

Analysis of the convergence rate for the cyclic projection algorithm applied to basic semi-algebraic convex sets

Jonathan M. Borwein*, Guoyin Li[†] and Liangjin Yao[‡]

Revised version: November 16 2013

Abstract

In this paper, we study the rate of convergence of the cyclic projection algorithm applied to finitely many basic semi-algebraic convex sets. We establish an explicit convergence rate estimate which relies on the maximum degree of the polynomials that generate the basic semi-algebraic convex sets and the dimension of the underlying space. We achieve our results by exploiting the algebraic structure of the basic semi-algebraic convex sets.

2010 Mathematics Subject Classification:

Primary 41A25, 90C25; Secondary 41A50, 90C31

Keywords: Cyclic projection algorithm, convex polynomial, distance function, Fejér monotone sequence, Hölderian regularity, Łojasiewicz's inequality, projector operator, basic semi-algebraic convex set, von Neumann alternating projection method.

1 Introduction

A very common problem in diverse areas of mathematics and engineering consists of trying to find a point in the intersection of closed convex sets C_i , $i = 1, \dots, m$. This problem is often referred to as the convex feasibility problem. One popular method for solving the convex feasibility problem is the so-called cyclic projection algorithm. Mathematically, the

*CARMA, University of Newcastle, Newcastle, New South Wales 2308, Australia. E-mail: jonathan.borwein@newcastle.edu.au. Laureate Professor at the University of Newcastle and Distinguished Professor at King Abdul-Aziz University, Jeddah.

[†]Department of Applied Mathematics, University of New South Wales, Sydney 2052, Australia. E-mail: g.li@unsw.edu.au.

[‡]CARMA, University of Newcastle, Newcastle, New South Wales 2308, Australia. E-mail: liangjin.yao@newcastle.edu.au.

cyclic projection algorithm is formulated as follows. Given finitely many closed convex sets C_1, C_2, \dots, C_m in \mathbb{R}^n with $\bigcap_{i=1}^m C_i \neq \emptyset$, let $x_0 \in \mathbb{R}^n$ and $P_i := P_{C_i}, i = 1, 2, \dots, m$, where P_{C_i} denotes the Euclidean projection to the set C_i . The sequence of *cyclic projections*, $(x_k)_{k \in \mathbb{N}}$, is defined by

$$(1.1) \quad x_1 := P_1 x_0, x_2 := P_2 x_1, \dots, x_m := P_m x_{m-1}, x_{m+1} := P_1 x_m \dots$$

When $m = 2$, the cyclic projection method reduces to the well known von Neumann alternating projection method (APM) (see [41] and also [8, 12, 13, 27] for some recent developments). The cyclic projection method has attracted much recent interest due to its simplicity and to numerous applications to diverse areas such as engineering and the physical sciences, see [7, 6, 10, 20, 21] and the references therein.

The convergence properties of cyclic projection methods have been examined by many researchers. In particular, Bregman [19] showed that the sequence $(x_k)_{k \in \mathbb{N}}$ generated by the cyclic projection algorithm, always converges to a point in C . Moreover, linear convergence of cyclic projection algorithm onto convex sets with regular intersections was shown in [23]. On the other hand, for convex sets with irregular intersections (for example, when the intersection is a singleton), the cyclic projection algorithm may not exhibit linear convergence even for simple two dimensional cases as observed by [9, Example 5.3] (see Section 5 for more examples). This then raises the following basic question:

Can we estimate the convergence rate of the cyclic projection algorithm for convex sets with possibly irregular intersections?

In this paper, we provide an answer for the above question by focusing on the case where each set C_i is a *basic semi-algebraic convex set* in \mathbb{R}^n in the sense that there exist $\gamma_i \in \mathbb{N}$ and convex polynomial functions, $g_{ij}, j = 1, \dots, \gamma_i$ such that

$$C_i = \{x \in \mathbb{R}^n \mid g_{ij}(x) \leq 0, j = 1, \dots, \gamma_i\}.$$

The main motivation for examining basic semi-algebraic convex sets lies with the following two facts. First, as recently established in [2, 3], optimization problems involving semi-algebraic structure have a number of remarkable properties (such as the celebrated Kurdyka-Łojasiewicz inequality) which enables us to obtain useful qualitative information of the problem. Second, the class of basic semi-algebraic convex sets is a broad class of convex sets which covers polyhedra and convex sets described by convex quadratic functions. Additionally, the structure can often be relatively easily identified [24].

By exploiting the precise algebraic structure, we are able to provide an explicit rate for the cyclic projection algorithm applied to finitely many basic semi-algebraic convex sets *without any regularity conditions*. More precisely, let C_i be basic semi-algebraic convex sets generated by polynomials in \mathbb{R}^n with degree at most $d \in \mathbb{N}$. We show that the sequence of cyclic projections $(x_k)_{k \in \mathbb{N}}$ (1.1) converges (at least) at the rate of $\frac{1}{k^\rho}$ when $d > 1$, where

$\rho := \frac{1}{\min\{(2d-1)^n-1, 2\beta(n-1)d^n-2\}}$ and $\beta(s)$ denotes the *central binomial coefficient* with respect to s —which is given by $\binom{s}{[s/2]}$.¹ When $d = 1$, the sequence of cyclic projections converges linearly.

The remainder of this paper is organized as follows. In Section 2, we collect notation and auxiliary results for future use and for the reader's convenience. In Section 3, we give a Hölderian regularity result for finitely many basic semi-algebraic convex sets. The proof of our main result (Theorem 4.2) forms the bulk of Section 4. In Section 5, we explore various concrete examples. Finally, we end the paper with some conclusions and open questions.

2 Preliminaries and auxiliary results

We assume throughout that \mathbb{R}^n is a Euclidean space with the norm $\|\cdot\|$ and inner product $\langle \cdot, \cdot \rangle$, where $n \in \mathbb{N} := \{1, 2, 3, \dots\}$. We reserve $d \in \mathbb{N}$. We denote by $\mathbb{B}(x, \varepsilon) := \{y \in \mathbb{R}^n \mid \|y - x\| < \varepsilon\}$. We adopt standard notation used in these books [12, 18, 37, 39, 42].

Given a subset C of \mathbb{R}^n , $\text{int } C$ is the *interior* of C , $\text{bd } C$ is the *boundary* of C , $\text{aff } C$ is the *affine hull* of C and \overline{C} is the norm *closure* of C . The *orthogonal set* is $C^\perp := \{x^* \in \mathbb{R}^n \mid (\forall c \in C) \langle x^*, c \rangle = 0\}$. The *distance function* to the set C , written $\text{dist}(\cdot, C)$, is defined by $x \mapsto \inf_{c \in C} \|x - c\|$. The *projector operator* to the set C , denoted by P_C , is defined by

$$P_C(x) := \{c \in C \mid \|x - c\| = \text{dist}(x, C)\}, \quad \forall x \in \mathbb{R}^n.$$

Let $D \subseteq \mathbb{R}^n$. The distance of two sets: C and D , is $\text{dist}(C, D) := \inf_{c \in C, d \in D} \|c - d\|$. Given $f: X \rightarrow]-\infty, +\infty]$, we set $\text{dom } f := f^{-1}(\mathbb{R})$. We say f is *proper* if $\text{dom } f \neq \emptyset$. Let f be a proper function on \mathbb{R}^n . Its associated *recession function* f^∞ is defined, for any $v \in \mathbb{R}^n$, by

$$f^\infty(v) := \liminf_{t \rightarrow \infty, v' \rightarrow v} \frac{f(tv')}{t}.$$

If f is further assumed to be lower semicontinuous and convex, one has (see [4, Proposition 2.5.2])

$$(2.1) \quad f^\infty(v) = \lim_{t \rightarrow \infty} \frac{f(x + tv) - f(x)}{t} = \sup_{t > 0} \frac{f(x + tv) - f(x)}{t} \quad \text{for all } x \in \text{dom } f.$$

2.1 Notation and facts on polynomials

Recall that $f: \mathbb{R}^n \rightarrow \mathbb{R}$ is a *polynomial* if there exists a number $r \in \mathbb{N}$ such that

$$f(x) := \sum_{0 \leq |\alpha| \leq r} \lambda_\alpha x^\alpha,$$

¹Here, $[a]$ denotes the integer part of a

where $\lambda_\alpha \in \mathbb{R}$, $x = (x_1, \dots, x_n)$, $x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, $\alpha_i \in \mathbb{N} \cup \{0\}$, and $|\alpha| := \sum_{j=1}^n \alpha_j$. The corresponding constant r is called the *degree* of f .

Next let us recall a useful property of polynomial functions.

Fact 2.1 (See [5, Remark 4]) *Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be polynomial, and $\{x_1, x_2\} \subseteq \mathbb{R}^n$. If f is constant on $D := [x_1, x_2]$, then f is constant on $\text{aff} D$.*

We now summarize some basic properties of convex polynomials that will be used later. The first property is a Frank-Wolfe type result for convex polynomial optimization problems while the second one is a directional-constancy property for a convex polynomial.

Fact 2.2 (Belousov) (See [15, Theorem 13, Section 4, Chapter II] or [16, Theorem 3] and [35].) *Let f be a convex polynomial on \mathbb{R}^n . Consider a set $D := \{x \mid g_i(x) \leq 0, i = 1, \dots, m\}$, where each g_i , $i = 1, \dots, m$, is a convex polynomial on \mathbb{R}^n . Suppose that $\inf_{x \in D} f(x) > -\infty$. Then f attains its minimum on D .*

Fact 2.3 (See [4, Proposition 3.2.1].) *Let f be a convex polynomial on \mathbb{R}^n and $v \in \mathbb{R}^n$. Assume that $f^\infty(v) = 0$. Then $f(x + tv) = f(x)$ for all $t \in \mathbb{R}$ and for all $x \in \mathbb{R}^n$.*

2.2 Notation and facts on semialgebraic sets/functions

Following [17], a set $D \subseteq \mathbb{R}^n$ is said to be *semi-algebraic* if

$$D := \bigcup_{i=1}^l \bigcap_{j=1}^s \{x \in \mathbb{R}^n \mid f_{ij}(x) = 0, h_{ij}(x) \leq 0\}$$

for some integers l, s and some polynomial functions f_{ij}, h_{ij} on \mathbb{R}^n ($1 \leq i \leq l, 1 \leq j \leq s$). Moreover, a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is said to be *semi-algebraic* if its *graph* $\text{gph} f := \{(x, f(x)) \mid x \in \mathbb{R}^n\}$ is semi-algebraic.

We now summarize below some basic properties of semi-algebraic sets and semi-algebraic functions. These properties will be useful for our later work.

Fact 2.4 *The following statements hold (the properties (P1) and (P4) are direct from the definitions).*

(P1) *Any polynomial is a semi-algebraic function.*

(P2) (See [17, Proposition 2.2.8].) *Let D be a semi-algebraic set. Then $\text{dist}(\cdot, D)$ is a semi-algebraic function.*

(P3) (See [17, Proposition 2.2.6].) *If f, g are semi-algebraic functions on \mathbb{R}^n and $\lambda \in \mathbb{R}$ then $f + g, \lambda f, \max\{f, g\}, fg$ are semi-algebraic.*

(P4) If f_i are polynomials, $i = 1, \dots, m$, and $\lambda \in \mathbb{R}$, then the sets $\{x \mid f_i(x) = \lambda, i = 1, \dots, m\}$, $\{x \mid f_i(x) \leq \lambda, i = 1, \dots, m\}$ are semi-algebraic sets.

(P5) (Łojasiewicz's inequality) (See [17, Corollary 2.6.7].) If ϕ, ψ are two continuous semi-algebraic functions on compact semi-algebraic set $K \subseteq \mathbb{R}^n$ such that $\emptyset \neq \phi^{-1}(0) \subseteq \psi^{-1}(0)$ then there exist constants $c > 0$ and $\tau \in (0, 1]$ such that

$$|\psi(x)| \leq c|\phi(x)|^\tau \quad \text{for all } x \in K.$$

Remark 2.5 As pointed out by [34], the corresponding exponent τ in Łojasiewicz's inequality (P5) is hard to determine and is typically not known. \diamond

Remark 2.6 Let $g_i, i = 1, \dots, m$ be polynomials on \mathbb{R}^n and set $S := \{x \in \mathbb{R}^n \mid g_i(x) \leq 0\}$. Let $\bar{x} \in S$. Then, (P2) and (P4) imply that $\phi = \max_{1 \leq i \leq m} [g_i]_+$ while $[g_i]_+ := \max\{g_i(\cdot), 0\}$ and $\psi = \text{dist}(\cdot, S)$ are semi-algebraic functions. Applying (P5) it follows that there exist $c, \varepsilon > 0$ and $\tau \in (0, 1]$ such that

$$(2.2) \quad \text{dist}(x, S) \leq c \max_{1 \leq i \leq m} [g_i(x)]_+^\tau \quad \text{for all } x \in \mathbb{B}(\bar{x}, \varepsilon).$$

\diamond

As explained in Remark 2.5, the exponent τ in (2.2) is hard to determine and is typically unknown. However, there are some special cases where we can provide some effective estimates on the exponent τ : To formulate these results, we introduce the following notation. Define

$$(2.3) \quad \kappa(n, d) := (d - 1)^n + 1.$$

We now present various results which show that the exponent τ in (2.2) can be effectively estimated when g_i has some appropriate extra structure.

Fact 2.7 (Gwoździewicz) (See [22, Theorem 3].) *Let g be a polynomial on \mathbb{R}^n with degree no larger than d . Suppose that $g(0) = 0$ and there exists $\varepsilon_0 > 0$ such that $g(x) > 0$ for all $x \in \mathbb{B}(0, \varepsilon_0) \setminus \{0\}$. Then there exist constants $c, \varepsilon > 0$ such that*

$$(2.4) \quad \|x\| \leq c g(x)^{\frac{1}{\kappa(n, d)}}, \quad \forall x \in \mathbb{B}(0, \varepsilon).$$

We denote by $\beta(s)$ the *central binomial coefficient* with respect to an integer s : $\binom{s}{\lfloor s/2 \rfloor}$ (with $\binom{0}{0} = 1$) [26].

Fact 2.8 (Kollár) (See [26, Theorem 3(i)].) *Let g_i be polynomials on \mathbb{R}^n with degree $\leq d$ for every $i = 1, \dots, m$. Let $g(x) := \max_{1 \leq i \leq m} g_i(x)$. Suppose that there exists $\varepsilon_0 > 0$ such that $g(x) > 0$ for all $x \in \mathbb{B}(0, \varepsilon_0) \setminus \{0\}$. Then there exist constants $c, \varepsilon > 0$ such that*

$$\|x\| \leq c g(x)^{\frac{1}{\beta(n-1)d^n}}, \quad \forall x \in \mathbb{B}(0, \varepsilon).$$

Fact 2.9 (See [29, Theorem 4.2].) *Let g be a convex polynomial on \mathbb{R}^n with degree at most d . Let $S := \{x \mid g(x) \leq 0\}$ and $\bar{x} \in S$. Then, g has a Hölder type local error bound with exponent $\kappa(n, d)^{-1}$, i.e., there exist constants $c, \varepsilon > 0$ such that*

$$\text{dist}(x, S) \leq c [g(x)]_+^{\frac{1}{\kappa(n, d)}} \quad \text{for all } x \in \mathbb{B}(\bar{x}, \varepsilon).$$

Fact 2.10 (See [29, Theorem 4.1].) *Let g be a convex polynomial on \mathbb{R}^n . Let $S := \{x \mid g(x) \leq 0\}$. Suppose that there exists $x_0 \in \mathbb{R}^n$ such that $g(x_0) < 0$. Then, g has a Lipschitz type global error bound, i.e., there exists a constant $c > 0$ such that*

$$\text{dist}(x, S) \leq c [g(x)]_+ \quad \text{for all } x \in \mathbb{R}^n.$$

Fact 2.11 (See [38, Theorem 2] Let g be a continuous convex function on \mathbb{R}^n . Let $S := \{x \in \mathbb{R}^n \mid g(x) \leq 0\}$. Suppose that there exists $x_0 \in \mathbb{R}^n$ such that $g(x_0) < 0$. Then, for every compact subset K of \mathbb{R}^n , there exists $c > 0$ such that

$$\text{dist}(x, S) \leq c [g(x)]_+ \quad \text{for all } x \in K.$$

The following example show us that the conclusion of Fact 2.11 can fail if we allow K to be noncompact.

