2\|\beta t^7 - \frac{\langle C_2, \widetilde{C}_3\rangle}{\|C_2\|\|C_3\|}\|C_3\|\gamma t^7\right)^2\\ &\qquad + \left(-\frac{\langle \widetilde{C}_2, \widetilde{C}_3\rangle}{\|C_2\|\|C_3\|}\|C_2\|\beta t^7 + \left(1-\frac{\|\widetilde{C}_3\|^2}{\|C_3\|^2}\right)\|C_3\|\gamma t^7\right)^2\\ &= \left\|R\begin{pmatrix}\|C_2\|\beta t^7\\ \|C_3\|\gamma t^7\end{pmatrix}\right\|_2^2. \end{aligned}$$

Then we obtain

$$N \le \|R\|_2 \cdot \left\|\begin{pmatrix}\|C_2\|\beta t^7\\ \|C_3\|\gamma t^7\end{pmatrix}\right\|_2 = \|R\|_2 \cdot \|\beta t^7 C_2 + \gamma t^7 C_3\|_F.$$

Next we estimate $\|R\|_2$. The characteristic polynomial of R is given by

$$p(\lambda) = \left(\lambda - \left(1-\frac{\|\widetilde{C}_2\|^2}{\|C_2\|^2}\right)\right)\left(\lambda - \left(1-\frac{\|\widetilde{C}_3\|^2}{\|C_3\|^2}\right)\right) - \left(\frac{\langle \widetilde{C}_2, \widetilde{C}_3\rangle}{\|C_2\|\|C_3\|}\right)^2$$

$$=\lambda^2 - \left(2 - \frac{\|\widetilde{C}_2\|^2}{\|C_2\|^2} - \frac{\|\widetilde{C}_3\|^2}{\|C_3\|^2}\right)\lambda + 1 - \frac{\|\widetilde{C}_2\|^2}{\|C_2\|^2} - \frac{\|\widetilde{C}_3\|^2}{\|C_3\|^2}$$
$$+ \frac{\|\widetilde{C}_2\|^2\|\widetilde{C}_3\|^2}{\|C_2\|^2\|C_3\|^2} - \left(\frac{\langle \widetilde{C}_2, \widetilde{C}_3\rangle}{\|C_2\|\|C_3\|}\right)^2.$$

Let

$$p_1(\lambda) = p(\lambda) + \left(\frac{\langle \widetilde{C}_2, \widetilde{C}_3\rangle}{\|C_2\|\|C_3\|}\right)^2, \quad p_2(\lambda) = p_1(\lambda) - \frac{\|\widetilde{C}_2\|^2\|\widetilde{C}_3\|^2}{\|C_2\|^2\|C_3\|^2}.$$

Recall that $c_4 \neq 0$ and C_1, C_2, C_3 form an orthogonal basis obtained by applying the Gram-Schmidt process to B_1, B_2, B_3 in this order. By the locations of the nonzero elements in B_i, we can easily see that $\widetilde{C}_2$ and $\widetilde{C}_3$ are linearly independent and hence $|\langle \widetilde{C}_2, \widetilde{C}_3\rangle| < \|C_2\|\|C_3\|$. Thus we obtain $p_2(\lambda) < p(\lambda) < p_1(\lambda)$. Here, we have

$$p_1(\lambda) = 0 \Longleftrightarrow \lambda = 1 - \frac{\|\widetilde{C}_2\|^2}{\|C_2\|^2},\ 1 - \frac{\|\widetilde{C}_3\|^2}{\|C_3\|^2}$$
$$p_2(\lambda) = 0 \Longleftrightarrow \lambda = 1,\ 1 - \frac{\|\widetilde{C}_2\|^2}{\|C_2\|^2} - \frac{\|\widetilde{C}_3\|^2}{\|C_3\|^2}.$$

Since $p(\lambda)$ is a convex quadratic function, each solution λ to $p(\lambda) = 0$ satisfies

$$-1 < 1 - \frac{\|\widetilde{B}_2\|^2}{\|B_2\|^2} - \frac{\|\widetilde{B}_3\|^2}{\|B_3\|^2} < \lambda < 1.$$

Therefore $\|R\|_2 < 1$. □

Now we can show the following proposition, which means that if we choose the initial point sufficiently close to the curve $G(t)$, then the AP sequence moves in the rate of $\Theta(t^7)$ towards the intersection point while remaining in a neighborhood of the curve $G(t)$.

