ON THE FINITARY CONTENT OF DYKSTRA'S CYCLIC PROJECTIONS ALGORITHM

PEDRO PINTO

Department of Mathematics, Technische Universität Darmstadt,
Schlossgartenstraße 7, 64289 Darmstadt, Germany,

E-mail: pinto@mathematik.tu-darmstadt.de

Abstract. We study the asymptotic behaviour of the well-known Dykstra's algorithm. We provide an elementary proof for the convergence of Dykstra's algorithm in which the standard argument is stripped to its central features and where the original compactness principles are circumvented, additionally providing highly uniform computable rates of metastability in a fully general setting. Moreover, under an additional assumption, we are even able to obtain effective general rates of convergence. We argue that such an additional condition is actually necessary for the existence of general uniform rates of convergence.

Keywords: Convex feasibility; projection methods; Dykstra's algorithm; rates of convergence; metastability; proof mining
MSC2020 Classification: 47H09; 41A65; 90C25; 03F10

1. Introduction

Many problems in convex optimization can be stated in terms of finding a point in the intersection of a family of convex and closed sets, what is known as the convex feasibility problem:

$$\text{(CFP)} \qquad \text{find some point } x \in \bigcap_{j\in I} C_j,$$

assuming *a priori* that $\bigcap_{j\in I} C_j \neq \emptyset$, i.e. the problem has a solution (is feasible). The study of such problems first appeared in connection with constraints defined by linear inequalities and where the feasibility set is the intersection of half-spaces. Since then the general problem has been the subject of much research due to its broad applicability in applied mathematics – e.g. in statistics, partial differential equations (Dirichlet problem over irregular regions), solving linear equations (Kaczmarz's method), image or signal restoration, and computed tomography. For further discussions we refer the reader to the surveys [2, 8].

One of the most successful and well-known techniques to iteratively approximate a solution to the CFP is von Neumann's method of alternating projections (MAP). For a subspace V, let P_V denote the orthogonal projection map onto V.

Theorem 1.1 (von Neumann [47])**.** *Let V_1, V_2 be two closed vector subspaces of a Hilbert space X. Then, for any point $x_0 \in X$ the iteration defined by*

$$x_{n+1} := P_{V_1}P_{V_2}(x_n)$$

converges strongly to $P_{V_1\cap V_2}(x_0)$.

The original proof by von Neumann doesn't generalize immediately to more than two subspaces. This was overcome by Halperin who extended the convergence result to a finite number of subspaces.

Date: July 6, 2024.

Theorem 1.2 (Halperin [18]). *Let $V_1, \cdots, V_m$ be $m \geq 2$ closed vector subspaces of a Hilbert space. Then, for any point $x_0 \in X$ the iteration defined by*

$$x_{n+1} := P_{V_1} \cdots P_{V_n}(x_n) \tag{MAP}$$

converges strongly to $P_{\bigcap_{j=1}^m V_j}(x_0)$.

The convergence of (MAP) holds more generally when the sets V_j are affine subspaces (i.e. translates of subspaces) provided that their intersection is nonempty. Elementary geometric proofs of Theorems 1.1 and 1.2 can be found in [33] and [34]. However, if the sets are just assumed to be closed and convex then the situation is more delicate. In 1965, Bregman established weak convergence of von Neuman's method in the general setting.

Theorem 1.3 (Bregman [6]). *Let $C_1, \cdots, C_m$ be $m \geq 2$ closed convex subsets of a Hilbert space such that $\bigcap_{j=1}^m C_j \neq \emptyset$. Then,* (MAP) *converges weakly to a point in the intersection.*

In 2004, Hundal [19] gave a counterexample (where the sets consist of a closed hyperplane and a closed convex cone in $\ell^2(\mathbb{N})$) for which the iteration indeed doesn't converge in norm (see also [36]). Moreover, there are very easy examples of convex sets where (MAP) converges strongly to a point in the intersection but such point is distinct from the projection of x_0. A different iterative scheme was proposed by Dykstra that does converge strongly to the projection of x_0 in the general setting of the convex feasibility problem.

