Concrete convergence rates for common fixed point problems under Karamata regularity

Tianxiang Liu* Bruno F. Lourenço†

July 18, 2024

Abstract

We introduce the notion of Karamata regular operators, which is a notion of regularity that is suitable for obtaining concrete convergence rates for common fixed point problems. This provides a broad framework that includes, but goes beyond, Hölderian error bounds and Hölder regular operators. By *concrete*, we mean that the rates we obtain are explicitly expressed in terms of a function of the iteration number k instead, of say, a function of the iterate x^k. While it is well-known that under Hölderian-like assumptions many algorithms converge linearly/sublinearly (depending on the exponent), little it is known when the underlying problem data does not satisfy Hölderian assumptions, which may happen if a problem involves exponentials and logarithms. Our main innovation is the usage of the theory of regularly varying functions which we showcase by obtaining concrete convergence rates for quasi-cylic algorithms in non-Hölderian settings. This includes certain rates that are neither sublinear nor linear but sit somewhere in-between, including a case where the rate is expressed via the Lambert W function. Finally, we connect our discussion to o-minimal geometry and show that definable operators in any o-minimal structure are always Karamata regular.

Keywords: common fixed point problem; concrete rates; Karamata regularity; quasi-cyclic algorithm, regular variation; Karamata theory, o-minimal structure.

1 Introduction

In this paper, we consider the following common fixed point problem:

$$\text{find} \quad x \in F := \bigcap_{i=1}^{m} \operatorname{Fix} T_i, \tag{1.1}$$

where each T_i $(i = 1, \ldots, m)$ is an α*-averaged* $(\alpha \in (0,1))$ operator (see definition in Section 2) on a finite dimensional real vector space $\mathcal{E}$. We assume that F is nonempty and $F \neq \mathcal{E}$. Many interesting problems can be reformulated as in (1.1) and two notable examples are convex feasibility problems and certain variational inequality problems. There are many algorithms for solving (1.1) and one particularly broad class of method correspond to the family of *quasi-cyclic algorithm*, considered in [4, Theorem 6.1] and further analyzed in [10].

Our main goal in this paper is to obtain *concrete* convergence rates for quasi-cyclic algorithms for (1.1). Here, we emphasize that, by *concrete* we mean that the convergence rate should be given

*School of Computing, Tokyo Institute of Technology, Japan. (liu@c.titech.ac.jp)

†Department of Fundamental Statistical Mathematics, Institute of Statistical Mathematics, Japan. This author was supported partly by the JSPS Grant-in-Aid for Early-Career Scientists 23K16844. (bruno@ism.ac.jp)

in terms of an explicit function *depending only on the iterate number*. More precisely, if we denote the k-th iterate of an algorithm by x^k our desired convergence rates should have the form

$$\mathrm{dist}(x^k,\, F) \le R(k),$$

where R is some function of k. We recall that if $R(k)$ is of the form c^{-k} for some $c > 1$ the convergence rate is often said to be *linear*. If $R(k)$ is of the form k^{-r} for some $r > 0$ the rate is said to be *sublinear*.

In order to obtain a concrete converge rate (i.e., to obtain $R(k)$) typically some assumptions on the operators T_i and their fixed point sets are required. As far as we know, the only way to obtain $R(k)$ so far is to make use of certain *Hölderian assumptions* such as assuming that the fixed point sets have a *Hölderian error bound* and the operators are *Hölder regular* (see [10] or Remark 3.3 below).

A Hölderian assumption for an operator typically takes the form of asking that over a bounded set, some *power* of the residual $\|x - Tx\|$ should be an upper bound to the true distance between x and the fixed point set of T. That is, given a bounded set B, there should exist $\rho \in (0,\, 1]$ and a constant $\kappa > 0$, such that $\mathrm{dist}(x,\, \mathrm{Fix}\, T) \le \kappa \|x - Tx\|^\rho$ holds, for $x \in B$. Or this assumption can be taken jointly, by requiring that $\mathrm{dist}(x,\, F) \le \kappa(\max_{i=1}^m \|x - T_i x\|)^\rho$ holds over B. In particular, this recovers the notion of Hölderian error bound when each T_i is a projection onto a given convex set C_i.

Under certain Hölderian assumptions, it was shown in [10, Theorem 3.1] that the quasi-cyclic algorithm converges at least linearly or sublinearly with a rate whose asymptotic behavior is controlled by the powers appearing in the assumed Hölderian conditions.

Here, however, we will consider the problem of obtaining concrete convergence rates when the underlying problem *does not* necessarily satisfy Hölderian assumptions. This is motivated by the fact that there are certain problems for which their regularity properties are better expressed under more general conditions. For example, in [23], it was shown that certain intersections of the exponential cone never admit Hölderian error bounds [23, Example 4.20]. Or, even when a Hölderian error bound holds it may be the case that a tighter error bound can be obtained by making use of a non-Hölderian error bound as in [23, Remark 4.14 (a)]. Other examples were found in [24, Section 5.1] and in the study of error bounds for log-determinant cones [22].

A situation where one *can* actually expect Hölderian assumptions to hold is when the problem data is semialgebraic, thanks to results such as the Łojasiewicz inequality as used in [10, Proposition 4.1]. Unfortunately, whenever the problem data is related to exponentials and logarithms (as it is in [23] and [22]) we cannot typically ensure that the underlying operators satisfy Hölderian assumptions. This boils down to the fact that functions involving exponentials and logarithms are typically not semialgebraic.

As we move away from Hölderian assumptions, obtaining concrete convergence rates becomes quite challenging. Here we should remark that, as discussed extensively in [24, Section 7.1], certain important previous works based on the Kurdya-Łojasiewicz property have results concerning convergence rates for certain algorithm under general desingularizing functions (e.g., [8, Theorem 24] and [9, Theorem 14]), but these results do not lead to concrete rates since they are expressed in terms of the iterates x^k instead of just as functions of k. Under KL theory, the cases where one obtains concrete convergence rates are typically restricted to the situation where the desingularizing function is a power function which implies the existence of the so-called *KL exponents*, see more details in [21, 33]. As such, the case of KL exponents can be seen as another kind of Hölder-type assumption. This is particularly more pronounced in the convex case, in view of results such as [9, Theorem 5] as used, say, in [23, Proposition 4.21] to connect a Hölderian error bound to the KL exponent of a certain function and vice-versa.

Similarly, in the analysis of set-valued mappings and fixed points of operators, several general notions of regularity have been proposed. For example, Ioffe suggested the usage of gauge functions[1] to express generalized versions of metric subregularity and other notions, see [20, Section 2]. This was also used in subsequent works, e.g., [29, 25]. But, again, a challenge that seems to remain is getting concrete convergence rates for algorithms when the considered regularity notion is no-longer Hölderian.

Part of the difficulty is that no matter how one frames a certain generalized notion of regularity, say through some function μ satisfying some key inequality, if we wish to obtain a convergence rate for some algorithm, we typically need to sum the effect of μ over the course of the algorithm and solve a recurrence inequality in k (the iteration number) in order to get a rate. Solving general recurrence relations is a notoriously ad-hoc endeavour and this also adds to the difficulty of reasoning about rates beyond the Hölderian case.

To the best of our knowledge, the first work to show examples of concrete rates under more general regularity conditions in a systematic way was [24], under the framework of *consistent error bound functions*. In particular, it was shown in [24, Proposition 6.9] that when the alternating projections algorithm is applied to the exponential cone and a certain subspace, it may happen that the rate is given by a function $R(k)$ that is proportional to $\frac{1}{\ln(k)}$. Also, for another choice of subspace the rate is "almost linear" in the sense that is faster than any sublinear rate but it may be slower than any linear rate. A key point in [24] was the notion of *regularly varying functions* [30, 6], a relatively known tool in probability theory but virtually unused in optimization theory. Again, to the best of our knowledge, [24] was the first work in optimization to make systematic use of regular variation to study convergence rates.

The analysis done in [24] concerns only convex feasibility problems and one of our goals here is to extend it to the more general problem class expressed in (1.1). For that, we will introduce the notion of *Karamata regular operator*, which is suitable for working with regular variation techniques. We will also deepen our usage of regular variation toolbox and address certain inelegancies and superfluous assumptions in [24]. Our results will also lead to sharper rates than the ones obtained in [24]. In particular, for certain convex feasibility problems our rates here will be better than the ones that could be obtained by invoking the results in [24].

The basic idea here is to obtain rates via a series of functional transformations performed onto the regularity function that appears in the definition of Karamata regularity. Then, regular variation will be useful because it will help us to analyze the asymptotic behaviour as we perform those functional transformations without the need of actually computing some of them. In this way, we will avoid, for instance, the need to evaluate certain hard integrals.

Under Karamata regularity, we will show how concrete convergence rates can be obtained through two different techniques. The first only requires knowledge of the index of regular variation, see Theorem 3.10. Here we remark that the index of regular variation is a quantity that can be easily obtained from a simple limit computation, provided that we have the underlying regularity function, see Definition 2.2. The second technique requires more work to be applied but leads to tighter rates as we will see in Theorem 3.14 and Section 4.

At the very end, we will connect the notion of Karamata regularity to definable sets in o-minimal structures and will show that operators that correspond to definable functions can always be taken to be Karamata regular and are, therefore, under the scope of the techniques described in this paper.

Besides the specific goal of obtaining convergence rates for algorithms for (1.1) and showing the applicability of Karamata regularity, we hope that this work will also inspire others to investigate

[1] One must be careful that gauge functions here are not the same gauge functions considered in, say, [28, Section 15].

other applications of regular variation in optimization, especially when the problem data is not semialgebraic and/or involves exponentials and logarithms. With this goal in mind we tried to present a survey-like overview of regular variation in Section 2 with a view towards optimization applications. Also, with the exception of a single result in [5], we confine all other references to results on regular variation to the classical textbook by Bingham, Goldie and Teugels [6].

1.1 Our contributions

Our contributions are as follows.

- We introduce the notion of (joint) Karamata regularity for operators in Definition 3.1, which is a regularity notion suitable for using tools from regular variation. We then discuss its connections with previous considered notions and prove a calculus rule in Proposition 3.4.

- We prove an abstract convergence rate result for quasi-cyclic algorithms for common fixed point problems in Theorem 3.8. Admittedly, applying directly Theorem 3.8 is hard because it requires *inverting* an already a relatively complicated integral. However, by making use of regular variation, we show how to bypass the evaluation of the complicated expression in Theorem 3.8. This is done through either the index of regular variation (Theorem 3.10) of the regularity function associated to Karamata regularity or through a sharper result that requires a bit more computation in Theorem 3.14. Several application examples are given in Section 4, with a focus on cases having non-Hölderian behavior.

- We explore the class of Karamata regular operators, and show that quasi-nonexpansive operators defined on an o-minimal structure can always be taken to be jointly Karamata regular, see Theorem 5.5. As this includes the case of certain large o-minimal structures containing the graph of the exponential function, this shows that theory developed in this paper is applicable quite broadly. We also explore certain consequences of Theorem 5.5 and show, for example, that the consistent error bound functions considered in [24] can also be taken to be regularly varying, provided that the problem data is definable.

This paper is organized as follows. In Section 2, we introduce the notation and discuss the necessary notion from the theory of regular variation. In Section 3, we introduce the notion of joint Karamata regularity and establish an abstract convergence result for quasi-cyclic algorithms. With the aid of regular variation, we further study the asymptotic properties of the convergence rates obtained. Later in Section 4, we establish explicit convergence rates under a number of scenarios. Finally, the class of Karamata regular operators is discussed in Section 5 in the context of o-minimal structures.

2 Preliminaries and basic notions from Karamata theory

Let $\mathcal{E}$ be a finite-dimensional Euclidean space equipped with an inner product $\langle\cdot,\cdot\rangle$ and a corresponding norm $\|\cdot\|$. We denote the ball of radius r centered in the origin by $\mathbb{B}_r := \{x \in \mathcal{E} \mid \|x\| \le r\}$. Given closed set $C \subseteq \mathcal{E}$ and $x \in \mathcal{E}$, we denote the projection of x onto C and the distance of x to C by $P_C(x)$ and $\operatorname{dist}(x, C)$, respectively.

We say an operator T is α*-averaged* ($\alpha \in (0,1)$) if there exists a nonexpansive operator R such that $T = (1-\alpha)I + \alpha R$, where I is the identity operator. In particular, since the T_i's in (1.1) are α-averaged, each T_i is nonexpansive and $\operatorname{Fix} T_i$ is convex, thanks to [3, Remark 4.24 and Proposition 4.13]. Moreover, the following property of α-averaged operators follows from [3, Proposition 4.25].

Lemma 2.1 (α-averaged operator). *Let T be an α-averaged ($\alpha \in (0,1)$) operator on $\mathcal{E}$. Then it satisfies*

$$\|T(x) - T(y)\|^2 + \frac{1-\alpha}{\alpha}\big\|(I-T)(x) - (I-T)(y)\big\|^2 \le \|x-y\|^2, \quad \forall\, x,\, y \in \mathcal{E}.$$

Next, we introduce some notation and preliminaries on the theory of regular variation, which we will use to conduct convergence analysis of algorithms for solving (1.1). More details on regular variation can be found in [30, 6]. We start with the notion of regularly varying functions.

Definition 2.2 (Regularly varying functions). *A function $f : [a,\, \infty) \to (0,\, \infty)$ $(a > 0)$ is said to be* regularly varying at infinity *if it is (Lebesgue) measurable and there exists a real number ρ such that*

$$\lim_{x\to\infty} \frac{f(\lambda x)}{f(x)} = \lambda^\rho, \quad \lambda > 0. \tag{2.1}$$

In this case, we write $f \in \mathrm{RV}$. Similarly, a measurable function $f : (0,\, a] \to (0,\, \infty)$ is said to be regularly varying at 0 if

$$\lim_{x\to 0_+} \frac{f(\lambda x)}{f(x)} = \lambda^\rho, \quad \lambda > 0, \tag{2.2}$$

in which case we write $f \in \mathrm{RV}^0$. The ρ in (2.1) and (2.2) is called the index *of regular variation.*

If the limit on the left hand side of (2.1) is 0, 1 and $+\infty$ for λ in $(0,\, 1)$, $\{1\}$ and $(1,\, \infty)$, respectively, then f is said to be a function of rapid variation of index ∞ *and we write $f \in \mathrm{RV}_\infty$. If $1/f \in \mathrm{RV}_\infty$, we say that f is a function of* rapid variation of index $-\infty$ *and write $f \in \mathrm{RV}_{-\infty}$. $\mathrm{RV}^0_{-\infty}$ and RV^0_∞ are defined analogously.*

We denote by RV^0_ρ the set of regularly varying functions at zero with index ρ. RV_ρ is defined analogously. The functions in RV^0_0, RV_0 are said to be *slowly varying*. Regular variation at 0 and at ∞ are naturally linked and we will use the following relation several times throughout this paper:

$$f \in \mathrm{RV}^0_\rho \quad \Longleftrightarrow \quad f(1/\cdot) \in \mathrm{RV}_{-\rho}, \tag{2.3}$$

where $\rho \in \mathbb{R} \cup \{-\infty,\, \infty\}$.

We also need some auxiliary definitions. We say that a nonnegative function f defined on a subset C of the real line is *locally bounded* if its restriction to each compact subset $K \subseteq C$ is bounded. If the restriction of f to each compact $K \subseteq C$ satisfies $\inf_{t\in K} f(t) > 0$, then we say that f is *locally bounded away from zero*. Finally, we say that f *locally integrable*, if $\int_K f$ is finite for each compact $K \subseteq C$.

An important fact is that we can always adjust the domain in order to ensure local boundedness. More precisely, if $f : [a,\, \infty) \to (0,\, \infty)$ belongs to RV, then there exists $b \ge a$ such that the restriction of f and $1/f$ to $[b,\, \infty)$ are both locally bounded, see [6, Corollary 1.4.2].

Analogously, if $f : (0,\, a] \to (0,\, \infty)$ belongs to RV^0, then $1/f(1/\cdot) : [1/a,\, \infty) \to (0,\, \infty)$ belongs to RV by (2.3). Then there exists some $b \in (0,\, a]$ (hence $1/b \ge 1/a$) such that the restriction of $1/f(1/\cdot)$ to $[1/b,\, \infty)$ is locally bounded. This implies that f is locally bounded away from zero over $(0,\, b]$.

We note that if f is a positive function on C and it is monotone (either nondecreasing or nonincreasing) then it is both *locally bounded* and *locally bounded away from zero*. For the sake of preciseness, we emphasize that f is nondecreasing (resp. increasing) if $f(t_1) \le f(t_2)$ (resp. $f(t_1) < f(t_2)$) holds when $t_1,\, t_2 \in C$ satisfies $t_1 < t_2$. Nonincreasing/decreasing are defined analogously.

Calculus rules For $f_1 \in \mathrm{RV}_{\rho_1}$, $f_2 \in \mathrm{RV}_{\rho_2}$ with $\rho_1, \rho_2, \alpha \in \mathbb{R}$ we have the following calculus rules, see [6, Proposition 1.5.7]:

$$f_1 f_2 \in \mathrm{RV}_{\rho_1+\rho_2}, \quad f_1 + f_2 \in \mathrm{RV}_{\max\{\rho_1, \rho_2\}}, \quad f_1^\alpha \in \mathrm{RV}_{\alpha\rho_1}, \quad f_1 \circ f_2 \in \mathrm{RV}_{\rho_1\rho_2}, \tag{2.4}$$

where the last relation requires the additional hypothesis that $f_2(x) \to \infty$ as $x \to \infty$. From (2.3) and (2.4) we see that if $f_1 \in \mathrm{RV}^0_{\rho_1}$, $f_2 \in \mathrm{RV}^0_{\rho_2}$, then:

$$f_1 f_2 \in \mathrm{RV}^0_{\rho_1+\rho_2}, \quad f_1 + f_2 \in \mathrm{RV}^0_{\min\{\rho_1, \rho_2\}}, \quad f_1^\alpha \in \mathrm{RV}^0_{\alpha\rho_1}, \quad f_1 \circ f_2 \in \mathrm{RV}^0_{\rho_1\rho_2}, \tag{2.5}$$

where the last relation requires the additional hypothesis that $f_2(x) \to 0$ as $x \to 0_+$.

Remark 2.3 (About the function domain and image)**.** *The literature on regular variation treats the domain of functions in a somewhat loose fashion. If $f : [a, \infty) \to (0, \infty)$ is in* RV*, we can freely restrict f to $[c, \infty)$ ($c > a$) or extend f to $[b, \infty)$ ($b < a$) by letting f take arbitrary positive values on $[b, a)$. More extremely, we can change the value of f in a single bounded interval $[b, c]$ and none of these operations would change the asymptotic properties of f at infinity nor the index of regular variation of f. So the calculus rules in* (2.4) *(and much of this paper, in fact) should be seen under this light: while $f_1, f_2 \in$* RV *might have different domains of definition, we can restrict/extend their domains until $f_1 + f_2$, $f_1 f_2$, $f_1 \circ f_2$ are well-defined. A similar comment applies to regular variation at* 0 *so that if $f \in \mathrm{RV}^0$ is defined over $(0, a]$, we can arbitrarily change the value of f in a single interval of the form $[b, c]$ with $b > 0$ and $c \in (b, \infty) \cup \{\infty\}$.*

Due to aforementioned flexibility of restricting the function domain, we also treat the image of functions in a loose fashion. For a function f whose image falls out of $(0, \infty)$, we still write $f \in$ RV *or $f \in \mathrm{RV}^0$, if there exists some $a > 0$ such that the image of the restriction $f|_{[a, \infty)}$ or $f|_{(0, a]}$ is contained in $(0, \infty)$, respectively. Again, this is because only the asymptotic property of f at infinity or* 0 *matters.*

Potter's bounds A great deal of information on the asymptotic behavior of a function can be extracted from its index of regular variation. A result known as *Potter bounds* states that if $f \in \mathrm{RV}_\rho$, then for every $A > 1$, $\epsilon > 0$, there exists a constant M such that $x \geq M$, $y \geq M$ implies

$$\frac{f(x)}{f(y)} \leq A \max\left\{ \left(\frac{x}{y}\right)^{\rho-\epsilon}, \left(\frac{x}{y}\right)^{\rho+\epsilon} \right\}, \tag{2.6}$$

see [6, Theorem 1.5.6]. Now, if $f \in \mathrm{RV}^0_\rho$, then $\hat{f}$ such that $\hat{f}(t) \coloneqq 1/f(1/t)$ belongs to RV_ρ, by (2.3) and (2.4). Applying (2.6), we see that for any $A > 1$, $\epsilon > 0$, there exists a constant M such that

$$\frac{f(t)}{f(s)} \leq A \max\left\{ \left(\frac{t}{s}\right)^{\rho-\epsilon}, \left(\frac{t}{s}\right)^{\rho+\epsilon} \right\}, \tag{2.7}$$

whenever $t \leq M$, $s \leq M$. We note that taking $\epsilon = |\rho|/2$ in (2.6), fixing x (if $\rho > 0$) or y (if $\rho < 0$) and taking limits, leads to the following conclusions:

$$f \in \mathrm{RV}_\rho \text{ and } \rho > 0 \quad \Rightarrow \quad \lim_{x\to\infty} f(x) = +\infty, \tag{2.8}$$

$$f \in \mathrm{RV}_\rho \text{ and } \rho < 0 \quad \Rightarrow \quad \lim_{x\to\infty} f(x) = 0, \tag{2.9}$$

see also [6, Proposition 1.3.6, item (v)]. Similarly, we have

$$f \in \mathrm{RV}^0_\rho \text{ and } \rho > 0 \quad \Rightarrow \quad \lim_{x\to 0_+} f(x) = 0, \tag{2.10}$$

$$f \in \mathrm{RV}_\rho^0 \text{ and } \rho < 0 \quad \Rightarrow \quad \lim_{x \to 0_+} f(x) = +\infty. \tag{2.11}$$

There is also an analogous result for rapidly varying function. Bingham, Goldie and Omey proved that if $f \in \mathrm{RV}_{-\infty}$, then given any $r > 0$ there exists a constant $M > 0$ such that $x \geq M$ implies

$$f(x) \leq x^{-r}, \tag{2.12}$$

see [5, Lemma 2.2], in particular $f(x) \to 0$ as $x \to \infty$. Therefore, if $f \in \mathrm{RV}_\infty$ (i.e., $1/f \in \mathrm{RV}_{-\infty}$) then, there exists $M > 0$ such that $x \geq M$ implies

$$x^r \leq f(x), \tag{2.13}$$

in particular $f(x) \to \infty$ as $x \to \infty$.

Generalized inverses Let $f : [a, \infty) \to \mathbb{R}$ be such that $f(x)$ tends to ∞ as $x \to \infty$. Then, we define its "arrow" generalized inverse as

$$f^{\leftarrow}(y) \coloneqq \inf\{x \in [a, \infty) \mid f(x) > y\}, \tag{2.14}$$

see [6, equation (1.5.10)]. The function $f^{\leftarrow}$ is well-defined over $(0, \infty)$ and nondecreasing.

Similarly, for a function $f : (0, a] \to \mathbb{R}$ with $\lim_{x \to 0_+} f(x) = 0$, we define its "minus" generalized inverse as

$$f^{-}(y) \coloneqq \sup\{x \in (0, a] \mid f(x) < y\}, \tag{2.15}$$

which is well-defined over $(0, \infty)$ and nondecreasing.

The two generalized inverses are related as follows. Suppose that $f : (0, a] \to (0, \infty)$ satisfies $\lim_{x \to 0_+} f(x) = 0$. Let $g \coloneqq 1/f(1/\cdot)$. Then $g(x)$ tends to ∞ as $x \to \infty$. Moreover,

$$\begin{aligned} f^{-}(y) = \sup\{x \in (0, a] \mid f(x) < y\} &= \frac{1}{\inf\{u \in [1/a, \infty) \mid f(1/u) < y\}} \\ = \frac{1}{\inf\{u \in [1/a, \infty) \mid g(u) > 1/y\}} &= \frac{1}{g^{\leftarrow}(1/y)}. \end{aligned} \tag{2.16}$$

In the spirit of Remark 2.3, we will observe that under local boundedness, the value of the generalized inverse $f^{\leftarrow}(y)$ does not depend on a for sufficiently large y. Also, it may happen that a non-monotone function $f(y)$ is increasing and continuous for sufficiently large y. Nevertheless, this will still be enough to conclude that the generalized inverse will eventually coincide with the usual inverse. For the sake of preciseness, in what follows, given a function $f : C \to \mathbb{R}$ and $S \subseteq \mathbb{R}$ we say that f^{-1} *is well-defined over* S if for every $y \in S$ there exists a unique $x \in C$ such that $f(x) = y$ holds. In this case, for $y \in S$, we can define $f^{-1}(y) \coloneqq x$ without ambiguity.