Proposition 7.6. *For each $\varepsilon > 0$, there exists δ, $K > 0$ such that if t, β, γ satisfy $0 < t < \delta$, $\|\beta C_2 + \gamma C_3\|_F < \varepsilon$, then there exist $\tilde{t}, \tilde{\beta}, \tilde{\gamma}$ such that*

$$\|\tilde{\beta} C_2 + \tilde{\gamma} C_3\|_F < \varepsilon,$$
$$0 < t - \frac{1}{4c_4^4\|C_1\|^2}t^7 - Kt^8 \leq \tilde{t} \leq t - \frac{1}{4c_4^4\|C_1\|^2}t^7 + Kt^8 < \delta,$$
$$P_E \circ P_{\mathbb{S}^3_+}\left(G(t) + \beta t^7 C_2 + \gamma t^7 C_3\right) = G(\tilde{t}) + \tilde{\beta}\tilde{t}^7 C_2 + \tilde{\gamma}\tilde{t}^7 C_3.$$

Proof. Let $c = 1/4c_4^4\|C_1\|^2$ and R be the matrix in Lemma 7.5. Note that Lemma 6.7 ensures $P_{\mathbb{S}^3_+}(G(t))$ has rank 1 for sufficiently small $t > 0$. Thus $P_{\mathbb{S}^3_+}(G(t))$ is calculated with the first eigenvalue and the associated eigenvector of $G(t)$. By Lemma 7.3 and Lemma 7.4, we have

$$\widetilde{G}(t, \beta, \gamma) := P_E \circ P_{\mathbb{S}^3_+}(G(t) + \beta t^7 C_2 + \gamma t^7 C_3) = P_E \circ P_{\mathbb{S}^3_+}((G + H)(t))$$
$$= G(t - ct^7) + \begin{pmatrix} C_2 & C_3 \end{pmatrix} R \begin{pmatrix} \beta t^7 \\ \gamma t^7 \end{pmatrix} + O(t^8).$$

Here $\begin{pmatrix} C_2 & C_3 \end{pmatrix} R \begin{pmatrix} \beta \\ \gamma \end{pmatrix} = \beta' C_2 + \gamma' C_3$ for some $\beta', \gamma' \in \mathbb{R}$. Since Lemma 7.5 implies that $\left\| \begin{pmatrix} C_2 & C_3 \end{pmatrix} R \begin{pmatrix} \beta \\ \gamma \end{pmatrix} \right\|_F \leq \|R\|_2 \|\beta C_2 + \gamma C_3\|_F$ and $\|R\|_2 < 1$, we have

$$
\begin{aligned}
&\|\beta C_2 + \gamma C_3\|_F < \varepsilon \\
&\Longrightarrow \exists \beta', \gamma' \text{ s.t } \|\beta' C_2 + \gamma' C_3\|_F < \|R\|_2 \cdot \varepsilon, \\
&\widetilde{G}(t) = G\left(t - ct^7\right) + t^7 \beta' C_2 + t^7 \gamma' C_3 + O(t^8). \qquad (14)
\end{aligned}
$$

For $U_* = \left(\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right)$, let $\varphi_C(p) := U_* + p_1 C_1 + p_2 C_2 + p_3 C_3$. Then $\varphi_C(p)$ is a diffeomorphism between $\mathbb{R}^3$ to E. Define $p(t) := \varphi_C^{-1}(G(t))$, $\tilde{p}(t, \beta, \gamma) := \varphi_C^{-1}(\widetilde{G}(t))$. Since the first eigenvalue is simple, we see that the eigenvalue and the associated eigenvectors of $G(t) + \beta t^7 C_2 + \gamma t^7 C_3$ are analytic in (t, β, γ) and hence $\widetilde{G}(t, \beta, \gamma) = U_* + \tilde{p}_1(t, \beta, \gamma) C_1 + \tilde{p}_2(t, \beta, \gamma) C_2 + \tilde{p}_3(t, \beta, \gamma) C_3$ is also analytic. Thus, by the Taylor expansion with respect to t about 0, we can actually rewrite (14) as

$$
\begin{aligned}
\widetilde{G}(t) &= G(t - ct^7) + t^7 \beta' C_2 + t^7 \gamma' C_3 + t^8 (r_1 C_1 + r_2 C_2 + r_3 C_3), \\
&= U_* + \left(p_1(t - ct^7) + r_1 t^8\right) C_1 + \left(p_2(t - ct^7) + \beta' t^7 + r_2 t^8\right) C_2 \\
&\qquad + \left(p_3(t - ct^7) + \gamma' t^7 + r_3 t^8\right) C_3 \qquad (15)
\end{aligned}
$$

where

$$
r_i(t, \beta, \gamma) = \frac{1}{8!} \frac{\partial^8 \tilde{p}_i}{\partial t^8} (\theta_i t, \beta, \gamma),
$$

for some $\theta_i \in (0, 1)$, $i = 1, 2, 3$. Let $r = (r_1, r_2, r_3)$ and

$$
D_\delta = \{(t, \beta, \gamma) \in \mathbb{R}^3 : 0 < t < \delta,\ \|\beta C_2 + \gamma C_3\|_F < \varepsilon\}.
$$