Example 2.12 (Shironin) (See [29, Example 4.1] or [40].) Let $g_1, g_2 : \mathbb{R}^4 \rightarrow \mathbb{R}$ be defined by

$$\begin{aligned} g_1(x_1, x_2, x_3, x_4) &:= x_1, \\ g_2(x_1, x_2, x_3, x_4) &:= x_1^{16} + x_2^8 + x_3^6 + x_1 x_2^3 x_3^3 + x_1^2 x_2^4 x_3^2 + x_2^2 x_3^4 + x_1^4 x_3^4 \\ &\quad + x_1^4 x_2^6 + x_1^2 x_2^6 + x_1^2 + x_2^2 + x_3^2 - x_4. \end{aligned}$$

Then g_1, g_2 are convex polynomials and

$$g_1(-k, 0, 0, k^{16} + k^2 + k) = g_2(-k, 0, 0, k^{16} + k^2 + k) = -k < 0, \quad \forall k \in \mathbb{N}.$$

Let $g(x) := \max\{g_1(x), g_2(x)\}$ for all $x \in \mathbb{R}^4$. Then, g is a continuous convex function and $g(-k, 0, 0, k^{16} + k^2 + k) < 0$ for all $k \in \mathbb{N}$. On the other hand, as shown in [40], there exists a sequence $(x_k)_{k \in \mathbb{N}}$ in \mathbb{R}^4 such that

$$[g(x_k)]_+ \leq 1, \quad \forall k \quad \text{but} \quad \text{dist}(x_k, S) \longrightarrow +\infty,$$

where $S := \{x \in \mathbb{R}^4 \mid g_1(x) \leq 0, g_2(x) \leq 0\}$. Let K be a noncompact set such that $\{x_k\} \subseteq K$. Then, the conclusion of Fact 2.11 “ $\text{dist}(x, S) \leq c [g(x)]_+$ for all $x \in K$ ” must fail in this case. \diamond

Recall that a set $C \subseteq \mathbb{R}^n$ is a *basic semi-algebraic convex* set if there exist $\gamma \in \mathbb{N}$ and convex polynomial functions, $g_j, j = 1, \dots, \gamma$ such that $C = \{x \in \mathbb{R}^n \mid g_j(x) \leq 0, j = 1, \dots, \gamma\}$. Clearly, any basic semi-algebraic convex set is convex and semi-algebraic. However, the following example shows that a convex and semi-algebraic set need not to be a basic semi-algebraic convex set.

Example 2.13 Consider the set $A := \{(x_1, x_2) \in \mathbb{R}^2 \mid 1 - x_1x_2 \leq 0, -x_1 \leq 0, -x_2 \leq 0\}$. Clearly, A is convex and semi-algebraic while the polynomial $(x_1, x_2) \mapsto 1 - x_1x_2$ is not convex. We now show that A is not a basic semi-algebraic convex set, i.e., it cannot be written as $\{x : g_i(x) \leq 0, i = 1, \dots, l\}$ for some convex polynomials $g_i, i = 1, \dots, l, l \in \mathbb{N}$. To see this, we proceed by the method of contradiction. Let $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ be defined by $f(x_1, x_2) := x_1$. Clearly $\inf_{x=(x_1, x_2) \in A} f(x) = 0$. Then, then by Fact 2.2, f should attain its minimum on A . This leads to a contradiction, and so, justifies the claim. \diamond

2.3 Notation and facts for projection methods

From now on, we assume that

$$\begin{aligned}
& m \in \mathbb{N}, \gamma_i \in \mathbb{N}, i = 1, \dots, m \\
& g_{i,1}, g_{i,2}, \dots, g_{i,\gamma_i} \text{ are convex polynomials on } \mathbb{R}^n, i = 1, \dots, m \\
& C_i := \left\{ x \in \mathbb{R}^n \mid g_{i,1}(x) \leq 0, g_{i,2}(x) \leq 0, \dots, g_{i,\gamma_i}(x) \leq 0 \right\}, i = 1, \dots, m \\
& P_i := P_{C_i}, \quad \forall i = 1, 2, \dots, m \\
& C := \bigcap_{i=1}^m C_i \neq \emptyset.
\end{aligned}$$

Fact 2.14 (Bregman) (See [19].) *Let $x_0 \in \mathbb{R}^n$. The sequence of cyclic projections, $(x_k)_{k \in \mathbb{N}}$, defined by*

$$(2.5) \quad x_1 := P_1 x_0, x_2 := P_2 x_1, \dots, x_m := P_m x_{m-1}, x_{m+1} := P_1 x_m \dots$$

converges to a point in C .

Fact 2.15 (Bauschke and Borwein) (See [9, Lemma 2.2 and Theorem 4.8], or [8, Fact 1.1(iii) and Fact 1.2(ii)].) *Let A, B be nonempty convex subsets of \mathbb{R}^n such that $A - B$ is closed. Let $b_0 \in X$ and $(a_k)_{k \in \mathbb{N}}, (b_k)_{k \in \mathbb{N}}$ be defined as below:*

$$\begin{cases} a_{k+1} &:= P_A b_k \\ b_{k+1} &:= P_B a_{k+1}. \end{cases}$$

Let $v := P_{A-B} 0$. Then, we have

- (i) $\|v\| = \text{dist}(A, B)$ and $a_k \longrightarrow a, \quad b_k \longrightarrow a + v$.
- (ii) $P_B x = P_{B \cap (A+v)} x = x + v, \forall x \in A \cap (B - v)$ and $P_A y = P_{A \cap (B-v)} y = y - v, \forall y \in B \cap (A + v)$

Definition 2.16 Let A be a nonempty convex subset of \mathbb{R}^n . We say the sequence $(x_k)_{k \in \mathbb{N}}$ in \mathbb{R}^n is Fejér monotone with respect to A if

$$\|x_{k+1} - a\| \leq \|x_k - a\|, \quad \forall k \in \mathbb{N}, a \in A.$$

Fact 2.17 (Bauschke and Borwein) (See [8, Theorem 3.3(iv)].) Let A be a nonempty closed convex subset of \mathbb{R}^n and let $(x_k)_{k \in \mathbb{N}}$ be Fejér monotone with respect to A , and $x_k \rightarrow x \in A$. Then $\|x_n - x\| \leq 2 \operatorname{dist}(x_n, A)$.

3 Hölderian regularity for basic semi-algebraic convex sets

In this section, we will establish Hölderian regularity for basic semi-algebraic convex sets and shall provide an effective estimate of the exponent in the regularity results. This result plays an important role in our following estimation of the convergence speed of the cyclic projection methods.

To do this, we first establish an error bound result which estimates the distance of a point to a basic semi-algebraic convex set S in terms of the polynomials which define S . More explicitly, we obtain an explicit exponent $\tau > 0$ such that there exist $c, \varepsilon > 0$,

$$\operatorname{dist}(x, S) \leq c \left(\max_{1 \leq i \leq m} [g_i(x)]_+ \right)^\tau \quad \text{whenever} \quad \|x - \bar{x}\| \leq \varepsilon,$$

where $S := \{x \in \mathbb{R}^n \mid g_i(x) \leq 0, i = 1, \dots, m\}$.

We note that this error bound property plays an important role in convergence analysis of many algorithms for optimization problems [36] as well as for the variational inequality problem [30], and that the exponent τ in the error bound property has a close relationship with the convergence rate of the algorithm. However, existing results such as the powerful Łojasiewicz's inequality do not provide any insight on how to explicitly estimate the exponent τ .

Before we proceed, let us use a simple example to illustrate that the exponent τ can be related to the maximum degree of the polynomials defined the basic semi-algebraic convex set and the dimension of the underlying space. This example is partially inspired by [26, Example 1].

Example 3.1 Let d be an even number. Consider convex polynomials $g_i, i = 1, \dots, n$ on \mathbb{R}^n given by $g_1(x) := x_1^d$ and $g_i, i = 2, \dots, n$ given by $g_i(x) := x_i^d - x_{i-1}, i = 2, \dots, n$. Then, direct verification gives us that $S := \{x \in \mathbb{R}^n \mid g_i(x) \leq 0, i = 1, \dots, n\} = \{0\}$, and so, $\operatorname{dist}(x, S) = \|x\|$. In this case, consider $x(t) = (t^{d^{n-1}}, t^{d^{n-2}}, \dots, t) \in \mathbb{R}^n, t \in (0, 1)$. Then $\operatorname{dist}(x(t), S) = O(t)$ and $\max_{1 \leq i \leq m} [g_i(x(t))]_+ = t^{d^n}$. Therefore, we see that if there exist $c, \varepsilon > 0$ and $\tau > 0$ such that

$$\operatorname{dist}(x, S) \leq c \left(\max_{1 \leq i \leq m} [g_i(x)]_+ \right)^\tau \quad \text{whenever} \quad \|x\| \leq \varepsilon,$$

then, $\tau \leq \frac{1}{d^n}$. Thus, we see that the exponent τ is related to the maximum degree of the polynomials defined the basic semi-algebraic convex set and the dimension of the underlying space. \diamond

We now introduce a decomposition of the index set.

Definition 3.2 For convex polynomials g_1, \dots, g_m on \mathbb{R}^n with $S = \{x \in \mathbb{R}^n \mid g_i(x) \leq 0, i = 1, \dots, m\}$, the index set $\{1, \dots, m\}$ can be decomposed as $J_0 \cup J_1$ with $J_0 \cap J_1 = \emptyset$ where

$$(3.1) \quad J_0 := \{i \in \{1, \dots, m\} \mid g_i(S) \equiv 0\} \quad \text{and} \quad J_1 := \{1, \dots, m\} \setminus J_0.$$

Now we come to our key technical result which provides a local error bound for convex polynomial systems. The main idea of the proof is to use the extreme rigidity of convex polynomial functions (see Fact 2.1 and 2.3) to reduce the problem to the one of comparing a (power of a) nonnegative convex polynomial vanishing only at 0 with the norm so that the estimate of the Lojasiewicz inequality that we introduced in Fact 2.7 and Fact 2.8 can be applied.

Theorem 3.3 (Local error bounds for convex polynomial systems) Let g_i be convex polynomials on \mathbb{R}^n with degree at most d for every $i = 1, \dots, m$. Let $S := \{x \in \mathbb{R}^n \mid g_i(x) \leq 0, i = 1, \dots, m\}$ and $\bar{x} \in S$. Then there exist $c, \varepsilon > 0$ such that

$$\text{dist}(x, S) \leq c \left(\max_{i \in J_1} [g_i(x)]_+ + \left(\max_{i \in J_0} [g_i(x)]_+ \right)^\tau \right) \quad \text{whenever} \quad \|x - \bar{x}\| \leq \varepsilon,$$

where $[a]_+ := \max\{a, 0\}$, $\tau := \max\left\{\frac{2}{\kappa(n, 2d)}, \frac{1}{\beta(n-1)d^n}\right\}$, $\kappa(n, 2d) := (2d-1)^n + 1$, $\beta(n-1)$ is the central binomial coefficient with respect to $n-1$ which is given by $\binom{n-1}{\lfloor (n-1)/2 \rfloor}$, and J_0, J_1 are defined as in (3.1).

Proof. We prove the desired conclusion by induction on the number of the polynomials m .

[**Trivial Case**] Suppose that $m = 1$. Then $J_0 = \{1\}$ or $J_0 = \emptyset$. If $J_0 = \{1\}$, then the conclusion follows by Fact 2.9 since $\max\left\{\frac{2}{\kappa(n, 2d)}, \frac{1}{\beta(n-1)d^n}\right\} \leq \frac{1}{\kappa(n, d)}$. If $J_0 = \emptyset$, then there exists x_0 such that $g_1(x_0) < 0$. In this case, the conclusion follows by Fact 2.10.

[**Reduction to the active cases**] Let us suppose that the conclusion is true for $m \leq p-1$, $p \in \mathbb{N}$, and look at the case for $m = p$. If $J_0 \neq \{1, \dots, m\}$, then $\{1, \dots, m\} \setminus J_0 \neq \emptyset$. Let $i_0 \notin J_0$. Then there exists $x_0 \in S$ such that $g_{i_0}(x_0) < 0$. Set $J := \{i \in \{1, 2, \dots, m\} \mid g_i(x_0) < 0\}$. Then $i_0 \in J \subseteq J_1$ and $J \cap J_0 = \emptyset$. Let A, B be defined by

$$\begin{aligned} A &:= \{x \in \mathbb{R}^n \mid g_i(x) \leq 0, \quad \forall i \in J\} \\ B &:= \{x \in \mathbb{R}^n \mid g_i(x) \leq 0, \quad \forall i \in \tilde{J}\}, \end{aligned}$$

where $\tilde{J} := \{1, 2, \dots, m\} \setminus J$. Thus, $J_0 \subseteq \tilde{J}$. Then we have $x_0 \in \text{int } A \cap B$ and $S = A \cap B$. Since $S \subseteq B$, we have

$$(3.2) \quad \tilde{J}_0 := \{i \in \{1, 2, \dots, m\} \setminus J \mid g_i(B) \equiv 0\} \subseteq J_0.$$

Since $\text{int } A \cap B \neq \emptyset$, [8, Corollary 4.5] implies that for every compact set K there exist $\gamma, \delta > 0$ such that

$$(3.3) \quad \text{dist}(x, S) = \text{dist}(x, A \cap B) \leq \gamma \max\{\text{dist}(x, A), \text{dist}(x, B)\} \quad \text{for all } x \in K.$$

Now applying Fact 2.11 with $f(x) = \max_{i \in J} [g_i(x)]_+$, there exists $c_1 > 0$ such that

$$(3.4) \quad \text{dist}(x, A) \leq c_1 \max_{i \in J} [g_i(x)]_+, \quad \forall x \in K.$$

From the induction hypothesis and (3.2), we see that there exist $\varepsilon > 0$ and $c_2 > 0$ such that for every $\|x - \bar{x}\| \leq \varepsilon$, $\max_{1 \leq i \leq m} [g_i(x)]_+ \leq 1$ such that

$$\begin{aligned} & \text{dist}(x, B) \\ & \leq c_2 \left(\max_{i \in \tilde{J} \setminus \tilde{J}_0} [g_i(x)]_+ + \left(\max_{i \in \tilde{J}_0} [g_i(x)]_+ \right)^\tau \right) \\ & = c_2 \left(\max_{i \in \tilde{J} \setminus J_0} \{ \max_{i \in \tilde{J} \setminus J_0} [g_i(x)]_+, \max_{i \in J_0 \setminus \tilde{J}_0} [g_i(x)]_+ \} + \left(\max_{i \in \tilde{J}_0} [g_i(x)]_+ \right)^\tau \right) \quad (\text{since } \tilde{J}_0 \subseteq J_0 \subseteq \tilde{J}) \\ & \leq c_2 \left(\max_{i \in \tilde{J} \setminus J_0} [g_i(x)]_+ + \max_{i \in J_0 \setminus \tilde{J}_0} [g_i(x)]_+ + \left(\max_{i \in \tilde{J}_0} [g_i(x)]_+ \right)^\tau \right) \\ & \leq c_2 \left(\max_{i \in \tilde{J} \setminus J_0} [g_i(x)]_+ + \left(\max_{i \in J_0 \setminus \tilde{J}_0} [g_i(x)]_+ \right)^\tau + \left(\max_{i \in \tilde{J}_0} [g_i(x)]_+ \right)^\tau \right) \\ & \leq c_2 \left(\max_{i \in \tilde{J} \setminus J_0} [g_i(x)]_+ + 2 \left(\max_{i \in J_0} [g_i(x)]_+ \right)^\tau \right) \\ & \leq 2c_2 \left(\max_{i \in \tilde{J} \setminus J_0} [g_i(x)]_+ + \left(\max_{i \in J_0} [g_i(x)]_+ \right)^\tau \right). \end{aligned}$$

Thus, the conclusion follows in this case by combining (3.3) and (3.4), and noting that $(\tilde{J} \setminus J_0) \cup J \subseteq J_1$.