Proposition 2.4. *Suppose that $f : [a, \infty) \to \mathbb{R}$ is locally bounded and $f(x) \to \infty$ as $x \to \infty$. Let $b \geq a$, $\widehat{f} \coloneqq f|_{[b, \infty)}$ and $M \coloneqq \sup_{x \in [a, b]} f(x)$. Then $f^{\leftarrow}(y) = \widehat{f}^{\leftarrow}(y)$ holds for $y \geq M$. In addition, if f is continuous and increasing on $[b, \infty)$, then f^{-1} is well-defined over (M, ∞) and $f^{-1}(y) = f^{\leftarrow}(y)$ holds for $y > M$.*

Proof. We first observe that M is finite because f is locally bounded. Also, for any $y \geq M$, if $f(x) > y$ holds we must have $x > b$, which together with $\widehat{f} \coloneqq f|_{[b, \infty)}$ implies the inclusion and the equality respectively:

$$\{x \in [a, \infty) \mid f(x) > y\} \subseteq \{x \in [b, \infty) \mid f(x) > y\} = \{x \in [b, \infty) \mid \widehat{f}(x) > y\}.$$

Since $b \geq a$, the converse inclusion $\{x \in [b, \infty) \mid f(x) > y\} \subseteq \{x \in [a, \infty) \mid f(x) > y\}$ holds, which together with the above relation implies that $\{x \in [a, \infty) \mid f(x) > y\} = \{x \in [b, \infty) \mid \widehat{f}(x) > y\}$. In view of the definition of the arrow inverse, we then have $f^{\leftarrow}(y) = \widehat{f}^{\leftarrow}(y)$. This proves the first half.

Now, we show the remaining half, where we assume that f is continuous and increasing on $[b, \infty)$. If x_1, x_2 and $y > M$ are such that $f(x_1) = y = f(x_2)$ holds, then, by the definition of M, we must have $x_1, x_2 \in (b, \infty)$. Since the restriction of f to $[b, \infty)$ is increasing, this implies that $x_1 = x_2$. Additionally, since f is continuous and goes to ∞ as $x \to \infty$, for every $y > M$ there exists at least one x satisfying $f(x) = y$, which is a consequence of $M \geq f(b)$ and the intermediate value theorem. We conclude that f^{-1} is well-defined over (M, ∞).

Finally, let $y > M$ be arbitrary. If $f(x) > y$ holds, then $x > b$ holds. Also, since $f(f^{-1}(y)) = y > M$, we have $f^{-1}(y) > b$ too. So $f(x) > y = f(f^{-1}(y))$ implies $x > f^{-1}(y)$, since f is increasing on (b, ∞). Therefore, we have the inclusion

$$\{x \in [a, \infty) \mid f(x) > y\} \subseteq \{x \in [b, \infty) \mid x > f^{-1}(y)\}.$$

However, if $x \in [b, \infty)$ and $x > f^{-1}(y)$ holds, since f is increasing on $[b, \infty)$, we have $f(x) > f(f^{-1}(y)) = y$. Therefore, both sets coincide and we have

$$f^{\leftarrow}(y) = \inf\{x \in [a, \infty) \mid f(x) > y\} = \inf\{x \in [b, \infty) \mid x > f^{-1}(y)\} = f^{-1}(y).$$

This completes the proof. $\square$

We observe that in the second half of the proof of Proposition 2.4, if $a = b$ holds, then f^{-1} is also well-defined at $f(b)$ and a direct computation shows that $f^{\leftarrow}(f(b)) = b = f^{-1}(f(b))$. In particular, if $f : [a, \infty) \to (0, \infty)$ is continuous, increasing and satisfies $f(x) \to \infty$ as $x \to \infty$, then $f^{-1} = f^{\leftarrow}$ holds over $[f(a), \infty)$. Similarly, if $f : (0, a] \to (0, \infty)$ is continuous, increasing and satisfies $f(x) \to 0$ as $x \to 0_+$, then $f^{-1} = f^{-}$ holds over $(0, f(a)]$.

Moving on, an important result is that if f is locally bounded on $[a, \infty)$ then:

$$f \in \mathrm{RV}_\rho,\ \rho > 0 \quad \Rightarrow \quad f^{\leftarrow} \in \mathrm{RV}_{1/\rho}, \tag{2.17}$$

see [6, Theorem 1.5.12] and this footnote[2].

In order to describe the behavior of the arrow inverse when the index is 0, we need an extra definition.

Definition 2.5. *A positive measurable function f belongs to the class of Karamata rapidly varying functions (denoted by KRV_∞) if and only if f can be restricted or extended to the interval $[1, \infty)$ in such a way that*

$$f(x) = \exp\left\{z(x) + \eta(x) + \int_1^x \xi(t)\frac{dt}{t}\right\}, \quad x \geq 1,$$

where z, η, ξ are measurable functions such that z is nondecreasing, $\eta(x) \to 0$ and $\xi(x) \to \infty$ as $x \to \infty$.

The definition of KRV_∞ in [6, Section 2.4] uses the so-called *Karamata indices*, but thanks to [6, Theorem 2.4.5] we can equivalently use Definition 2.5. We have the following implications:

$$f \in \mathrm{KRV}_\infty \quad \Rightarrow \quad f \in \mathrm{RV}_\infty, \tag{2.18}$$

[2]For an example of what can go awry if f is not locally bounded, let $f : [1, \infty) \to (0, \infty)$ be such that $f(t) \coloneqq t$, for $t \geq 2$ and $f(t) \coloneqq 1/(2 - t)$ for $t \in [1, 2)$. We have $f \in \mathrm{RV}_1$, but $f^{\leftarrow}(x) = \inf\{y \in [1, \infty) \mid f(y) > x\}$ does not go to ∞ as $x \to \infty$, so, in particular, $f^{\leftarrow} \notin \mathrm{RV}_1$ (in view of (2.8)). Still, we can always adjust the domain of a regularly varying function in order to ensure local boundedness as discussed previously. In this case, it is enough to restrict f to $[2, \infty)$.

$$f \in \mathrm{RV}_\infty \text{ and } f \text{ is nondecreasing} \quad \Rightarrow \quad f \in \mathrm{KRV}_\infty, \tag{2.19}$$

$$f \in \mathrm{KRV}_\infty,\ g(x) := x^\alpha,\ \alpha \in \mathbb{R} \quad \Rightarrow \quad gf \in \mathrm{KRV}_\infty, \tag{2.20}$$

where (2.18) and (2.19) follows from [6, Proposition 2.4.4, item(iv)]. We now check (2.20). Writing $g(x)$ as $\exp\left\{\int_1^x \alpha/t dt\right\}$, we see that gf admits a representation as in Definition 2.5 where $\xi(t)+\alpha$ appears instead of just $\xi(t)$. Since $\xi(t)+\alpha$ still goes to ∞ as $x \to \infty$, this shows that $gf \in \mathrm{KRV}_\infty$.

With that we have the following results. If f is locally bounded and $f(x)$ goes to ∞ as $x \to \infty$, then

$$f \in \mathrm{RV}_0 \quad \Rightarrow \quad f^{\leftarrow} \in \mathrm{KRV}_\infty, \tag{2.21}$$

$$f \in \mathrm{RV}_\infty \quad \Rightarrow \quad f^{\leftarrow} \in \mathrm{RV}_0, \tag{2.22}$$

see [6, Theorem 2.4.7].

For regular varying functions at 0 it will be more convenient to use the minus inverse. Suppose that $f : (0, a] \to (0, \infty) \in \mathrm{RV}^0_\rho$ is such that $f(x)$ goes to 0 as $x \to 0_+$, $\rho \geq 0$ and f is locally bounded away from zero. Then, $g := 1/f(1/\cdot)$ belongs to RV_ρ, $g(t)$ goes to ∞ as $t \to \infty$ and g is locally bounded on its domain $[1/a, \infty)$. This together with (2.16), (2.17) and (2.21) allows us to conclude that if f is bounded away from zero, then

$$f \in \mathrm{RV}^0_\rho,\ \rho > 0 \quad \Rightarrow \quad f^- \in \mathrm{RV}^0_{1/\rho}, \tag{2.23}$$

and in case of $\rho = 0$ we have

$$f \in \mathrm{RV}^0_0 \quad \Rightarrow \quad f^- \in \mathrm{RV}^0_\infty. \tag{2.24}$$

Asymptotic equivalence We say that two functions $f : (0, a] \to (0, \infty)$ and $g : (0, a] \to (0, \infty)$ are *asymptotically equivalent up to a constant* if there is a constant $\mu > 0$ such that

$$f(t) - \mu g(t) = o(g(t)), \text{ as } t \to 0_+. \tag{2.25}$$

In this case, we write $f(t) \overset{c}{\sim} g(t)$ *as* $t \to 0_+$, or we may simply write $f \overset{c}{\sim} g$ if it is clear from context what is meant. If $\mu = 1$, we say that f and g are *asymptotically equivalent* and write $f(t) \sim g(t)$ *as* $t \to 0_+$ or $f \sim g$. Then, for measurable functions f and g we have the following implication:

$$f \overset{c}{\sim} g, \quad f \in \mathrm{RV}^0_\rho \quad \Rightarrow \quad g \in \mathrm{RV}^0_\rho. \tag{2.26}$$

Indeed, by the definition one has $\lim_{x\to 0_+} \frac{f(x)}{g(x)} = \mu + \lim_{x\to 0_+} \frac{f(x)-\mu g(x)}{g(x)} = \mu$ and therefore,

$$\lim_{x\to 0_+} \frac{g(\lambda x)}{g(x)} = \lim_{x\to 0_+} \frac{g(\lambda x)}{f(\lambda x)} \frac{f(\lambda x)}{f(x)} \frac{f(x)}{g(x)} = \lim_{x\to 0_+} \frac{g(\lambda x)}{f(\lambda x)} \cdot \lim_{x\to 0_+} \frac{f(\lambda x)}{f(x)} \cdot \lim_{x\to 0_+} \frac{f(x)}{g(x)} = \frac{1}{\mu}\lambda^\rho \mu = \lambda^\rho.$$

For multiple functions, we called them *pairwise asymptotically equivalent up to a constant* (resp. *pairwise asymptotically equivalent*) if any two functions among them are *asymptotically equivalent up to a constant* (resp. *asymptotically equivalent*).

The notion in (2.25) corresponds to asymptotic equivalence at 0_+, but similarly, we can define asymptotic equivalence at infinity, e.g., f, g are *asymptotically equivalent up to a constant* (at infinity) if $f(t) - \mu g(t) = o(g(t))$, as $t \to +\infty$. For simplicity, we will use the same notation as it will be clear from context if asymptotic equivalence is meant at 0_+ or at ∞. Similarly, for measurable functions f and g we have

$$f \overset{c}{\sim} g, \quad f \in \mathrm{RV}_\rho \quad \Rightarrow \quad g \in \mathrm{RV}_\rho. \tag{2.27}$$

3 Karamata regularity and convergence rates

In this section, we will explore the convergence of a family of algorithms for the common fixed point problem (1.1). Naturally, this will be done under certain assumptions on the operators T_i. We start by introducing the following definition.

Definition 3.1 (Karamata regularity)**.** *Let $L_i : \mathcal{E} \to \mathcal{E}$ $(i = 1, \dots, n)$ be operators with $C \coloneqq \bigcap_{i=1}^{n} \operatorname{Fix} L_i \neq \emptyset$ and $B \subset \mathcal{E}$ a given bounded set. The L_i are said to be* jointly Karamata regular (JKR) over B *if there exists a function $\psi_B : \mathbb{R}_+ \to \mathbb{R}_+$ such that the following properties are satisfied.*

(i) The following error bound condition holds:

$$\operatorname{dist}(x, C) \le \psi_B\Big(\max_{1\le i\le n} \|x - L_i(x)\|\Big), \quad \forall\, x \in B. \tag{3.1}$$

(ii) ψ_B is nondecreasing and satisfies $\lim_{t\to 0_+} \psi_B(t) = \psi_B(0) = 0$.

(iii) For some $a > 0$, it holds that $\psi_B|_{(0,\, a]} \in \mathrm{RV}^0_\rho$ with $\rho \in [0,\, 1]$.

We will refer to ψ_B as a regularity function for the L_i's over B. *If the operators L_i are JKR over all bounded sets B in such a way that the regularity functions ψ_B can be taken to be pairwise asymptotically equivalent up to a constant, we call them* uniformly jointly Karamata regular (UJKR)*.*

In particular, when $n = 1$, we will drop the qualifier "jointly" and call the single operator L Karamata regular (KR) over B *and* uniformly Karamata regular (UKR)*, respectively.*

Remark 3.2 (Domain of ψ_B and positivity)**.** *In item (iii) of Definition 3.1, as far as regular variation at zero is concerned, the actual value of a does not matter since only the behavior of ψ_B as it approaches zero is relevant. Nevertheless, even if we are flexible with the domain and image as in Remark 2.3, a function in RV^0_ρ must, at the very least, be positive close to zero. Therefore, the requirement that the restriction of ψ_B to some $(0,\, a]$ is in RV^0_ρ together with monotonicity implies $\psi_B(t) > 0$ for $t \neq 0$.*

Remark 3.3 (Connection with existing concepts)**.** *Definition 3.1 is closely related to several existing definitions.*

(i) (Bounded Hölder regular intersection) *Definition 3.1 extends the definition of bounded Hölder regular intersection in [10, Definition 2.2] as follows. Let $C_1, \dots, C_n \subseteq \mathcal{E}$ be convex sets and let $L_i \coloneqq P_{C_i}$ denote the projection operator onto C_i. With that, $C_1, \dots, C_n \subseteq \mathcal{E}$ has a bounded Hölder regular intersection if and only if for every bounded set B the L_i's are JKR over B and there exists $c_B > 0$ and $\gamma_B \in (0,\, 1]$ such that (3.1) holds with regularity function $\psi_B(\cdot) \coloneqq c_B(\cdot)^{\gamma_B}$. In particular, if $\gamma_B \equiv \gamma$ does not depend on B (in this case, we have that operators L_i are UJKR), then the collection $\{C_i\}$ is bounded Hölder regular with uniform exponent γ. We remark that the notion of bounded Hölder regularity coincides with the notion of Hölderian error bound.*

(ii) (Bounded Hölder regular operators) *An operator L is is a bounded Hölder regular operator (as in [10, Definition 2.4]) if and only if for every bounded set B, L is KR over B and the regularity function ψ_B can be taken to be of the form $\psi_B(\cdot) = c_B(\cdot)^{\gamma_B}$ with $c_B > 0$ and $\gamma_B \in (0,\, 1]$. The exponent $\gamma_B \equiv \gamma$ does not depend on B (in this case, we have that L is UKR) if and only if L is bounded Hölder regular with uniform exponent γ.*

(iii) (Consistent error bounds) *For closed convex sets* $C_i \subseteq \mathcal{E}$ $(i = 1, \ldots, n)$ *with non-empty intersection, a consistent error bound function* Φ *([24, Definition 3.1]) is a two-parameter function on* $\mathbb{R}^2_+$*, which is nondecreasing with respect to each variable and satisfies* $\lim_{a \to 0_+} \Phi(a, b) = \Phi(0, b) = 0$ *for all* $b \geq 0$ *and*

$$\operatorname{dist}\Big(x, \bigcap_{i=1}^n C_i\Big) \leq \Phi\Big(\max_{1 \leq i \leq n} \operatorname{dist}(x, C_i), \|x\|\Big), \quad \forall\ x \in \mathcal{E}. \tag{3.2}$$

If (3.2) holds, and for any $b > 0$ *there exists some* $\rho \in [0, 1]$ *such that* $\Phi(\cdot, b)|_{(0, a]} \in \mathrm{RV}^0_\rho$ *for some* $a > 0$*, then the operators* $L_i := P_{C_i}$ *are JKR over the ball* $\mathbb{B}_b$ *of radius* b *for any* b*. In addition, for any bounded set* B*, there exists some* $r_B > 0$ *such that* $B \subseteq \mathbb{B}_{r_B}$*. Let* $\psi_B(\cdot) := \Phi(\cdot, r_B)$*. Then we have from (3.2) and the monotonicity of* Φ *with respect to the second variable that*

$$\begin{aligned}\operatorname{dist}\Big(x, \bigcap_{i=1}^n \operatorname{Fix} L_i\Big) = \operatorname{dist}\Big(x, \bigcap_{i=1}^n C_i\Big) &\leq \Phi\Big(\max_{1 \leq i \leq n} \operatorname{dist}(x, C_i), r_B\Big) \\ &= \psi_B\Big(\max_{1 \leq i \leq n} \operatorname{dist}(x, C_i)\Big) = \psi_B\Big(\max_{1 \leq i \leq n} \|x - L_i(x)\|\Big), \ \forall\ x \in B.\end{aligned}$$

Note that $\lim_{t \to 0_+} \psi_B(t) = \lim_{t \to 0_+} \Phi(t, r_B) = 0$ *and* $\psi_B(0) = \Phi(0, r_B) = 0$*. In addition,* ψ_B *is nondecreasing, thanks to the monotonicity of* Φ *with respect to the first variable. Moreover, the property of* ψ_B *in item (iii) of Definition 3.1 directly follows from the assumption on* Φ*.*

The summary is that if we have a consistent error bound function Φ *for the* C_i*'s where for sufficiently large* $b > 0$ *the functions* $\Phi(\cdot, b)$ *are regularly varying with index* $\rho_b \in [0, 1]$*, then the* P_{C_i}*'s are JKR over any bounded set* B *and the regularity function can be taken to be* $\psi_B(\cdot) := \Phi(\cdot, b)$ *for any* b *satisfying* $b \geq r_B$*. Later in Corollary 5.7, we will see that if the* C_i*'s are definable over an o-minimal structure and have non-empty intersection, we can always construct such a consistent error bound function.*

A typical situation in applications is having some regularity condition on each *individual operator* (e.g., as in Remark 3.3 (ii)) and some error bound condition on the fixed point sets (e.g., as in Remark 3.3 (iii)). The next results indicates how to aggregate these individual results and establish joint Karamata regularity for the operators.

Proposition 3.4 (Calculus of ψ_B). *Suppose that* $L_i : \mathcal{E} \to \mathcal{E}$ $(i = 1, \ldots, n)$ *are nonexpansive, closed operators such that* $\bigcap_{i=1}^n \operatorname{Fix} L_i \neq \emptyset$ *holds. Let* Φ *be a consistent error bound function for the sets* $\operatorname{Fix} L_i$*,* $B \subseteq \mathcal{E}$ *a bounded set and suppose that* Φ *and* L_i *are as follows:*

(i) *for any* $b > 0$*, there exists* $\theta_b \in [0, 1]$ *such that* $\Phi(\cdot, b)|_{(0, a]} \in \mathrm{RV}^0_{\theta_b}$ *for some* $a > 0$*;*

(ii) *each* L_i *is Karamata regular over* B*.*

Then, the following statements hold.

(*a*) *The operators* L_i $(i = 1, \ldots, n)$ *are jointly Karamata regular over* B*.*

(*b*) *Let* Γ^i_B *be the regularity function for each* L_i *over* B*. Assume that* $\Gamma^i_B|_{(0, a]} \in \mathrm{RV}^0_{\rho_i}$*. Then, the function defined by*

$$\psi_B := \Theta_B \circ \Gamma_B$$

satisfies (3.1) and $\psi_B|_{(0, a]} \in \mathrm{RV}^0_\rho$*, where* $\rho = \theta_b \min_{1 \leq i \leq n} \rho_i$*,* $\Theta_B(\cdot) := \Phi(\cdot, b)$*,* $\Gamma_B := \sum_{i=1}^n \Gamma^i_B$ *and* b *is such that* $B \subseteq \mathbb{B}_b$*.*

Proof. We will prove item (a) and (b) together. Since each L_i is nonexpansive and closed, each $\operatorname{Fix} L_i$ must be closed and convex, e.g., [3, Proposition 4.13]. By assumption, $\bigcap_{i=1}^n \operatorname{Fix} L_i \neq \emptyset$ and Φ is a consistent error bound function for the sets $\operatorname{Fix} L_i$ $(i = 1, \ldots, n)$. Therefore,

$$\operatorname{dist}\Big(x, \bigcap_{i=1}^{n} \operatorname{Fix} L_i\Big) \leq \Phi\Big(\max_{1\leq i\leq n} \operatorname{dist}(x, \operatorname{Fix} L_i), \|x\|\Big), \ \forall\, x \in \mathcal{E}.$$

This together with $B \subseteq \mathbb{B}_b$, the monotonicity of Φ and the definition of Θ_B further implies

$$\operatorname{dist}\Big(x, \bigcap_{i=1}^{n} \operatorname{Fix} L_i\Big) \leq \Phi\Big(\max_{1\leq i\leq n} \operatorname{dist}(x, \operatorname{Fix} L_i), b\Big) = \Theta_B\Big(\max_{1\leq i\leq n} \operatorname{dist}(x, \operatorname{Fix} L_i)\Big), \ \forall\, x \in B. \tag{3.3}$$

On the other hand, we see from assumption (ii) and the assumption on the Γ_B^i's that

$$\operatorname{dist}(x, \operatorname{Fix} L_i) \leq \Gamma_B^i\left(\|x - L_i(x)\|\right), \ \forall\, x \in B. \tag{3.4}$$

Combining (3.3) and (3.4), we have from the monotonicity of Θ_B and the nonnegativity of Γ_B^i that for all $x \in B$,

$$\begin{aligned}\operatorname{dist}\Big(x, \bigcap_{i=1}^{n} \operatorname{Fix} L_i\Big) &\leq \Theta_B\Big(\max_{1\leq i\leq n} \Gamma_B^i\left(\|x - L_i(x)\|\right)\Big) \leq \Theta_B\Big(\max_{1\leq i\leq n} \Gamma_B\left(\|x - L_i(x)\|\right)\Big)\\ &= \Theta_B\Big(\Gamma_B\Big(\max_{1\leq i\leq n} \|x - L_i(x)\|\Big)\Big) = \psi_B\Big(\max_{1\leq i\leq n} \|x - L_i(x)\|\Big).\end{aligned}$$

Note that $\Theta_B(\cdot) = \Phi(\cdot, b)$ is nondecreasing and

$$\lim_{t\to 0_+} \Theta_B(t) = \lim_{t\to 0_+} \Phi(t, b) = 0 = \Theta_B(0).$$

Also, Γ_B is nondecreasing and it holds that $\lim_{t\to 0_+} \Gamma_B(t) = \Gamma_B(0) = 0$, thanks to $\Gamma_B = \sum\limits_{i=1}^{n} \Gamma_B^i$ and the properties of Γ_B^i. Consequently, $\psi_B = \Theta_B \circ \Gamma_B$ is nondecreasing and satisfies $\lim_{t\to 0_+} \psi_B(t) = \psi_B(0) = 0$. Moreover, since $\Theta_B|_{(0,\, a]} \in \mathrm{RV}^0_{\theta_b}$ and $\Gamma_B^i|_{(0,\, a]} \in \mathrm{RV}^0_{\rho_i}$, using (2.5) we have $\psi_B|_{(0,\, a]} = (\Theta_B \circ \Gamma_B)|_{(0,\, a]} \in \mathrm{RV}^0_{\rho}$ with $\rho = \theta_b \min\limits_{1\leq i\leq n} \rho_i$. This completes the proof. ☐

3.1 General convergence theory

The goal of this section is to present a general result that connects the error bound function ψ_B appearing in Definition 3.1 to the convergence rate of a sequence generated by the quasi-cyclic algorithm described in [10]. This will be accomplished by applying a series of functional transformation to ψ_B which will culminate in Theorem 3.8, the main result of this section. Theorem 3.8 is an abstract result that is hard to apply directly, but in later sections we will show how to use regular variation to better estimate the asymptotic properties of the function appearing in Theorem 3.8 without explicitly computing it.

Quasi-cylic algorithms We recall the common fixed point problem in (1.1). In [10, Section 3], Borwein, Li and Tam analysed the framework of *quasi-cyclic algorithms* which was considered earlier by Bauschke, Noll and Phan in [4]. A quasi-cylic algorithm is given by iterations that are as follows:

$$x^{k+1} = \sum_{i=1}^{m} w_i^k T_i(x^k), \tag{3.5}$$

where the weight parameters $w_i^k \geq 0$ satisfy $\sum_{i=1}^m w_i^k = 1$ and $\nu := \inf_{k\in\mathbb{N}} \min_{i\in I_+(k)} w_i^k > 0$, where $I_+(k) := \{1 \leq i \leq m \mid w_i^k > 0\}$. Later, in Theorem 3.8 we will impose additional conditions on the $I_+(k)$ in order to ensure convergence. The algorithm framework (3.5) covers a number of projection algorithms, the Douglas-Rachford splitting method as well as the forward-backford splitting method, see [10].

A general convergence rate result In [10], the authors derived convergence rate results for the iteration (3.5) under certain Hölderian assumptions. Here, one of our main goals is to prove convergence rates under the more general Karamata regularity condition as in Definition 3.1.

Before we proceed, we need to go through a few technical lemmas. We start with a result regarding the minus inverse defined in (2.15).