Now we have

$$
\begin{aligned}
&\sup\{\|r(t, \beta, \gamma)\| : (t, \beta, \gamma) \in D_\delta\} \\
&\leq \sup \left\{ \left(\sum_{i=1}^{3} \left| \frac{1}{8!} \frac{\partial^8 \tilde{p}_i}{\partial t^8} (t, \beta, \gamma) \right|^2 \right)^{\frac{1}{2}} : (t, \beta, \gamma) \in D_\delta \right\} =: K_\delta.
\end{aligned}
$$

First, we show that there exists $\tilde{t}$ such that $p_1(t - ct^7) + r_1 t^8 = p_1(\tilde{t})$ and $|\tilde{t} - (t - ct^7)| \leq 2K_\delta t^8$. Since $G(t) = U_* + tB_1 + g_2(t) B_2 + g_3(t) B_3 = U_* + p_1(t) C_1 + p_2(t) C_2 + p_3(t) C_3$ and $B_1 = C_1$, we see that $p_1(t) = t + \frac{\langle C_1, B_2 \rangle}{\|C_1\|^2} g_2(t) + \frac{\langle C_1, B_3 \rangle}{\|C_1\|^2} g_3(t)$ and hence $p_1'(0) = 1$. By taking δ smaller if necessary, we may assume that for $0 < t < \delta$, $p_1'(t - ct^7) > 1/2$ and that $p_1(t - ct^7) + r_1 t^8$ is in the range of the inverse function g of p_1 around $t - ct^7$. Then the Taylor's theorem implies that

$$
\begin{aligned}
\tilde{t} &:= g\left(p_1(t - ct^7) + r_1 t^8\right) \\
&= g\left(p_1(t - ct^7)\right) + g'\left(p_1(t - ct^7)\right) r_1 t^8 + \frac{1}{2} g''(t - ct^7 + \theta r_1 t^8)(r_1 t^8)^2, \\
&= t - ct^7 + \left(g'\left(p_1(t - ct^7)\right) + \frac{1}{2} g''(t - ct^7 + \theta r_1 t^8) r_1 t^8\right) r_1 t^8,
\end{aligned}
$$

for some $\theta \in (0,1)$. Since $g'(p_1(t-ct^7)) < 2$, we have $|t-ct^7-\tilde{t}| \le 2K_\delta t^8$. Next, the equation (15) and $\|\beta' C_2 + \gamma' C_3\|_F < \|R\|_2 \varepsilon$ imply that

$$\begin{aligned}&\|\tilde{p}_2(t)C_2 + \tilde{p}_3(t)C_3 - (p_2(t-ct^7)C_2 + p_3(t-ct^7)C_3)\|_F \\ &= \|t^7\beta' C_2 + t^7\gamma' C_3 + t^8(r_1C_1 + r_2C_2 + r_3C_3)\|_F \le \|R\|_2\varepsilon t^7 + \max_{i\in[3]} \|C_i\|_F K_\delta t^8.\end{aligned}$$

Let $g(t) = (t, g_2(t), g_3(t))$ be the functions defining the slowest curve (5). Recall that $g_2(t) = O(t^2)$, $g_3(t) = O(t^3)$. Since $G(t) = U_* + tB_1 + g_2(t)B_2 + g_3(t)B_3 = U_* + p_1(t)C_1 + p_2(t)C_2 + p_3(t)C_3$ and $B_1 = C_1$, we see that $g_2(t)\langle B_2, C_2\rangle + g_3(t)\langle B_3, C_2\rangle = p_2(t)\|C_2\|^2$ and hence $p_2(t) = O(t^2)$. Similarly, $p_3(t) = O(t^2)$. Thus we have