From now on, we may assume that $J_0 = \{1, \dots, m\}$. That is,

$$\{x \mid g_i(x) \leq 0, i = 1, \dots, m\} = \{x \mid g_i(x) = 0, i = 1, \dots, m\}.$$

This implies that $\inf_{x \in \mathbb{R}^n} \max_{1 \leq i \leq m} \{g_i(x)\} = 0$. Then,

$$0_{\mathbb{R}^m} \notin \{(g_1(x), \dots, g_m(x)) \mid x \in \mathbb{R}^n\} + \text{int } \mathbb{R}_+^m.$$

Hence the convex separation theorem ensures that there exist $\alpha_i \geq 0$ with $\sum_{i=1}^m \alpha_i = 1$ such that $\sum_{i=1}^m \alpha_i g_i(x) \geq 0$ for all $x \in \mathbb{R}^n$. Denote $I := \{i \mid \alpha_i > 0\} \neq \emptyset$. Then, we have $\sum_{i \in I} \alpha_i = 1$ and

$$(3.5) \quad \sum_{i \in I} \alpha_i g_i(x) \geq 0 \quad \text{for all } x \in \mathbb{R}^n.$$

[**Decompose the underlying space into sum of two subspaces M and M^\perp**] Consider $D := \{x \in \mathbb{R}^n \mid g_i(x) \leq 0, i \in I\}$. Clearly, D is a convex set and $\bar{x} \in D$. Moreover, for any $v \in D$, (3.5) implies that

$$(3.6) \quad g_i(v) = 0, \quad \forall i \in I.$$

In other words, g_i takes constant value 0 on D . Then, Fact 2.1 implies that D is either a singleton or an affine set with dimension larger than one.

Let $M := D - \bar{x}$. Then M is a subspace. We may decompose $\mathbb{R}^n = M + M^\perp$. Denote $\dim M = k$ ($k \leq n$).

We now see that

$$(3.7) \quad \sum_{i \in I} g_i^2(x) > 0 \text{ for all } x - \bar{x} \in M^\perp \setminus \{0\}.$$

Otherwise, there exists $x_0 \in \mathbb{R}^n$ such that $x_0 - \bar{x} \in M^\perp \setminus \{0\}$ and $g_i(x_0) = 0$ for all $i \in I$. This shows that $x_0 \in D$. Thus $x_0 - \bar{x} \in M$ and hence $x_0 - \bar{x} \in M \cap M^\perp$. This contradicts the fact that $x_0 - \bar{x} \neq 0$.

Similarly, we have

$$(3.8) \quad \max_{i \in I} g_i(x) > 0 \text{ for all } x - \bar{x} \in M^\perp \setminus \{0\}.$$

[**Distance estimation on M^\perp**] We first show that there exist $\varepsilon_0, \gamma_0 > 0$ such that

$$(3.9) \quad \|x - \bar{x}\| \leq \gamma_0 \left(\sum_{i \in I} g_i^2(x) \right)^{\frac{1}{\kappa(n-k, 2d)}}, \quad \forall x - \bar{x} \in M^\perp \cap \mathbb{B}(0, \varepsilon_0).$$

Since $\dim M^\perp = n - k$, there exists an $n \times (n - k)$ matrix Q_0 with the rank $n - k$ such that $Q_0(\mathbb{R}^{n-k}) = M^\perp$. Then Q_0 is a bijective operator from \mathbb{R}^{n-k} to M^\perp . Then (3.7) shows that

$$(3.10) \quad \sum_{i \in I} g_i^2(\bar{x} + Q_0 b) > 0, \quad \forall b \in \mathbb{R}^{n-k} \setminus \{0\}.$$

Define $h : \mathbb{R}^{n-k} \rightarrow \mathbb{R}$ by $h(b) := \sum_{i \in I} g_i^2(\bar{x} + Q_0 b)$. Then $h(0) = \sum_{i \in I} g_i^2(\bar{x}) = 0$ by (3.6). Thus by (3.10) and Fact 2.7, there exist $\varepsilon_1, \gamma_1 > 0$ such that for all

$$\|b\| \leq \gamma_1 h(b)^{\frac{1}{\kappa(n-k, 2d)}} = \gamma_1 \left(\sum_{i \in I} g_i^2(\bar{x} + Q_0 b) \right)^{\frac{1}{\kappa(n-k, 2d)}}, \quad \forall \|b\| \leq \varepsilon_1.$$

Setting $x := \bar{x} + Q_0 b$, it follows that

$$\begin{aligned} \|x - \bar{x}\| &= \|Q_0(Q_0^{-1}(x - \bar{x}))\| \leq \|Q_0\| \cdot \|Q_0^{-1}(x - \bar{x})\| \\ &\leq \|Q_0\| \gamma_1 \left(\sum_{i \in I} g_i^2(x) \right)^{\frac{1}{\kappa(n-k, 2d)}}, \quad \forall \|Q_0^{-1}(x - \bar{x})\| \leq \varepsilon_1. \end{aligned}$$

Hence, there exist $\varepsilon_1, \gamma_1 > 0$ such that

$$\|x - \bar{x}\| \leq \gamma_0 \left(\sum_{i \in I} g_i^2(x) \right)^{\frac{1}{\kappa(n-k, 2d)}}, \quad \forall x - \bar{x} \in M^\perp \cap \mathbb{B}(0, \varepsilon_0).$$

Thence (3.9) holds.

By Fact 2.8 and (3.8), there exist $\tilde{\varepsilon}_0, \tilde{\gamma}_0 > 0$ such that $\tilde{\varepsilon}_0 \leq \varepsilon_0$ and, for all x with $x - \bar{x} \in M^\perp \cap \mathbb{B}(0, \tilde{\varepsilon}_0)$,

$$\max_{i \in I} g_i(x) \leq 1$$

and

$$\begin{aligned} \|x - \bar{x}\| &\leq \tilde{\gamma}_0 \left(\max_{i \in I} g_i(x) \right)^{\frac{1}{\beta(n-k-1)d^{n-k}}} \\ (3.11) \quad &= \tilde{\gamma}_0 \left(\max_{i \in I} [g_i(x)]_+ \right)^{\frac{1}{\beta(n-k-1)d^{n-k}}}. \end{aligned}$$

[Distance estimation on M] Set $r := \begin{cases} \max\{\frac{\sum_{j \in I \setminus \{i\}} \alpha_j}{\alpha_i} \mid i \in I\} > 0, & \text{if } |I| \geq 2; \\ 1, & \text{otherwise.} \end{cases}$

Thus $r \geq 1$. Note that $\sum_{i \in I} \alpha_i g_i(x) \geq 0$, we have for each $i \in I$

$$\max_{i \in I} [g_i(x)]_+ \geq g_i(x) \geq -\frac{\sum_{j \in I \setminus \{i\}} \alpha_j g_j(x)}{\alpha_i} \geq -r \max_{i \in I} [g_i(x)]_+.$$

Hence we have $|g_i(x)| \leq r \max_{i \in I} [g_i(x)]_+$. This together with (3.9) implies that

$$\|x - \bar{x}\| \leq \gamma_0 r^2 |I| \left(\max_{i \in I} [g_i(x)]_+ \right)^{\frac{2}{\kappa(n-k, 2d)}}, \quad \forall x - \bar{x} \in M^\perp \cap \mathbb{B}(0, \varepsilon_0).$$

Combining this with (3.11), we see that, for every $x - \bar{x} \in M^\perp \cap \mathbb{B}(0, \tilde{\varepsilon}_0)$,

$$(3.12) \quad \|x - \bar{x}\| \leq (\gamma_0 r^2 |I| + \tilde{\gamma}_0) \left(\max_{i \in I} [g_i(x)]_+ \right)^{\max\{\frac{2}{\kappa(n-k, 2d)}, \frac{1}{\beta(n-k-1)d^{n-k}}\}}.$$

We now consider two cases.

Case 1: $\dim M = \{0\}$.

We have $D = S = \{\bar{x}\}$. Thus $M = 0$ and $M^\perp = \mathbb{R}^n$. We can assume that $\max_{i \in I} [g_i(x)]_+ \leq 1$ for all $x - \bar{x} \in \mathbb{B}(0, \tilde{\varepsilon}_0)$. Then by (3.12), we have

$$\text{dist}(x, S) = \|x - \bar{x}\| \leq (\gamma_0 r^2 |I| + \tilde{\gamma}_0) \left(\max_{i \in I} [g_i(x)]_+ \right)^{\max\{\frac{2}{\kappa(n, 2d)}, \frac{1}{\beta(n-1)d^n}\}}, \quad \forall \|x - \bar{x}\| \leq \tilde{\varepsilon}_0.$$

Case 2: $k = \dim M \geq 1$.

Since $\dim M = k$, there exists a full rank matrix $Q \in \mathbb{R}^{n \times k}$ such that $Q(\mathbb{R}^k) = M$. For each $u \in M$ and $i \in I$, (3.6) implies that

$$g_i^\infty(u) = \lim_{t \rightarrow \infty} \frac{g_i(\bar{x} + tu) - g_i(\bar{x})}{t} = 0.$$

Then, Fact 2.3 implies that

$$(3.13) \quad g_i(x + u) = g_i(x) \quad \text{for all } x \in \mathbb{R}^n, u \in M, i \in I.$$

Since $S \subseteq D = \bar{x} + M$, it follows that

$$S = \{x \in \bar{x} + M \in \mathbb{R}^n \mid g_i(x) \leq 0, i \notin I\} = \bar{x} + Q(\widehat{S}),$$

where $\widehat{S} := \{a \in \mathbb{R}^k \mid g_i(\bar{x} + Qa) \leq 0, i \notin I\}$.

Note that $0 \in \widehat{S}$. The induction hypothesis implies that there exist $\tilde{\varepsilon}_1, \tilde{\gamma}_1 > 0$ such that $\max_{i \notin I} [g_i(\bar{x} + Qa)]_+ \leq 1$ and

$$\text{dist}(a, \widehat{S}) \leq \tilde{\gamma}_1 \left(\max_{i \notin I} [g_i(\bar{x} + Qa)]_+ \right)^{\max \left\{ \frac{2}{\kappa(k, 2d)}, \frac{1}{\beta(k-1)d^k} \right\}} \quad \text{for all } \|a\| \leq \tilde{\varepsilon}_1.$$

This implies that there exist $\varepsilon_2, \gamma_2 > 0$ such that

$$(3.14) \quad \text{dist}(x, S) \leq \gamma_2 \left(\max_{i \notin I} [g_i(x)]_+ \right)^{\max \left\{ \frac{2}{\kappa(k, 2d)}, \frac{1}{\beta(k-1)d^k} \right\}} \quad \text{for all } x - \bar{x} \in M \cap \mathbb{B}(0, \varepsilon_2).$$

[Combining the estimation and simplification] Now let $\varepsilon \leq \min\{\tilde{\varepsilon}_0, \varepsilon_2\}$ be such that $\max_{1 \leq i \leq m} [g_i(x)]_+ \leq 1$ for all $x \in \mathbb{B}(\bar{x}, \varepsilon)$. Let K be a compact set containing $\mathbb{B}(\bar{x}, \varepsilon) \cup \mathbb{B}(0, \varepsilon)$. Denote the Lipschitz constant of g_i over K by L_i , i.e., $|g_i(x_1) - g_i(x_2)| \leq L_i \|x_1 - x_2\|$ for all $x_1, x_2 \in K$. Set $L := \max_{1 \leq i \leq m} L_i$ and $\gamma := \max\{\gamma_0 r^2 |I| + \tilde{\gamma}_0, \gamma_2\}$.