Lemma 3.5. *Let $f : (0,\, a] \to (0,\, \infty)$ satisfy $\lim_{x\to 0_+} f(x) = 0$. Then f^- is nondecreasing. If f is nondecreasing, then $0 < s \leq f(t)$ implies that $f^-(s) \leq t$.*

Proof. The monotonicity of f^- follows directly from its definition. Next, let $0 < s \leq f(t)$ and suppose that f is nondecreasing. For all $y \in (0,\, a]$ satisfying $f(y) < s$, we have $f(y) < s \leq f(t)$. Then, the monotonicity of f implies $y < t$. Therefore, $f^-(s) = \sup\{y \in (0,\, a] \mid f(y) < s\} \leq t$. This completes the proof. ☐

Let $f : (0,\, a] \to (0,\, \infty)$ satisfy $\lim_{x\to 0_+} f(x) = 0$. Fix $\delta > 0$ and consider the following integral.

$$\Phi_f(x) := \int_x^\delta \frac{1}{f^-(t)} dt, \quad x > 0, \tag{3.6}$$

which is well-defined for $x \in (0, \infty)$. The integral in (3.6) plays a fundamental role in this paper and its *inverse* is related to the convergence rate of algorithms as will be described in Theorem 3.8. In the next lemma, we will check some properties of Φ_f.

Lemma 3.6. *Let $f : (0,\, a] \to (0,\, \infty)$ satisfy $\lim_{x\to 0_+} f(x) = 0$ and let Φ_f be defined as in (3.6). Then the following statements hold.*

(i) Φ_f is continuous and decreasing;

(ii) Suppose that f is nondecreasing and one of the conditions below is satisfied:

  (a) $f \in \mathrm{RV}_\rho^0$ with $\rho \in [0,\, 1)$;

  (b) $f(x) \geq cx$ holds for some $c > 0$ as $x \to 0_+$.

  Then, $\Phi_f(x) \to \infty$ as $x \to 0_+$.

Proof. First, we see from the definition of f^- in (2.15) that $f^-(x) > 0$ for all $x > 0$. Then Φ_f is decreasing. Next, we show the continuity of Φ_f on $(0,\, \infty)$. For any fixed $\widehat{x} \in (0,\, \infty)$, one can find a compact interval $[c,\, d]$ such that $\widehat{x} \in [c,\, d]$ and $c > 0$. Then for any $x \in [c,\, d]$, one has

$$\Phi_f(x) = \int_x^\delta \frac{1}{f^-(t)} dt = \int_c^\delta \frac{1}{f^-(t)} dt + \int_x^c \frac{1}{f^-(t)} dt = \Phi_f(c) + \int_c^x \frac{-1}{f^-(t)} dt.$$

Let $g(t) := \frac{-1}{f^-(t)}$. By Lemma 3.5, we see that g is nondecreasing, and therefore $g|_{[c,\, d]}$ is measurable and integrable. Using [1, Theorem 4.4.1], we then have that $\Phi_f|_{[c,\, d]}$ is absolutely continuous. This together with the arbitrariness of $\widehat{x}$ proves the continuity of Φ_f on $(0,\, \infty)$ and completes proof of item *(i)*.

Now we prove $\Phi_f(x) \to \infty$ as $x \to 0_+$ in two cases.

- In case *(a)*, from (2.5) we have $f(t)/t \in \mathrm{RV}^0_{\rho-1}$. Since $\rho - 1 < 0$, we see from (2.11) that $f(t)/t \to \infty$ as $t \to 0_+$, which further implies $f(t) \geq t$ when $t \in (0, \varepsilon]$ for some $\varepsilon \in (0, \delta)$. By the monotonicity of f and Lemma 3.5, it holds $f^-(t) \leq t$ for $t \in (0, \varepsilon]$. Using this and $f^- > 0$, we have for all $x \in (0, \varepsilon)$ that

$$\Phi_f(x) := \int_x^\delta \frac{1}{f^-(t)} dt \geq \int_x^\varepsilon \frac{1}{f^-(t)} dt \geq \int_x^\varepsilon \frac{1}{t} dt = \ln(\varepsilon) - \ln(x),$$

  which proves $\Phi_f(x) \to \infty$ as $x \to 0_+$.

- In case *(b)*, we see that there exists some $\varepsilon \in (0, \delta/c]$ such that $cx \leq f(x)$ for $x \in (0, \varepsilon]$. By the monotonicity of f and Lemma 3.5, it holds $f^-(cx) \leq x$ for $x \in (0, \varepsilon]$. Thus, for all $x \in (0, c\varepsilon)$, it holds

$$\Phi_f(x) := \int_x^\delta \frac{1}{f^-(t)} dt \geq \int_x^{c\varepsilon} \frac{1}{f^-(t)} dt = c \int_{x/c}^\varepsilon \frac{1}{f^-(cy)} dy \geq c \int_{x/c}^\varepsilon \frac{1}{y} dy = c\ln(\varepsilon) - c\ln(x/c),$$

  which proves $\Phi_f(x) \to \infty$ as $x \to 0_+$.

This completes the proof. $\square$

Remark 3.7 (Condition in Lemma 3.6)**.** *We list two special conditions contained in Lemma 3.6 (ii) (b):*

(i) $\lim_{x\to 0_+} \frac{f(x)}{x} = \infty$. *This includes the entropic error bound function in [23, Section 4.2.1] and [24, Section 6.2]:* $f(x) := -x\ln(x)$, $x \in (0, a]$ *for some* $a > 0$ *and we note that* f *belongs to* RV^0_1.

(ii) $\lim_{x\to 0_+} \frac{f(x)}{x} = \mu$ *for some* $\mu > 0$. *This corresponds to* $f(x) \overset{c}{\sim} x$, *and further implies that* $f \in \mathrm{RV}^0_1$, *thanks to* (2.26).

One the other hand, the condition $f \in \mathrm{RV}^0_1$ *alone is not enough to guarantee* $\Phi_f(x) \to \infty$ *as* $x \to 0_+$. *Consider the following function:*

$$g(x) := x(1+x)\left(\ln(1+1/x)\right)^2, \quad x \in (0, 1/(e^2-1)].$$

We have $g \in \mathrm{RV}^0_1$ *and* g *is increasing and continuous. Then the usual inverse* g^{-1} *exists. Let* $f := g^{-1}$. *We then have* $f^- = f^{-1} = g$. *Moreover, we have from* (2.23) *that* $f \in \mathrm{RV}^0_1$ *and* f *is increasing. However, when* $x, \delta \in (0, 1/(e^2-1)]$,

$$\Phi_f(x) := \int_x^\delta \frac{1}{f^-(t)} dt = \int_x^\delta \frac{1}{t(1+t)\left(\ln(1+1/t)\right)^2} dt = \frac{1}{\ln(1+1/\delta)} - \frac{1}{\ln(1+1/x)},$$

which implies that $\Phi_f(x) \not\to \infty$ *as* $x \to 0_+$.

All pieces are now in place for the main result of this subsection.

Theorem 3.8. *Let sequence* $\{x^k\}$ *be generated by quasi-cyclic algorithm* (3.5). *Then* $\{x^k\}$ *is bounded. Let* B *be a bounded set containing* $\{x^k\}$ *and suppose that the following assumptions hold:*

(a) T_i $(i = 1, \ldots, m)$ *are jointly Karamata regular (*resp. *Karamata regular when* $m = 1$*) over* B *with regularity function* ψ_B *as in Definition 3.1;*

(b) *there exists some* $s > 0$ *such that for each* $k \in \mathbb{N}$,

$$I_+(k) \cup I_+(k+1) \cup \cdots \cup I_+(k+s-1) = \{1, \ldots, m\}.$$

Fix any $\widehat{a} > \operatorname{dist}^2(x^0, F)$ and $\delta > 0$ and define

$$\widehat{\phi}(u) := \psi_B^2\Big(\sqrt{\frac{2\alpha(1+4\nu s)}{\nu(1-\alpha)}u}\Big), \quad \phi(u) = \widehat{\phi}(u)|_{(0,\,\widehat{a}]}, \quad \Phi_\phi(u) := \int_u^\delta \frac{1}{\phi^-(t)}dt, \;\; u > 0. \tag{3.7}$$

Then $\{x^k\}$ converges to some $x^* \in \bigcap_{i=1}^m \operatorname{Fix} T_i$ finitely or the convergence rate is given by

$$\operatorname{dist}(x^k, F) \leq \sqrt{\Phi_\phi^{-1}\left(\Phi_\phi\left(\operatorname{dist}^2(x^0, F)\right) + \lfloor k/s \rfloor\right)}, \quad \forall\, k \in \mathbb{N}. \tag{3.8}$$

Proof. Recall that $F = \bigcap_{i=1}^m \operatorname{Fix} T_i$ and let $y \in F$. Then for all $k \in \mathbb{N}$, the nonexpansiveness of T_i (see from Lemma 2.1) gives

$$\|x^{k+1} - y\| = \Big\|\sum_{i=1}^m w_i^k T_i(x^k) - y\Big\| \leq \sum_{i=1}^m w_i^k \|T_i(x^k) - y\| \leq \sum_{i=1}^m w_i^k \|x^k - y\| = \|x^k - y\|,$$

which simultaneously proves that $\{x^k\}$ is bounded and Fejér monotone. In particular, whenever $r \geq k$ and $y \in F$ we have

$$\|x^r - y\| \leq \|x^k - y\|, \tag{3.9}$$

$$\operatorname{dist}(x^r, F) \leq \operatorname{dist}(x^k, F). \tag{3.10}$$

If there exists some $\widehat{k}$ such that $x^{\widehat{k}} \in F$, we then see from (3.5) and $\sum_{i=1}^m w_i^k = 1$ that $x^k = x^{\widehat{k}}$ for all $k \geq \widehat{k}$. In this case, $\{x^k\}$ converges to some $x^* \in F = \cap_{i=1}^m \operatorname{Fix} T_i$ finitely. Next, we consider the case that $x^k \notin F$ holds for all $k \in \mathbb{N}$.

By assumption (a) and $\{x^k\} \subseteq B$, we know that for all $k \in \mathbb{N}$,

$$\operatorname{dist}^2(x^{ks}, F) \leq \psi_B^2\Big(\sqrt{\max_{1\leq i\leq m} \|x^{ks} - T_i(x^{ks})\|^2}\Big). \tag{3.11}$$

Fix any $t \in \{1, \ldots, m\}$. By assumption (b), for any k there exists some $t_k \in \{ks, \ldots, (k+1)s - 1\}$ such that $t \in I_+(t_k)$. Thus, we have from the nonexpansiveness of T_t that

$$\begin{aligned}
\|x^{ks} - T_t(x^{ks})\|^2 &\leq \left(\|x^{ks} - x^{t_k}\| + \|x^{t_k} - T_t(x^{t_k})\| + \|T_t(x^{t_k}) - T_t(x^{ks})\|\right)^2 \\
&\leq \left(\|x^{t_k} - T_t(x^{t_k})\| + 2\|x^{ks} - x^{t_k}\|\right)^2 \\
&\overset{(a)}{\leq} \Big(\|x^{t_k} - T_t(x^{t_k})\| + 2\sum_{j=ks}^{t_k - 1} \|x^j - x^{j+1}\|\Big)^2 \\
&\overset{(b)}{\leq} 2\|x^{t_k} - T_t(x^{t_k})\|^2 + 8(t_k - ks)\sum_{j=ks}^{t_k-1} \|x^j - x^{j+1}\|^2 \\
&\leq 2\|x^{t_k} - T_t(x^{t_k})\|^2 + 8s \sum_{j=ks}^{(k+1)s-1} \|x^j - x^{j+1}\|^2,
\end{aligned} \tag{3.12}$$

where (a) follows from repeated applications of the triangle inequality. For (b), we consider two cases. If $t_k = ks$, then (b) holds. For $t_k > ks$, we use the convexity of the square function so that $(a_1 + \cdots + a_r)^2 \leq r\sum_{i=1}^r a_i^2$ holds for arbitrary $a_i \in \mathbb{R}$. This inequality is first applied with $r := 2$, $a_1 := 2\|x^{t_k} - T_t(x^{t_k})\|^2$, a_2 being the remaining sum and, then, it is applied once more with $r := t_k - ks$ to bound the remaining terms.

Next, we bound the two terms in the right-hand side of the last inequality of (3.12). First, since each T_i is α-averaged, we have from Lemma 2.1 that for all $x \in \mathcal{E}$ and $y \in \operatorname{Fix} T_i$,

$$\|T_i(x) - y\|^2 + \frac{1-\alpha}{\alpha}\big\|(I - T_i)(x)\big\|^2 \leq \|x - y\|^2$$

holds[3]. Therefore, for all $r \in \mathbb{N}$, $x \in \mathcal{E}$ and $y \in F$,

$$\begin{aligned}\Big\| \sum_{i=1}^{m} w_i^r T_i(x) - y \Big\|^2 = \Big\| \sum_{i=1}^{m} w_i^r \left(T_i(x) - y\right) \Big\|^2 &\le \sum_{i=1}^{m} w_i^r \big\|T_i(x) - y\big\|^2 \\ &\le \|x-y\|^2 - \frac{1-\alpha}{\alpha} \sum_{i=1}^{m} w_i^r \|x - T_i(x)\|^2.\end{aligned} \tag{3.13}$$

Some extra algebraic acrobatics leads to

$$\begin{aligned}\frac{\nu(1-\alpha)}{\alpha}\big\|x^{t_k} - T_t(x^{t_k})\big\|^2 &\overset{(a)}{\le} \frac{1-\alpha}{\alpha} \sum_{i=1}^{m} w_i^{t_k} \big\|x^{t_k} - T_i(x^{t_k})\big\|^2 \\ &\overset{(b)}{\le} \big\|x^{t_k} - P_F(x^{ks})\big\|^2 - \Big\| \sum_{i=1}^{m} w_i^{t_k} T_i(x^{t_k}) - P_F(x^{ks}) \Big\|^2 \\ &= \big\|x^{t_k} - P_F(x^{ks})\big\|^2 - \big\|x^{t_k+1} - P_F(x^{ks})\big\|^2 \\ &\overset{(c)}{\le} \big\|x^{ks} - P_F(x^{ks})\big\|^2 - \big\|x^{(k+1)s} - P_F(x^{ks})\big\|^2 \\ &\le \operatorname{dist}^2(x^{ks}, F) - \operatorname{dist}^2(x^{(k+1)s}, F),\end{aligned} \tag{3.14}$$

where (a) follows from $t \in I_+(t_k)$ and the definition of ν which implies that $w_t^{t_k} \ge \nu$, (b) follows from (3.13) by letting $r := t_k$, $x := x^{t_k}$ and $y := P_F(x^{ks})$ in (3.13). Then, (c) follows from $ks \le t_k$, $t_k + 1 \le (k+1)s$ and two applications of (3.9): first with x^{t_k}, x^{ks} and $P_F(x^{ks})$ and second with $x^{(k+1)s}, x^{t_k+1}$ and $P_F(x^{ks})$.

For each $j \in \{ks, \dots, (k+1)s-1\}$, we let $r := j$, $x := x^j$ and $y := P_F(x^{ks})$ in (3.13), and obtain

$$\begin{aligned}\|x^j - x^{j+1}\|^2 = \Big\|x^j - \sum_{i=1}^{m} w_i^j T_i(x^j)\Big\|^2 &\le \sum_{i=1}^{m} w_i^j \|x^j - T_i(x^j)\|^2 \\ &\le \frac{\alpha}{1-\alpha}\big(\|x^j - P_F(x^{ks})\|^2 - \|x^{j+1} - P_F(x^{ks})\|^2\big).\end{aligned}$$

This further implies that

$$\begin{aligned}\sum_{j=ks}^{(k+1)s-1} \|x^j - x^{j+1}\|^2 &\le \frac{\alpha}{1-\alpha}\big(\|x^{ks} - P_F(x^{ks})\|^2 - \|x^{(k+1)s} - P_F(x^{ks})\|^2\big) \\ &\le \frac{\alpha}{1-\alpha}\Big(\operatorname{dist}^2(x^{ks}, F) - \operatorname{dist}^2(x^{(k+1)s}, F)\Big).\end{aligned} \tag{3.15}$$

Let $\triangle_k := \operatorname{dist}^2\left(x^{ks}, F\right) - \operatorname{dist}^2\left(x^{(k+1)s}, F\right)$. Then we have from (3.10) that $\triangle_k \le \widehat{a}$. Now, we combine (3.11), (3.12), (3.14), (3.15) and the arbitrariness of $t \in \{1, \dots, m\}$ to obtain

$$\begin{aligned}\operatorname{dist}^2(x^{ks}, F) &\le \psi_B^2\Big(\max_{1\le t\le m} \sqrt{\|x^{ks} - T_t(x^{ks})\|^2}\Big) \\ &\le \psi_B^2\Big(\max_{1\le t\le m} \sqrt{2\|x^{t_k} - T_t(x^{t_k})\|^2 + \frac{8s\alpha\triangle_k}{1-\alpha}}\Big) \\ &\le \psi_B^2\Big(\sqrt{\frac{2\alpha\triangle_k}{\nu(1-\alpha)} + \frac{8s\alpha\triangle_k}{1-\alpha}}\Big) \\ &\le \psi_B^2\Big(\sqrt{\frac{2\alpha(1+4\nu s)}{\nu(1-\alpha)}\triangle_k}\Big) = \widehat{\phi}(\triangle_k) = \phi(\triangle_k).\end{aligned} \tag{3.16}$$

[3]Each T_i may be in fact α_i-averaged with a different $\alpha_i \in (0,1)$, but without loss of generality we may assume that the inequality holds with $\alpha = \max_{1\le i\le m}\{\alpha_i\}$.

We see from (3.10) and the nonnegativity of $\mathrm{dist}(x^k, F)$ that the sequence $\{\mathrm{dist}(x^k, F)\}$ converges to some $c^* \geq 0$. Letting $k \to \infty$ on both sides of (3.16), recalling $\lim_{x\to 0_+} \phi(x) = 0$ (due to $\lim_{x\to 0_+} \psi_B(x) = 0$) we have $\mathrm{dist}(x^k, F) \to c^* = 0$. Since $\{x^k\}$ is bounded, there exists a subsequence $\{x^{k_i}\}$ which converges to some point $x^* \in \mathcal{E}$. Therefore, $\mathrm{dist}(x^{k_i}, F) \to 0$ together with the closedness of F implies that $x^* \in F$. We note from the Fejér monotonicity of $\{x^k\}$ with respect to F (or (3.9)) that the nonnegative sequence $\{\|x^k - x^*\|\}$ is nonincreasing and thus convergent. This together with $\|x^{k_i} - x^*\| \to 0$ implies that $\{x^k\}$ converges to $x^* \in F = \bigcap_{i=1}^m \mathrm{Fix}\, T_i$.

We note that $\lim_{x\to 0_+} \phi(x) = 0$ and ϕ is nondecreasing, thanks to the same properties of the regularity function ψ_B. Consequently,

$$\begin{aligned}&\Phi_\phi\left(\mathrm{dist}^2\big(x^{(k+1)s}, F\big)\right) - \Phi_\phi\left(\mathrm{dist}^2\left(x^{ks}, F\right)\right)\\ &= \int_{\mathrm{dist}^2\left(x^{(k+1)s}, F\right)}^{\mathrm{dist}^2\left(x^{ks}, F\right)} \frac{1}{\phi^-(t)}dt \geq \frac{\triangle_k}{\phi^-\left(\mathrm{dist}^2\left(x^{ks}, F\right)\right)} \geq 1,\end{aligned} \tag{3.17}$$

where the first inequality follows from the monotonicity of ϕ^- and the second inequality follows from (3.16) and Lemma 3.5. Moreover, for any $\ell > 0$, summing both sides of (3.17) for $k = 0, \ldots, \ell - l$, we obtain

$$\Phi_\phi\left(\mathrm{dist}^2\left(x^{\ell s}, F\right)\right) - \Phi_\phi\left(\mathrm{dist}^2\left(x^0, F\right)\right) \geq \ell.$$

This together with the monotonicity of Φ_ϕ and the Fejér monotonicity of $\{x^k\}$ further implies that for any $k \in \mathbb{N}$,

$$\Phi_\phi\left(\mathrm{dist}^2\left(x^k, F\right)\right) \geq \Phi_\phi\left(\mathrm{dist}^2\big(x^{\lfloor k/s\rfloor\cdot s}, F\big)\right) \geq \Phi_\phi\left(\mathrm{dist}^2\left(x^0, F\right)\right) + \lfloor k/s\rfloor.$$

Notice from Lemma 3.6 that Φ_ϕ is continuous and decreasing, and therefore the usual inverse Φ_ϕ^{-1} exists. Also, $\Phi_\phi\left(\mathrm{dist}^2\left(x^0, F\right)\right) + \lfloor k/s\rfloor$ is in the interval $\left[\Phi_\phi\left(\mathrm{dist}^2\left(x^0, F\right)\right), \Phi_\phi\left(\mathrm{dist}^2\left(x^k, F\right)\right)\right]$, so continuity implies that it is also in the domain of Φ_ϕ^{-1}. Then, applying Φ_ϕ^{-1} to the above inequality and rearranging the terms finally leads to (3.8). This completes the proof. $\square$

3.2 Rates based on the index of regular variation

Theorem 3.8 is a general result on convergence rates. However, typically we would like to obtain more concrete results and say, for example, whether the rate is linear, sublinear and etc. A direct application of Theorem 3.8 would require one to compute the function Φ_ϕ and its inverse, which can be both highly nontrivial and devoid of closed forms.

In this paper, we will show two techniques for obtaining "concrete" convergence rates and avoid the direct computation Φ_ϕ and Φ_ϕ^{-1}. The first, presented in this subsection, is based solely on the index of regular variation of the error bound function ψ_B. We note that if we are given ψ_B, computing the index of regular variation is a significantly easier task, since it is just a limit computation as in Definition 2.2. However, even if we have access to ψ_B, computing Φ_ϕ^{-1} can be highly nontrivial.

Before we proceed we need to review more tools from regular variation. First, we need a result that is a part of *Karamata's theorem*, which tells us about the behavior of regular varying functions under taking integrals. Suppose that $f : [a, \infty) \to (0, \infty) \in \mathrm{RV}_\rho$ is locally bounded. If $\sigma \geq -(\rho + 1)$, then

$$\lim_{x\to\infty} \frac{x^{\sigma+1} f(x)}{\int_a^x t^\sigma f(t)dt} = \sigma + 1 + \rho, \tag{3.18}$$

see [6, Theorem 1.5.11]. Karamata's theorem is a crucial result of this body of theory and the case $\sigma = 0$ and $\rho > -1$ represents a remarkable property of regularly varying functions: as far as the behavior at infinity is concerned, functions in RV_ρ behave as polynomials of degree ρ in that $\int_a^x f(t)dt$ is asymptotically equivalent to $\frac{x}{\rho+1}f(x)$. This foreshadows why this will be useful for us and immediately suggests how to bypass the computation of the hard integral that appears in Theorem 3.8.

Next, suppose that $f : (0, a] \to (0, \infty) \in \mathrm{RV}^0_\rho$ is locally bounded away from zero. Then $f(1/\cdot) \in \mathrm{RV}_{-\rho}$ is locally bounded on $[\widehat{a}, \infty)$ for $\widehat{a} := 1/a$. In order to derive analogous statements for regular variation at 0, we make the substitution $x = 1/y$ in (3.18), and change the variable inside the integrals. Recalling that $f(1/\cdot) \in \mathrm{RV}_{-\rho}$, we conclude that if $\sigma \geq -1 + \rho$, then

$$\lim_{y\to 0_+} \frac{y^{-(\sigma+1)}f(y)}{\int_y^a t^{-\sigma-2}f(t)dt} = \sigma + 1 - \rho. \tag{3.19}$$

We note that $y^{-(\sigma+1)}f(y) \in \mathrm{RV}^0_{\rho-\sigma-1}$ and that (2.26) and (3.19) imply that if $\sigma + 1 - \rho > 0$, then g defined by $g(y) := \int_y^a t^{-\sigma-2}f(t)dt$ belongs to $\mathrm{RV}^0_{\rho-\sigma-1}$ as well. When $\sigma = -2$, we have the following important special case:

$$f \in \mathrm{RV}^0_\rho \text{ and } -1 > \rho \quad \Rightarrow \quad \int_y^a f(t)dt \overset{c}{\sim} yf(y) \in \mathrm{RV}^0_{\rho+1}. \tag{3.20}$$

A similar result holds for $\sigma = -2$, $\rho = -1$. Let $\ell(s) := f(1/s)s^{-1}$, so that $\ell \in \mathrm{RV}_0$, i.e., is a function of slow variation. Then, since f is measurable and locally bounded away from zero, ℓ is measurable, locally bounded over $[1/a, \infty)$. In particular, ℓ is locally integrable over $[1/a, \infty)$ and we can invoke [6, Proposition 1.5.9a] to conclude that $\int_{1/a}^x s^{-1}\ell(s)ds \in \mathrm{RV}_0$. Observe that

$$\int_{1/x}^a f(t)dt = \int_{1/a}^x s^{-2}f(1/s)ds = \int_{1/a}^x s^{-1}\ell(s)ds. \tag{3.21}$$

Therefore, as a function of x, the left-hand-side of (3.21) belongs to RV_0. Performing the substitution $x = 1/y$, we obtain the following implication:

$$f \in \mathrm{RV}^0_{-1} \quad \Rightarrow \quad \int_y^a f(t)dt \in \mathrm{RV}^0_0. \tag{3.22}$$

Finally, we need a result on the integral of rapidly varying functions. Suppose that $f : [a, \infty) \to (0, \infty)$ belongs to KRV_∞ and is locally bounded, then

$$\lim_{x\to\infty} \frac{f(x)}{\int_a^x f(t)dt/t} \to \infty, \tag{3.23}$$

see [6, Proposition 2.6.9][4]. With that, we have the following proposition, see also [15, Corollary 1.1] for a related result.