$$\begin{aligned}&\left\|\tilde{p}_2(t)C_2 + \tilde{p}_3(t)C_3 - (p_2(\tilde{t})C_2 + p_3(\tilde{t})C_3)\right\|_F \\ &\le \left\|\tilde{p}_2(t)C_2 + \tilde{p}_3(t)C_3 - (p_2(t-ct^7)C_2 + p_3(t-ct^7)C_3)\right\|_F \\ &\qquad + \left\|p_2(t-ct^7)C_2 + p_3(t-ct^7)C_3 - (p_2(\tilde{t})C_2 + p_3(\tilde{t})C_3)\right\|_F \\ &\le \|R\|_2\varepsilon t^7 + \max_{i\in[3]}\|C_i\|_F(K_\delta + K'_\delta)t^8,\end{aligned}$$

for some $K'_\delta = O(\delta)$. Since $\|R\|_2 < 1$, by taking δ smaller if necessary, we obtain

$$\begin{aligned}&0 < t - ct^7 - 2K_\delta t^8 \le \tilde{t} \le t - ct^7 + 2K_\delta t^8 < \delta, \\ &\frac{1}{\tilde{t}^7}\left\|(\tilde{p}_2(t)C_2 + \tilde{p}_3(t)C_3) - (p_2(\tilde{t})C_2 + p_3(\tilde{t})C_3)\right\|_F < \varepsilon,\end{aligned}$$

for all $(t, \beta, \gamma) \in D_\delta$. The second inequality implies that there exist $\tilde{\beta}, \tilde{\gamma}$ such that

$$\begin{aligned}&\|\tilde{\beta}C_2 + \tilde{\gamma}C_3\|_F < \varepsilon, \\ &\tilde{p}_2(t)C_2 + \tilde{p}_3(t)C_3 = p_2(\tilde{t})C_2 + p_3(\tilde{t})C_3 + \tilde{\beta}\tilde{t}^7C_2 + \tilde{\gamma}\tilde{t}^7C_3.\end{aligned}$$

Therefore we have

$$\begin{aligned}\widetilde{G}(t) &= U_* + p_1(\tilde{t})C_1 + p_2(\tilde{t})C_2 + p_3(\tilde{t})C_3 + \tilde{\beta}\tilde{t}^7C_2 + \tilde{\gamma}\tilde{t}^7C_3 \\ &= G(\tilde{t}) + \tilde{\beta}\tilde{t}^7C_2 + \tilde{\gamma}\tilde{t}^7C_3.\end{aligned}$$

□

We use the following lemma on a recursive sequence.

Lemma 7.7. *Suppose that the sequence $\{x_k\}$ satisfies $(q+1)C - (q+2)Kx_0 > 0$ and*

$$0 < x_{k-1}(1 - Cx_{k-1}^q - Kx_{k-1}^{q+1}) \le x_k \le x_{k-1}(1 - Cx_{k-1}^q + Kx_{k-1}^{q+1}) \quad (k = 1, 2, \ldots)$$

for some $C, K > 0$, $q \in \mathbb{N}$. Then

$$\lim_{k\to\infty} (qC)^{\frac{1}{q}} k^{\frac{1}{q}} x_k = 1.$$

Proof. We show $x_k \to 0$. Suppose $\alpha := \inf_k x_k > 0$. Let $f(x) = -Cx^{q+1} + Kx^{q+2}$ and $M = (q+1)C/((q+2)K)$. Then $f'(x) = -(q+1)Cx^q + (q+2)Kx^{q+1} = (-(q+1)C + (q+2)Kx)x^q < 0$ for $0 < x < M$. Since $0 < x_0 < M < C/K$, we see $x_1 \le x_0 + f(x_0) < x_0$, and hence we obtain inductively that $x_k \le x_{k-1} + f(x_{k-1}) < x_{k-1}$ for all k. Since $f(x)$ is decreasing for $0 < x < M$, we see $\alpha \le x_k \le x_{k-1} + f(x) < x_{k-1} - C\alpha^{q+1} + K\alpha^{q+2}$. Thus we have $\alpha < \alpha + C\alpha^{q+1} - K\alpha^{q+2} \le x_{k-1}$ for all k. Therefore, a contradiction occurs and hence $\inf_k x_k = 0$. Since x_k is decreasing, we obtain $x_k \to 0$. Then almost identical arguments in the proof of [11, Lemma 3.1] gives the result. $\square$