To arrive at the conclusion, we only need to show that for any $x \in \mathbb{B}(\bar{x}, \varepsilon)$,

$$\text{dist}(x, S) \leq c \left(\max_{1 \leq i \leq m} [g_i(x)]_+ \right)^{\max \left\{ \frac{2}{\kappa(n, 2d)}, \frac{1}{\beta(n-1)d^n} \right\}},$$

where $c := 2\gamma + L\gamma^2$. To see this, let us fix an arbitrary $x \in \mathbb{B}(\bar{x}, \varepsilon)$. Note that $\mathbb{R}^n = M + M^\perp$. Then, one can decompose $x - \bar{x} = u + v$ for some $u \in M \cap \mathbb{B}(0, \varepsilon)$ and $v \in M^\perp \cap \mathbb{B}(0, \varepsilon)$. This together with (3.14) and (3.12) implies that

$$(3.15) \quad \text{dist}(u + \bar{x}, S) \leq \gamma \left(\max_{i \notin I} [g_i(u + \bar{x})]_+ \right)^{\max \left\{ \frac{2}{\kappa(k, 2d)}, \frac{1}{\beta(k-1)d^k} \right\}}$$

$$(3.16) \quad \|v\| \leq \gamma \left(\max_{i \in I} [g_i(v + \bar{x})]_+ \right)^{\max \left\{ \frac{2}{\kappa(n-k, 2d)}, \frac{1}{\beta(n-k-1)d^{n-k}} \right\}}.$$

Therefore,

$$\text{dist}(x, S) \leq \text{dist}(u + \bar{x}, S) + \|x - (u + \bar{x})\|$$

$$\begin{aligned}
&= \text{dist}(u + \bar{x}, S) + \|v\| \\
&\leq \gamma \left(\max_{i \notin I} [g_i(u + \bar{x})]_+ \right)^{\max \left\{ \frac{2}{\kappa(k, 2d)}, \frac{1}{\beta(k-1)d^k} \right\}} + \gamma \left(\max_{i \in I} [g_i(v + \bar{x})]_+ \right)^{\max \left\{ \frac{2}{\kappa(n-k, 2d)}, \frac{1}{\beta(n-k-1)d^{n-k}} \right\}} \\
&= \gamma \left(\max_{i \notin I} [g_i(u + \bar{x})]_+ \right)^{\max \left\{ \frac{2}{\kappa(k, 2d)}, \frac{1}{\beta(k-1)d^k} \right\}} + \gamma \left(\max_{i \in I} [g_i(x)]_+ \right)^{\max \left\{ \frac{2}{\kappa(n, 2d)}, \frac{1}{\beta(n-1)d^n} \right\}} \\
&\leq \gamma \left(\max_{1 \leq i \leq m} [g_i(u + \bar{x})]_+ \right)^{\max \left\{ \frac{2}{\kappa(k, 2d)}, \frac{1}{\beta(k-1)d^k} \right\}} + \gamma \left(\max_{1 \leq i \leq m} [g_i(x)]_+ \right)^{\max \left\{ \frac{2}{\kappa(n, 2d)}, \frac{1}{\beta(n-1)d^n} \right\}},
\end{aligned}$$

where the second equality follows by (3.13) and $v + \bar{x} + u = x$. Note that

$$\begin{aligned}
&\left| \max_{1 \leq i \leq m} [g_i(u + \bar{x})]_+ - \max_{1 \leq i \leq m} [g_i(x)]_+ \right| \leq \max_{1 \leq i \leq m} |g_i(u + \bar{x}) - g_i(x)| \\
&\leq L \|u + \bar{x} - x\| = L \|v\| \\
&\leq L \gamma \left(\max_{i \in I} [g_i(v + \bar{x})]_+ \right)^{\max \left\{ \frac{2}{\kappa(n-k, 2d)}, \frac{1}{\beta(n-k-1)d^{n-k}} \right\}} \quad (\text{by (3.16)}) \\
(3.17) \quad &= L \gamma \left(\max_{i \in I} [g_i(x)]_+ \right)^{\max \left\{ \frac{2}{\kappa(n-k, 2d)}, \frac{1}{\beta(n-k-1)d^{n-k}} \right\}} \quad (\text{by 3.13 and } v + \bar{x} + u = x).
\end{aligned}$$

As $\max_{1 \leq i \leq m} [g_i(x)]_+ \leq 1$ for all $x \in \mathbb{B}(\bar{x}, \varepsilon)$, it follows that

$$\begin{aligned}
\text{dist}(x, S) &\leq \gamma \left(\max_{1 \leq i \leq m} [g_i(u + \bar{x})]_+ \right)^{\frac{2}{\kappa(k, 2d)}} + \gamma \left(\max_{1 \leq i \leq m} [g_i(x)]_+ \right)^{\max \left\{ \frac{2}{\kappa(n, 2d)}, \frac{1}{\beta(n-1)d^n} \right\}} \\
&\leq \gamma \left(\max_{1 \leq i \leq m} [g_i(x)]_+ + L \gamma \left(\max_{i \in I} [g_i(x)]_+ \right)^{\frac{2}{\kappa(n-k, 2d)}} \right)^{\frac{2}{\kappa(k, 2d)}} \quad (\text{by (3.17)}) \\
&\quad + \gamma \left(\max_{1 \leq i \leq m} [g_i(x)]_+ \right)^{\max \left\{ \frac{2}{\kappa(n, 2d)}, \frac{1}{\beta(n-1)d^n} \right\}} \\
&\leq \gamma \left(\max_{1 \leq i \leq m} [g_i(x)]_+^{\frac{2}{\kappa(n-k, 2d)}} + L \gamma \left(\max_{i \in I} [g_i(x)]_+^{\frac{2}{\kappa(n-k, 2d)}} \right) \right)^{\frac{2}{\kappa(k, 2d)}} \\
&\quad + \gamma \left(\max_{1 \leq i \leq m} [g_i(x)]_+ \right)^{\max \left\{ \frac{2}{\kappa(n, 2d)}, \frac{1}{\beta(n-1)d^n} \right\}} \\
&\leq (\gamma + L \gamma^2) \max_{1 \leq i \leq m} [g_i(x)]_+^{\frac{4}{\kappa(n-k, 2d) \cdot \kappa(k, 2d)}} + \gamma \left(\max_{1 \leq i \leq m} [g_i(x)]_+ \right)^{\max \left\{ \frac{2}{\kappa(n, 2d)}, \frac{1}{\beta(n-1)d^n} \right\}} \\
&= (\gamma + L \gamma^2) \max_{1 \leq i \leq m} [g_i(x)]_+^{\frac{4}{((2d-1)^{n-k}+1) \cdot ((2d-1)^k+1)}} + \gamma \left(\max_{1 \leq i \leq m} [g_i(x)]_+ \right)^{\max \left\{ \frac{2}{\kappa(n, 2d)}, \frac{1}{\beta(n-1)d^n} \right\}} \\
(3.18) \quad &\leq (\gamma + L \gamma^2) \max_{1 \leq i \leq m} [g_i(x)]_+^{\frac{2}{(2d-1)^{n+1}}} + \gamma \left(\max_{1 \leq i \leq m} [g_i(x)]_+ \right)^{\max \left\{ \frac{2}{\kappa(n, 2d)}, \frac{1}{\beta(n-1)d^n} \right\}}.
\end{aligned}$$

Similarly, we also have

$$\text{dist}(x, S) \leq \gamma \left(\max_{1 \leq i \leq m} [g_i(u + \bar{x})]_+ \right)^{\frac{1}{\beta(k-1)d^k}} + \gamma \left(\max_{1 \leq i \leq m} [g_i(x)]_+ \right)^{\max \left\{ \frac{2}{\kappa(n, 2d)}, \frac{1}{\beta(n-1)d^n} \right\}}$$

$$\begin{aligned}
&\leq \gamma \left(\max_{1 \leq i \leq m} [g_i(x)]_+ + L\gamma (\max_{i \in I} [g_i(x)]_+^{\frac{1}{\beta(n-k-1)d^{n-k}}}) \right)^{\frac{1}{\beta(k-1)d^k}} \quad (\text{by (3.17)}) \\
&\quad + \gamma \left(\max_{1 \leq i \leq m} [g_i(x)]_+ \right)^{\max \left\{ \frac{2}{\kappa(n,2d)}, \frac{1}{\beta(n-1)d^n} \right\}} \\
&\leq \gamma \left(\max_{1 \leq i \leq m} [g_i(x)]_+^{\frac{1}{\beta(n-k-1)d^{n-k}}} + L\gamma (\max_{i \in I} [g_i(x)]_+^{\frac{1}{\beta(n-k-1)d^{n-k}}}) \right)^{\frac{1}{\beta(k-1)d^k}} \\
&\quad + \gamma \left(\max_{1 \leq i \leq m} [g_i(x)]_+ \right)^{\max \left\{ \frac{2}{\kappa(n,2d)}, \frac{1}{\beta(n-1)d^n} \right\}} \\
&= (\gamma + L\gamma^2) \max_{1 \leq i \leq m} [g_i(x)]_+^{\frac{1}{(\beta(n-k-1)d^{n-k}) \cdot (\beta(k-1)d^k)}} + \gamma \left(\max_{1 \leq i \leq m} [g_i(x)]_+ \right)^{\max \left\{ \frac{2}{\kappa(n,2d)}, \frac{1}{\beta(n-1)d^n} \right\}} \\
&\leq (\gamma + L\gamma^2) \max_{1 \leq i \leq m} [g_i(x)]_+^{\frac{1}{\beta(n-1)d^n}} + \gamma \left(\max_{1 \leq i \leq m} [g_i(x)]_+ \right)^{\max \left\{ \frac{2}{\kappa(n,2d)}, \frac{1}{\beta(n-1)d^n} \right\}},
\end{aligned}$$

where the last inequality was obtained by the Chu-Vandermonde identity.

In combination with (3.18) we obtain

$$\begin{aligned}
\text{dist}(x, S) &\leq (\gamma + L\gamma^2) \max_{1 \leq i \leq m} [g_i(x)]_+^{\max \left\{ \frac{2}{(2d-1)^n+1}, \frac{1}{\beta(n-1)d^n} \right\}} + \gamma \left(\max_{1 \leq i \leq m} [g_i(x)]_+ \right)^{\max \left\{ \frac{2}{\kappa(n,2d)}, \frac{1}{\beta(n-1)d^n} \right\}} \\
&= (2\gamma + L\gamma^2) \max_{1 \leq i \leq m} [g_i(x)]_+^{\max \left\{ \frac{2}{(2d-1)^n+1}, \frac{1}{\beta(n-1)d^n} \right\}}.
\end{aligned}$$

This completes the proof. \square

As a corollary, we obtain a local error bound result which is independent of the partition of the index set.

Corollary 3.4 *Let g_i be convex polynomials on \mathbb{R}^n with degree at most d for every $i = 1, \dots, m$. Let $S := \{x \in \mathbb{R}^n \mid g_i(x) \leq 0, i = 1, \dots, m\}$ and $\bar{x} \in S$. Then there exist $c, \varepsilon > 0$ such that*

$$\text{dist}(x, S) \leq c \left(\max_{1 \leq i \leq m} [g_i(x)]_+ \right)^\tau \quad \text{whenever} \quad \|x - \bar{x}\| \leq \varepsilon,$$

where $[a]_+ := \max\{a, 0\}$, $\tau := \max \left\{ \frac{2}{\kappa(n,2d)}, \frac{1}{\beta(n-1)d^n} \right\} = \frac{1}{\min \left\{ \frac{(2d-1)^n+1}{2}, \beta(n-1)d^n \right\}}$ and $\beta(n-1)$ is the central binomial coefficient with respect to $n-1$.

Proof. Choose ε small enough so that $\max_{1 \leq i \leq m} [g_i(x)]_+ \leq 1$. Then, the conclusion follows immediately from the preceding Theorem 3.3 by noting that $[g_i(x)]_+ \leq (\max_{1 \leq i \leq m} [g_i(x)]_+)^{\frac{2}{\kappa(n,2d)}}$ for each $i = 1, \dots, m$. \square

Remark 3.5 (Discussion of the exponent) Let g_i be convex polynomials on \mathbb{R}^n with degree at most d for every $i = 1, \dots, m$. Let $S := \{x \in \mathbb{R}^n \mid g_i(x) \leq 0, i = 1, \dots, m\}$ and $\bar{x} \in S$. We now make some discussion on the exponent in our local error bound results.

- (1) Theorem 3.3 shows that in the case when $d = 1$ or if there exists $x_0 \in \mathbb{R}^n$ such that $g_i(x_0) < 0$, $i = 1, \dots, m$, we indeed obtain a Lipschitz type local error bound. That is to say, in these cases, there exist $c, \varepsilon > 0$ such that

$$\text{dist}(x, S) \leq c \max_{1 \leq i \leq m} [g_i(x)]_+ \quad \text{whenever} \quad \|x - \bar{x}\| \leq \varepsilon,$$

where $[a]_+ := \max\{a, 0\}$. To see this, if $d = 1$, then $\frac{2}{\kappa(n, 2d)} = 1$ and $\frac{1}{\beta(n-1)d^n} \geq 1$. So, $\tau = \max\left\{\frac{2}{\kappa(n, 2d)}, \frac{1}{\beta(n-1)d^n}\right\} = 1$ and the conclusion follows immediately from Theorem 3.3. On the other hand, if there exists $x_0 \in \mathbb{R}^n$ such that $g_i(x_0) < 0$, $i = 1, \dots, m$, then, $J_0 = \emptyset$. So, the conclusion follows immediately from the same theorem.

- (2) In the case when $n = 1$, we see that $\frac{2}{\kappa(n, 2d)} = \frac{1}{\beta(n-1)d^n} = \frac{1}{d}$. So, when each g_i is a univariate convex polynomial, then there exist $c, \varepsilon > 0$ such that

$$\text{dist}(x, S) \leq c \left(\max_{1 \leq i \leq m} [g_i(x)]_+ \right)^{\frac{1}{d}} \quad \text{whenever} \quad \|x - \bar{x}\| \leq \varepsilon.$$

Note that, for the naive simple example $g_1(x) := x^d$, local error bound holds at 0 with exponent $\frac{1}{d}$. This suggests that our result matches what one might expect in the univariate case.

- (3) On the other hand, in general, our estimation on the exponent will not be optimal.

For example, if the inequality system consists of one single convex polynomial, Fact 2.9 shows that the exponent can be set as $\frac{1}{(d-1)^{n+1}}$ while our results produce a weaker exponent $\max\left\{\frac{2}{(2d-1)^{n+1}}, \frac{1}{\beta(n-1)d^n}\right\}$. An interesting feature of the exponent $\frac{1}{(d-1)^{n+1}}$ in Fact 2.9 is that, in the convex quadratic case, it collapses to $\frac{1}{2}$ which is independent of the dimension of the underlying space and which agrees with the known result presented in [32]. By contrast, our estimate $\max\left\{\frac{2}{3^{n+1}}, \frac{1}{\beta(n-1)2^n}\right\}$ depends heavily on the dimension n .

Moreover, as indicated in Example 3.1, the best possible exponent might be $\frac{1}{d^n}$ (see [26] for some relevant discussion regarding the best possible exponent for general nonconvex polynomial system). It would be interesting to find how to could improve our estimate here.

Making better sense of these estimates will be one of our future research topics. \diamond

Given $D \subseteq \mathbb{R}^n$, we set $\text{dist}^r(\cdot, D) := (\text{dist}(\cdot, D))^r$ for every $r \in \mathbb{R}$.

Theorem 3.6 (Hölderian regularity) *Let $\gamma_i \in \mathbb{N}$, $i = 1, \dots, m$, and $g_{i,j}$ be are convex polynomials on \mathbb{R}^n with degree $d \in \mathbb{N}$, $j = 1, \dots, \gamma_i$, $i = 1, \dots, m$. Recall that*

$$C_i = \left\{x \in \mathbb{R}^n \mid g_{i,j}(x) \leq 0, j = 1, \dots, \gamma_i\right\} \text{ and } C = \bigcap_{i=1}^m C_i.$$

Let $\theta > 0$ and $K \subseteq \mathbb{R}^n$ be a compact set. Then there exists $c > 0$ such that

$$\text{dist}^\theta(x, C) \leq c \left(\sum_{i=1}^m \text{dist}^\theta(x, C_i) \right)^\tau, \quad \forall x \in K,$$

where $\tau := \frac{1}{\min \left\{ \frac{(2d-1)^n+1}{2}, \beta(n-1)d^n \right\}}$ and $\beta(n-1)$ is the central binomial coefficient with respect to $n-1$ which is given by $\binom{n-1}{\lfloor (n-1)/2 \rfloor}$.