Proposition 3.9. *Let $f : [a, \infty) \to (0, \infty) \in \mathrm{KRV}_\infty$ (where $a > 0$) be a nondecreasing function, then*

$$\int_a^x f(t)dt \in \mathrm{KRV}_\infty.$$

Proof. Let $\lambda > 1$ and let $F(x) := \int_a^x f(t)dt$, $(x > a)$. Then

$$F(\lambda x) - F(x) = \int_x^{\lambda x} f(t)dt \geq (\lambda x - x)f(x),$$

[4]Strictly speaking, Proposition 2.6.9 is about function classes BD and MR_∞ which are defined in Chapter 2 therein. However, KRV_∞ functions are both BD and MR_∞ by [6, Equation (2.4.3) and Proposition 2.4.4].

where the inequality follows from the monotonicity of f. Dividing by $F(x)$ and readjusting the terms, we have

$$\frac{F(\lambda x)}{F(x)} \geq 1 + (\lambda - 1)\frac{xf(x)}{F(x)} \tag{3.24}$$

Since $1/t \geq 1/x$ for $t \in [a, x]$, we have $\int_a^x f(t)dt/t \geq \int_a^x f(t)dt/x = \frac{F(x)}{x}$ and therefore,

$$\frac{xf(x)}{F(x)} \geq \frac{f(x)}{\int_a^x f(t)dt/t}. \tag{3.25}$$

By (3.23) the right-hand-side of (3.25) goes to ∞ as $x \to \infty$. Thus, the right-hand-side of (3.24) goes to ∞ as $x \to \infty$. Consequently, $F(\lambda x)/F(x)$ goes to ∞ when $x \to \infty$ and $\lambda > 1$. For $\lambda < 1$, we have $1/\lambda > 1$ and $\lim_{x\to\infty} F(\lambda x)/F(x) = \lim_{y\to\infty} F(y)/F(y/\lambda) = 0$ by what was just proved. This shows that $F \in \mathrm{RV}_\infty$. Since F is the integral of a positive monotone function, it must be nondecreasing as well[5], so, in fact, $F \in \mathrm{KRV}_\infty$ by (2.19). $\square$

We now have all pieces for our first results on the asymptotic properties of Φ_f.

Theorem 3.10 (Index of regular variation and asymptotic behavior of Φ_f)**.** *Let $f : (0, a] \to (0, \infty)$ be a nondecreasing function in RV_ρ^0 for $\rho \in [0, 1]$ such that $\lim_{t\to 0_+} f(t) = 0$ holds and consider the function Φ_f in (3.6). Then the following statements hold.*

(i) *If $\rho = 0$, then $\Phi_f \in \mathrm{RV}_{-\infty}^0$. In particular, $\Phi_f^{-1} \in \mathrm{RV}_0$ and for every $r > 0$, $x^{-r} = o(\Phi_\phi^{-1}(x))$ as $x \to \infty$.*

(ii) *If $\rho \in (0, 1)$, then $\Phi_f \in \mathrm{RV}_{1-1/\rho}^0$. In particular, $\Phi_f^{-1} \in \mathrm{RV}_{\rho/(\rho-1)}$ and $\Phi_f^{-1}(x) = o(x^{-r})$ as $x \to \infty$ for every $0 < r < -\rho/(\rho - 1)$.*

(iii) *If $\rho = 1$ and there exists some $c > 0$ such that $f(x) \geq cx$ as $x \to 0_+$, then $\Phi_f \in \mathrm{RV}_0^0$. In particular, $\Phi_f^{-1} \in \mathrm{RV}_{-\infty}$ and for every $r > 0$, $\Phi_f^{-1}(x) = o(x^{-r})$ as $x \to \infty$.*

Proof. Let $F : [1, \infty) \to (0, \infty)$ be such that

$$F(x) \coloneqq \Phi_f(1/x). \tag{3.26}$$

By Lemma 3.6, F is increasing and continuous, and thus locally bounded. Furthermore, $F^{-1}(y) = 1/(\Phi_f^{-1}(y))$ holds for $y \in [1, \infty)$.

We start with item (i), where $\rho = 0$. Let $g \coloneqq 1/f(1/\cdot)$. Then, we see from (2.3) and (2.4) that $g \in \mathrm{RV}_0$. Due to the monotonicity of f, both f and g are locally bounded. Moreover, $g(t)$ goes to ∞ as $t \to \infty$. By (2.21), $g^{\leftarrow} \in \mathrm{KRV}_\infty$ and by (2.16) we have

$$f^{-}(x) = \frac{1}{g^{\leftarrow}(1/x)}.$$

We note that F can be written as follows

$$F(y) = \int_{1/y}^{\delta} \frac{1}{f^{-}(s)} ds = \int_{1/\delta}^{y} \frac{t^{-2}}{f^{-}(1/t)} dt = \int_{1/\delta}^{y} t^{-2} g^{\leftarrow}(t) dt,$$

where the second equality is obtained by the substitution $s = t^{-1}$. Then, by (2.20), $(\cdot)^{-2}g^{\leftarrow}(\cdot)$ belongs to KRV_∞ and invoking Proposition 3.9, we conclude that F belongs to KRV_∞ which implies that $\Phi_f \in \mathrm{RV}_{-\infty}^0$ by (2.3) and (3.26). This proves the first half of item (i).

[5] As a reminder, $F(y) - F(x) = \int_x^y f(t)dt \geq f(x)(y - x) > 0$, if $y > x$.

Due to (2.13) and (2.18), $F(x)$ goes to ∞ as $x \to \infty$. Since F belongs to KRV_∞ we further conclude that $F^{\leftarrow} \in \mathrm{RV}_0$ by (2.22). Since $F^{\leftarrow}(y) = F^{-1}(y)$ holds for large y by Proposition 2.4, we see that $\Phi_f^{-1} \in \mathrm{RV}_0$ by (2.4).

Next, let $r > 0$ and let h be the function given by $h(x) = x^{-r}/\Phi_f^{-1}(x)$. By the calculus rules in (2.4), $h \in \mathrm{RV}_{-r}$. Since the index of h is negative, $h(x)$ goes to 0 as $x \to \infty$ by (2.9). That is, $x^{-r} = o(\Phi_f^{-1}(x))$ as $x \to \infty$. This concludes the proof of item (i).

We now move on to the case $\rho \in (0, 1]$. By (2.23), $1/f^-$ is regularly varying at 0 with index $-1/\rho$. We note $1/f^-$ is monotone by Lemma 3.5 and hence locally bounded away from zero. Since $-1 \geq -1/\rho$ and $\Phi_f(x) = \int_x^\delta \frac{1}{f^-(s)} ds$, we have $\Phi_f \in \mathrm{RV}^0_{1-1/\rho}$, which follows from (3.20) and (3.22). This proves the first halves of items (ii) and (iii).

By (3.26) and (2.3), $F \in \mathrm{RV}_{1/\rho - 1}$. We now verify that $F \to \infty$ as $x \to \infty$. If $\rho \in (0, 1)$ it follows from (2.8) and if $\rho = 1$ it follows from the assumption in item (ii) and Lemma 3.6 (b). The conclusion is that in either case, we have $F^{\leftarrow}(y) = F^{-1}(y)$ for large y, by Proposition 2.4. By the calculus rule for the inverse in (2.17) and (2.21), together with (2.4) and the definition of $\mathrm{RV}_{-\infty}$ in Definition 2.2, we conclude that $\Phi_f^{-1} = 1/F^{-1}$ belongs to $\mathrm{RV}_{\rho/(\rho-1)}$ if $\rho \in (0, 1)$ and to $\mathrm{RV}_{-\infty}$ if $\rho = 1$.

Suppose that $\rho \in (0, 1)$ and let r be such that $0 < r < -\rho/(\rho - 1)$. Let $\tau > 0$ be such that $r + \tau < -\rho/(\rho - 1)$. The index of Φ_f^{-1} is $\rho/(\rho - 1)$ and we apply Potter's bounds (see (2.6)) to Φ_f^{-1} with $\epsilon = -\rho/(\rho - 1) - r - \tau$. Fixing y, we see that for large x, we have

$$\Phi_f^{-1}(x) \leq \hat{A} x^{-r-\tau},$$

where $\hat{A}$ is some fixed constant. This implies that $\Phi_f^{-1}(x)/(x^{-r}) \leq \hat{A} x^{-\tau}$ and that the quotient goes to 0 as $x \to \infty$. Therefore, $\Phi_f^{-1}(x) = o(x^{-r})$. This concludes the proof of item (ii).

Next, suppose that $\rho = 1$ and let $r > 0$. We apply (2.12) to Φ_f^{-1}. Since (2.12) is valid for all $r > 0$, it is valid in particular for $2r$. So for sufficiently large x, we have $\Phi_f^{-1}(x)/x^{-r} \leq x^{-r}$, which implies that the quotient goes to 0 as $x \to \infty$. Therefore, $\Phi_f^{-1}(x) = o(x^{-r})$ and this proves item (iii). This completes the proof. $\square$

Using Theorem 3.10 we can analyze the quasi-cyclic iteration as follows.

Theorem 3.11 (Index of regular variation and convergence rates)**.** *Under the setting of Theorem 3.8, let ρ denote the index of $\psi_B|_{(0,\,a]}$ and suppose that the convergence is not finite. Then the following statements hold.*

(i) *If $\rho = 0$, then $\Phi_\phi \in \mathrm{RV}^0_{-\infty}$. In particular, $\Phi_\phi^{-1} \in \mathrm{RV}_0$ and for every $r > 0$, $k^{-r} = o(\Phi_\phi^{-1}(k))$ as $k \to \infty$.*

(ii) *If $\rho \in (0, 1)$, then $\Phi_\phi \in \mathrm{RV}^0_{1-1/\rho}$. In particular, $\Phi_\phi^{-1} \in \mathrm{RV}_{\rho/(\rho-1)}$ and $\Phi_\phi^{-1}(k) = o(k^{-r})$ as $k \to \infty$ for every $0 < r < -\rho/(\rho - 1)$.*

(iii) *Suppose that $\rho = 1$ and B is closed and connected. Then, $\Phi_\phi \in \mathrm{RV}^0_0$. In particular, $\Phi_\phi^{-1} \in \mathrm{RV}_{-\infty}$ and for every $r > 0$, $\Phi_\phi^{-1}(k) = o(k^{-r})$ as $k \to \infty$.*

Proof. By the calculus rules in (2.5), we see from $\psi_B|_{(0,\,a]} \in \mathrm{RV}^0_\rho$ and the definition of ϕ in (3.7) that ϕ has index ρ. Moreover, ϕ is nondecreasing with $\lim_{t\to 0_+} \phi(t) = 0$, thanks to the monotonicity of ψ_B and $\lim_{t \to 0_+} \psi_B(t) = 0$ in Definition 3.1. With that, items (i), (ii) are a consequence of Theorem 3.10 applied to $f := \phi$.

Item (iii) also follows from Theorem 3.10, but we need to check that the assumption that there exists $c > 0$ such that $\phi(t) \geq ct$ holds as t goes to 0_+. Let $d(y) := \max_{1 \leq i \leq m} \|y - T_i(y)\|$. Fixing any $y \in B$, we let $z := P_F(y)$ and note from $F := \bigcap_{i=1}^m \operatorname{Fix} T_i$ that $z = T_i(z)$ holds for all

$i \in \{1, \ldots, m\}$, which together with the nonexpansiveness of T_i (see Lemma 2.1) implies that for all i,

$$\|y - T_i(y)\| = \|y - z + z - T_i(y)\| \leq 2\|y - z\| = 2\,\mathrm{dist}(y, F).$$

This together with the arbitrariness of i and the definition of ψ_B further implies that

$$d(y)/2 \leq \mathrm{dist}(y, F) \leq \psi_B(d(y)), \ \forall\ y \in B.$$

By assumption, B is closed, so it contains the limiting point x^* which satisfies $d(x^*) = 0$. Also because the convergence is assumed to be not finite, no x^k can be a common fixed point. In particular, since B contains the sequence $\{x^k\}$, there exists at least another $\bar{y} \in B$ with $d(\bar{y}) > 0$. Now, $d(\cdot)$ is a continuous function and B is connected so $d(B)$ is a connected subset of $\mathbb{R}$, i.e., $d(B)$ is an interval. In particular, over B, d assumes all values between 0 and $d(\bar{y})$. This tells us that

$$t/2 \leq \psi_B(t)$$

holds for all sufficiently small t. Letting $\kappa \coloneqq \sqrt{2\alpha(1 + 4\nu s)/(\nu(1 - \alpha))}$ and recalling the definition of ϕ, we have

$$\sqrt{\phi(t)} = \psi_B(\kappa\sqrt{t}) \geq \kappa\sqrt{t}/2,$$

for small t. This implies that $\phi(t) \geq \kappa^2 t/4$ holds for small t, which is what we wanted to show. $\square$

The three cases in Theorem 3.11 can be interpreted as follows. When $\rho = 1$, the convergence of the quasi-cyclic algorithm is *almost linear*, which means that, for any r, the iterates of algorithm converges to a fixed point faster than k^{-r} goes to 0 as $k \to \infty$. This includes cases where the convergence is, in fact, linear but it also includes the possibility that the convergence is slower than linear (i.e., slower than c^{-k} for any c) as observed empirically in [24, Figure 1]. When $\rho \in (0, 1)$, the convergence rate is at least sublinear and is faster than s^{-r} for all r with $0 < r < -\rho/2(\rho - 1)$, which it means that the converge rate is *almost* the rate that would be afforded if ψ_B were "purely Hölderian" of the form t^ρ as we will discuss in Section 4.3.

The case $\rho = 0$ is the least informative, which only tells us that Φ_ϕ^{-1} is going to 0 slower than any sublinear rate, where we recall that Φ_ϕ^{-1} gives an *upper bound* on the true convergence rate. So, in essence, we are getting a lower bound to an upper bound, which is only helpful if we suspect that the upper bound is actually tight. In order to get more information on the case $\rho = 0$, we need extra assumptions as in Theorem 3.14.

3.3 Tighter rates

In this section, we discuss a tool to obtain tighter rates than the one described in Theorem 3.10. The drawback is that although we do not need to compute the integral appearing in (3.6), we still need to be able to say something about a certain function g that will appear in Theorem 3.14. We start with the following lemma about preservation of asymptotic equivalence under the arrow inverse.

Lemma 3.12. *Suppose that $f_1 : [a, \infty) \to (0, \infty)$ and $f_2 : [b, \infty) \to (0, \infty)$ are measurable locally bounded functions such that $f_1 \in \mathrm{RV}_\rho$ with $\rho > 0$. If $f_1(t) \sim f_2(t)$ as $t \to \infty$, then $f_1^{\leftarrow}(t) \sim f_2^{\leftarrow}(t)$ as $t \to \infty$. Similarly, if $f_1 \overset{c}{\sim} f_2$, then $f_1^{\leftarrow} \overset{c}{\sim} f_2^{\leftarrow}$.*

Proof. Let $\ell(t) \coloneqq f_1(t)/t^\rho$, so that $\ell \in \mathrm{RV}_0$. By assumption, $f_2 \sim f_1$ holds, so we have $f_2(t) \sim t^\rho \ell(t)$ as well. Then, [6, Proposition 1.5.15] implies that $f_1^{\leftarrow}$ and $f_2^{\leftarrow}$ both satisfy

$$f_1^{\leftarrow}(t) \sim t^{1/\rho}\ell^\sharp(t^{1/\rho}) \text{ and } f_2^{\leftarrow}(t) \sim t^{1/\rho}\ell^\sharp(t^{1/\rho}),$$

for a certain function $\ell^\sharp \in \mathrm{RV}_0$, which leads to $f_1^{\leftarrow} \sim f_2^{\leftarrow}$.

Now, if $f_1 \overset{c}{\sim} f_2$, then there exists a positive constant μ such that $f_1 \sim \mu f_2$. By what we just proved, we have $f_1^{\leftarrow} \sim (\mu f_2)^{\leftarrow}$ and, by the definition of the arrow inverse (2.14), $(\mu f_2)^{\leftarrow}(t) = f_2^{\leftarrow}(t/\mu)$ holds. Since, $f_2^{\leftarrow} \in \mathrm{RV}_{1/\rho}$, the definition of regular variation implies $f_2^{\leftarrow}(t/\mu)/f_2^{\leftarrow}(t) \to (1/\mu)^{1/\rho}$ as $t \to \infty$ so that $f_2^{\leftarrow}(t/\mu) \overset{c}{\sim} f_2^{\leftarrow}(t)$ as $t \to \infty$. We conclude that $f_1^{\leftarrow} \overset{c}{\sim} f_2^{\leftarrow}$ holds. ☐

If $f,\, g \in \mathrm{RV}_\rho^0$ are measurable, locally bounded away from zero and satisfy $f \sim g$ then, letting $\bar{f} \coloneqq 1/f(1/\cdot)$ and $\bar{g} \coloneqq 1/g(1/\cdot)$, the functions $\bar{f},\, \bar{g}$ are locally bounded, belong to RV_ρ and satisfy $\bar{f} \sim \bar{g}$. Recalling (2.16) and applying Lemma 3.12 to $\bar{f},\, \bar{g}$, we conclude that $f^- \sim g^-$. With that, we conclude that, under the stated hypothesis, if $f,\, g : (0,\, a] \to (0,\, \infty)$ are such that $f,\, g \in \mathrm{RV}_\rho^0$ with $\rho > 0$ then

$$f \overset{c}{\sim} g \Rightarrow f^- \overset{c}{\sim} g^-. \tag{3.27}$$

The following lemma is inspired by [14, Theorem 1], but here we relax the monotonicity assumption imposed therein.

Lemma 3.13. *Suppose that $f_1 : [a,\, \infty) \to (0,\, \infty)$ and $f_2 : [b,\, \infty) \to (0,\, \infty)$ are locally bounded functions such that $f_1 \in \mathrm{RV}_\rho$ with $\rho \geq 0$, $f_1(t) \to \infty$, $f_2(t) \to \infty$ and $f_1(t) = o(f_2(t))$ as $t \to \infty$, then $f_2^{\leftarrow}(t) = o(f_1^{\leftarrow}(t))$ as $t \to \infty$.*

Proof. Because $f_1(t) = o(f_2(t))$ holds, for every $\alpha \in (0,\, 1)$, there exists $t_\alpha \geq \max\{a,\, b\}$ such that

$$f_1(t) \leq \alpha f_2(t), \qquad \forall t \geq t_\alpha.$$

Now, because f_1 is locally bounded, $y_\alpha \coloneqq 1 + \sup\{f_1(t) \mid a \leq t \leq t_\alpha\}$ is finite. In particular, if $f_1(t) > y$ holds with $y \geq y_\alpha$, we must have $t > t_\alpha$, which further leads to $\alpha f_2(t) \geq f_1(t) > y$. Consequently, for $y \geq y_\alpha$, the inequality $f_1(t) > y$ implies $f_2(t) > \frac{y}{\alpha}$. In view of the definition of the arrow inverse (2.14), we have that if $y \geq y_\alpha$, then

$$\inf\{t \in [a,\, \infty) \mid f_1(t) > y\} = f_1^{\leftarrow}(y) \geq f_2^{\leftarrow}(y/\alpha) = \inf\{t \in [b,\, \infty) \mid f_2(t) > y/\alpha\}.$$

Put otherwise, $f_1^{\leftarrow}(\alpha u) \geq f_2^{\leftarrow}(u)$ holds for sufficiently large u which leads to

$$\limsup_{u\to\infty} \frac{f_2^{\leftarrow}(u)}{f_1^{\leftarrow}(u)} \leq \limsup_{u\to\infty} \frac{f_1^{\leftarrow}(\alpha u)}{f_1^{\leftarrow}(u)}.$$

Next, we divide in two cases.

$\boxed{\rho = 0}$ Since $f_1 \in \mathrm{RV}_0$, we have $f_1^{\leftarrow} \in \mathrm{RV}_\infty$ (see (2.21) and (2.18)) which implies that the lim sup on the right-hand-side is zero for $\alpha \in (0,1)$ (see Definition 2.2). That is, $f_2^{\leftarrow}(t) = o(f_1^{\leftarrow}(t))$ holds.

$\boxed{\rho > 0}$ We have $f_1^{\leftarrow} \in \mathrm{RV}_{1/\rho}$ (see (2.17)). By the definition of regular variation, we have

$$\limsup_{u\to\infty} \frac{f_2^{\leftarrow}(u)}{f_1^{\leftarrow}(u)} \leq \limsup_{u\to\infty} \frac{f_1^{\leftarrow}(\alpha u)}{f_1^{\leftarrow}(u)} = \alpha^{1/\rho}$$

and since $\alpha \in (0,\, 1)$ is arbitrary, we conclude that the lim sup on the left-hand-side is zero and $f_2^{\leftarrow}(t) = o(f_1^{\leftarrow}(t))$ holds. ☐

We are now positioned to state the main result of this subsection. In what follows, we use the notation $f(s) = 1/o(g(s))$ as $s \to \infty$ to indicate that $1/(f(s)g(s)) \to 0$ as $s \to \infty$.

Theorem 3.14. *Suppose that $f : (0,\, a] \to (0,\, \infty) \in \mathrm{RV}_\rho^0$ with $\rho \in [0,\, 1]$ is nondecreasing and $\lim_{x\to 0_+} f(x) = 0$. Let $g(x) \coloneqq \frac{1}{xf^-(1/x)}$ for $x \in [1/\delta,\, \infty)$ and Φ_f be defined as in (3.6). Then g is locally bounded. Moreover, the following statements hold.*

(i) *If* $\rho = 0$ *and* $\ln(g) \in \mathrm{RV}_q$ *with* $q > 0$ *then letting* $\alpha \in \mathbb{R}$ *and* $\widehat{g}(x) := x^\alpha g(x)$*, we have* $\Phi_f^{-1}(s) \sim \frac{1}{\widehat{g}^{\leftarrow}(s)}$ *as* $s \to \infty$*.*

(ii) *If* $\rho \in (0, 1)$*, then* $\Phi_f^{-1}(s) \sim \frac{1}{g^{\leftarrow}\left((1/\rho - 1)s\right)}$ *as* $s \to \infty$*.*

(iii) *In case of* $\rho = 1$*: if* $f(t) \overset{c}{\sim} t$ *as* $t \to 0_+$*, then there exist* $\tau_1, \tau_2 > 0$ *and* $0 < c_1 < c_2 < 1$ *such that* $\tau_1 c_1^s \le \Phi_f^{-1}(s) \le \tau_2 c_2^s$ *whenever* s *is large enough; if* $t = o(f(t))$ *as* $t \to 0_+$ *then* $\Phi_f^{-1}(s) = \frac{1}{o(g^{\leftarrow}(s))}$ *as* $s \to \infty$*.*

Proof. Initially, we need some bookkeeping to verify that the several functions involved in this theorem satisfy certain desirable properties.

First, we see from (2.3), (2.4), (2.23) and (2.24) that $g \in \mathrm{RV}_{1/\rho - 1}$ when $\rho \in (0, 1]$ and $g \in \mathrm{RV}_\infty$ when $\rho = 0$. Also, since local boundedness is preserved by taking products and $1/x$ and $1/f^-(1/x)$ are locally bounded (both by monotonicity) over $[1/\delta, \infty)$, g is also locally bounded. Over $[1/\delta, \infty)$ the functions $1/x$ and $1/f^-(1/x)$ are positive and monotone, so they are also locally bounded away from zero. So, similarly, g is locally bounded away from zero over $[1/\delta, \infty)$.

We let $F(x) := \Phi_f(1/x)$ and then have

$$F(x) = \Phi_f(1/x) = \int_{1/x}^{\delta} \frac{1}{f^-(t)} dt = \int_{1/\delta}^{x} \frac{1}{s^2 f^-(1/s)} ds = \int_{1/\delta}^{x} \frac{g(s)}{s} ds. \tag{3.28}$$

Since $\lim_{x \to 0_+} f(x) = 0$, we invoke Lemma 3.6 to conclude that Φ_f is continuous and decreasing. This together with $F(\cdot) = \Phi_f(1/\cdot)$ implies that F is continuous and increasing.

We now consider (3.28) in three cases as follows.

$\boxed{\rho = 0 \text{ and } \ln(g) \in \mathrm{RV}_q \text{ with } q > 0.}$ In this case, first we observe that the same argument that showed that g is locally bounded and locally bounded away from zero also applies to $\widehat{g}$. This implies that $\ln(\widehat{g})$ is locally bounded[6].

Next, let $h(x) := \ln(g(x)) - \ln(x)$. Since $\ln(g) \in \mathrm{RV}_q$ with $q > 0$ holds, we have $\ln(x)/\ln(g(x)) \to 0$ as $x \to \infty$, which follows from (2.4) and (2.9). Similarly, $\alpha \ln(x)/\ln(g(x)) \to 0$ as $x \to \infty$. This implies that

$$h(x) \sim \ln(g(x)) \sim \ln(\widehat{g}(x)) \quad \text{as} \quad x \to \infty. \tag{3.29}$$

Since $\ln(g(x)) \to \infty$ as $x \to \infty$ (see (2.8)), (3.29) implies that there exists $\widehat{a} > 1/\delta$ such that $h(x)$ is positive over $[\widehat{a}, \infty)$ and the restriction of h to $[\widehat{a}, \infty)$ belongs to RV_q.