Proof of Theorem 7.1. Let $\varepsilon > 0$. By Lemma 6.7, for each β, γ with $\|\beta C_2 + \gamma C_3\|_F < \varepsilon$, there exists $\delta' > 0$ such that for $t \in [0, \delta')$, $\det(G(t) + \beta t^7 C_2 + \gamma t^7 C_3) > 0$ and hence $P_{\mathbb{S}^3_+}(G(t) + \beta t^7 C_2 + \gamma t^7 C_3)$ has rank 1. Let $c := \frac{1}{4c_4^4\|C_1\|^2} = \frac{1}{4c_4^4(4c_1^2+2)}$. By iteratively applying Proposition 7.6, there exists $\delta > 0$ such that for $\beta_0 = \gamma_0 = 0$ and t_0 with $0 < t_0 < \delta$, we can construct (t_k, β_k, γ_k), $k = 1, 2, \ldots$, satisfying

$$\|\beta_k C_2 + \gamma_k C_3\|_F < \varepsilon,$$
$$0 < t_{k-1} - ct_{k-1}^7 - Kt_{k-1}^8 \le t_k \le t_{k-1} - ct_{k-1}^7 + Kt_{k-1}^8 < \delta,$$
$$P_E \circ P_{\mathbb{S}^3_+}\left(G(t_{k-1}) + \beta_{k-1}t_{k-1}^7 C_2 + \gamma_{k-1}t_{k-1}^7 C_3\right) = G(t_k) + \beta_k t_k^7 C_2 + \gamma_k t_k^7 C_3,$$

for some $K > 0$. Then Lemma 7.7 implies $\lim_{k\to\infty}(6c)^{\frac{1}{6}}k^{\frac{1}{6}}t_k = 1$. Since $\|U_k - U_*\| = (4c_1^2+2)^{\frac{1}{2}}t_k + O(t_k^2)$, we obtain

$$\begin{aligned} 1 = \lim_{k\to\infty}(6c)^{\frac{1}{6}}k^{\frac{1}{6}}t_k &= \lim_{k\to\infty}(6c)^{\frac{1}{6}}k^{\frac{1}{6}}\left((4c_1^2+2)^{-\frac{1}{2}}\|U_k - U_*\| + O(t_k^2)\right) \\ &= \lim_{k\to\infty}\left(\frac{3}{32c_4^4(2c_1^2+1)^4}\right)^{\frac{1}{6}}k^{\frac{1}{6}}\|U_k - U_*\|. \end{aligned}$$

$\square$

Numerical experiments. Figure 5 is consistent with our claim that the convergence rate of $\|U_k - U_*\|$ is $\Theta(k^{-1/6})$ in the case that $c_1 = c_4 = 1, c_2 = c_3 = c_5 = 0$ as in Example 6.1 and the initial point is taken from the slowest curve. We observe from the right of Figure 5 that the plot of $1/\|U_k - U_*\|^6$ approximately coincides with the line $196.0 + 0.0098k$. Hence $\|U_k - U_*\| \approx (196.0 + 0.0098k)^{-1/6} \approx 2.16k^{-1/6}$ for sufficiently large k. The estimate in Theorem 7.1 gives $\|U_k - U_*\| \approx (32\cdot 27)^{1/6}k^{-1/6} \approx 3.08k^{-1/6}$. This discrepancy between the coefficient given in Theorem 7.1 and the results of the numerical experiments is likely due to the slow convergence of the limit in the estimate.

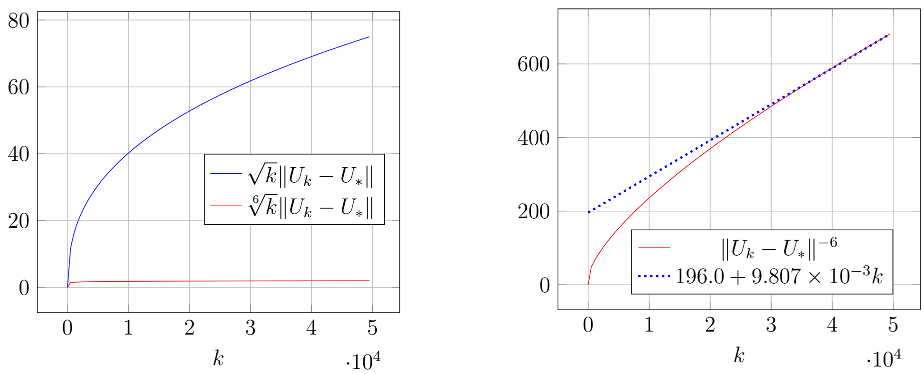

FIGURE 5. The left figure displays the plots of $\sqrt{k}\|U_k - U_*\|$ and $\sqrt[6]{k}\|U_k - U_*\|$ in Example 6.1 with the initial point on the slowest curve, and the right figure displays the plots of $\|U_k - U_*\|^{-6}$ and the line fitting.