Proof. To see the conclusion, we only need to show that for each $\bar{x} \in \mathbb{R}^n$, there exist $c, \varepsilon > 0$ such that

$$(3.19) \quad \text{dist}^\theta(x, C) \leq c \left(\sum_{i=1}^m \text{dist}^\theta(x, C_i) \right)^\tau, \quad \text{for all } \|x - \bar{x}\| \leq \varepsilon.$$

Indeed, granting this and fixing a compact set K , then for any $\bar{x} \in K$ there exist $c_{\bar{x}}, \varepsilon_{\bar{x}} > 0$ such that

$$\text{dist}^\theta(x, C) \leq c_x \left(\sum_{i=1}^m \text{dist}^\theta(x, C_i) \right)^\tau, \quad \text{for all } \|x - \bar{x}\| \leq \varepsilon_x.$$

As K is compact and $\bigcup_{\bar{x} \in K} \mathbb{B}(\bar{x}; \varepsilon_{\bar{x}}) \supseteq K$, we can find finitely many points $\bar{x}_1, \dots, \bar{x}_s \in K$, $s \in \mathbb{N}$, such that $\bigcup_{i=1}^s \mathbb{B}(\bar{x}_i; \varepsilon_{\bar{x}_i}) \supseteq K$. Let $c := \max\{c_{\bar{x}_1}, \dots, c_{\bar{x}_s}\}$. Then, for any $x \in K$, there exists $i_0 \in \{1, \dots, s\}$ such that $x \in \mathbb{B}(\bar{x}_{i_0}; \varepsilon_{\bar{x}_{i_0}})$, and hence

$$\text{dist}^\theta(x, C) \leq c_{\bar{x}_{i_0}} \left(\sum_{i=1}^m \text{dist}^\theta(x, C_i) \right)^\tau \leq c \left(\sum_{i=1}^m \text{dist}^\theta(x, C_i) \right)^\tau.$$

We now show (3.19) holds. Fix $\bar{x} \in \mathbb{R}^n$. We consider two cases.

Case 1: $\bar{x} \notin C$.

Then there exist $\varepsilon_1, \eta, M > 0$ such that

$$\sum_{i=1}^m \text{dist}^\theta(x, C_i) \geq \eta \text{ and } \text{dist}^\theta(x, C) \leq M \quad \text{for all } \|x - \bar{x}\| \leq \varepsilon_1.$$

Therefore, $\text{dist}^\theta(x, C) \leq M = \frac{M}{\eta^\tau} \eta^\tau \leq \frac{M}{\eta^\tau} \left(\sum_{i=1}^m \text{dist}^\theta(x, C_i) \right)^\tau$ for all $\|x - \bar{x}\| \leq \varepsilon_1$ and hence, (3.19) holds.

Case 2: $\bar{x} \in C$.

We have

$$C = \left\{ x \in \mathbb{R}^n \mid g_{i,1}(x) \leq 0, g_{i,2}(x) \leq 0, \dots, g_{i,\gamma_i}(x) \leq 0, \quad i = 1, \dots, m \right\}.$$

By Corollary 3.4, there exist positive constants c_0 and δ such that

$$\text{dist}(x, C) \leq c_0^{\frac{1}{\theta}} \left(\max_{1 \leq i \leq m} \{ [g_{i,1}(x)]_+, \dots, [g_{i,\gamma_i}(x)]_+ \} \right)^\tau, \quad \forall \|x - \bar{x}\| \leq \delta.$$

Hence

$$(3.20) \quad \text{dist}^\theta(x, C) \leq c_0 \left(\max_{1 \leq i \leq m} \{[g_{i,1}(x)]_+, \dots, [g_{i,\gamma_i}(x)]_+\} \right)^{\theta\tau}, \quad \forall \|x - \bar{x}\| \leq \delta.$$

Now we claim that there exists $\beta > 0$ such that

$$(3.21) \quad \left(\max_{1 \leq i \leq m} \{[g_{i,1}(x)]_+, \dots, [g_{i,\gamma_i}(x)]_+\} \right)^\theta \leq \beta \sum_{i=1}^m \text{dist}^\theta(x, C_i), \quad \forall \|x - \bar{x}\| \leq \delta.$$

Suppose to the contrary that there exists a sequence $(x_k)_{k \in \mathbb{N}}$ in $\mathbb{B}(\bar{x}, \delta)$ such that

$$(3.22) \quad \left(\max_{1 \leq i \leq m} \{[g_{i,1}(x_k)]_+, \dots, [g_{i,\gamma_i}(x_k)]_+\} \right)^\theta > k \sum_{i=1}^m \text{dist}^\theta(x_k, C_i), \quad \forall k \in \mathbb{N}.$$

Without loss of generality, we can assume that $g_{i,1}, g_{i,2}, \dots, g_{i,\gamma_i}$ have the Lipschitz constant $L > 0$ on $\mathbb{B}(\bar{x}, \delta)$ for every $i = 1, 2, \dots, m$. Then, there exists a subsequence $(x_{k_l})_{l \in \mathbb{N}}$ of $(x_k)_{k \in \mathbb{N}}$, $1 \leq i_0 \leq m$ and $1 \leq j_0 \leq \gamma_{i_0}$ such that

$$\max_{1 \leq i \leq m} \{[g_{i,1}(x_{k_l})]_+, \dots, [g_{i,\gamma_i}(x_{k_l})]_+\} = [g_{i_0,j_0}(x_{k_l})]_+, \quad \forall l \in \mathbb{N}.$$

It follows from (3.22) that

$$(3.23) \quad ([g_{i_0,j_0}(x_{k_l})]_+)^\theta > k_l \sum_{i=1}^m \text{dist}^\theta(x_{k_l}, C_i), \quad \forall l \in \mathbb{N}.$$

Then $[g_{i_0,j_0}(x_{k_l})]_+ = g_{i_0,j_0}(x_{k_l})$ and hence for every $l \in \mathbb{N}$,

$$(3.24) \quad (g_{i_0,j_0}(x_{k_l}))^\theta > k_l \sum_{i=1}^m \text{dist}^\theta(x_{k_l}, C_i) = k_l \sum_{i=1}^m \|x_{k_l} - P_i(x_{k_l})\|^\theta \geq k_l \|x_{k_l} - P_{i_0}(x_{k_l})\|^\theta.$$

Since $P_{i_0}(x_{k_l}) \in C_{i_0}$, we have $g_{i_0,j_0}(P_{i_0}(x_{k_l})) \leq 0$ and $\|x_{k_l} - P_{i_0}(x_{k_l})\| \leq \|x_{k_l} - \bar{x}\| < \delta$ by $\bar{x} \in C_{i_0}$. Combining this with (3.24), we have

$$L^\theta \|x_{k_l} - P_{i_0}(x_{k_l})\|^\theta \geq (g_{i_0,j_0}(x_{k_l}) - g_{i_0,j_0}(P_{i_0}(x_{k_l})))^\theta > k_l \|x_{k_l} - P_{i_0}(x_{k_l})\|^\theta, \quad \forall l \in \mathbb{N}.$$

Hence we have $L^\theta > k_l$ for every $l \in \mathbb{N}$, this contradicts the fact that $k_l \rightarrow +\infty$. Thus, (3.21) holds.

Combining (3.21) and (3.20), we see that

$$\text{dist}^\theta(x, C) \leq c_0 \beta^\tau \left(\sum_{i=1}^m \text{dist}^\theta(x, C_i) \right)^\tau, \quad \forall \|x - \bar{x}\| \leq \delta,$$

and so the conclusion follows. \square

4 Convergence rate for the cyclic projection algorithm

In this section, we derive explicit convergence rate of the cyclic projection algorithm applied to finite intersections of basic semi-algebraic convex sets.

Before we come to our main result, we need the following useful lemma, Lemma 4.1, which is a special case of Alber and Reich's result in [1]. For the reader's convenience, we provide a direct and self-contained proof.

Lemma 4.1 (Recurrence relationships) *Let $p > 0$, and let $\{\delta_k\}_{k=0}^\infty$ and $\{\beta_k\}_{k=0}^\infty$ be two sequences of nonnegative numbers satisfying the conditions*

$$\beta_{k+1} \leq \beta_k(1 - \delta_k \beta_k^p) \text{ as } k = 0, 1, \dots$$

Then, we have

$$(4.1) \quad \beta_k \leq \left(\beta_0^{-p} + p \sum_{i=0}^{k-1} \delta_i \right)^{-\frac{1}{p}} \text{ for all } k \in \mathbb{N}.$$

We use the convention that $\frac{1}{0} = +\infty$. In particular, we have $\lim_{k \rightarrow \infty} \beta_k = 0$ whenever $\sum_{k=0}^\infty \delta_k = \infty$.

Proof. It follows from our assumption that

$$0 \leq \beta_{i+1} \leq \beta_i \leq \dots \leq \beta_0 \text{ and } \delta_i \beta_i^{p+1} \leq \beta_i - \beta_{i+1} \text{ as } i \in \mathbb{N}.$$

Fix $k \in \mathbb{N}$. We consider two cases.

Case 1: $\beta_k = 0$.

Clearly, (4.1) holds.

Case 1: $\beta_k \neq 0$.

Thus $\beta_k > 0$ and hence $\beta_i > 0$ for every $i \leq k$. Define the nonincreasing function $h : \mathbb{R}_{++} \rightarrow]-\infty, +\infty]$ by $h(x) := x^{-(p+1)}$. As $\delta_i h(\beta_i)^{-1} = \delta_i \beta_i^{p+1} \leq \beta_i - \beta_{i+1}$, then we get

$$\delta_i \leq (\beta_i - \beta_{i+1})h(\beta_i) \leq \int_{\beta_{i+1}}^{\beta_i} h(x)dx = \frac{\beta_{i+1}^{-p} - \beta_i^{-p}}{p}.$$

This implies that

$$(4.2) \quad \beta_{i+1}^{-p} - \beta_i^{-p} \geq p\delta_i \text{ for all } i \in \mathbb{N} \cup \{0\}.$$

Now fix any $k \in \mathbb{N}$ and, summing (4.2) from $i = 0$ to $i = k - 1$, we get

$$\beta_k^{-p} - \beta_0^{-p} \geq p \sum_{i=0}^{k-1} \delta_i.$$

which implies the conclusion in (4.1). \square

We also need the following technical result. The proof of it follows in part that of [23, Lemmas 3&4], one may also consult [10].

Proposition 4.2 (Cyclic convergence rate) *Let $D_i \subseteq \mathbb{R}^n$ be a closed convex set, $\forall i = 1, \dots, m$, and $D := \bigcap_{i=1}^m D_i \neq \emptyset$. Let $x_0 \in \mathbb{R}^n$ and the sequence of cyclic projections, $(x_k)_{k \in \mathbb{N}}$, be defined by*

$$x_1 := P_1 x_0, x_2 := P_2 x_1, \dots, x_m := P_m x_{m-1}, x_{m+1} := P_1 x_m, \dots,$$

where we set $P_i := P_{D_i}$ for the convenience. Suppose that Hölderian regularity with exponent τ ($0 < \tau \leq 1$) holds: for any compact set $K \subseteq \mathbb{R}^n$ and $\theta > 0$, there exists $c_0 > 0$ such that

$$\text{dist}^\theta(x, D) \leq c_0 \left(\sum_{i=1}^m \text{dist}^\theta(x, D_i) \right)^\tau, \quad \forall x \in K.$$

Then x_k converges to $x_\infty \in D$. Moreover, there exist $M > 0$ and $r_0 \in]0, 1[$ such that

$$\|x_k - x_\infty\| \leq \begin{cases} M k^{-\frac{1}{2\tau-1-2}} & \text{if } \tau \in]0, 1[, \\ M r_0^k & \text{if } \tau = 1, \end{cases} \quad \forall k \in \mathbb{N},$$

Proof. We denoted by $\alpha_i := (i \bmod m) + 1, \forall i \in \mathbb{N}$. Thus $x_{k+1} = P_{D_{\alpha_k}} x_k$. By Fact 2.14, there exists $x_\infty \in D$ such that $x_k \rightarrow x_\infty$.

We first follow closely the proofs of [23, Lemmas 3&4] to get that

$$(4.3) \quad \text{dist}^2(x_k, D) - \text{dist}^2(x_{k+1}, D) \geq \text{dist}^2(x_k, D_{\alpha_k}), \quad \forall k \in \mathbb{N}.$$

Indeed, using the definition of projection operator, we have

$$\begin{aligned} \text{dist}^2(x_k, D) - \text{dist}^2(x_{k+1}, D) &\geq \|x_k - P_D x_k\|^2 - \|x_{k+1} - P_D x_k\|^2 \\ &= \|x_k - P_D x_k\|^2 - \|P_{\alpha_k} x_k - P_D x_k\|^2 = \|x_k - P_D x_k\|^2 - \|P_{\alpha_k} x_k - x_k + x_k - P_D x_k\|^2 \\ &= \|x_k - P_D x_k\|^2 - \|P_{\alpha_k} x_k - x_k\|^2 - \|x_k - P_D x_k\|^2 + 2\langle x_k - P_{\alpha_k} x_k, x_k - P_D x_k \rangle \\ &\geq -\|x_k - P_{\alpha_k} x_k\|^2 + 2\langle x_k - P_{\alpha_k} x_k, x_k - P_{\alpha_k} x_k + P_{\alpha_k} x_k - P_D x_k \rangle \\ &= -\|x_k - P_{\alpha_k} x_k\|^2 + 2\|x_k - P_{\alpha_k} x_k\|^2 + 2\langle x_k - P_{\alpha_k} x_k, P_{\alpha_k} x_k - P_D x_k \rangle \\ &\geq \|x_k - P_{\alpha_k} x_k\|^2 = \text{dist}^2(x_k, D_{\alpha_k}), \quad \forall k \in \mathbb{N}. \end{aligned}$$

Hence (4.3) holds.

Next we claim that for every $i \in \mathbb{N}$

$$(4.4) \quad \text{dist}(x_k, D_{\alpha_i}) \leq \text{dist}(x_k, D_{\alpha_k}) + \text{dist}(x_{k+1}, D_{\alpha_{k+1}}) + \dots + \text{dist}(x_{k+m-1}, D_{\alpha_{k+m-1}}).$$

To see this, note that there exists $i_0 \leq m-1$ such that $\alpha_{i_0+k} = \alpha_i$. Then, we have

$$\begin{aligned} \text{dist}(x_k, D_{\alpha_i}) &= \text{dist}(x_k, D_{\alpha_{i_0+k}}) \leq \|x_k - x_{i_0+k}\| + \|x_{i_0+k} - P_{\alpha_{i_0+k}} x_{i_0+k}\| \\ &\leq \|x_k - x_{k+1}\| + \dots + \|x_{i_0+k-1} - x_{i_0+k}\| + \|x_{i_0+k} - P_{\alpha_{i_0+k}} x_{i_0+k}\| \\ &= \|x_k - P_{\alpha_k} x_k\| + \dots + \|x_{i_0+k-1} - P_{\alpha_{i_0+k-1}} x_{i_0+k-1}\| + \|x_{i_0+k} - P_{\alpha_{i_0+k}} x_{i_0+k}\| \\ &= \text{dist}(x_k, D_{\alpha_k}) + \text{dist}(x_{k+1}, D_{\alpha_{k+1}}) + \dots + \text{dist}(x_{i_0+k}, D_{\alpha_{i_0+k}}) \\ &\leq \text{dist}(x_k, D_{\alpha_k}) + \text{dist}(x_{k+1}, D_{\alpha_{k+1}}) + \dots + \text{dist}(x_{k+m-1}, D_{\alpha_{k+m-1}}) \quad (\text{by } i_0 \leq m-1). \end{aligned}$$

Hence (4.4) holds.

Thus by (4.4),

$$\begin{aligned} \text{dist}^2(x_k, D_{\alpha_i}) &\leq \left(m \max_{k \leq i \leq k+m-1} \text{dist}(x_i, D_{\alpha_i}) \right)^2 \\ (4.5) \quad &\leq m^2 \left(\text{dist}^2(x_k, D_{\alpha_k}) + \text{dist}^2(x_{k+1}, D_{\alpha_{k+1}}) + \dots + \text{dist}^2(x_{k+m-1}, D_{\alpha_{k+m-1}}) \right). \end{aligned}$$

By the assumption, there exists $c_0 > 0$ such that

$$\text{dist}^2(x_k, D) \leq c_0 \left(\sum_{i=1}^m \text{dist}^2(x_k, D_i) \right)^\tau, \quad \forall k \in \mathbb{N}.$$

By enlarging c_0 if necessary, we may assume that $c_0 > 1$.