Letting $b := \int_{1/\delta}^{\widehat{a}} e^{h(s)} ds$ and recalling (3.28), we apply [24, Lemma 5.10] to conclude that

$$\ln(F(x) - b) = \ln\left(\int_{1/\delta}^{x} e^{h(s)} ds - b\right) = \ln\left(\int_{\widehat{a}}^{x} e^{h(s)} ds\right) \sim h(x) \quad \text{as} \quad x \to \infty. \tag{3.30}$$

This implies that $F(x) \to \infty$ as $x \to \infty$. Consequently,

$$\lim_{x \to \infty} \frac{\ln(F(x) - b)}{\ln(F(x))} = \lim_{x \to \infty} \frac{\ln(F(x)) + \ln(1 - b/F(x))}{\ln(F(x))} = 1 + \lim_{x \to \infty} \frac{\ln(1 - b/F(x))}{\ln(F(x))} = 1$$

holds, namely, $\ln(F(x) - b) \sim \ln(F(x))$ holds. From (3.29) and (3.30) we obtain

$$\ln(F(x)) \sim \ln(F(x) - b) \sim h(x) \sim \ln(\widehat{g}(x)) \quad \text{as} \quad x \to \infty. \tag{3.31}$$

Since $\ln(\widehat{g}(x)) \to \infty$ as $x \to \infty$, the same is true of $\ln(F(x))$ and there exists $\tilde{a} > 1/\delta$ such that both functions are positive on $[\tilde{a}, \infty)$. Let

$$F_r(\cdot) := \ln(F(\cdot))|_{[\tilde{a}, \infty)}, \qquad \widehat{g}_r(\cdot) := \ln(\widehat{g}(\cdot))|_{[\tilde{a}, \infty)}. \tag{3.32}$$

[6] Direct from definition. Over any compact set C, there are positive constants M_1, M_2 such that $M_1 \le g(t) \le M_2$ for $t \in C$. Therefore, $\ln(M_1) \le \ln(g(t)) \le \ln(M_2)$, so that $|\ln(g(t))| \le \max\{|\ln(M_1)|, |\ln(M_2)|\}$.

We then have from (3.31) and $\ln(\widehat{g}) \in \mathrm{RV}_q$ that $F_r(x) \sim \widehat{g}_r(x)$ as $x \to \infty$ and $\widehat{g}_r \in \mathrm{RV}_q$ with $q > 0$. In addition, F_r is locally bounded (due to the continuity of F), and $\widehat{g}_r$ is locally bounded, thanks to the local boundedness of $\ln(\widehat{g})$. We apply Lemma 3.12 to the two functions in (3.32) and obtain

$$F_r^{\leftarrow}(s) \sim \widehat{g}_r^{\leftarrow}(s) \quad \text{as} \quad x \to \infty. \tag{3.33}$$

We recall that both $\ln(F(x))$ and $\ln(\widehat{g}(x))$ go to ∞ as $x \to \infty$ and F_r is continuous and increasing. We apply Proposition 2.4 to them and see from (3.32) and (3.33) that

$$F^{-1}(e^s) = (\ln(F))^{-1}(s) = (\ln(F))^{\leftarrow}(s) = F_r^{\leftarrow}(s) \sim \widehat{g}_r^{\leftarrow}(s) = (\ln(\widehat{g}))^{\leftarrow}(s) = \widehat{g}^{\leftarrow}(e^s) \quad \text{as} \quad s \to \infty,$$

where the first equation follows from the fact that F is invertible and the last equation follows from the definition of arrow generalized inverse. This further gives

$$\Phi_f^{-1}(s) = 1/F^{-1}(s) \sim 1/\widehat{g}^{\leftarrow}(s) \quad \text{as} \quad s \to \infty,$$

which proves item (i).

$\boxed{\rho \in (0,\, 1)}$ In this case, we have $g \in \mathrm{RV}_{1/\rho - 1}$ with $1/\rho - 1 > 0$. Then, invoking (3.18) with g and $\sigma = -1$, we have from (3.28) that

$$F(x) \sim \frac{\rho}{1-\rho} g(x) \quad \text{as} \quad x \to \infty. \tag{3.34}$$

Recalling that g and F are locally bounded Applying Lemma 3.12 to (3.34) and Proposition 2.4 that

$$\Phi_f^{-1}(s) = 1/F^{-1}(s) \sim 1/\Big(\frac{\rho}{1-\rho} g\Big)^{\leftarrow}(s) = 1/g^{\leftarrow}\big((1/\rho - 1)s\big) \quad \text{as} \quad s \to \infty,$$

which proves statement (ii).

$\boxed{\rho = 1}$ We first prove the first half of the statement. Since f is nondecreasing, it is locally bounded away from zero, so $f(t) \overset{c}{\sim} t$ implies $f^{-}(t) \overset{c}{\sim} t$ as $t \to 0_+$. In particular, there exists $\mu > 0$ such that $1/f^{-}(1/s) \sim s/\mu$ as $s \to \infty$. This further implies

$$\frac{g(s)}{s} = \frac{1/f^{-}(1/s)}{s^2} \sim \frac{1}{\mu s} \quad \text{as} \quad s \to \infty.$$

Therefore, fix any $\epsilon \in (0,\, 1)$, there exists $M > 0$ such that for all $s \geq M$ we have

$$\frac{1-\epsilon}{\mu s} \leq \frac{g(s)}{s} \leq \frac{1+\epsilon}{\mu s}.$$

This together with (3.28) implies that when $x \geq M$,

$$F(x) = \int_{1/\delta}^{x} \frac{g(s)}{s} ds = \int_{1/\delta}^{M} \frac{g(s)}{s} ds + \int_{M}^{x} \frac{g(s)}{s} ds$$

Let $\kappa_1 := \int_{1/\delta}^{M} \frac{g(s)}{s} ds - (1-\epsilon)\ln(M)/\mu$ and $\kappa_2 = \int_{1/\delta}^{M} \frac{g(s)}{s} ds - (1+\epsilon)\ln(M)/\mu$. Then for $x \geq M$,

$$(1-\epsilon)\ln(x)/\mu + \kappa_1 \leq F(x) \leq (1+\epsilon)\ln(x)/\mu + \kappa_2. \tag{3.35}$$

Due to Lemma 3.6, we see that Φ_f and F are continuous, and F is increasing. We recall that if f_1, f_2 are increasing function satisfying $f_1 \leq f_2$ we have $f_2^{-1} \leq f_1^{-1}$. Applying this principle to (3.35), we conclude that whenever s is large enough, F^{-1} is sandwiched between the inverses of the functions appearing on the right-hand-side and the left-hand-side of (3.35), which leads to

$$\Phi_f^{-1}(s) = 1/F^{-1}(s) \in \left[\tau_1 c_1^s,\, \tau_2 c_2^s\right],$$

where $\tau_1 := e^{\mu\kappa_1/(1-\epsilon)}$, $\tau_2 := e^{\mu\kappa_2/(1+\epsilon)}$, $c_1 := e^{-\mu/(1-\epsilon)}$ and $c_2 := e^{-\mu/(1+\epsilon)}$. This proves the first half of statement (iii).

Now, it remains the case of $t = o(f(t))$ as $t \to 0_+$. In this case, we have $g \in \mathrm{RV}_0$. Using this, the local boundedness of g and the definition of F, we invoke [6, Proposition 1.5.9a] which tells us that

$$F \in \mathrm{RV}_0 \quad \text{and} \quad g(x) = o(F(x)) \;\; \text{as} \;\; x \to \infty. \tag{3.36}$$

Due to $t = o(f(t))$ as $t \to 0_+$, it holds that $\frac{1/f(1/x)}{x} = \frac{1/x}{f(1/x)} \to 0$ as $x \to \infty$. Namely, we have $1/f(1/x) = o(x)$. We note from the monotonicity of f and $\lim_{x\to 0_+} f(x) = 0$ that $1/f(1/x)$ is locally bounded and goes to ∞ as $x \to \infty$. Therefore, applying Lemma 3.13 to $1/f(1/x)$, x and recalling (2.16), we obtain

$$\lim_{x\to\infty} g(x) = \lim_{x\to\infty} \frac{1}{xf^-(1/x)} = \lim_{x\to\infty} \frac{1/f^-(1/x)}{x} = \lim_{x\to\infty} \frac{(1/f(1/\cdot))^{\leftarrow}(x)}{x} = \infty.$$

This together with (3.36) implies that $F(x) \to \infty$ as $x \to \infty$. Note that F is locally bounded (due to its continuity) and g is locally bounded. Applying Lemma 3.13 to g and F by recalling (3.36), we have

$$\Phi_f^{-1}(s) = 1/F^{-1}(s) = 1/o(g^{\leftarrow}(s)) \;\; \text{as} \;\; s \to \infty,$$

which proves the latter half of statement (iii). This completes the proof. ☐

Typically Theorem 3.14 would be invoked in the context of Theorem 3.8 with $f = \phi$. However, ϕ may have terms of different orders so it will be helpful to verify that we can focus on the important terms only, especially in item (ii).

Proposition 3.15. *Suppose that f and $\hat{f}$ satisfy the assumptions in Theorem 3.10 and $f \overset{c}{\sim} \hat{f} \in \mathrm{RV}_\rho^0$ with $\rho \in (0, 1)$. Then we have $\Phi_f \overset{c}{\sim} \Phi_{\hat{f}}$ and $\Phi_f^{-1} \overset{c}{\sim} \Phi_{\hat{f}}^{-1}$.*

Proof. By (2.23) and (3.27), we have $f^-, \hat{f}^- \in \mathrm{RV}_{1/\rho}^0$ and

$$f^- \overset{c}{\sim} \hat{f}^-.$$

Now, $1/f^-$ and $1/\hat{f}^-$ both belong to $\mathrm{RV}_{-1/\rho}^0$. Moreover, thanks to Lemma 3.5, $1/f^-$ and $1/\hat{f}^-$ are monotone and therefore locally bounded away from zero. Applying (3.20)[7] (i.e., Karamata's theorem) to both Φ_f and $\Phi_{\hat{f}}$ we have

$$\Phi_f \overset{c}{\sim} \frac{y}{f^-(y)} \overset{c}{\sim} \frac{y}{\hat{f}^-(y)} \overset{c}{\sim} \Phi_{\hat{f}} \in \mathrm{RV}_{1-1/\rho}^0.$$

Defining F and $\hat{F}$ by $F(x) := \Phi_f(1/x)$ and $\hat{F}(x) := \Phi_{\hat{f}}(1/x)$, we have $F, \hat{F} \in \mathrm{RV}_{1/\rho-1}$ and $1/\rho - 1 > 0$ because $\rho \in (0, 1)$. Then, we have $F \overset{c}{\sim} \hat{F}$. By Lemma 3.6, F and $\hat{F}$ are increasing, continuous and both go to ∞ as $x \to \infty$, since $\rho \in (0,1)$. By Lemma 3.12 and Proposition 2.4, we conclude that $F^{-1} \overset{c}{\sim} \hat{F}^{-1}$. Since $F^{-1} = 1/\Phi_f^{-1}$ and $\hat{F}^{-1} = 1/\Phi_{\hat{f}}^{-1}$, we conclude that $\Phi_f^{-1} \overset{c}{\sim} \Phi_{\hat{f}}^{-1}$. ☐

4 Applications and examples

In this section we take a look at some examples of non-Hölderian behavior that is covered by the definition of Karamata regularity. In sections 4.1 and 4.2 we will focus on new results that are obtainable based on the techniques developed in this paper and we will contrast with the results described in [10] and [24].

[7] We restrict the domain of $1/f^-$ and $1/\hat{f}^-$ to $(0, \delta]$.

4.1 Exotic error bounds and convergence rate of alternating projections

First, we will see examples of error bounds among convex sets. We will consider two closed convex sets C_1, C_2 with non-empty intersection and apply Definition 3.1 to the projection operators $L_1 \coloneqq P_{C_1}$, $L_2 \coloneqq P_{C_2}$. Then, in each case we examine the corresponding convergence rate of the alternating projections (AP) algorithm for solving the common fixed point problem

$$\text{find} \quad x \in C \coloneqq (\operatorname{Fix} L_1) \cap (\operatorname{Fix} L_2) = C_1 \cap C_2. \tag{4.1}$$

As a reminder, AP for (4.1) is obtained from the quasi-cyclic iteration described in (3.5) by letting $T_i \coloneqq L_i$, $w_i^k \coloneqq (k+i \mod 2)$, so that w_1^k is $(0,1,0,1,\ldots)$, w_2^k is $(1,0,1,0,\ldots)$ and therefore $\nu = 1$. With that, item (b) of Theorem 3.8 is satisfied with $s = 2$. Then, the assumption in item (a) of Theorem 3.8, i.e., joint Karamata regularity of P_{C_1}, P_{C_2} over a bounded set B containing $\{x^k\}$, corresponds to an error bound condition on C_1, C_2 over B, namely,

$$\operatorname{dist}(x,\, C) \leq \psi_B\Big(\max_{1\leq i\leq 2} \|x - L_i(x)\|\Big) = \psi_B\Big(\max_{1\leq i\leq 2} \operatorname{dist}(x,\, C_i)\Big), \quad \forall\, x \in B, \tag{4.2}$$

where ψ_B satisfies items *(ii)* and *(iii)* in Definition 3.1. Since projections are 1/2-averaged, the function $\hat{\phi}$ in Theorem 3.8 is of the form

$$\hat{\phi}(u) = \psi_B^2(\sqrt{18u}). \tag{4.3}$$

In summary, in order to estimate the convergence rate of AP applied to C_1 and C_2, we have to estimate the function $\Phi_{\hat{\phi}}^{-1}$. We will show how to do using the results developed so far.

We emphasize that there are many different algorithms that are covered by the quasi-cyclic algorithm and this was extensively showcased in [10]. Although initially we will focus on the particular case of AP, we recall that Theorems 3.10 and 3.14 apply in general to quasi-cyclic iterations, since the aforementioned results are also valid under the setting of Theorem 3.8. Our choice of using the AP algorithm is only to better emphasize the analysis technique. We will also focus on new results and insights that were not obtainable under our previous work [24]. In particular, later in Section 4.2 we will see an example of convergence rate for the DR algorithm, which is something that was not possible under the framework developed in [24] and is also beyond the analysis done in [10] since one of the sets is not semialgebraic.

Before we move on we recall some properties of the Lambert W function which denotes the converse relation of the function $f(w) \coloneqq we^w$. It has two real branches W_0 and W_{-1}. The principal branch $W_0(x)$ is continuous and increasing on its domain $[-e^{-1},\, \infty)$, and satisfies $W_0(x) \geq -1$, $W_0(0) = 0$, with $W_0(x) \to \infty$ as $x \to \infty$. The branch $W_{-1}(x)$ is continuous and decreasing on its domain $[-e^{-1},\, 0)$, and satisfies $W_{-1}(x) \leq -1$, $W_{-1}(-e^{-1}) = -1$, with $W_{-1}(x) \to -\infty$ as $x \to 0_-$. More details on the Lambert W function and its applications to optimization can be seen on [11].

An example with $\rho \in (0,\, 1)$: a Hölder-entropic error bound First we will construct two convex sets satisfying a non-Hölderian error bound with index $\frac{1}{2}$. We start by letting the function $\gamma : [-0.5,\, 0.5] \to \mathbb{R}$ be such that $\gamma(0) \coloneqq 0$ and

$$\gamma(x) \coloneqq e^{2W_{-1}\left(-\frac{|x|}{2}\right)}, \qquad \forall\, x \in [-0.5,\, 0.5] \setminus \{0\}.$$

The inverse of the restriction of γ to $[0,\, 0.5]$ exists and, with a slight abuse of notation, we will denote it by γ^{-1}. The domain of γ^{-1} is $[0,\, \gamma(0.5)]$, where $\gamma(0.5) < 1$, and its expression on $(0,\, \gamma(0.5)]$ is given by

$$\gamma^{-1}(y) = -\sqrt{y}\ln(y). \tag{4.4}$$

The function γ^{-1} is increasing and concave, and satisfies $\gamma^{-1}(0) = 0$, which implies that for $x \geq 0$ and $\alpha > 0$ such that $(1+\alpha)x$ is in the domain of γ^{-1} we have

$$\gamma^{-1}(x) = \gamma^{-1}\Big(\frac{1}{1+\alpha} \cdot (1+\alpha)x + \frac{\alpha}{1+\alpha} \cdot 0\Big) \geq \frac{\gamma^{-1}((1+\alpha)x)}{1+\alpha}. \tag{4.5}$$

Define

$$C_1 \coloneqq \{(x,\, \mu) \mid \gamma(x) \leq \mu\}, \quad C_2 \coloneqq \{(x,\, 0) \mid x \in \mathbb{R}\}. \tag{4.6}$$

Let $C \coloneqq C_1 \cap C_2$, with that we have $C = \{(0,\, 0)\}$. Our goal is to show the following theorem.

Theorem 4.1 (A Hölder-entropic error bound)**.** *Let C_1, C_2 be defined as in (4.6) and let $B_b \coloneqq \{(x,\, \mu) \mid |x| + |\mu| \leq b\}$. Then for small enough $b > 0$ there exists $\kappa > 0$ such that*

$$\operatorname{dist}(w,\, C) \leq \kappa\, \gamma^{-1}\Big(\max_{1 \leq i \leq 2} \operatorname{dist}(w,\, C_i)\Big), \quad \forall\, w \in B_b, \tag{4.7}$$

where γ^{-1} is defined as in (4.4). Furthermore, γ^{-1} belongs to $\mathrm{RV}^0_{1/2}$ and (4.7) is optimal in the following sense: over those B_b with b sufficiently small, C_1, C_2 do not satisfy a Hölderian error bound with exponent $1/2$ nor (4.7) holds if γ^{-1} is substituted with an RV^0_ρ function with $\rho > 1/2$. [8]

Proof. First, we can extend γ from its domain $[-0.5,\, 0.5]$ to $(-2e^{-1},\, 2e^{-1})$: let $\widehat{\gamma} : (-2e^{-1},\, 2e^{-1}) \to \mathbb{R}$ be such that $\widehat{\gamma}(0) \coloneqq 0$ and

$$\widehat{\gamma}(x) \coloneqq e^{2W_{-1}\left(-\frac{|x|}{2}\right)}, \qquad \forall\, x \in (-2e^{-1},\, 2e^{-1}) \setminus \{0\}.$$

Note from the properties of W_{-1} on $(-e^{-1},\, 0)$ that $\widehat{\gamma}$ is continuous and decreasing on $(-2e^{-1},\, 0)$. Hence, the inverse of the restriction of $\widehat{\gamma}$ to $(-2e^{-1},\, 0)$ exists and, with a slight abuse of notation, we call it $\widehat{\gamma}^{-1}$. By calculation we have $\widehat{\gamma}^{-1}(y) = \sqrt{y}\ln(y)$ for $y \in (0,\, e^{-2})$, which is convex and decreasing. By [27, Proposition 2] we have $\widehat{\gamma}$ is convex and decreasing on $(-2e^{-1},\, 0)$. Since $\widehat{\gamma}(x) = \widehat{\gamma}(-x)$ and $\widehat{\gamma}$ is continuous at 0, we have that $\widehat{\gamma}$ is convex on its domain[9]. Given that γ is the restriction of $\widehat{\gamma}$ on $[-0.5,\, 0.5]$ and recalling that a convex function is Lipschitz continuous over any compact set contained in the interior of its domain (e.g., [17, Theorem 3.1.1]), we conclude that γ is Lipschitz continuous on its domain.

Let L denote the Lipschitz constant of γ on its domain, and define $\widehat{L} \coloneqq \max\{L,\, 1\}$, $\widehat{b} \coloneqq \frac{\gamma(0.5)}{2\sqrt{2}\widehat{L}}$. For any $w = (x,\, \mu) \in B_{\widehat{b}}$,

$$2\sqrt{2}\widehat{L} \max_{1 \leq i \leq 2} \operatorname{dist}(w,\, C_i) \leq 2\sqrt{2}\widehat{L} \operatorname{dist}(w,\, C) \leq 2\sqrt{2}\widehat{L}(|x| + |\mu|) \leq 2\sqrt{2}\widehat{L}\widehat{b} = \gamma(0.5). \tag{4.8}$$

Moreover, let $(\bar{x},\, \gamma(\bar{x}))$ denote the projection of $(x,\, 0)$ onto C_1[10], in particular $\bar{x} \in \operatorname{dom}\gamma$. From $\widehat{L} \geq 1$ and $\gamma(0.5) < 1$ we have $\widehat{b} \leq 0.5$, and since $w \in B_{\widehat{b}}$, we have $x,\, \mu \in \operatorname{dom}\gamma$.

---

[8] That means, one cannot expect to obtain an error bound function which improves the current one γ^{-1} by either increasing its exponent or keeping the exponent but removing the slowly varying term $-\ln(\cdot)$.

[9] It suffices to prove that the epigraph $S \coloneqq \{(x,\, u) \mid \widehat{\gamma}(x) \leq u\}$ is convex, i.e., for any $(x,\, u)$, $(y,\, v) \in S$ we have $\widehat{\gamma}(\alpha x + (1-\alpha)y) \leq \alpha u + (1-\alpha)v$ holds for all $\alpha \in (0,\, 1)$. Without loss of generality, we assume that $x < 0 < y$ and $z \coloneqq \alpha x + (1-\alpha)y < 0$. The convexity of $\widehat{\gamma}$ on $(-2e^{-1},\, 0)$, together with $\widehat{\gamma}(0) = 0$ and its continuity at 0, gives $\widehat{\gamma}(z) \leq \frac{z}{x}\widehat{\gamma}(x)$. Combing this with $\frac{z}{x} \leq \frac{z-y}{x-y} = \alpha$, we obtain $\widehat{\gamma}(z) \leq \alpha\widehat{\gamma}(x) \leq \alpha u \leq \alpha u + (1-\alpha)v$, which completes the proof.

[10] If the projection were of the form $(\bar{x},\, \bar{\mu})$ with $\gamma(\bar{x}) < \bar{\mu}$, we would be able to get a point closer to $(x,\, 0)$ by replacing $\bar{\mu}$ with $\gamma(\bar{x})$.

Using the Lipschitz continuity of γ on its domain, we then obtain

$$\begin{aligned}
\gamma(x) &\le L|x-\bar{x}| + \gamma(\bar{x}) \le \widehat{L}(|x-\bar{x}| + \gamma(\bar{x})) \le \sqrt{2}\widehat{L}\,\mathrm{dist}\left((x,\,0),\,C_1\right) \\
&\le \sqrt{2}\widehat{L}\left(\mathrm{dist}(w,\,C_1) + \|w-(x,\,0)\|\right) = \sqrt{2}\widehat{L}\left(\mathrm{dist}(w,\,C_1) + \mathrm{dist}(w,\,C_2)\right) \\
&\le 2\sqrt{2}\widehat{L}\max_{1\le i\le 2}\mathrm{dist}(w,\,C_i), \\
\gamma(\mu) &\le L|\mu| \le \widehat{L}|\mu| = \widehat{L}\,\mathrm{dist}(w,\,C_2) \le \widehat{L}\max_{1\le i\le 2}\mathrm{dist}(w,\,C_i).
\end{aligned} \tag{4.9}$$

Due to (4.8), the two right-hand side terms of (4.9) are in the domain of γ^{-1}. Thus, for each inequality in (4.9), we can apply γ^{-1} at both sides obtain

$$|x| \le \gamma^{-1}\Big(2\sqrt{2}\widehat{L}\max_{1\le i\le 2}\mathrm{dist}(w,\,C_i)\Big), \quad |\mu| \le \gamma^{-1}\Big(\widehat{L}\max_{1\le i\le 2}\mathrm{dist}(w,\,C_i)\Big).$$

This together with (4.5) further implies

$$\begin{aligned}
\mathrm{dist}(w,\,C) \le |x| + |\mu| &\le \gamma^{-1}\Big(2\sqrt{2}\widehat{L}\max_{1\le i\le 2}\mathrm{dist}(w,\,C_i)\Big) + \gamma^{-1}\Big(\widehat{L}\max_{1\le i\le 2}\mathrm{dist}(w,\,C_i)\Big) \\
&\le \big(2\sqrt{2}+1\big)\widehat{L}\gamma^{-1}\Big(\max_{1\le i\le 2}\mathrm{dist}(w,\,C_i)\Big).
\end{aligned}$$

This means that (4.7) holds for any $b \in (0,\,\widehat{b})$ and $\kappa := \big(2\sqrt{2}+1\big)\,\widehat{L}$.

Next, we show that the error bound we have in (4.7) is optimal in the sense that no error bound function in RV^0 with an index greater than $1/2$ is admissible. We will also show that no Hölderian error bound with exponent $1/2$ holds for C_1 and C_2.