8. Conclusion

In this paper, we derived three new analytic formulas for sequences constructed by the alternating projection method applied to an affine space and the cone of positive semidefinite matrices. In particular, using the first formula, we presented examples that demonstrate gaps between the actual convergence rates and the upper bounds based on singularity degrees. The second formula was used to construct the slowest curve for a concrete instance of a 3-plane. The generalization of the slowest curve for the parametric family of 3-planes gives rise to the third formula. The third formula was applied to show the tightness of the convergence rate of the alternating projection method when applied to a 3-plane and $\mathbb{S}^3_+$ whose intersection is a singleton and has singularity degree 2.

We formulate our results in this paper only for cases where the intersection is a singleton, for simplicity of the argument. However, under certain conditions, the argument can be extended to the non-singleton intersection case, which will be the subject of future study.

9. Acknowledgments

The first author was supported by JSPS KAKENHI Grant Number JP17K18726 and JSPS Grant-in-Aid for Transformative Research Areas (A) (22H05107). The second author was supported by JSPS KAKENHI Grant Number JP19K03631 and JP24K06841. The third author was supported by JSPS KAKENHI Grant Number JP20K11696, JP24K14843 and ERATO HASUO Metamathematics for Systems Design Project (No.JPMJER1603), JST.

References

[1] H.H. Bauschke, J.M. Borwein, *On projection algorithms for solving convex feasibility problems*, SIAM Rev. 38 (1996), pp. 367–426.
[2] J.M. Borwein, H. Wolkowicz, *Facial reduction for a cone-convex programming problem*, J. Aust. Math. Soc. 30(3) pp. 369–380 (1980/81).
[3] J.M. Borwein, G. Li, L. Yao, *Analysis of the convergence rate for the cyclic projection algorithm applied to basic semialgebraic convex sets*, SIAM, J. Optim 24 (2014), pp. 498–527.
[4] P.L. Combettes,, *The convex feasibility problem in image recovery*, Advances in imaging and electron physics 95 (1996), pp. 155–270.
[5] F. Deutsch, *The method of alternating orthogonal projections*, in Approximation theory, spline functions and applications. Springer, 1992, pp. 105–121.
[6] D. Drusvyatskiy, G. Li, H. Wolkowicz, *A note on alternating projections for ill-posed semidefinite feasibility problems*, Math. Program 162 (2017), pp. 537–548.
[7] D. Drusvyatskiy, H. Wolkowicz, *The many faces of degeneracy in conic optimization*, Foundations and Trends in Optimization, Vol. 3, No. 2, pp. 77–170, 2017.
[8] H. Federer, *Curvature measures*, Trans. Amer. Math. Soc. 93 (1959), pp. 418–491.
[9] J.-B. Hiriart-Urruty, J. Malick, *A Fresh Variational-Analysis Look at the Positive Semidefinite Matrices World*, J. Optim. Theory Appl. 153 (2012), pp. 551–557.
[10] S. Marchesini, Y.-C. Tu, H.-T. Wu, *Alternating projection, ptychographic imaging and phase synchronization*, Applied and Computational Harmonic Analysis, Vol. 41, Issue 3, pp. 815–851 (2016).
[11] H. Ochiai, Y. Sekiguchi, H. Waki, *Exact Convergence Rates of Alternating Projections for Nontransversal Intersections*, Jpn. J. Ind. Appl. Math 41, pp. 57–83 (2024).
[12] H. Ochiai, Y. Sekiguchi, H. Waki, *Expansions of the Characteristic Polynomial of a Perturbed Positive Semidefinite Matrix and Convergence Analysis of Alternating Projections*, arXiv:2401.16689 [math.OC].

[13] G. Pataki, *Strong duality in conic linear programming: facial reduction and extended dual*, in Bailey, D. et al. (eds.): Computational and Analytical Mathematics, Springer Proceedings in Mathematics & statistics, 50, pp. 613–634. Springer, New York (2013).

(H. Ochiai, H. Waki) Institute of Mathematics for Industry, Kyushu University, 744 Motooka, Nishi-ku, Fukuoka 819-0395, Japan

Email address: ochiai@imi.kyushu-u.ac.jp, waki@imi.kyushu-u.ac.jp

(Y. Sekiguchi) Graduate School of Marine Science and Technology, Etchujima 2-1-8, Koto-ku, Tokyo 135-8533, Japan

Email address, Corresponding author: yoshi-s@kaiyodai.ac.jp