Let $r := \tau^{-1}$. Then by (4.5), for every $k \in \mathbb{N}$,

$$\begin{aligned} &\frac{1}{c_0^r} \text{dist}^{2r}(x_k, D) \\ &\leq \sum_{i=1}^m \text{dist}^2(x_k, D_i) \leq m \max_{1 \leq i \leq m} \text{dist}^2(x_k, D_i) \\ &\leq m^3 \left(\text{dist}^2(x_k, D_{\alpha_k}) + \text{dist}^2(x_{k+1}, D_{\alpha_{k+1}}) + \dots + \text{dist}^2(x_{k+m-1}, D_{\alpha_{k+m-1}}) \right) \\ &\leq m^3 \sum_{i=k}^{k+m-1} \text{dist}^2(x_i, D) - \text{dist}^2(x_{i+1}, D) \quad (\text{by (4.3)}) \\ (4.6) \quad &= m^3 \left(\text{dist}^2(x_k, D) - \text{dist}^2(x_{k+m}, D) \right). \end{aligned}$$

Thus we have

$$(4.7) \quad \text{dist}^2(x_{k+m}, D) \leq \text{dist}^2(x_k, D) - \frac{1}{m^3 c_0^r} \text{dist}^{2r}(x_k, D).$$

Now we consider two cases.

Case 1: $\tau \in]0, 1[$.

Thus $r > 1$. Fix $k_0 \in \mathbb{N}$. Let $\beta_i := \text{dist}^2(x_{k_0+im}, D)$, $\forall i \in \mathbb{N} \cup \{0\}$. Then (4.7) shows that

$$(4.8) \quad \beta_{i+1} \leq \beta_i - \frac{1}{m^3 c_0^r} \beta_i^r = \beta_i \left(1 - \frac{1}{m^3 c_0^r} \beta_i^{r-1}\right).$$

By Lemma 4.1,

$$\text{dist}^2(x_{k_0+im}, D) = \beta_i \leq \left(\beta_0^{1-r} + (r-1)i \frac{1}{m^3 c_0^r} \right)^{-\frac{1}{r-1}}, \quad \forall i \in \mathbb{N}.$$

Thus, there exists $M_0 > 0$ such that

$$\text{dist}(x_{k_0+im}, D) \leq M_0 \frac{1}{\sqrt[r-1]{i}}, \quad \forall i \in \mathbb{N}.$$

Hence there exists $M_1 > 0$ such that

$$(4.9) \quad \text{dist}(x_k, D) \leq M_1 \frac{1}{\sqrt[r-1]{k}} = M_1 k^{-\frac{1}{2\tau-1-2}}, \quad \forall k \in \mathbb{N}.$$

So, we have

$$(4.10) \quad \|x_k - x_\infty\| \leq \|x_k - P_D(x_k)\| + \|P_D(x_k) - x_\infty\| = \text{dist}(x_k, D) + \|P_D(x_k) - x_\infty\|.$$

By [23, Lemma 3],

$$\|x_{k+l} - P_D(x_k)\| \leq \|x_k - P_D(x_k)\| = \text{dist}(x_k, D), \quad \forall l \in \mathbb{N}.$$

Letting $l \rightarrow \infty$ in the above inequality, we obtain that

$$(4.11) \quad \|x_\infty - P_D(x_k)\| \leq \text{dist}(x_k, D).$$

Combining (4.10), (4.11) and (4.9),

$$(4.12) \quad \|x_k - x_\infty\| \leq 2 \text{dist}(x_k, D) \leq 2M_1 k^{-\frac{1}{2\tau-1-2}}, \quad \forall k \in \mathbb{N}.$$

Thus, the conclusion follows by letting $M := 2M_1$.

Case 2: $\tau = 1$.

Then we have $r = \tau^{-1} = 1$, and so, (4.7) implies that for all $k \in \mathbb{N}$

$$\text{dist}(x_{k+m}, D) \leq \sqrt{1 - \frac{1}{m^3 c_0^r}} \text{dist}(x_k, D).$$

Hence there exist $M'_1 > 0$ and $r_0 \in]0, 1[$ such that $\text{dist}(x_k, D) \leq M'_1 r_0^k$, $\forall k \in \mathbb{N}$. Then, using a similar method of proof as in Case 1, we obtain that

$$(4.13) \quad \|x_k - x_\infty\| \leq 2 \text{dist}(x_k, D) \leq 2M'_1 r_0^k, \quad \forall k \in \mathbb{N}.$$

Thus, the conclusion follows by letting $M := 2M'_1$. □

Remark 4.3 (Connection to the existing result on linear convergence) In the case where there exists $i_0 \in \{1, \dots, m\}$ such that $D_{i_0} \cap \text{int}(\bigcap_{i \neq i_0} D_i) \neq \emptyset$ (in this case, we say the intersection is regular), then the Hölderian regularity result holds with exponent $\tau = 1$. So, the preceding proposition implies that the cyclic projection algorithm converges linearly in the regular intersection case. Thus, this recovers the linear convergence result for cyclic projection algorithm established in [23]. \diamond

We are now ready for one of our main results.

Theorem 4.4 (Estimate of the cyclic convergence rate) *Let $x_0 \in \mathbb{R}^n$ and the sequence of cyclic projections, $(x_k)_{k \in \mathbb{N}}$, be defined by*

$$x_1 := P_1 x_0, x_2 := P_2 x_1, \dots, x_m := P_m x_{m-1}, x_{m+1} := P_1 x_m \dots$$

Then x_k converges to $x_\infty \in C$, and there exist $M > 0$ and $r_0 \in]0, 1[$ such that

$$\|x_k - x_\infty\| \leq \begin{cases} M^{\frac{1}{k^\rho}}, & \text{if } d > 1; \\ M r_0^k, & \text{if } d = 1 \end{cases}, \quad \forall k \in \mathbb{N},$$

where $\rho := \frac{1}{\min\{(2d-1)^n - 1, 2\beta(n-1)d^n - 2\}}$ and $\beta(n-1)$ is the central binomial coefficient with respect to $n-1$ which is given by $\binom{n-1}{\lfloor (n-1)/2 \rfloor}$.

Proof. Combining Theorem 3.6 and Proposition 4.2, we directly obtain $\tau := \frac{1}{\min\{\frac{(2d-1)^{n+1}}{2}, \beta(n-1)d^n\}}$. Note that $2\tau^{-1} - 2 = \min\{(2d-1)^n - 1, 2\beta(n-1)d^n - 2\}$. Thus the conclusion follows from the preceding proposition. \square

Remark 4.5 (Discussion on our estimation of the convergence rate) Although our estimate of the convergence rate works for cyclic projection algorithm with finitely many basic semialgebraic convex sets without any regularity condition, the estimated convergence rate is quite poor when the dimension n of the underlying space and the maximal degree d are large. This is mainly due to the fact that the estimated convergence rate is derived by using the local error bound result for general convex polynomial systems. It would be interesting to see how one could improve the estimation of the convergence rate by either adopting other approaches or by further exploiting the structure of the underlying convex sets. For example, one possibility would be to examine problems involving some suitable additional curvature or uniform convexity assumptions. This will be another of our future research topics.

4.1 Alternating projection algorithm

In this subsection, we discuss the convergence rate of the alternating projection algorithm. We assume throughout this subsection that

$$\begin{aligned} &g_i, h_j \text{ are convex polynomials with degree at most } d, \forall i = 1, 2, \dots, m, j = 1, 2, \dots, l \\ &A := \{x \in \mathbb{R}^n \mid g_i(x) \leq 0, i = 1, \dots, m\} \\ &B := \{x \in \mathbb{R}^n \mid h_j(x) \leq 0, j = 1, \dots, l\} \\ &b_0 \in \mathbb{R}^n, \quad a_{k+1} := P_A b_k, \quad b_{k+1} := P_B a_{k+1}. \end{aligned}$$

As an immediate corollary of Theorem 4.4, we first obtain the following estimate on the convergence rate of the alternating projection algorithm in the case where the two sets have nonempty intersection.

The case of two sets with nonempty intersection

Corollary 4.6 (Alternating convergence rate) *Suppose that $A \cap B \neq \emptyset$. Let the sequence $\{(a_k, b_k)\}$ be generated by the alternating projection algorithm. Then, $a_k, b_k \rightarrow c \in A \cap B$. Moreover, there exist $M > 0$ and $r_0 \in]0, 1[$ such that for every $k \in \mathbb{N}$*

$$\|a_k - c\| \leq \begin{cases} M \frac{1}{k^\rho}, & \text{if } d > 1; \\ Mr_0^k, & \text{if } d = 1 \end{cases}, \quad \text{and} \quad \|b_k - c\| \leq \begin{cases} M \frac{1}{k^\rho}, & \text{if } d > 1; \\ Mr_0^k, & \text{if } d = 1 \end{cases},$$

where $\rho := \frac{1}{\min\{(2d-1)^n - 1, 2\beta(n-1)d^n - 2\}}$ and $\beta(n-1)$ is the central binomial coefficient with respect to $n-1$ which is given by $\binom{n-1}{\lfloor (n-1)/2 \rfloor}$.

Recently, [2] established a local convergence rate analysis for proximal alternating projection methods for very general nonconvex problems, and where the corresponding convergence rate involves the exponent of the Kurdyka-Łojasiewicz inequality. The *proximal alternating projection method* is a variant of the alternating projection algorithm we discussed here. On the other hand, as we discussed before, in general, the actual exponent of the Kurdyka-Łojasiewicz inequality is typically unknown and hard to estimate. Corollary 4.6 above complements the result of [2] in the case of basic convex semialgebraic cases by providing an explicit estimate of the convergence rate.

The case of two sets with empty intersection

In this part, we consider the general case where the intersection of these two sets is (possibly) empty. We first need the following lemma.

Lemma 4.7 *The difference $B - A$ of two basic semi-algebraic sets A, B is closed.*

Proof. Let $b_k \in B$ and $a_k \in A$ be such that $b_k - a_k \rightarrow c$. We now show that $c \in B - A$. Consider the following convex polynomial optimization problem

$$(P) \quad \min_{x \in \mathbb{R}^n, y \in \mathbb{R}^n} \quad \|(y - x) - c\|^2$$

$$\text{s.t.} \quad g_i(x) \leq 0, i = 1, \dots, m,$$

$$h_j(y) \leq 0, j = 1, \dots, l.$$

Note that (a_k, b_k) are feasible for (P). Hence we see that $\inf(P) = 0$. By Fact 2.2, the optimal solution of (P) exists. Thus there exists $x \in A$ and $y \in B$ such that $c = y - x \in B - A$. Hence the conclusion follows. \square

Remark 4.8 With A and B defined as above, Fact 2.2 implies that $B - A$ is closed convex. Hence $P_{B-A}0 \neq \emptyset$. Let $v := P_{B-A}0$. Then, there exist $a \in A$ and $b \in B$ such that $v = b - a$ and hence $\text{dist}(A, B) = \|v\|$. \diamond

Remark 4.9 In general, the distance between two convex and semi-algebraic sets need not be attained. For instance, consider $D := \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 x_2 \geq 1, x_1 \geq 0, x_2 \geq 0\}$ and $E := \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 = 0\}$. It is clear that D, E are both convex and semialgebraic; while D is not a basic semi-algebraic convex set (as explained in Example 2.13). Clearly, $\text{dist}(D, E) = 0$ but $D \cap E = \emptyset$. Thus, the distance is not attained in this case. \diamond

The proof of Theorem 4.10 partially follows that of [8, Theorem 3.12].

Theorem 4.10 (Convergence rate in the infeasible case) *Let the sequence $\{(a_k, b_k)\}$ be generated by the alternating projection algorithm. Then $a_k \rightarrow \tilde{a} \in A$ and $b_k \rightarrow \tilde{b} \in B$ with $\tilde{b} - \tilde{a} = v$ where $v := P_{B-A}0$. Moreover, there exist $M > 0$ and $r_0 \in]0, 1[$ such that for every $k \in \mathbb{N}$*

$$(4.14) \quad \|a_k - \tilde{a}\| \leq \begin{cases} M \frac{1}{k^\rho}, & \text{if } d > 1; \\ Mr_0^k, & \text{if } d = 1 \end{cases} \quad \text{and} \quad \|b_k - \tilde{b}\| \leq \begin{cases} M \frac{1}{k^\rho}, & \text{if } d > 1; \\ Mr_0^k, & \text{if } d = 1 \end{cases},$$

where $\rho := \frac{1}{\min\{(2d-1)^{n-1}, 2\beta(n-1)d^{n-2}\}}$ and $\beta(n-1)$ is the central binomial coefficient with respect to $n-1$ which is given by $\binom{n-1}{\lfloor (n-1)/2 \rfloor}$.

Proof. Lemma 4.7 implies that $B - A$ is closed. Then by Fact 2.15(i), there exist $\tilde{a} \in A, \tilde{b} \in B$ such that $a_k \rightarrow \tilde{a} \in A$ and $b_k \rightarrow \tilde{b} \in B$ with $\tilde{b} - \tilde{a} = v$. By Theorem 3.6, there exists $c_0 > 1$ such that

$$(4.15) \quad \text{dist}(a_k, A \cap (B - v)) \leq c_0 (\text{dist}(a_k, A) + \text{dist}(a_k, B - v))^{\frac{1}{r}} = c_0 \text{dist}^{\frac{1}{r}}(a_k, B - v),$$

where $r := \min \left\{ \frac{(2d-1)^n+1}{2}, \beta(n-1)d^n \right\}$. Fix $x \in A \cap (B-v)$. Note that $v = P_{B-A}0$ we have $P_B(x) = x + v$ by Fact 2.15(ii). This implies that

$$\begin{aligned}
\text{dist}^2(a_k, B-v) &\leq \|a_k - (b_k - v)\|^2 = \|(a_k - x) - (b_k - (v+x))\|^2 \\
&= \|(a_k - x) - (P_B a_k - P_B x)\|^2 \\
&\leq \|a_k - x\|^2 - \|P_B a_k - P_B x\|^2 \quad (\text{by [12, Proposition 4.8]}) \\
&= \|a_k - x\|^2 - \|b_k - (x+v)\|^2 \\
&\leq \|a_k - x\|^2 - \|P_A b_k - P_A(x+v)\|^2 \\
&= \|a_k - x\|^2 - \|a_{k+1} - x\|^2 \quad (\text{by Fact 2.15(ii)}).
\end{aligned}$$

In particular, choose $x = P_{A \cap (B-v)} a_k$. Then, we have

$$\begin{aligned}
\text{dist}^2(a_k, B-v) &\leq \text{dist}^2(a_k, A \cap (B-v)) - \|a_{k+1} - P_{A \cap (B-v)} a_k\|^2 \\
&\leq \text{dist}^2(a_k, A \cap (B-v)) - \text{dist}^2(a_{k+1}, A \cap (B-v)).
\end{aligned}$$

Combining with (4.15), we have

$$\begin{aligned}
\frac{1}{c_0^{2r}} \text{dist}^{2r}(a_k, A \cap (B-v)) &\leq \text{dist}^2(a_k, (B-v)) \\
(4.16) \quad &\leq \text{dist}^2(a_k, A \cap (B-v)) - \text{dist}^2(a_{k+1}, A \cap (B-v)).
\end{aligned}$$

Thus

$$\text{dist}(a_{k+1}, A \cap (B-v))^2 \leq \text{dist}(a_k, A \cap (B-v))^2 - \frac{1}{c_0^{2r}} \text{dist}(a_k, A \cap (B-v))^{2r}.$$

Now, let $\beta_k := \text{dist}(a_k, A \cap (B-v))^2$, $k \in \mathbb{N}$. Then, we have

$$(4.17) \quad \beta_{k+1} \leq \beta_k \left(1 - \frac{1}{c_0^{2r}} \beta_k^{r-1}\right).$$

Now we consider two cases.