First, suppose that $\psi \in \mathrm{RV}^0_\rho$ is a nondecreasing function such that for sufficiently small $b > 0$,

$$\mathrm{dist}(w,\,C) \le \psi\Big(\max_{1\le i\le 2}\mathrm{dist}(w,\,C_i)\Big), \qquad \forall\, w = (x,\,\mu) \in B_b.$$

Let $w_k := (t_k,\,0)$ with $t_k \to 0_+$ and $t_k \in (0,\,\min(b,\,0.5))$. Then we would have

$$t_k = \mathrm{dist}(w_k,\,C) \le \psi\left(\mathrm{dist}((t_k,\,0),\,C_1)\right) \le \psi\left(\|(t_k,\,0) - (t_k,\,\gamma(t_k))\|\right) = \psi\left(\gamma(t_k)\right). \tag{4.10}$$

Since $\gamma^{-1} \in \mathrm{RV}^0_{1/2}$, the restriction of γ to $(0,\,0.5]$ belongs to RV^0_2 by (2.23). With that, the composition $\psi\circ\gamma$ belongs to RV^0 and has index 2ρ, see (2.5). If $\rho > 1/2$, we then have $\frac{t}{\psi(\gamma(t))} \in \mathrm{RV}^0$ with a negative index by (2.5). By (2.11), this implies that $t/\psi(\gamma(t)) \to +\infty$ as $t \to 0_+$, which contradicts (4.10). Therefore, ρ must be in $[0,\,1/2]$.

Also, if there exists a Hölderian error bound with exponent $1/2$ for C_1 and C_2, then (4.10) holds with $\psi(t) = c\cdot t^{1/2}$, where $c > 0$. Consequently, recalling that $e^{W_{-1}(t)} = t/W_{-1}(t)$ holds, we have $t_k/\psi(\gamma(t_k)) = 2|W_{-1}(-|t_k|/2)|/c \to \infty$, which contradicts (4.10). This completes the proof. ☐

Next, we consider the problem of estimating the convergence rate of the AP algorithm when applied to C_1 and C_2 as defined in (4.6). Denote the sequence generated by the AP algorithm by $\{x^k\}$. Let B be a bounded set containing $\{x^k\}$ and $r > 0$ be such that $B \subseteq \mathbb{B}_r$. Let $b > 0$ and $\kappa > 0$ be given as in Theorem 4.1 such that (4.7) holds. One can see that there exists some $c \in (0,\,e^{-2})$ such that $w \in B_b$ whenever $w \in B$ and $\max_{1\le i\le 2}\mathrm{dist}(w,\,C_i) \le c$ hold[11]. With that, we

[11] Suppose that such a constant c does not exist. Then we can construct a sequence $\{c_k\}$ with $c_k \to 0$, a sequence $\{w^k\} \subseteq B$ satisfying $\max_{1\le i\le 2}\mathrm{dist}(w^k,\,C_i) \le c_k$ but $w^k \notin B_b$, for all k. Since B is bounded, without loss of generality, we may assume that w^k converges to some w^*, which further leads to $w^* \in C_1 \cap C_2 = (0,\,0)$. Since B_b contains a neighborhood of $(0,\,0)$, we have $w^k \in B_b$ for large k which is a contradiction.

let ψ_B be as follows

$$\psi_B(t) := \begin{cases} 0 & \text{if} \ \ t = 0, \\ -\kappa\sqrt{t}\ln(t) & \text{if} \ \ 0 < t \le c, \\ \max\{r,\ -\kappa\sqrt{c}\ln(c)\} & \text{if} \ \ t > c. \end{cases} \tag{4.11}$$

The function ψ_B satisfies items *(ii)* and *(iii)* of Definition 3.1. Now, we show that error bound condition (4.2) holds. Given any $x \in B$, we consider three cases: if $\max_{1\le i\le 2} \operatorname{dist}(x, C_i) = 0$, then $x \in C$ and hence (4.2) holds, thanks to $\psi_B(0) = 0$; if $0 < \max_{1\le i\le 2} \operatorname{dist}(x, C_i) \le c$, then $x \in B_b$, which together with (4.7) and the second case in (4.11) implies that (4.2) holds; if $\max_{1\le i\le 2} \operatorname{dist}(x, C_i) > c$, we see from $C = \{(0, 0)\}$, $x \in B \subseteq \mathbb{B}_r$ and the third case in (4.11) that (4.2) holds.

From what is discussed above, we then have that $T_1 = P_{C_1}$ and $T_2 = P_{C_2}$ are jointly Karamata regular over B with regularity function ψ_B defined as in (4.11). We will analyze the convergence rate as follows. We will compute the function ϕ appearing in Theorem 3.8 and use Theorem 3.14 to estimate the asymptotic properties of Φ_ϕ^{-1}. Recalling (4.3) and the relation between ϕ and $\hat{\phi}$ in (3.7), we observe that when t is small enough,

$$\phi(t) = \psi_B^2(\sqrt{18t}) = 3\sqrt{2}\kappa^2\sqrt{t}(\ln(\sqrt{18t}))^2 \overset{c}{\sim} \sqrt{t}(\ln(t))^2 \in \mathrm{RV}^0_{1/2}. \tag{4.12}$$

Let $\hat{f}(t) := \sqrt{t}(\ln(t))^2$. Note that both ϕ and $\hat{f}$ are continuous and increasing on $(0, a]$ for some small enough a. After restricting them on $(0, a]$, we have that both $\phi(t)$ and $\hat{f}(t)$ satisfy the assumptions in Theorem 3.10. On the other hand, let $\hat{g}(s) := \frac{1}{s\hat{f}^-(1/s)}$ for $s \in [1/\delta, \infty)$. Now, we first apply Proposition 3.15 to ϕ and $\hat{f}$ by recalling (4.12), and then apply Theorem 3.14 (ii) by letting $f = \hat{f}$ and $g = \hat{g}$, and therefore obtain

$$\sqrt{\Phi_\phi^{-1}(s)} \overset{c}{\sim} \sqrt{\Phi_{\hat{f}}^{-1}(s)} \sim \sqrt{\frac{1}{\hat{g}^{\leftarrow}(s)}} \quad \text{as} \quad s \to \infty. \tag{4.13}$$

That is, instead of evaluating $\sqrt{\Phi_\phi^{-1}(s)}$ directly which may be quite cumbersome, we may estimate the convergence rate through the (relatively) simpler expression $\sqrt{\frac{1}{\hat{g}^{\leftarrow}(s)}}$. So our next task is evaluating $\hat{g}^{\leftarrow}(s)$.

Since $\hat{f}$ is continuous and increasing on its domain $(0, a]$, its usual inverse exists with $\hat{f}^{-1}(u) = \hat{f}^-(u)$ for small enough positive u. Let $h := \hat{f}^{-1}$. For small enough u, we have

$$u = \hat{f}(h(u)) = \sqrt{h(u)}\,(\ln(h(u)))^2. \tag{4.14}$$

Let

$$z_u := \frac{\ln(h(u))}{4}.$$

Hence, $z_u \to -\infty$ as $u \to 0_+$. We also have $e^{z_u} = h(u)^{1/4}$. Then (4.14) together with $z_u < 0$ as $u \to 0_+$ implies that

$$-\frac{\sqrt{u}}{4} = z_u e^{z_u}, \tag{4.15}$$

i.e., for sufficiently small positive u,

$$W_{-1}\left(-\frac{\sqrt{u}}{4}\right) = z_u,$$

where we note that this must be indeed the W_{-1} branch because the W_0 branch is lower bounded by -1 and z_u goes to $-\infty$ as $u \to 0_+$. Consequently, using (4.15) we have for sufficiently small positive u,

$$h(u) = (e^{z_u})^4 = \left(-\frac{\sqrt{u}}{4z_u}\right)^4 = \frac{u^2}{256z_u^4} = \frac{u^2}{256\big[W_{-1}(-\frac{\sqrt{u}}{4})\big]^4}.$$

Therefore, for large enough s, we have

$$\hat{g}(s) = \frac{1}{s\hat{f}^{-}(1/s)} = \frac{1}{sh(1/s)} = 256s\left[W_{-1}\left(-\frac{1}{4\sqrt{s}}\right)\right]^4. \tag{4.16}$$

By the properties of W_{-1}, we have that $\hat{g}$ is increasing and continuous on $[M, \infty)$ for some large enough $M > 1/\delta$ and $\hat{g}(s) \to \infty$ as $s \to \infty$. We note from Theorem 3.14 that $\hat{g}$ is locally bounded. By Proposition 2.4, $\hat{g}^{\leftarrow}(s) = \hat{g}^{-1}(s)$ holds for large enough s. Let $w(s) := \hat{g}^{-1}(s)$ for large enough s. Let $t_s := W_{-1}\big(-\frac{1}{4\sqrt{w(s)}}\big)$. Then we obtain from (4.16) that

$$s = \hat{g}(w(s)) = 256w(s)t_s^4. \tag{4.17}$$

By definition of the Lambert W function, we have $t_s e^{t_s} = -\frac{1}{4\sqrt{w(s)}}$, that is, $\sqrt{w(s)} = -\frac{e^{-t_s}}{4t_s}$. Since $t_s < 0$, in combination with (4.17), we obtain

$$\frac{\sqrt{s}}{4} = 4\left(-\frac{e^{-t_s}}{4t_s}\right) \cdot \sqrt{t_s^4} = -t_s e^{-t_s},$$

which implies that

$$W_0\left(\frac{\sqrt{s}}{4}\right) = -t_s.$$

This together with (4.13) and (4.17) implies that

$$\sqrt{\Phi_\phi^{-1}(s)} \overset{c}{\sim} \sqrt{\frac{1}{\hat{g}^{\leftarrow}(s)}} = \sqrt{\frac{1}{\hat{g}^{-1}(s)}} = \sqrt{\frac{1}{w(s)}} = \sqrt{\frac{256t_s^4}{s}} = \frac{16t_s^2}{\sqrt{s}} = \frac{16\big[W_0(\frac{\sqrt{s}}{4})\big]^2}{\sqrt{s}} \quad \text{as} \quad s \to \infty.$$

Now, se^s is a rapidly varying function, so W_0 is slowly varying by (2.22), which implies that $W_0(\lambda s) \sim W_0(s)$ for every $\lambda > 0$. Therefore,

$$\sqrt{\Phi_\phi^{-1}(s)} \overset{c}{\sim} \frac{\big[W_0(\sqrt{s})\big]^2}{\sqrt{s}} \quad \text{as} \quad s \to \infty. \tag{4.18}$$

This implies that when the AP algorithm is used to find a feasible point in the intersection of C_1 and C_2 defined as in (4.6), the sequence $\{x^k\}$ converges at least at a rate proportional to $\frac{[W_0(\sqrt{k})]^2}{\sqrt{k}}$. Using our previous techniques in [24] we would only be able to conclude that, up to multiplicative constants, the convergence rate is at least as fast as $(1/k)^{\frac{1}{2}-\varepsilon}$ for all positive ε satisfying $1/2 - \varepsilon > 0$, see [24, item (ii) Theorem 5.7]. Notably, ε cannot be taken to be zero in the context of [24, Theorem 5.7].

The rate in (4.18), however, is more explicit and gives the asymptotic equivalence class of the estimate $\sqrt{\Phi_\phi^{-1}(k)}$. It is interesting to note that $\big[W_0(\sqrt{s})\big]^2/\sqrt{s}$ does not correspond to a sublinear rate, however, we can see from

$$\lim_{s\to\infty} \frac{\big[W_0(\sqrt{s})\big]^2/\sqrt{s}}{1/\sqrt{s}} = \infty \quad \text{and} \quad \lim_{s\to\infty} \frac{\big[W_0(\sqrt{s})\big]^2/\sqrt{s}}{(1/s)^{1/2-\varepsilon}} = 0$$

that $\big[W_0(\sqrt{s})\big]^2/\sqrt{s}$ is squeezed between $1/\sqrt{s}$ and $(1/s)^{1/2-\varepsilon}$ for all positive ε such that $1/2-\varepsilon > 0$. In this way, (4.18) gives a finer evaluation of the convergence rate.

An example with $\rho = 1$: Better estimates for the exponential cone We denote by $K_{\exp}$ the exponential cone which is given by

$$K_{\exp} := \left\{x = (x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_2 > 0, x_3 \geq x_2 e^{x_1/x_2}\right\} \cup \{x = (x_1, x_2, x_3) \mid x_1 \leq 0, x_3 \geq 0, x_2 = 0\}.$$

As shown in [23, Section 4.2.1] and [24, Section 6.2], a so-called *entropic error bound* holds between $C_1 := K_{\exp}$ and $C_2 := \{x \in \mathbb{R}^3 \mid x_2 = 0\}$. More precisely, there exists a constant $\kappa_B > 0$ and a function ψ_B such that (4.2) holds and ψ_B satisfies

$$\psi_B(a) = -\kappa_B a \ln(a) \tag{4.19}$$

for sufficiently small a.

Consider the AP algorithm applied to the feasibility problem of finding $x \in C_1 \cap C_2$. We previously analyzed the convergence rate in [24, Section 6.2] and we were able to conclude the rate is faster than any sublinear rate, but the estimate is worse than any linear rate. This was also observed empirically, see Figure 1 in [24]. Here, however, we obtain a finer result that indicates where the rate is located in the gap between linear and sublinear rates.

Continuing our analysis, from (4.19) and (4.3), we conclude that the function ϕ in Theorem 3.8 takes the form

$$\phi(t) = \psi_B^2(\sqrt{18t}) = \kappa t \left(\ln(t) + \ln(18)\right)^2 \overset{c}{\sim} t(\ln(t))^2 \in \mathrm{RV}_1^0,$$

for sufficiently small t and $\kappa := \frac{9}{2}\kappa_B^2$. Note that $t = o(\phi(t))$ as $t \to 0_+$. Moreover, ϕ is continuous and increasing on $(0,\, a]$ for some small enough a. So, we restrict ϕ on $(0,\, a]$. Let $g(s) := \frac{1}{s\phi^{-}(1/s)}$ for $s \in [1/\delta,\, \infty)$. Applying the second half of Theorem 3.14 (iii) and letting $f := \phi$, we obtain

$$\sqrt{\Phi_\phi^{-1}(s)} = \frac{1}{o\left(\sqrt{g^{\leftarrow}(s)}\right)} \quad \text{as} \quad s \to \infty. \tag{4.20}$$

Let $h(s) := \phi^{-1}(1/s)$ be defined for large enough s. Since ϕ is increasing and continuous on $(0,\, a]$, h is continuous and decreasing for large enough s, e.g., see [16, Remark 1 and Proposition 1]. Moreover, for large s, it holds that $\phi^{-1}(1/s) = \phi^{-}(1/s)$ and

$$1/s = \phi(h(s)) = \kappa\, h(s) \left(\ln(h(s)) + \ln(18)\right)^2.$$

Consequently, for large enough s,

$$g(s) = \frac{1}{s\phi^{-}(1/s)} = \frac{1}{s\phi^{-1}(1/s)} = \frac{1/s}{h(s)} = \kappa \left(\ln(h(s)) + \ln(18)\right)^2. \tag{4.21}$$

Note that h is decreasing and $h(s) \to 0_+$ as $s \to \infty$. This implies that g is increasing and continuous on $[M,\, \infty)$ for some large enough $M > 1/\delta$. Moreover, $g(s) \to \infty$ as $s \to \infty$. Note from Theorem 3.14 that g is locally bounded. By Proposition 2.4, we conclude that $g^{\leftarrow}(s) = g^{-1}(s)$ holds for large enough s. Define

$$w(s) := g^{-1}(s)$$

for large enough s. From (4.21) we have

$$s = g(w(s)) = \kappa \left(\ln\left(h(w(s))\right) + \ln(18)\right)^2. \tag{4.22}$$

Note that when $s \to \infty$, we have $w(s) \to \infty$, $h(w(s)) \to 0_+$ and therefore $\ln(h(w(s))) \to -\infty$. Recalling that $h^{-1}(u) = \frac{1}{\phi(u)}$ holds for small enough u, we solve (4.22) in terms of $w(s)$ and conclude that for large enough s we have

$$w(s) = h^{-1}\left(e^{-\sqrt{\frac{s}{\kappa}} - \ln(18)}\right) = \frac{1}{\phi\left(e^{-\sqrt{\frac{s}{\kappa}} - \ln(18)}\right)} = \frac{18}{se^{-\sqrt{\frac{s}{\kappa}}}}.$$

Let $c := e^{\frac{1}{2\sqrt{\kappa}}}$. We then have $c > 1$. Combining the above equation with (4.20) and recalling that for large s we have $w(s) = g^{-1}(s) = g^{\leftarrow}(s)$, we obtain

$$\sqrt{\Phi_\phi^{-1}(s)} = \frac{1}{o\left(\sqrt{18}/\left(\sqrt{s}c^{-\sqrt{s}}\right)\right)} \quad \text{as} \quad s \to \infty.$$

This is further equivalent to

$$\sqrt{s}c^{-\sqrt{s}} = o\left(\sqrt{\Phi_\phi^{-1}(s)}\right) \quad \text{as} \quad s \to \infty. \tag{4.23}$$

As a reminder, we already knew from [24, Theorem 4.7 and Proposition 6.9 (iii)] that the convergence rate of AP applied to C_1 and C_2 is faster than any sublinear rate and the estimate obtained therein is slower than any linear rate. We called such a rate *almost linear*. The relation (4.23) refines this result and tells us that the predicted rate is not only slower than any linear rate but, up to multiplicative constants, it is slower than $\sqrt{k}c^{-\sqrt{k}}$. This is a stronger statement because $\sqrt{k}c^{-\sqrt{k}}$ goes to zero slower than d^{-k} for any $d > 1$, i.e.,

$$\lim_{k\to\infty} \frac{d^{-k}}{\sqrt{k}c^{-\sqrt{k}}} = 0.$$

4.2 A Douglas-Rachford splitting method example

In this subsection, we will demonstrate an example of convergence rate for the DR algorithm. Specifically, we consider the two sets C_1 and C_2 defined in (4.6), i.e.,

$$C_1 \coloneqq \{(x,\, \mu) \mid \gamma(x) \le \mu\}, \quad C_2 \coloneqq \{(x,\, 0) \mid x \in \mathbb{R}\},$$

and apply the DR algorithm to C_1 and C_2 as follows:

$$w^{k+1} = T_{\rm DR}(w^k) \coloneqq w^k + P_{C_1}\left(2P_{C_2}(w^k) - w^k\right) - P_{C_2}(w^k). \tag{4.24}$$

Note that C_1 is not semialgebraic, so the convergence rate analysis in [10] is not applicable. However, we can obtain the convergence rate of $\{w^k\}$ by using the techniques developed in this paper.

To proceed, we first compute $\operatorname{Fix} T_{\rm DR}$. Using the definition of $\operatorname{Fix} T_{\rm DR}$ and the closed-form expression of $P_{C_2}(\cdot)$, we have

$$\begin{aligned} w \coloneqq (x,\, \mu) \in \operatorname{Fix} T_{\rm DR} &\iff P_{C_1}\left(2P_{C_2}(w) - w\right) = P_{C_2}(w) \\ &\iff P_{C_1}\left((x,\, -\mu)\right) = (x,\, 0) \\ &\iff x = 0,\, \mu \ge 0, \end{aligned}$$

where we check the third equivalence as follows. Suppose that $P_{C_1}((x,\, -\mu)) = (x,\, 0) = P_{C_2}(w)$ holds. Then, $(x,\, 0) \in C_1 \cap C_2$, therefore $x = 0$. Now, $\gamma(0) = 0$, so if $\mu < 0$ holds, then $(x,\, -\mu) = (0,\, -\mu) \in C_1$, which would lead to $P_{C_1}((x,\, -\mu)) = (x,\, -\mu)$. This contradicts with $P_{C_1}((x,\, -\mu)) = (x,\, 0)$, so μ must be nonnegative. Conversely, if $x = 0$ and $\mu \ge 0$, then $P_{C_2}(w) = (0,\, 0)$ and $P_{C_1}((0,\, -\mu)) = (0,\, 0)$ holds because γ is nonnegative. This proves that

$$\operatorname{Fix} T_{\rm DR} = \{(0,\, \mu) \mid \mu \ge 0\}. \tag{4.25}$$

We will now discuss how to apply Theorem 3.8 to $T_{\rm DR}$. The DR algorithm (4.24) can be obtained from the quasi-cyclic iteration described in (3.5) by letting $m = 1$, $T_1 = T_{\rm DR}$ and $w_i^k \equiv 1$ (thus $\nu = 1$). With that, item (b) of Theorem 3.8 is satisfied with $s = 1$.

In order to verify the assumption in item (a) of Theorem 3.8, we need to show the Karamata regularity of T_{DR} over a bounded set B containing $\{w^k\}$, namely, the existence of ψ_B satisfying the last two items of Definition 3.1 and the first item as below

$$\operatorname{dist}\big(w,\, \operatorname{Fix} T_{\mathrm{DR}}\big) \le \psi_B\Big(\, \|T_{\mathrm{DR}}(w) - w\|\,\Big), \quad \forall\, w \in B. \tag{4.26}$$

Once this is done and recalling that T_{DR} is 1/2-averaged [10, Lemma 4.1], the function $\hat{\phi}$ in Theorem 3.8 is of the form

$$\hat{\phi}(u) = \psi_B^2(\sqrt{10u}) \tag{4.27}$$

and ϕ is obtained by restricting $\hat{\phi}$ to some interval of the form $(0,\, a]$.

So the first step in order to obtain the convergence rate of the sequence $\{w^k\}$ is to compute ψ_B. For this goal, we will first show the following theorem. In what follows, we consider the following notation for $r > 0$:

$$\mathcal{B}_r := \big\{ w \in \mathbb{R}^2 \mid \operatorname{dist}(w,\, \operatorname{Fix} T_{\mathrm{DR}}) \le r \big\}\,.$$

Theorem 4.2. *Let T_{DR} be given in (4.24). There exist $r > 0$ and $\kappa > 0$ such that for all $w \in \mathcal{B}_r$,*

$$\operatorname{dist}\big(w,\, \operatorname{Fix} T_{\mathrm{DR}}\big) \le \kappa\gamma^{-1}\Big(\, \|T_{\mathrm{DR}}(w) - w\|\,\Big). \tag{4.28}$$

Proof. For any $w = (x,\, \mu)$, we have from (4.25) that

$$\begin{aligned} \operatorname{dist}\big(w,\, \operatorname{Fix} T_{\mathrm{DR}}\big) &= \|(x,\, \mu) - (0,\, \max(\mu,\, 0))\| = \|(x,\, \min(\mu,\, 0))\|\,, \\ \|T_{\mathrm{DR}}(w) - w\| &= \|P_{C_1}(2P_{C_2}(w) - w) - P_{C_2}(w)\| = \|P_{C_1}((x,\, -\mu)) - (x,\, 0)\|\,. \end{aligned} \tag{4.29}$$

Since $\lim_{t \to 0+} \frac{t}{\gamma^{-1}(t)} = \lim_{t \to 0+} \frac{t}{-\sqrt{t}\ln(t)} = 0$ and $\gamma^{-1}(0) = 0$, there exists some $c_1 > 0$ such that

$$t \le \gamma^{-1}(t), \qquad \forall\, t \in [0,\, c_1]. \tag{4.30}$$

Due to the continuity of γ and $\gamma(0) = 0$, there exists some $c_2 > 0$ such that

$$\sqrt{2}\widehat{L}\sqrt{4t^2 + \gamma^2(t)} \le \min\left(\gamma(0.5),\, c_1\right) \tag{4.31}$$

holds for all $t \in [0,\, c_2]$, where $\widehat{L} := \max(L,\, 1)$ and L is the Lipschitz constant of γ on its domain, as shown in Theorem 4.1. Let $r := \min\{c_1,\, c_2,\, \gamma(0.5),\, 0.5\}$. For any $w = (x,\, \mu) \in \mathcal{B}_r$, we consider the following two cases.

(a) $(x,\, -\mu) \in C_1$. In this case, $\gamma(x) \le -\mu$ holds. Therefore, we have $\mu \le 0$ and $\gamma(x) \le -\mu = |\mu|$. Furthermore, we note from (4.29) and $w \in \mathcal{B}_r$ that $\operatorname{dist}\big(w,\, \operatorname{Fix} T_{\mathrm{DR}}\big) = \|(x,\, \mu)\| \le r$. This implies $|\mu| \le r \le \gamma(0.5)$ and $|\mu| \le r \le c_1$. Consequently,

$$\operatorname{dist}\big(w,\, \operatorname{Fix} T_{\mathrm{DR}}\big) \le |x| + |\mu| \le \gamma^{-1}(|\mu|) + |\mu| \le 2\gamma^{-1}(|\mu|) = 2\gamma^{-1}\Big(\, \|T_{\mathrm{DR}}(w) - w\|\,\Big), \tag{4.32}$$

where the last inequality follows from (4.30) and the equality follows from (4.29).

(b) $(x,\, -\mu) \notin C_1$. In this case, we have $\gamma(x) > -\mu$. Note from (4.29) that $|x| \le \operatorname{dist}\big(w,\, \operatorname{Fix} T_{\mathrm{DR}}\big) \le r \le 0.5$. Let $(\bar{x},\, \bar{\mu})$ denote the projection of $(x,\, -\mu)$ onto C_1. Since γ is defined on $[-0.5,\, 0.5]$, $|x| \le 0.5$, and $\gamma(a) < \gamma(b)$ whenever $|a| < |b| \le 0.5$, we then have $|\bar{x}| \le |x| \le 0.5$ and $\bar{\mu} = \gamma(\bar{x})$.[12]

[12] Suppose that $|\bar{x}| > |x|$. We then have $\bar{\mu} \ge \gamma(\bar{x}) > \gamma(x)$, which together with $\gamma(x) > -\mu$ implies that $0 < \gamma(x) + \mu = \|(x,\, -\mu) - (x,\, \gamma(x))\| < \mu + \bar{\mu} < \|(x,\, -\mu) - (\bar{x},\, \bar{\mu})\|$. That means we can find $(x,\, \gamma(x)) \in C_1$ which is closer to $(x,\, -\mu)$ than $(\bar{x},\, \bar{\mu})$. This leads to a contradiction, thus we conclude that $|\bar{x}| \le |x|$. Next we prove that $\bar{\mu} = \gamma(\bar{x})$. If this does not hold, we must have $\bar{\mu} > \gamma(\bar{x})$ since $(\bar{x},\, \bar{\mu}) \in C_1$. Now, we consider two cases. If $-\mu \le \gamma(\bar{x})$, then $0 \le \gamma(\bar{x}) + \mu < \bar{\mu} + \mu$. Thus, the point $(\bar{x},\, \gamma(\bar{x})) \in C_1$ satisfies $\|(x,\, -\mu) - (\bar{x},\, \gamma(\bar{x}))\| < \|(x,\, -\mu) - (\bar{x},\, \bar{\mu})\|$, which leads to a contradiction. If $-\mu > \gamma(\bar{x})$, since $-\mu < \gamma(x)$ and γ is continuous with $\gamma(u) = \gamma(-u)$ for $u \in [-0.5,\, 0.5]$, we can find some $\widehat{x}$ with the same sign with x such that $\gamma(\widehat{x}) = -\mu$. We then have $\gamma(\bar{x}) < \gamma(\widehat{x}) < \gamma(x)$, so $|\bar{x}| < |\widehat{x}| < |x|$. Thus, the point $(\widehat{x},\, \gamma(\widehat{x})) \in C_1$ satisfies $\|(x,\, -\mu) - (\widehat{x},\, \gamma(\widehat{x}))\| = |x - \widehat{x}| < |x - \bar{x}| \le \|(x,\, -\mu) - (\bar{x},\, \bar{\mu})\|$, which leads to a contradiction. This proves $\bar{\mu} = \gamma(\bar{x})$.

We then obtain from the Lipschitz continuity of γ on its domain that

$$\gamma(x) \le \gamma(\bar{x}) + L|x-\bar{x}| \le \widehat{L}\left(|x-\bar{x}| + \gamma(\bar{x})\right) \le \sqrt{2}\widehat{L}\,\|(x-\bar{x},\,\gamma(\bar{x}))\|\,. \tag{4.33}$$

Moreover, we see from $|\bar{x}| \le |x| \le r \le c_2$ that

$$\sqrt{2}\widehat{L}\,\|(x-\bar{x},\,\gamma(\bar{x}))\| \le \sqrt{2}\widehat{L}\sqrt{4r^2+\gamma^2(r)} \le \min\left(\gamma(0.5),\,c_1\right), \tag{4.34}$$

where the last inequality follows from (4.31). This means that the right-hand-side of (4.33) is in the domain of γ^{-1}. Then, applying γ^{-1} at both sides of (4.33) and invoking (4.5), we obtain

$$|x| \le \gamma^{-1}\left(\sqrt{2}\widehat{L}\,\|(x-\bar{x},\,\gamma(\bar{x}))\|\right) \le \sqrt{2}\widehat{L}\gamma^{-1}\left(\|(x-\bar{x},\,\gamma(\bar{x}))\|\right). \tag{4.35}$$

Next, we consider two cases. If $\mu \ge 0$, combining (4.29) and (4.35) we obtain

$$\operatorname{dist}\big(w,\,\operatorname{Fix}T_{\mathrm{DR}}\big) = |x| \le \sqrt{2}\widehat{L}\gamma^{-1}\left(\|(x-\bar{x},\,\gamma(\bar{x}))\|\right) = \sqrt{2}\widehat{L}\gamma^{-1}\left(\|T_{\mathrm{DR}}(w)-w\|\right). \tag{4.36}$$

Finally, suppose that $\mu < 0$. We obtain from $\widehat{L} \ge 1$, (4.30) and (4.34) that

$$\sqrt{2}\widehat{L}\,\|(x-\bar{x},\,\gamma(\bar{x}))\| \le \sqrt{2}\widehat{L}\gamma^{-1}(\|(x-\bar{x},\,\gamma(\bar{x}))\|). \tag{4.37}$$

We have from $(x,\,-\mu) \notin C_1$ that $|\mu| = -\mu < \gamma(x)$. In view of (4.33) and (4.37), we obtain

$$|\mu| < \gamma(x) \le \sqrt{2}\widehat{L}\,\|(x-\bar{x},\,\gamma(\bar{x}))\| \le \sqrt{2}\widehat{L}\gamma^{-1}(\|(x-\bar{x},\,\gamma(\bar{x}))\|). \tag{4.38}$$

Then, (4.29), (4.35), (4.38) imply that

$$\begin{aligned}\operatorname{dist}\big(w,\,\operatorname{Fix}T_{\mathrm{DR}}\big) &\le |x| + |\mu| \\ &\le 2\sqrt{2}\widehat{L}\gamma^{-1}\left(\|(x-\bar{x},\,\gamma(\bar{x}))\|\right) = 2\sqrt{2}\widehat{L}\gamma^{-1}\left(\|T_{\mathrm{DR}}(w)-w\|\right).\end{aligned} \tag{4.39}$$

Finally, (4.32), (4.36) and (4.39) taken together imply that (4.28) holds with $\kappa = 2\sqrt{2}\widehat{L}$. This completes the proof. ☐

Next, we will use Theorem 4.2 to obtain ψ_B and further estimate the convergence rate of $\{w^k\}$ in (4.24). Let B be a bounded set containing $\{w^k\}$ and R be such that $B \subseteq \mathcal{B}_R$. Let $r > 0$ and $\kappa > 0$ be given as in Theorem 4.2. One can see that there exists some $c \in (0,\,e^{-2})$ such that $w \in \mathcal{B}_r$ whenever $w \in B$ and $\|T_{\mathrm{DR}}(w) - w\| \le c$ hold[13]. Let

$$\psi_B(t) := \begin{cases} 0 & \text{if } \ t = 0, \\ -\kappa\sqrt{t}\ln(t) & \text{if } \ 0 < t \le c, \\ \max\{R,\,-\kappa\sqrt{c}\ln(c)\} & \text{if } \ t > c. \end{cases} \tag{4.40}$$

With that, ψ_B satisfies items *(ii)* and *(iii)* of Definition 3.1. Now we check item *(i)* of Definition 3.1, i.e., condition (4.26). Indeed, given any $w \in B$, we consider three cases: if $\|T_{\mathrm{DR}}(w) - w\| = 0$, then $w \in \operatorname{Fix}T_{\mathrm{DR}}$ and hence (4.26) holds, thanks to $\psi_B(0) = 0$; if $0 < \|T_{\mathrm{DR}}(w) - w\| \le c$, then $w \in \mathcal{B}_r$, which together with (4.28) and the second case in (4.40) implies that (4.26) holds; if $\|T_{\mathrm{DR}}(w) - w\| > c$, we see from $w \in B \subseteq \mathcal{B}_R$ and the third case in (4.40) that (4.26) holds.

[13]Suppose that such a constant c does not exists. Then we can construct a sequence $\{c_k\}$ with $c_k \to 0$, a sequence $\{u^k\} \subseteq B$ satisfying $\|T_{\mathrm{DR}}(u^k) - u^k\| \le c_k$ but $u^k \notin \mathcal{B}_r$. Since B is bounded, without loss of generality, we may assume that u^k converges to some u^*. Now, T_{DR} is continuous, so $\|T_{\mathrm{DR}}(u^k) - u^k\| \le c_k$ and $c_k \to 0$ leads to $u^* \in \operatorname{Fix}T_{\mathrm{DR}}$. Now, $\mathcal{B}_r$ contains a neighbourhood of $\operatorname{Fix}T_{\mathrm{DR}}$ and, in particular, a neighbourhood of u^*, so $u^k \in \mathcal{B}_r$ for large k which is a contradiction.

Recalling (4.27), we observe that when t is small enough,

$$\phi(t) = \hat{\phi}(t) = \psi_B^2(\sqrt{10t}) = \sqrt{10}\kappa^2\sqrt{t}(\ln(\sqrt{10t}))^2 \overset{c}{\sim} \sqrt{t}(\ln(t))^2 \in \mathrm{RV}^0_{1/2}.$$

So we are back to the same situation as in (4.12). Following the same exact computations that culminate in (4.18), we then obtain

$$\sqrt{\Phi_\phi^{-1}(s)} \overset{c}{\sim} \frac{\left[W_0(\sqrt{s})\right]^2}{\sqrt{s}} \quad \text{as} \quad s \to \infty. \tag{4.41}$$

As a reminder, the set C_1 defined in (4.6) is not semialgebraic and hence the convergence rate of the DR algorithm (4.24) cannot be obtained from [10, Proposition 4.1], which requires all sets involved to be semialgebraic. To the best of our knowledge, the rate obtained in (4.41) is new for the DR algorithm.

4.3 Recovering previous results

In this quick subsection we reap some extra fruits of the theory developed so far.

Sublinear/linear rates under Hölder regularity. We briefly sketch how to recover the main convergence rate result in [10] as follows. For the iteration in (3.5) suppose that the T_j's are bounded Hölder regular with uniform exponent $\gamma_{1,j} \in (0,\, 1]$ and the intersection of the fixed point sets is nonempty and has a Hölderian error bound of uniform exponent $\gamma_2 \in (0,\, 1]$, see Remark 3.3 for a review of these notions. Let $\gamma_1 \coloneqq \min_{1\le j\le m} \gamma_{1,j}$, $\gamma \coloneqq \gamma_1\gamma_2$ and suppose that we are under the assumptions of Theorem 3.8.

The fact that the intersection of the fixed point sets has a Hölderian error bound of uniform exponent $\gamma_2 \in (0,\, 1]$ implies that they have a consistent error bound function of the form $\Phi(a,\, b) \coloneqq \sigma(b)a^{\gamma_2}$ for some some nondecreasing function σ, see [24, Theorem 3.5]. Also, each T_j is uniformly Karamata regular and, given a bounded set B, T_j has a regularity function of the form $\Gamma_B^j(a) = \kappa_B a^{\gamma_{1,j}}$ for some $\kappa_B > 0$.

Then, applying Proposition 3.4, we conclude that the T_j are uniformly jointly Karamata regular and the regularity function ψ_B can be taken to be asymptotically equivalent (up to a constant) to the composition of $\sigma(b)a^{\gamma_2}$ with the sum of all the Γ_B^j's. By the regular variation calculus rules we have $\psi_B \in \mathrm{RV}^0_\gamma$ (see (2.5)) and $\psi_B(t) \overset{c}{\sim} t^\gamma$ as $t \to 0_+$, since, intuitively, only the terms with smallest exponent matter as $t \to 0_+$. Similarly, the ϕ in Theorem 3.8 also satisfies $\phi(t) \in \mathrm{RV}^0_\gamma$ and $\phi(t) \overset{c}{\sim} t^\gamma$ as $t \to 0_+$.

Next, we consider two cases. First, if $\gamma \in (0,\, 1)$, then applying item (ii) of Theorem 3.14 and Proposition 3.15, we conclude that $\Phi_\phi^{-1}(s) \overset{c}{\sim} \frac{1}{g^{\leftarrow}\left((1/\gamma-1)s\right)}$ as $s \to \infty$, where $g(t) \coloneqq \frac{1}{t\sqrt[\gamma]{1/t}} = t^{\frac{1-\gamma}{\gamma}}$. Now, g is invertible over $(0,\, \infty)$, so we get $g^{\leftarrow}(s) = g^{-1}(s) = s^{\frac{\gamma}{1-\gamma}}$ for large s by Proposition 2.4. Overall, we obtain

$$\sqrt{\Phi_\phi^{-1}(k)} \overset{c}{\sim} k^{-\frac{\gamma}{2(1-\gamma)}} \quad \text{as} \quad k \to \infty. \tag{4.42}$$

If $\gamma = 1$, then $\phi(t) \overset{c}{\sim} t$ as $t \to 0_+$ so the first half of item (iii) of Theorem 3.14 implies a linear convergence rate. Both rates match what is given in [10, Theorem 3.1].

AP and logarithmic error bounds We say that convex sets $C_1, \ldots, C_m$ have a logarithmic error bound with exponent γ if they admit a consistent error bound function Φ such that for every $b > 0$, there exist $\kappa_b > 0$ and $a_b > 0$ with $\Phi(a,\, b) = \kappa_b\left(-\frac{1}{\ln(a)}\right)^\gamma$ for $a \in (0,\, a_b)$, see [24, Definition 5.8].

This corresponds to a very pathological kind of error bound that is worse than any Hölderian error bound, see [24, Example 5.9] for a family of examples. Another example of logarithmic error bound was found in [23, Section 4.2.3] in the study of exponential cones, a highly expressive class of cones that can be used to model convex problems that require the exponential and logarithm functions, see [12] and [26, Chapter 5]. In particular, a logarithmic error bound of exponent 1 holds for the exponential cone and a certain subspace, see [24, Section 6.2] for more details. Other examples of logarithmic error bounds in the context of optimization over log-determinant cones can also be found in [22].

For simplicity, let C_1 and C_2 be two closed convex sets which satisfy a logarithmic consistent error bound Φ with exponent $\gamma > 0$ and consider the method of alternating projections. Let B be a bounded set containing the iterates. Then the operators P_{C_1} and P_{C_2} are jointly Karamata regular over B, and the regularity function ψ_B can be taken to be such that $\psi_B(u) = \Phi(u, b)$ for some large b (see Remark 3.3), which implies that $\psi_B(u) = \kappa\left(-\frac{1}{\ln(u)}\right)^{\gamma}$ for small u and some constant $\kappa > 0$. Recalling (4.3) and the relation between ϕ and $\hat{\phi}$ in (3.7), we observe that when u is small enough,

$$\phi(u) = \psi_B^2(\sqrt{18u}) = \kappa^2\left(-\frac{1}{\ln(\sqrt{18u})}\right)^{2\gamma} \overset{c}{\sim} \left(-\frac{1}{\ln(u)}\right)^{2\gamma} \in \mathrm{RV}_0^0. \tag{4.43}$$

Note that ϕ is continuous and increasing on $(0, a]$ for some small enough a. So we can restrict ϕ on $(0, a]$. Let $g(s) := \frac{1}{s\phi^{-}(1/s)}$ and $\widehat{g}(s) := sg(s)$ for $s \in [1/\delta, \infty)$. Then $\widehat{g}(s) = \frac{1}{\phi^{-}(1/s)} = \frac{1}{\phi^{-1}(1/s)}$ holds for large enough s. We see from this and the explicit expression of ϕ in (4.43) that $\widehat{g}$ is continuous and increasing on $[M, \infty)$ for some large enough $M > 1/\delta$. Moreover, due to $\lim_{t\to 0_+} \phi(t) = 0$, we have $\widehat{g}(s) \to \infty$ as $s \to \infty$. By Theorem 3.14, we have that g is locally bounded and thus so is $\widehat{g}$. Consequently, $\widehat{g}^{-1}(s) = \widehat{g}^{\leftarrow}(s)$ for large enough s, thanks to Proposition 2.4. Now, for large enough s we let $y_s := \phi^{-1}(1/s)$. Then, $1/s = \phi(y_s) = \psi_B^2(\sqrt{18y_s})$, which implies that $1/\sqrt{s} = \psi_B(\sqrt{18y_s}) = \kappa\left(-\frac{1}{\ln(\sqrt{18y_s})}\right)^{\gamma}$. Solving this to obtain $\ln(y_s)$, we have

$$\ln(y_s) = -\ln(18) - 2\kappa^{1/\gamma}s^{1/(2\gamma)}$$

for large s, which leads to

$$\ln(g(s)) = \ln(\widehat{g}(s)) - \ln(s) = -\ln\left(\phi^{-1}(1/s)\right) - \ln(s) = -\ln(y_s) - \ln(s) \in \mathrm{RV}_{\frac{1}{2\gamma}}.$$

Applying Theorem 3.14 (i) by letting $f = \phi$ and $\alpha = 1$, we have

$$\sqrt{\Phi_\phi^{-1}(s)} \sim \sqrt{\frac{1}{\widehat{g}^{\leftarrow}(s)}} = \sqrt{\frac{1}{\widehat{g}^{-1}(s)}} = \sqrt{\phi(1/s)} \overset{c}{\sim} \left(\frac{1}{\ln(s)}\right)^{\gamma} \quad \text{as} \quad s \to \infty, \tag{4.44}$$

where the last equivalence follows from (4.43). The rate in (4.44) matches what is given in [24, Theorem 5.12].

5 Definable operators and jointly Karamata regular operators

In this section, we further explore the class of jointly Karamata regular operators proposed in Definition 3.1. Our main goal is to show that operators that are definable in some o-minimal structure are always jointly Karamata regular, provided that their fixed points intersect.

Next, we recall the definition of o-minimal structure, which is somewhat technical. That said, the reader may take heart from the fact that, besides some basic notions, the sole tool we need from this body of theory is the so-called *monotonicity lemma*. In view of this situation, we defer more detailed explanations to [7, Section 4], [2, pg. 452] or to [32]. We define o-minimal structures and definable set/functions following [7, Section 4].

Definition 5.1 (o-minimal structure and definable sets/functions)**.** *An o-minimal structure on* $(\mathbb{R}, +, .)$ *is a sequence of Boolean algebras $\mathcal{O}_n$ of subsets of $\mathbb{R}^n$ such that for each n we have*

(i) if A belongs to $\mathcal{O}_n$ then $A \times \mathbb{R}$ and $\mathbb{R} \times A$ belong to $\mathcal{O}_{n+1}$;

(ii) if $\pi : \mathbb{R}^{n+1} \to \mathbb{R}^n$ is the projection map such that $\pi(x_1, \ldots, x_{n+1}) = (x_1, \ldots, x_n)$, then $\pi(A) \in \mathcal{O}_n$ for all $A \in \mathcal{O}_{n+1}$;

(iii) $\mathcal{O}_n$ contains the family of algebraic subsets of $\mathbb{R}^n$, i.e., every set of the form $\{x \in \mathbb{R}^n \mid p(x) = 0\}$ where $p : \mathbb{R}^n \to \mathbb{R}$ is a polynomial function;

(iv) the elements of $\mathcal{O}_1$ are exactly the finite unions of intervals and points.

A subset $A \in \mathbb{R}^n$ is said to be definable *if $A \in \mathcal{O}_n$. A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is said to be* definable *if its graph $\{(y, z) \in \mathbb{R}^n \times \mathbb{R}^m \mid z = f(y)\}$ belongs to $\mathcal{O}_{n+m}$. An extended-valued function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is definable if its graph $\{(y, \alpha) \in \mathbb{R}^n \times \mathbb{R} \mid f(y) = \alpha\}$ is in $\mathcal{O}_{n+1}$.*

Key examples of o-minimal structures are given by the semialgebraic sets and the globally subanalytics sets. Here we recall that a set is semialgebraic if it can be written as a finite union of solution sets of polynomial equalities and polynomial inequalities. In particular, the fact that the coordinate projection of a semialgebraic set is also a semialgebraic (item (ii) of Definition 5.1) corresponds to the celebrated Tarski-Seidenberg theorem.

Many important optimization problems are defined over semialgebraic sets and this is useful in several ways. For example, as a consequence of the so-called Łojasiewicz inequality, the error bounds governing the intersection of convex semialgebraic sets can always be taken to be Hölderian. Similarly, if the underlying operators describing an algorithm are semialgebraic, they must be bounded Hölder regular. For a proof, see, for example, [10, Proposition 4.1].

Unfortunately, the exponential function is not semialgebraic nor global subanalytic. Therefore, many problems that require exponentials and logarithms need to be defined over a larger o-minimal structure such as the *log-exp structure*, which contains the aforementioned structures together with the graph of the exponential function.

A troublesome aspect of the log-exp structure is that we can no longer expect that the underlying definable sets/functions have Hölder-like properties. Indeed, in the study of the error bounds appearing in problems defined over the exponential cone there is an example of an intersection having no Hölderian error bounds, see [23, Example 4.20]. A similar phenomenon occurs in the context of log-determinant cones [22].

In spite of this difficulty, in this section our goal is to show that continuous quasi-nonexpansive definable operators in any o-minimal structure are jointly Karamata regular as in Definition 3.1. This serves as a counterpart to the fact that Hölderian behavior can no longer be expected over general o-minimal structures and also shows that theory developed in this paper applies quite generally.

We start with a natural extension of [24, Proposition 3.3], which essentially says that there is always some function that describes the joint level of regularity of some given operators. Its proof follows the same line of arguments as [24, Proposition 3.3], but we show the details here for the sake of self-containment.

Lemma 5.2. *Let $L_i : \mathcal{E} \to \mathcal{E}$ $(i = 1, \dots, n)$ be continuous operators such that $\bigcap_{i=1}^n \operatorname{Fix} L_i \neq \emptyset$. For all $a,\, b \geq 0$, let $\Omega_{a,b} := \left\{ y \in \mathcal{E} \mid \max\limits_{1 \leq i \leq n} \|y - L_i(y)\| \leq a,\, \|y\| \leq b \right\}$. Define*

$$\Phi(a,\, b) := \begin{cases} \max\limits_{y \in \Omega_{a,b}} \operatorname{dist}(y,\, \bigcap_{i=1}^n \operatorname{Fix} L_i) & if\ \Omega_{a,b} \neq \emptyset; \\ 0 & otherwise. \end{cases} \tag{5.1}$$

Then Φ satisfies the following conditions:

(i) *for any $b \geq 0$, function $\Phi(\cdot,\, b)$ is nondecreasing, $\lim\limits_{a \to 0+} \Phi(a,\, b) = 0$ and $\Phi(0,\, b) = 0$;*

(ii) *for any $a \geq 0$, function $\Phi(a,\, \cdot)$ is nondecreasing;*

(iii) *for any $x \in \mathcal{E}$, we have $\operatorname{dist}(x,\, \bigcap_{i=1}^n \operatorname{Fix} L_i) \leq \Phi\Big(\max\limits_{1 \leq i \leq n} \|x - L_i(x)\|,\, \|x\| \Big)$.*

Proof. Except for the continuity requirement in item (i), all the other properties of Φ in items (i), (ii), (iii) follow directly from the definition of Φ. We only need to show that $\lim_{a \to 0+} \Phi(a,\, b) = 0$ holds for every $b \geq 0$.

Suppose on the contrary that there exist some $\delta > 0$ and a sequence $\{a_k\} \subseteq [0, \infty)$ converging to 0 such that $\Phi(a_k,\, b) \geq \delta$ holds for every k. We then see from (5.1) that $\Omega_{a_k,b} \neq \emptyset$. By the continuity of each L_i, we have that $\Omega_{a_k,b}$ is compact. Since the distance function to $\bigcap_{i=1}^n \operatorname{Fix} L_i$ is continuous, for each a_k, there exists $y^k \in \Omega_{a_k,b}$ such that

$$\operatorname{dist}\Big(y^k,\, \bigcap_{i=1}^n \operatorname{Fix} L_i\Big) = \Phi(a_k,\, b) \geq \delta.$$

Note that $\{y^k\}$ is bounded. Then there exists a subsequence $\{y^{k_j}\}$ which converges to some limit $\bar{y}$. Since $a_k \to 0$, we see from $\max\limits_{1 \leq i \leq n} \|y^k - L_i(y^k)\| \leq a_k$ that $\bar{y} \in \bigcap_{i=1}^n \operatorname{Fix} L_i$. Consequently, $\operatorname{dist}(y^{k_j},\, \bigcap_{i=1}^n \operatorname{Fix} L_i) \to \operatorname{dist}(\bar{y},\, \bigcap_{i=1}^n \operatorname{Fix} L_i) = 0$, which contradicts the fact that $\operatorname{dist}(y^{k_j},\, \bigcap_{i=1}^n \operatorname{Fix} L_i) = \Phi(a_{k_j},\, b) \geq \delta$ should hold for every j. This proves the continuity requirement in item (i) and completes the proof. $\square$

Next, we state formally the monotonicity lemma, which, in particular, ensures that a real definable function defined over an interval cannot oscillate infinitely over its domain.

Lemma 5.3 (Monotonicity lemma, [32, Chapter 3])**.** *Let $f : (a,\, b) \to \mathbb{R}$ be a definable function in some o-minimal structure. Then, there exists a finite subdivision $a = a_0 < a_1 < \ldots < a_k = b$ such that on each subinterval $(a_i,\, a_{i+1})$, f is continuous and either constant or strictly monotone (i.e., increasing or decreasing).*

Before we move on to the main theorem of this section, we recall some basic properties of definable sets and functions, see [31, Section 2.1], [13, Section 1.3] and [19] for more details. In summary, the class of definable functions/sets is remarkably stable: if f, g are definable functions in some *o*-minimal structure, then whenever the functions $f \pm g$, $f \circ g$, f^{-1}, f/g are well-defined, they must be definable over the same *o*-minimal structure. In addition, finite unions and intersections of definable sets are definable as well. Naturally, inverse images of definable sets through definable functions are definable as well.