Case 1: $d > 1$.

In this case, we have $r > 1$. Applying the preceding Lemma 4.1 with $\delta_k := \frac{1}{c_0^{2r}}$ and $p := r-1$, by (4.17), we see that

$$\text{dist}^2(a_k, A \cap (B-v)) = \beta_k \leq \left(\beta_0^{1-r} + \frac{(r-1)}{c_0^{2r}} k \right)^{-\frac{1}{r-1}} \quad \text{for all } k \in \mathbb{N}.$$

Thence there exists $M_0 > 0$ such that

$$\text{dist}(a_k, A \cap (B-v)) \leq M_0 \frac{1}{k^\rho}, \quad \forall k \in \mathbb{N}$$

where $\rho := \frac{1}{\min \left\{ (2d-1)^n-1, 2\beta(n-1)d^n-2 \right\}}$. Then, [8, Example 3.2] shows that $(a_k)_{k \in \mathbb{N}}$ is Fejér monotone with respect to $A \cap (B-v)$. Thus, by Fact 2.17,

$$\|a_k - \tilde{a}\| \leq 2 \text{dist}(a_k, A \cap (B-v)) \leq 2M_0 \frac{1}{k^\rho}, \quad \forall k \in \mathbb{N}.$$

Case 2: $d = 1$.

Thus $r = 1$. Then by (4.17), $\text{dist}(a_{k+1}, A \cap (B - v)) \leq \theta \text{dist}(a_k, A \cap (B - v))$, where $\theta := \sqrt{1 - \frac{1}{c_0^2}}$. Then $\theta < 1$ since $c_0 > 1$. Hence there exists $M_1 > 0$ such that

$$\text{dist}(a_k, A \cap (B - v)) \leq M_1 \theta^k.$$

From [8, Example 3.2], we see that $(a_k)_{k \in \mathbb{N}}$ is Fejér monotone with respect to $A \cap (B - v)$. Thus, by Fact 2.17,

$$\|a_k - \tilde{a}\| \leq 2 \text{dist}(a_k, A \cap (B - v)) \leq 2M_1 \theta^k.$$

Set $M_2 := \max\{2M_0, 2M_1\}$. Combining the above two cases, we have

$$\|a_k - \tilde{a}\| \leq \begin{cases} M_2 \frac{1}{k^\rho}, & \text{if } d > 1; \\ M_2 \theta^k, & \text{if } d = 1 \end{cases}, \quad \forall k \in \mathbb{N}.$$

Similarly, we can show that there exist $L > 0$ and $\eta \in]0, 1[$ such that

$$\|b_k - \tilde{b}\| \leq \begin{cases} L \frac{1}{k^\rho}, & \text{if } d > 1; \\ L \eta^k, & \text{if } d = 1 \end{cases}, \quad \forall k \in \mathbb{N}.$$

Therefore, the conclusion follows by taking $M := \max\{M_2, L\}$ and $r_0 := \max\{\theta, \eta\}$. \square

5 Examples and remarks

In this section, we will provide several examples of the rates of convergence of the cyclic projection algorithm and the von Neumann alternating projection algorithm. We first start with some examples where the basic semialgebraic convex sets are described by convex quadratic functions. Subsequently, we will examine examples where the basic semialgebraic convex sets are described by higher degree convex polynomials.

Basic semialgebraic convex sets described by convex quadratic functions

Example 5.1 Let

$$C_1 := \{(x, y) \in \mathbb{R}^2 \mid (x + 1)^2 + y^2 - 1 \leq 0\}$$

$$C_2 := \{(x, y) \in \mathbb{R}^2 \mid x + y - 1 \leq 0\}$$

$$C_3 := \{(x, y) \in \mathbb{R}^2 \mid (x - 1)^2 + y^2 - 1 \leq 0\}$$

$$C_4 := \{(x, y) \in \mathbb{R}^2 \mid x + (y + 2)^2 - 4 \leq 0\}.$$

Take $x_0 \in \mathbb{R}^2$. Let $(x_k)_{k \in \mathbb{N}}$ be defined by

$$x_1 := P_1 x_0, x_2 := P_2 x_1, x_3 := P_3 x_2, x_4 := P_4 x_3, x_5 := P_1 x_4 \dots$$

Then $\|x_k\| = O(\frac{1}{k^{\frac{1}{6}}})$. \diamond

Proof. Clearly, $\bigcap_{i=1}^4 C_i = \{0\}$. Then apply $n = 2$ and $d = 2$ to Theorem 4.2. \square

Example 5.2 Let $\alpha \geq 0$ and

$$\begin{aligned} A &:= \{(x, y) \in \mathbb{R}^2 \mid (x+1)^2 + y^2 - 1 \leq 0\} = (-1, 0) + \overline{\mathbb{B}}(0, 1) \\ B &:= \{(x, y) \in \mathbb{R}^2 \mid -x + \alpha \leq 0\}. \end{aligned}$$

Let $(a_k)_{k \in \mathbb{N}} := (u_k, v_k)_{k \in \mathbb{N}}$ and $(b_k)_{k \in \mathbb{N}} := (s_k, t_k)_{k \in \mathbb{N}}$ be defined by

$$b_0 \in \mathbb{R}^2, \quad a_{k+1} := P_A b_k, \quad b_{k+1} := P_B a_{k+1}.$$

Then for every $k \geq 2$

$$\begin{aligned} b_k &= \left(\alpha, \frac{t_1}{\sqrt{(1+\alpha)^{2(k-1)} + t_1^2 \sum_{i=0}^{k-2} ((1+\alpha)^{2i})}} \right) \\ a_{k+1} &= \left(-1 + \frac{\alpha + 1}{\sqrt{(\alpha + 1)^2 + \frac{t_1^2}{(1+\alpha)^{2(k-1)} + t_1^2 \sum_{i=0}^{k-2} ((1+\alpha)^{2i})}}}, \frac{t_1}{\sqrt{(1+\alpha)^{2k} + t_1^2 \sum_{i=0}^{k-1} ((1+\alpha)^{2i})}} \right). \end{aligned}$$

Consequently, $a_k \rightarrow 0$ and $b_k \rightarrow (\alpha, 0)$ at the rate of $k^{-\frac{1}{2}}$ when $\alpha = 0$. When $\alpha \neq 0$ (then $A \cap B = \emptyset$), $a_k \rightarrow 0$ and $b_k \rightarrow (\alpha, 0)$ at the rate of $(1+\alpha)^{-k}$. \diamond

Proof. We first claim that

$$(5.1) \quad b_{k+1} = (\alpha, t_{k+1}) = \left(\alpha, \frac{t_k}{\sqrt{(1+\alpha)^2 + t_k^2}} \right), \quad \forall k \geq 1.$$

By [12, Examples 3.17&3.21 and Proposition 3.17], we have

$$\begin{aligned} (5.2) \quad P_A(x, y) &= (-1, 0) + \frac{(x+1, y)}{\max\{1, \|(x+1, y)\|\}}, \quad \forall (x, y) \in \mathbb{R}^2 \\ P_B(x, y) &= (\alpha, y), \quad \forall (x, y) \notin \text{int } B. \end{aligned}$$

Let $k \geq 1$. Since $A \cap B = \{0\}$ or $A \cap B = \emptyset$, $a_k \notin \text{int } B$. Then by (5.2), $b_k = (\alpha, v_k)$ and then

$$\begin{aligned} a_{k+1} &= P_A b_k = (-1, 0) + \frac{(1+\alpha, v_k)}{\max\{1, \|(1+\alpha, v_k)\|\}} = (-1, 0) + \frac{(1+\alpha, v_k)}{\sqrt{(1+\alpha)^2 + v_k^2}} \\ b_{k+1} &= P_B(a_{k+1}) = \left(\alpha, \frac{v_k}{\sqrt{(1+\alpha)^2 + v_k^2}} \right) = \left(\alpha, \frac{t_k}{\sqrt{(1+\alpha)^2 + t_k^2}} \right). \end{aligned}$$

Hence (5.1) holds. Next we show that

$$(5.3) \quad b_k = \left(\alpha, \frac{t_1}{\sqrt{(1+\alpha)^{2(k-1)} + t_1^2 \sum_{i=0}^{k-2} ((1+\alpha)^{2i})}} \right), \quad \forall k \geq 2.$$

We prove (5.3) by the induction on k .

By (5.1), (5.3) holds when $k = 2$. Now assume that (5.3) holds when $k = p$, where $p \geq 2$. Now we consider the case of $k = p + 1$. By the assumption, we have

$$(5.4) \quad b_p = \left(\alpha, \frac{t_1}{\sqrt{(1+\alpha)^{2(p-1)} + t_1^2 \sum_{i=0}^{p-2} ((1+\alpha)^{2i})}} \right).$$

Then by (5.1), we have

$$\begin{aligned} b_{p+1} &= \left(\alpha, \frac{t_p}{\sqrt{(1+\alpha)^2 + t_p^2}} \right) \\ &= \left(\alpha, \frac{\frac{t_1}{\sqrt{(1+\alpha)^{2(p-1)} + t_1^2 \sum_{i=0}^{p-2} ((1+\alpha)^{2i})}}}{\sqrt{(1+\alpha)^2 + \frac{t_1^2}{(1+\alpha)^{2(p-1)} + t_1^2 \sum_{i=0}^{p-2} ((1+\alpha)^{2i})}}} \right) \\ &= \left(\alpha, \frac{t_1}{\sqrt{(1+\alpha)^{2p} + t_1^2 \sum_{i=0}^{p-1} ((1+\alpha)^{2i})}} \right). \end{aligned}$$

Hence (5.3) holds.

Combining (5.2) and (5.3), we have for every $k \geq 2$

$$\begin{aligned} a_{k+1} &= P_A b_k \\ &= \left(-1 + \frac{\alpha + 1}{\sqrt{(\alpha + 1)^2 + \frac{t_1^2}{(1+\alpha)^{2(k-1)} + t_1^2 \sum_{i=0}^{k-2} ((1+\alpha)^{2i})}}}, \frac{t_1}{\sqrt{(1+\alpha)^{2k} + t_1^2 \sum_{i=0}^{k-1} ((1+\alpha)^{2i})}} \right). \end{aligned}$$

Hence $a_k \rightarrow 0$ and $b_k \rightarrow (\alpha, 0)$ at the rate of $k^{-\frac{1}{2}}$ when $\alpha = 0$. When $\alpha \neq 0$, $a_k \rightarrow 0$ and $b_k \rightarrow (\alpha, 0)$ at the rate of $(1+\alpha)^{-k}$. \square

Remark 5.3 According to Theorem 4.10, we can only deduce that $(a_k)_{k \in \mathbb{N}}$ in Example 5.2 converges to $(0, 0)$ and $(b_k)_{k \in \mathbb{N}}$ converge to $(\alpha, 0)$ at the rate of at least of $k^{-\frac{1}{6}}$. \diamond

Example 5.4 Let

$$\begin{aligned} A &:= \{(x, y) \in \mathbb{R}^2 \mid (x+1)^2 + y^2 - 1 \leq 0\} \\ B &:= \{(x, y) \in \mathbb{R}^2 \mid (x-1)^2 + y^2 - 1 \leq 0\}. \end{aligned}$$

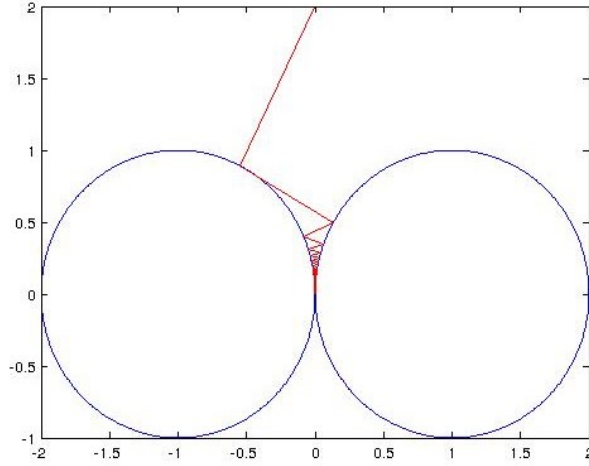


Figure 1: The iteration commencing at $(0, 2)$.

Let $(x_k, y_k)_{k \in \mathbb{N}}$ be defined by

$$(x_0, y_0) \in \mathbb{R}^2, (x_1, y_1) := P_A(x_0, y_0), (x_2, y_2) := P_B(x_1, y_1), (x_3, y_3) := P_A(x_2, y_2), \quad \dots$$

Note that

$$P_A(x, y) = \left(-1 + \frac{x-1}{\sqrt{(x+1)^2 + y^2}}, \frac{y}{\sqrt{(x+1)^2 + y^2}} \right), \quad \forall (x, y) \in \mathbb{R}_+ \times \mathbb{R}_{++}$$

$$P_B(x, y) = \left(1 + \frac{x+1}{\sqrt{(x-1)^2 + y^2}}, \frac{y}{\sqrt{(x-1)^2 + y^2}} \right), \quad \forall (x, y) \in \mathbb{R}_- \times \mathbb{R}_{++}.$$

Figure 1 depicts the algorithm's trajectory with starting point $(0, 2)$.

Suppose, without loss of generality, that one starts on a point on one the half-circles nearest the other circle. Then the distance from zero (for every $k \in \mathbb{N}$), $r_k := \sqrt{x_k^2 + y_k^2}$ satisfies $r_k^2 = 2\alpha_k$ where $\alpha_k := |x_k|$ since $(x_k, y_k) \in \text{bd } A \cup \text{bd } B$. Hence

$$1 - \alpha_{k+1} = \frac{1 + \alpha_k}{\sqrt{1 + 4\alpha_k}}.$$

Linearizing, we obtain that $w_k := 4\alpha_k$ approximately satisfies the logistics equation

$$w_{k+1} \approx w_k(1 - w_k)$$

This can be explicitly solved by writing

$$\frac{1}{w_{k+1}} - \frac{1}{w_k} = \frac{1}{1 - w_k}$$

When summing and dividing by N , leads to

$$\begin{aligned}
(5.5) \quad \lim_{N \rightarrow \infty} \frac{1}{Nw_N} &= \lim_{N \rightarrow \infty} \left(\frac{1}{Nw_N} - \frac{1}{Nw_0} \right) \\
&= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=0}^{N-1} \frac{1}{1-w_k} \\
&= \lim_{N \rightarrow \infty} \frac{1}{1-w_N} = 1,
\end{aligned}$$

since Césaro summability is conservative and $w_N \rightarrow 0$. Hence $\alpha_k \sim 1/(4k)$ and so

$$\sqrt{x_k^2 + y_k^2} = r_k \sim \frac{1}{\sqrt{2k}}.$$

For instance, with $\alpha_0 = 1, N = 10^6$, we obtain $\alpha_N \approx 0.0000002499992442$. A similar analysis can be performed in the previous example. \diamond

Remark 5.5 According to Theorem 4.10, we can only deduce that $(a_k)_{k \in \mathbb{N}}$ and $(b_k)_{k \in \mathbb{N}}$ in Example 5.4 converge to $(0, 0)$ at the rate of at least of $k^{-\frac{1}{6}}$. \diamond

Basic semialgebraic convex sets described by convex polynomials

In general, identifying the exact convergence rate of the cyclic projection method in a direct way can be quite arduous when applied to finitely many basic semialgebraic convex sets. Below, we provide two simple examples to illustrate how the convergence rate depends on the maximum degree of the polynomials that described the basic semialgebraic convex sets and on the dimension of the underlying space.

Example 5.6 Let A, B be defined by

$$\begin{aligned}
A &:= \{(x, y) \in \mathbb{R}^2 \mid x \leq 0\} \\
B &:= \{(x, y) \in \mathbb{R}^2 \mid y^d - x \leq 0\},
\end{aligned}$$

where d is an even number. Let $(a_k)_{k \in \mathbb{N}}, (b_k)_{k \in \mathbb{N}} := (x_k, y_k)_{k \in \mathbb{N}}$ be defined by

$$b_0 := (x_0, y_0) \in \mathbb{R} \times \mathbb{R}_{++} \text{ with } \|b_0\| \leq 1, \quad a_{k+1} := P_A b_k, \quad b_{k+1} := P_B a_{k+1}.$$

Then for every $k \in \mathbb{N}$

$$a_{k+1} = (0, y_k), \quad b_k = (y_k^d, y_k), \quad \text{and} \quad d y_{k+1}^{2d-1} + y_{k+1} = y_k.$$

Consequently, $(a_k)_{k \in \mathbb{N}}$ and $(b_k)_{k \in \mathbb{N}}$ converge to 0 at the rate of at least of $k^{-\frac{1}{2d-2}}$.

Proof. Let $k \in \mathbb{N}$. Since $A \cap B = \{0\}$, $b_k \notin \text{int } A$. Since $b_k \in \text{bd } B$ and then $x_k = y_k^d$. Thus $a_{k+1} = P_A b_k = P_A(x_k, y_k) = (0, y_k)$. Then we have $y_k \rightarrow 0$,

$$b_{k+1} = (x_{k+1}, y_{k+1}) = (y_{k+1}^d, y_{k+1}) = P_B a_{k+1} = P_B(0, y_k).$$

Thus y_{k+1} is a minimizer of the function

$$y \mapsto \|(y^d, y) - (0, y_k)\|^2 = \|(y^d, y - y_k)\|^2 = y^{2d} + y^2 - 2yy_k + y_k^2.$$

Thus $(dy_{k+1}^{2d-2} + 1)y_{k+1} - y_k = dy_{k+1}^{2d-1} + y_{k+1} - y_k = 0$. Then we have

$$\frac{1}{(dy_{k+1}^{2d-2} + 1)^{2d-2}} \frac{1}{y_{k+1}^{2d-2}} = \frac{1}{y_k^{2d-2}}$$

Now, for the function $h(x) := \frac{1}{(x+1)^{2d-2}}$, we have

$$h'(x) = -(2d-2) \frac{1}{(x+1)^{2d-1}} \text{ for all } x \in \mathbb{R}.$$

Note that $y_k \rightarrow 0$. So, for all large k ,

$$\frac{1}{(dy_{k+1}^{2d-2} + 1)^{2d-2}} = h(dy_{k+1}^{2d-2}) \approx h(0) + h'(0)dy_{k+1}^{2d-2} = 1 - d(2d-2)y_{k+1}^{2d-2}.$$

This gives us that for all large k ,

$$\frac{1}{y_k^{2d-2}} \approx \left(1 - d(2d-2)y_{k+1}^{2d-2}\right) \frac{1}{y_{k+1}^{2d-2}} = \frac{1}{y_{k+1}^{2d-2}} - d(2d-2).$$

In other words, there exists $l_0 \in \mathbb{N}$ such that for all $l \geq l_0$

$$(5.6) \quad \frac{1}{y_{l+1}^{2d-2}} - \frac{1}{y_l^{2d-2}} \approx d(2d-2).$$

Let $k > l_0$. Summing the above relation from $l = l_0$ to k , we have

$$\frac{1}{y_k^{2d-2}} - \frac{1}{y_{l_0}^{2d-2}} \approx d(2d-2)(k - l_0 + 1)$$

and so,

$$y_k \approx \left(\frac{1}{y_{l_0}^{2d-2}} + d(2d-2)(k - l_0 + 1) \right)^{-\frac{1}{2d-2}} = O(k^{-\frac{1}{2d-2}}).$$

Thence $(a_k)_{k \in \mathbb{N}}$ and $(b_k)_{k \in \mathbb{N}}$ converge to 0 at that rate. \square

Remark 5.7 Similarly, according to Theorem 4.10, we can only deduce that $(a_k)_{k \in \mathbb{N}}$ and $(b_k)_{k \in \mathbb{N}}$ in Example 5.6 converge to $(0, 0)$ at the rate of at least of $k^{-\frac{1}{2d^2-2}}$.

Example 5.8 Let

$$A := \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid (x_1 + 1)^4 + \sum_{i=2}^n x_i^4 - 1 \leq 0\}$$

$$B := \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid (x_1 - 2)^4 + \sum_{i=2}^n x_i^4 - 1 \leq 0\}.$$

Let $(a_k)_{k \in \mathbb{N}}, (b_k)_{k \in \mathbb{N}}$ be defined by

$$b_0 \in \mathbb{R}^2, \quad a_{k+1} := P_A b_k, \quad b_{k+1} := P_B a_{k+1}.$$

Then $\|a_k\| = O(\frac{1}{k^{\rho_n}})$, $\|b_k - (1, 0, \dots, 0)\| = O(\frac{1}{k^{\rho_n}})$ with $\rho_n = \frac{1}{\min \{7^n - 1, 2\beta(n-1)4^n - 2\}}$. \diamond

Proof. By the assumption, there exist unique points $a_0 \in \text{bd } A$, $b_0 \in \text{bd } B$ such that $1 = \text{dist}(A, B) = \|a_0 - b_0\|$. Clearly, $a_0 = (0, 0, \dots, 0)$ and $b_0 = (1, 0, \dots, 0)$. Then, the conclusion follows by applying Theorem 4.10 with $d = 4$. \square

6 Conclusion and Open Questions

In this paper, we studied the rate of convergence of the cyclic projection algorithm applied to finitely many basic semi-algebraic convex sets. We established an explicit convergence rate estimate which relies on the maximum degree of the polynomials that generate the basic semi-algebraic convex sets and the dimension of the underlying space. We also examined some concrete examples and compared the actual convergence rate with our estimate.

Although our estimate of the convergence rate works for cyclic projection algorithm with finitely many basic semialgebraic convex sets without any regularity condition, the limitation of our approach is that the estimated convergence rate behaves quite badly when the dimension n of the underlying space and the maximal degree d are large. Our results have suggested the following future research topics and open questions:

- The explicit examples (Examples 5.2, 5.4 and 5.6) show that, in general, our estimate of the convergence rate of the cyclic projection algorithm will not be tight. It would be interesting to see how one can sharpen the estimate obtained in this paper and get a tight estimate for the cyclic projection algorithm. In particular, finding the right exponent when each set is defined by convex quadratic functions would be a good starting point.
- Can we extend the approach here to analyze the convergence rate of the Douglas-Rachford algorithm? Almost nothing is known except for affine sets.

These will be our future research topics and will be examined later on.

Acknowledgments

The authors are grateful to Dr. Simeon Reich and the two anonymous referees and the editor for their pertinent and constructive comments. Jonathan Borwein, Guoyin Li and Liangjin Yao were partially supported by various Australian Research Council grants.

References

- [1] Y. Alber and S. Reich, An iterative method for solving a class of nonlinear operator equations in Banach spaces, *Panamerican Mathematical Journal*, vol. 4, pp. 39–54, 1994.
- [2] H. Attouch, J. Bolte, P. Redont and A. Soubeyran, Proximal alternating minimization and projection methods for nonconvex problems: an approach based on the Kurdyka-Łojasiewicz inequality, *Mathematics of Operation Research*, Vol. 35, no. 2, pp. 438–457, 2010.
- [3] H. Attouch, J. Bolte and B. Svaiter, Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods. *Mathematical Programming* Vol. 137, no. 1–2, Ser. A, pp. 91–129, 2013.
- [4] A. Auslender and M. Teboulle, *Asymptotic cones and functions in optimization and variational inequalities*. Springer Monographs in Mathematics. Springer-Verlag, New York, 2003.
- [5] A. Auslender, Existence of optimal solutions and duality results under weak conditions, *Mathematical Programming* Vol. 88, pp 45–59, 2000.
- [6] J.-B. Baillon, P. L. Combettes, and R. Cominetti, There is no variational characterization of the cycles in the method of periodic projections, *Journal of Functional Analysis*, vol. 262, pp. 400–408, 2012.
- [7] H.H. Bauschke, Projection algorithms: results and open problems, *Inherently parallel algorithms in feasibility and optimization and their applications*, pp. 11–22, 2001.
- [8] H.H. Bauschke and J.M. Borwein, On the convergence of von Neumanns alternating projection algorithm for two sets, *Set-Valued Analysis*, vol. 1, pp. 185–212, 1993.
- [9] H.H. Bauschke and J.M. Borwein, Dykstra’s alternating projection algorithm for two sets, *Journal of Approximation Theory*, vol. 79, pp. 418–443, 1994.
- [10] H.H. Bauschke and J.M. Borwein: On projection algorithms for solving convex feasibility problems, *SIAM Review*, vol. 38, pp. 367–426, 1996.
- [11] H.H. Bauschke, J.M. Borwein, and A.S. Lewis, The method of cyclic projections for closed convex sets in Hilbert space, *Recent developments in optimization theory and nonlinear analysis*, Contemporary Mathematic, pp. 1–38, 1997
- [12] H.H. Bauschke and P.L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, 2011.
- [13] H.H. Bauschke, D.R. Luke, H. M. Phan, and X. Wang, Restricted normal cones and the method of alternating projections, *Set-Valued and Variational Analysis*, in press, <http://arxiv.org/abs/1205.0318v1>.
- [14] E.G. Belousov, *Introduction to convex analysis and integer programming*, Izdat. Moskov. Univ., Moscow, 1977
- [15] E.G. Belousov, On types of Hausdorff discontinuity from above for convex closed mappings, *Optimization*, vol. 49, pp. 303–325, 2001.
- [16] E. G. Belousov and D. Klatte, A Frank-Wolfe type theorem for convex polynomial programs, *Computational Optimization and Applications*, vol. 22, pp. 37–48, 2002.
- [17] J. Bochnak, M. Coste and M.F. Roy, *Real algebraic geometry*, Springer-Verlag, Berlin, 1998.

- [18] J.M. Borwein and J.D. Vanderwerff, *Convex Functions*, Cambridge University Press, 2010.
- [19] L.M. Brègman, Finding the common point of convex sets by the method of successive projection. *Doklady Akademii Nauk SSSR*, vol. 162, pp. 487–490, 1965.
- [20] R.E. Bruck and S. Reich, Nonexpansive projections and resolvents of accretive operators in Banach spaces, *Houston Journal of Mathematics*, vol. 3, pp. 459–470, 1977.
- [21] F. Deutsch, *Best approximation in inner product spaces*, CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, 7, Springer-Verlag, New York, 2001.
- [22] J. Gwoździwicz, The Łojasiewicz exponent of an analytic function at an isolated zero, *Commentarii Mathematici Helvetici*, vo. 74, pp. 364–375, 1999.
- [23] L.G. Gubin; B.T. Polyak, and E.V. Raik, The method of projections for finding the common point of convex sets, *USSR Computational Mathematics and Mathematical Physics*, vol. 7, pp. 1–24, 1967.
- [24] J. W. HELTON, AND J. W. NIE, Semidefinite representation of convex sets, *Mathematical Programming*, Vol 122, Ser. A, pp. 21–64, 2010.
- [25] D. Klatte, Hoffman’s error bound for systems of convex inequalities. *Mathematical programming with data perturbations*, 185–199, Lecture Notes in Pure and Appl. Math., Dekker, New York, 1998.
- [26] J. Kollár, An effective Łojasiewicz inequality for real polynomials, *Periodica Mathematica Hungarica*, vol. 38, pp. 213–221, 1999.
- [27] E. Kopecká and S. Reich, A note on the von Neumann alternating projections algorithm, *Journal of Nonlinear and Convex Analysis*, vol. 5, pp. 379–386, 2004.
- [28] A.S. Lewis and J. S. Pang, Error bounds for convex inequality systems Generalized Convexity, Generalized Monotonicity J.P. Crouzeix, J.E. Martinez-Legaz and M. Volle (eds), pp. 75–110, 1998.
- [29] G. Li, On the asymptotic well behaved functions and global error bound for convex polynomials, *SIAM Journal on Optimization*, vol. 20, pp. 1923–1943, 2010.
- [30] G. Li and K.F. Ng, Error bounds of generalized D-gap functions for nonsmooth and nonmonotone variational inequality problems, *SIAM Journal on Optimization*, vol. 20, pp. 667–690, 2009.
- [31] G. Li and B.S. Mordukhovich, Hölder metric subregularity with applications to proximal point method, *SIAM Journal on Optimization*, vol. 22, pp. 1655–1684, 2012.
- [32] W. Li, Error bounds for piecewise convex quadratic programs and applications, *SIAM J. on Control and Optim.*, vol. 33, pp. 1510–1529, 1995.
- [33] X.D. Luo and Z.Q. Luo, Extension of Hoffman’s error bound to polynomial systems, *SIAM Journal on Optimization*, vol. 4, pp. 383–392, 1994.
- [34] Z.Q. Luo, J.S. Pang, and D. Ralph, *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press, Cambridge, 1996.
- [35] W.T. Obuchowska, On generalizations of the Frank-Wolfe theorem to convex and quasi-convex programmes, *Computational Optimization and Applications*, vol. 33, pp. 349–364, 2006.
- [36] J.S. Pang, Error bounds in mathematical programming. *Mathematical Programming*, vol. 79, pp. 299–332, 1997.
- [37] R.R. Phelps, *Convex Functions, Monotone Operators and Differentiability*, 2nd Edition, Springer-Verlag, 1993.
- [38] S.M. Robinson, *Regularity and Stability for Convex Multivalued Functions*, *Mathematics of Operations Research*, vol. 1, pp. 130–143, 1976.
- [39] R.T. Rockafellar and R.J-B Wets, *Variational Analysis*, 3rd Printing, Springer-Verlag, 2009.
- [40] V.M. Shironin, *On Hausdorff continuity of convex and convex polynomial mappings*, in *Mathematical Optimization: Questions of Solvability and Stability*, E.G. Belousov and B. Bank, eds., Moscow University Publishing, Moscow, 1986.
- [41] J. von Neumann, *Functional Operators, Vol.II*, Princeton University Press, 1950.
- [42] C. Zălinescu, *Convex Analysis in General Vector Spaces*, World Scientific Publishing, 2002.