A very powerful technique to show that a given set A is definable is to express A as a solution of a "first-order formula quantified over definable sets and functions". A detailed proof is given in [13, Theorem 1.13], but this principle is referenced throughout the literature, e.g., [31, Appendix A] and [19, Section 2]. A simple application is that if A_i are definable sets, $Q_i \in \{\exists, \forall\}$ are quantifiers, $\Delta \in \{<, \leq, =, \neq\}$ and f is a definable function then the set of x satisfying

$$Q_1 y_1 \in A_1,\ Q_2 y_2 \in A_2, \ldots, Q_m y_m \in A_m, \qquad f(x, y_1, \ldots, y_m)\Delta 0 \text{ holds.} \tag{5.2}$$

is definable. This principle gives, for instance, an easy proof of the fact that the set C_f of continuity points of a definable function f is definable as well. After all, C_f is exactly the set of solutions of the formula "$\forall \epsilon > 0, \exists \delta > 0, \forall y \in \{z \mid \|x - z\| \leq \delta\}, \qquad \|f(x) - f(y)\| - \epsilon \leq 0$ holds". This principle also leads to an easy proof of the following well-known lemma regarding partial minimization.

Lemma 5.4 (Partial minimization preserves definability)**.** *Let $A \subset \mathbb{R}^n$ be a definable set and let $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ be a definable function, all over the same o-minimal structure. Then, the function $\varphi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ given by $\varphi(x) = \inf_{y \in A} f(x, y)$ is definable.*

Proof. The graph of φ is the intersection of the sets of solutions (x, α) of two first-order formulae:

$$\forall\, y \in A,\ f(x, y) \geq \alpha \quad \text{ and } \quad \forall\, \epsilon > 0, \exists\, y \in A,\ f(x, y) < \alpha + \epsilon.$$

Due to (5.2), these two sets are definable. Consequently, the graph of φ is definable and φ is definable. This completes the proof. ☐

We move on to the main result of this section and we recall that an operator $L : \mathcal{E} \to \mathcal{E}$ is said to be *quasi-nonexpansive* if $\|Lx - y\| \leq \|x - y\|$ holds for every $x \in \mathcal{E}$ and $y \in \operatorname{Fix} L$.

Theorem 5.5. *Let $L_i : \mathcal{E} \to \mathcal{E}$ $(i = 1, \ldots, n)$ be continuous operators such that $\bigcap_{i=1}^n \operatorname{Fix} L_i \neq \emptyset$. If all the L_i are quasi-nonexpansive and definable in some o-minimal structure, then $L_1, \ldots, L_n$ are jointly Karamata regular (*resp. *Karamata regular when $n = 1$) over any bounded set $B \subseteq \mathcal{E}$. In particular, the corresponding regularity function ψ_B can be taken to be $\Phi(\cdot, r)$ defined in Lemma 5.2 with sufficiently large r. Furthermore, $\Phi(\cdot, r) \in \mathrm{RV}_\rho^0$ with index $\rho \in [0, 1]$ when restricted to $(0, 1]$ (ρ may depend on r).*

Proof. For simplicity, let $C := \bigcap_{i=1}^n \operatorname{Fix} L_i$. First, we get rid of a trivial case. If $C = \mathcal{E}$ and B is an arbitrary bounded set, the conclusion holds if we take ψ_B to be the identity map restricted to the interval $[0, \infty)$. Henceforth, we next suppose that $C \neq \mathcal{E}$.

Fix any bounded set $B \subseteq \mathcal{E}$. Then there exists some $r_B > 0$ such that $B \subseteq \mathbb{B}_{r_B}$, where we recall that $\mathbb{B}_{r_B}$ is the closed ball centered on the origin with radius r_B. We choose some $\widehat{x} \notin C$ and let r be such that

$$r \geq \max\{r_B, \operatorname{dist}(0, C), \|\widehat{x}\|\}.$$

Let Φ be defined as in (5.1) and let

$$\psi_B(t) := \Phi(t, r), \quad t \geq 0. \tag{5.3}$$

We then know from Lemma 5.2 (ii) and (iii) that for any $x \in \mathbb{B}_r$ (in particular, $x \in B$),

$$\begin{aligned} \operatorname{dist}(x, C) &\leq \Phi\Big(\max_{1 \leq i \leq n} \|x - L_i(x)\|, \|x\|\Big) \\ &\leq \Phi\Big(\max_{1 \leq i \leq n} \|x - L_i(x)\|, r\Big) = \psi_B\Big(\max_{1 \leq i \leq n} \|x - L_i(x)\|\Big). \end{aligned} \tag{5.4}$$

One can see from Lemma 5.2 (i) that ψ_B is nondecreasing and satisfies

$$\psi_B(0) = \lim_{t \to 0_+} \psi_B(t) = \lim_{t \to 0_+} \Phi(t, r) = 0.$$

It then remains to show that $\psi_B|_{(0,\,1]} \in \mathrm{RV}^0_\rho$ with $\rho \in [0,\,1]$.

First, we show that $\psi_B(t) > 0$ holds for any $t > 0$. Let $d(x) := \max_{1 \le i \le n} \|x - L_i(x)\|$. Notice that $d(P_C(0)) = 0$ and $d(\widehat{x}) > 0$. Since all the L_i are continuous, d is continuous as well. By the intermediate value theorem, for any $t > 0$, there exists some $x_t = \alpha P_C(0) + (1-\alpha)\widehat{x}$ with $\alpha \in [0,\,1)$ such that $d(x_t) = \min\{t,\, d(\widehat{x})\}$. Moreover, we have

$$\|x_t\| = \|\alpha P_C(0) + (1-\alpha)\widehat{x}\| \le \alpha\|P_C(0)\| + (1-\alpha)\|\widehat{x}\| = \alpha\,\mathrm{dist}(0,\,C) + (1-\alpha)\|\widehat{x}\| \le r,$$

which together with $0 < d(x_t) \le t$ implies that $x_t \notin C$ and $x_t \in \Omega_{t,r}$, where $\Omega_{t,r}$ is defined as in Lemma 5.2. As a result, for any $t > 0$ we have

$$\psi_B(t) = \max_{x \in \Omega_{t,r}} \mathrm{dist}(x,\,C) > 0. \tag{5.5}$$

Next, we prove $\psi_B \in \mathrm{RV}^0$. Let

$$h(x,\,t) := \mathrm{dist}(x,\,C) + \delta_\Omega(x,\,t) \quad \text{with} \quad \Omega := \Big\{(x,\,t) \mid \max_{1 \le i \le n} \|x - L_i(x)\| \le t,\ \|x\| \le r\Big\},$$

where δ_Ω is the indicator function of Ω and we recall that r is fixed. Since all the L_i are definable, each $\mathrm{Fix}\, L_i$ is definable[14] and thus $C = \bigcap_{i=1}^n \mathrm{Fix}\, L_i$ is definable. By Lemma 5.4 applied to $\|x - y\|$ and C we see that $\mathrm{dist}(x,\,C)$ is definable, see also [18, Proposition 2.8].

Note that Ω can be written as the intersection of $n+1$ sets, each of which is definable[15]. Consequently, Ω is definable and $h(x,\,t)$ is definable as well, since it is the sum of two definable functions. From (5.5) we have $\psi_B(t) = \max_x h(x,\,t)$. Then $\psi_B(t)$ is definable by Lemma 5.4.

In particular, the restriction ψ_B to the interval $(0,\,1)$ is a definable positive function. Next, for every $\mu \in (0,1)$, we define the function

$$\psi_B^\mu(t) := \frac{\psi_B(\mu t)}{\psi_B(t)},$$

for $t \in (0,\,1)$. Because compositions and quotients of definable functions are definable, we have that $\psi_B^\mu(t)$ is definable. By Lemma 5.3, there is an interval $(0,\,c) \subseteq (0,\,1)$ over which $\psi_B^\mu(t)$ is either strictly monotone or constant, in particular

$$\lim_{t \to 0_+} \psi_B^\mu(t) \ \text{ exists in } \mathbb{R} \cup \{-\infty,\,\infty\}, \tag{5.6}$$

see also [13, Exercise 2.3]. Now, we let $f(x) := \psi_B(1/x)$, $x \in [1,\,\infty)$. Then we see from the monotonicity of ψ_B (recall (5.3) and Lemma 5.2) that f is nonincreasing and therefore measurable. Moreover, for all $x \ge 1$ and $\lambda \ge 1$ we have

$$0 < \frac{f(\lambda x)}{f(x)} \le 1. \tag{5.7}$$

On the other hand, using (5.6) we obtain for all $\lambda \in (1,\infty)$,

$$\lim_{x \to \infty} \frac{f(\lambda x)}{f(x)} = \lim_{x \to \infty} \frac{\psi_B\left(1/(\lambda x)\right)}{\psi_B(1/x)} = \lim_{t \to 0_+} \frac{\psi_B\left(\frac{1}{\lambda} t\right)}{\psi_B(t)} = \lim_{t \to 0_+} \psi_B^{1/\lambda}(t) \ \text{ exists in } \mathbb{R} \cup \{-\infty,\,\infty\}.$$

[14] Note that $\mathrm{Fix}\, L_i$ is the inverse image of $\{0\}$ of the definable map $L_i - I$, where I is the identity map. Therefore, it is definable, thanks to [31, B.3].

[15] Again, this can be proved in multiple ways. For example, by observing that the set of x satisfying $\|x - L_i(x)\| \le t$ is the inverse image of the inverval $(-\infty,\,0]$ by the definable map $\|x - L_i(x)\| - t$.

This together with (5.7) implies that for all $\lambda \geq 1$, $\lim_{x\to\infty} \frac{f(\lambda x)}{f(x)}$ exists in $\mathbb{R}$. Now let Λ_0 be the set of $\lambda \geq 1$ for which $\lim_{x\to\infty} \frac{f(\lambda x)}{f(x)} = 0$ holds. Λ_0 must be definable because it is the set of $\lambda \geq 1$ satisfying

$$\forall \epsilon > 0, \exists M > 0, \forall x \geq M \qquad \left| \frac{f(\lambda x)}{f(x)} \right| \leq \epsilon.$$

Let $\Lambda_>$ be the complement of Λ_0 intersected with $[1, \infty)$. With that, $\Lambda_>$ correspond to the $\lambda \geq 1$ for which $\lim_{x\to\infty} \frac{f(\lambda x)}{f(x)} > 0$ holds. Since definability is preserved by complements and intersections, $\Lambda_>$ is definable and, in view of Definition 5.1, must be a finite union of intervals and points.

Next we consider two cases. If $\Lambda_>$ contains an interval, then $\lim_{x\to\infty} \frac{f(\lambda x)}{f(x)}$ exists, is finite and positive for all λ in a set of positive measure. Since f is measurable, this is enough to conclude that $f \in \mathrm{RV}$, see [6, Theorem 1.4.1 (ii)]. In view of (2.3) and the definition of f, we conclude that $\psi_B|_{(0,\,1]} \in \mathrm{RV}^0_\rho$ with some $\rho \in \mathbb{R}$.

If $\Lambda_>$ does not contain an interval, then it must be a union of finitely many points. However, one may verify that if $\lambda \in \Lambda_>$ then $\lambda^n \in \Lambda_>$ for all $n \in \mathbb{N}$, e.g., see [6, Section 1.4] or this footnote[16]. So the only way that $\Lambda_>$ can be a finite union of points is if $\Lambda_> = \{1\}$ holds. In this case, we have $\Lambda_0 = (1, \infty)$ so $f \in \mathrm{RV}_{-\infty}$ and ψ_B belongs to RV^0_∞.

Overall, we conclude that $\psi_B \in \mathrm{RV}^0 \cup \mathrm{RV}^0_\infty$. Our next step is to show that ψ_B cannot be in RV^0_∞. For that, given any x and every i, we let $y \coloneqq P_C(x)$ and we see from the quasi-nonexpansiveness of L_i that

$$\|x - L_i(x)\| = \|x - y + y - L_i(x)\| \leq 2\|x - y\| = 2\,\mathrm{dist}(x, C).$$

This together with (5.4) and $d(x) \coloneqq \max_{1\leq i\leq n} \|x - L_i(x)\|$ implies that for all $x \in \mathbb{B}_r$,

$$\frac{d(x)}{2} \leq \mathrm{dist}(x, C) \leq \psi_B(d(x)). \tag{5.8}$$

Recall that $P_C(0)$ and $\widehat{x}$ are in $\mathbb{B}_r$ and they satisfy $d(P_C(0)) = 0$ and $d(\widehat{x}) > 0$, respectively. Since $d(\cdot)$ is a continuous function, $d(\cdot)$ takes all values between 0 and $d(\widehat{x})$ over the ball $\mathbb{B}_r$. This together with (5.8) implies that for sufficiently small t we have

$$t/2 \leq \psi_B(t). \tag{5.9}$$

That is, $1/2 \leq \psi_B(t)/t$ holds for all sufficiently small t. Or, put otherwise,

$$\frac{1}{2} \leq \frac{\psi_B(1/x)}{1/x} = \frac{f(x)}{x^{-1}}$$

holds for all sufficiently large $x > 0$. This implies that $x^{-1}\frac{1}{f(x)} \leq 2$ for x sufficiently large. For the sake of obtaining a contradiction, suppose that $\psi_B \in \mathrm{RV}^0_\infty$. Then $1/f \in \mathrm{RV}_\infty$ and is nondecreasing, since f is nonincreasing. With that $1/f \in \mathrm{KRV}_\infty$ by (2.19). Therefore, the function that maps x to $x^{-1}\frac{1}{f(x)}$ is in RV_∞, by (2.18) and (2.20). However this implies that $x^{-1}\frac{1}{f(x)} \to \infty$ as $x \to \infty$, e.g., see (2.13). This is a contradiction and it shows that ψ_B cannot be rapidly varying, so it must be in RV^0_ρ for some $\rho \in \mathbb{R}$.

For the last part, we will prove that $\rho \in [0, 1]$ holds. First, for $\lambda > 1$, we see from the monotonicity of ψ_B that $\lambda^\rho = \lim_{t\to 0_+} \psi_B(\lambda t)/\psi_B(t) \geq 1$, which proves that $\rho \geq 0$.

It remains to show that $\rho \leq 1$. Suppose to the contrary that $\rho > 1$. We let $\delta \in (0, \rho - 1)$ and conclude from Potter's bound (2.7) (set $A = 2$) that for sufficiently small t and s,

$$\psi_B(t) \leq 2\psi_B(s) \max\left\{ \left(\frac{t}{s}\right)^{\rho-\delta}, \left(\frac{t}{s}\right)^{\rho+\delta} \right\}. \tag{5.10}$$

[16] It can also be proved by induction by observing that letting $g_\lambda \coloneqq \lim_{x\to\infty} \frac{f(\lambda x)}{f(x)}$ and assuming that $g_\lambda \in (0, \infty)$, we have $\frac{f(\lambda^2 x)}{f(x)} = \frac{f(\lambda^2 x)}{f(\lambda x)} \frac{f(\lambda x)}{f(x)} \to g_\lambda^2$.

In view of (5.9) and (5.10), the following inequalities hold for sufficiently small t and s,

$$\frac{1}{2} \leq \frac{\psi_B(t)}{t} \leq 2\psi_B(s) \max\left\{t^{\rho-\delta-1}\left(\frac{1}{s}\right)^{\rho-\delta},\ t^{\rho+\delta-1}\left(\frac{1}{s}\right)^{\rho+\delta}\right\}.$$

If we fix s and let $t \to 0_+$, the right-hand side will converge to 0 (because $\rho - \delta - 1 > 0$), which gives a contradiction. Consequently, we must have $\rho \leq 1$. This completes the proof. $\square$

Theorem 5.5 has some useful corollaries that we will discuss next. The first is that in the setting of quasi-cyclic algorithms as in (3.5), we can always obtain joint Karamata regularity.

Corollary 5.6. *Let $T_i : \mathcal{E} \to \mathcal{E}$ $(i = 1, \ldots, m)$ be given as in problem (1.1). If T_i are definable in some o-minimal structure, then they are jointly Karamata regular (*resp. *Karamata regular when $m = 1$) over any bounded set $B \subseteq \mathcal{E}$.*

Proof. Each T_i is α-averaged, so, in particular, continuous and quasi-nonexpansive. Moreover, $\bigcap_{i=1}^m \operatorname{Fix} T_i \neq \emptyset$. Then applying Theorem 5.5 with $L_i = T_i$ and $n = m$, we conclude that T_i $(i = 1, \ldots, m)$ are jointly Karamata regular (*resp.* Karamata regular when $m = 1$) over any bounded set $B \subseteq \mathcal{E}$. $\square$

As a consequence of Corollary 5.6, whenever the problem is definable in some *o*-minimal structure, we can in principle use Theorem 3.10 and Theorem 3.14 to analyze convergence rates of algorithms for solving the common fixed point problem (1.1).

Another consequence of Theorem 5.5 is that the error bounds that describe intersections of definable convex sets can be taken to be regularly varying, which is a result we were not aware when [24] was written. In what follows, we will use the notion of consistent error bound see Remark 3.3 and [24].

Corollary 5.7. *Let $C_1, \ldots, C_m \subseteq \mathcal{E}$ be closed convex sets definable in some o-minimal structure with nonempty intersection. Then, there exists a consistent error bound function Φ for $C_1, \ldots, C_m$ such that for all sufficiently large $r > 0$, we have $\Phi(\cdot, r) \in \mathrm{RV}_\rho^0$ with index $\rho \in [0, 1]$ when restricted to $(0, 1]$ (ρ may depend on r). In particular, for every bounded set $B \subseteq \mathcal{E}$, there exists a nondecreasing function $\psi_B \in \mathrm{RV}_\rho^0$ with $\rho \in [0, 1]$ when restricted to $(0, 1]$ and $\lim_{t\to 0_+} \psi_B(t) = 0$ such that*

$$\operatorname{dist}\Big(x, \bigcap_{i=1}^m C_i\Big) \leq \psi_B\Big(\max_{1\leq i\leq m} \operatorname{dist}(x, C_i)\Big), \quad \forall\ x \in B.$$

Proof. Apply Theorem 5.5 by taking $m = n$ and L_i to be the projection operator onto C_i. With that the Φ in Lemma 5.2 becomes a consistent error bound function for the $C_1, \ldots, C_m$. Then we see from Theorem 5.5 that for sufficiently large $r > 0$, we have $\Phi(\cdot, r) \in \mathrm{RV}_\rho^0$ with $\rho \in [0, 1]$ when restricted to $(0, 1]$. Finally, let B be any bounded set and let b be large enough so that $\sup_{x\in B} \|x\| \leq b$ and $\Phi(\cdot, b) \in \mathrm{RV}_\rho^0$ with $\rho \in [0, 1]$ when restricted to $(0, 1]$. Define $\psi_B \coloneqq \Phi(\cdot, b)$. Then it follows from Lemma 5.2 (i) that ψ_B is nondecreasing and it satisfies $\lim_{t\to 0_+} \psi_B(t) = 0$. Moreover, for every $x \in B$, letting $d(x) \coloneqq \max_{1\leq i\leq m} \operatorname{dist}(x, C_i)$ we have from Lemma 5.2 (iii) that

$$\operatorname{dist}\Big(x, \bigcap_{i=1}^m C_i\Big) \leq \Phi(d(x), \|x\|) \leq \Phi(d(x), b) = \psi_B(d(x)).$$

This completes the proof. $\square$

Corollary 5.7 implies that the results of this paper and [24] can be used to analyze several different types of algorithms for feasibility problems over definable convex sets.

References

[1] K. B. Athreya and S. N. Lahiri. *Measure Theory and Probability Theory*, volume 19. Springer, 2006.

[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 Operations Research*, 35(2):438–457, 2010.

[3] H. Bauschke and P. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. CMS Books in Mathematics. Springer International Publishing, 2017.

[4] H. H. Bauschke, D. Noll, and H. M. Phan. Linear and strong convergence of algorithms involving averaged nonexpansive operators. *Journal of Mathematical Analysis and Applications*, 421(1):1–20, 2015.

[5] N. H. Bingham, C. M. Goldie, and E. Omey. Regularly varying probability densities. *Publications de l'Institut Mathématique. Nouvelle Série*, 80:47–57, 2006.

[6] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular Variation*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1987.

[7] J. Bolte, A. Daniilidis, A. Lewis, and M. Shiota. Clarke subgradients of stratifiable functions. *SIAM Journal on Optimization*, 18(2):556–572, Jan. 2007.

[8] J. Bolte, A. Daniilidis, O. Ley, and L. Mazet. Characterizations of Łojasiewicz inequalities: subgradient flows, talweg, convexity. *Transactions of the American Mathematical Society*, 362(6):3319–3363, 2010.

[9] J. Bolte, T. P. Nguyen, J. Peypouquet, and B. W. Suter. From error bounds to the complexity of first-order descent methods for convex functions. *Mathematical Programming*, 165(2):471–507, 2017.

[10] J. M. Borwein, G. Li, and M. K. Tam. Convergence rate analysis for averaged fixed point iterations in common fixed point problems. *SIAM Journal on Optimization*, 27(1):1–33, 2017.

[11] J. M. Borwein and S. B. Lindstrom. Meetings with Lambert W and other special functions in optimization and analysis. *Pure and Applied Functional Analysis*, 1(3):361–396, 2016.

[12] V. Chandrasekaran and P. Shah. Relative entropy optimization and its applications. *Mathematical Programming*, 161(1):1–32, 2017.

[13] M. Coste. *An Introduction to O-minimal Geometry*. Pisa: Istituti editoriali e poligrafici internazionali., 2000.

[14] D. Djurčić and A. Torgašev. Some asymptotic relations for the generalized inverse. *Journal of Mathematical Analysis and Applications*, 335(2):1397–1402, 2007.

[15] N. Elez and D. Djurčić. Some properties of rapidly varying functions. *Journal of Mathematical Analysis and Applications*, 401(2):888–895, 2013.

[16] P. Embrechts and M. Hofert. A note on generalized inverses. *Mathematical Methods of Operations Research*, 77(3):423–432, 2013.

[17] J.-B. Hiriart-Urruty and C. Lemaréchal. *Convex Analysis and Minimization Algorithms I*. Springer Berlin Heidelberg, 1993. doi:10.1007/978-3-662-02796-7.

[18] P. D. Hoàng. Łojasiewicz-type inequalities and global error bounds for nonsmooth definable functions in o-minimal structures. *Bulletin of the Australian Mathematical Society*, 93(1):99–112, 2016.

[19] A. D. Ioffe. An invitation to tame optimization. *SIAM Journal on Optimization*, 19(4):1894–1917, Jan. 2009.

[20] A. D. Ioffe. Nonlinear regularity models. *Mathematical Programming*, 139(1–2):223–242, Mar. 2013.

[21] G. Li and T. K. Pong. Calculus of the exponent of Kurdyka–Łojasiewicz inequality and its applications to linear convergence of first-order methods. *Foundations of Computational Mathematics*, 18(5):1199–1232, 2018.

[22] Y. Lin, S. B. Lindstrom, B. F. Lourenço, and T. K. Pong. Tight error bounds for log-determinant cones without constraint qualifications. *ArXiv e-prints*, 2024. arXiv:2403.07295.

[23] S. B. Lindstrom, B. F. Lourenço, and T. K. Pong. Error bounds, facial residual functions and applications to the exponential cone. *Mathematical Programming*, 200(1):229–278, 2023.

[24] T. Liu and B. F. Lourenço. Convergence analysis under consistent error bounds. *Foundations of Computational Mathematics*, 24(2):429–479, 2024.

[25] D. R. Luke, N. H. Thao, and M. K. Tam. Implicit error bounds for Picard iterations on Hilbert spaces. *Vietnam Journal of Mathematics*, 46(2):243–258, Feb. 2018.

[26] MOSEK ApS. *MOSEK Modeling Cookbook Release 3.2.3*, 2021. URL: https://docs.mosek.com/modeling-cookbook/index.html.

[27] M. Mršević. Convexity of the inverse function. *The Teaching of Mathematics*, (20):21–24, 2008.

[28] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, 1997.

[29] D. Russell Luke, N. H. Thao, and M. K. Tam. Quantitative convergence analysis of iterated expansive, set-valued mappings. *Mathematics of Operations Research*, 43(4):1143–1176, Nov. 2018.

[30] E. Seneta. *Regularly Varying Functions*. Lecture Notes in Mathematics. Springer Berlin Heidelberg, 1976.

[31] L. van den Dries and C. Miller. Geometric categories and o-minimal structures. *Duke Mathematical Journal*, 84(2), Aug. 1996.

[32] L. P. D. van den Dries. *Tame Topology and O-minimal Structures*. Cambridge University Press, May 1998.

[33] P. Yu, G. Li, and T. K. Pong. Kurdyka–Łojasiewicz exponent via inf-projection. *Foundations of Computational Mathematics*, 22(4):1171–1217, July 2021.