Consider $C_1, \cdots, C_m$ to be $m \geq 2$ closed convex subsets of a Hilbert space with nonempty intersection. For $n \geq 1$, let C_n denote the set C_{j_n} where $j_n := [n-1]+1$ with $[r] :\equiv r \mod m$, and let P_n denote the metric projection onto C_n. For $x_0 \in X$ an initial point, Dykstra's cyclic projections algorithm is defined recursively by the equations

$$\begin{cases} x_0 \in X \\ q_{-(m-1)} = \cdots = q_0 = 0 \end{cases} \quad \text{and } \forall n \geq 1 \begin{cases} x_n := P_n(x_{n-1} + q_{n-m}) \\ q_n := x_{n-1} + q_{n-m} - x_n \end{cases} \tag{D}$$

In 1983, Dykstra [15] proved the strong convergence in the particular case when all the sets are closed convex cones of a finite dimensional Hilbert space. The result was later extended to the general setting by Boyle and Dykstra [5].

Theorem 1.4 (Boyle-Dykstra [5]). *Let $C_1, \cdots, C_m$ be $m \geq 2$ closed convex subsets of a Hilbert space such that $C := \bigcap_{j=1}^m C_j \neq \emptyset$. Then, for any point $x_0 \in X$ the iteration (x_n) generated by* (D) *converges strongly to $P_C(x_0)$.*

When the C_i's are closed vector subspaces (or more generally, closed affine subspaces), the projection is a linear map and it is easy to see that the scheme (D) reduces to (MAP). It may be helpful to think of Dykstra's algorithm as operating in stages: it starts with some initial guess $x_0 \in X$ and by setting auxiliary terms $q_{-(m-1)}, \cdots, q_{-1}, q_0$ to zero. One is then able to compute $x_1, \cdots, x_m$. Using these points, we can now *update* the values of the auxiliary terms, namely computing $q_1, \cdots, q_m$. After this, the process repeats and will approximate in norm the projection of x_0 onto the feasibility set.

Dykstra's cyclic projections algorithm is an attractive method for strongly approximating the intersection of closed convex sets in a Hilbert space as, while it converges to the feasible point closest to x_0, it only requires knowledge about the projections onto the individual convex sets C_j. However, contrary to (MAP) not much is known regarding its quantitative information. For the particular case when the sets C_j are closed half-spaces, i.e. the intersection is a polyhedral subset of X, Deutsch and Hundal [10] obtained rates of convergence for (D) (see [9] and the references therein for further discussions on rates of convergence and their applications). Strikingly, in the case of $m = 2$ the rate doesn't depend on the initial point but only on an upper bound on its distance to the intersection.

In this paper, we analyse the asymptotic behaviour of Dykstra's algorithm and provide highly uniform quantitative information in the general setting of Theorem 1.4.

When analysing the convergence of the sequence (x_n), one considers the equivalent Cauchy formulation

$$\forall \varepsilon > 0\ \exists n \in \mathbb{N}\ \forall i, j \geq n\, (\|x_i - x_j\| \leq \varepsilon)\,. \tag{$\dagger$}$$

Effective rates of convergence, i.e. a computable function giving a witness to the value of n in ($\dagger$), are in general excluded (as per the fundamental work of Specker [43]). Proof-theoretical considerations (see [23]) guide us to the next best possible thing

$$\forall \varepsilon > 0\ \forall f : \mathbb{N} \to \mathbb{N}\ \exists n \in \mathbb{N}\ \forall i, j \in [n; n + f(n)]\, (\|x_i - x_j\| \leq \varepsilon)\,, \tag{$\ddagger$}$$

where $[n; n+f(n)] = \{n, n+1, \cdots, n+f(n)\}$. The statement ($\ddagger$) is noneffectively equivalent to the Cauchy property of the sequence. This reformulation of ($\dagger$) is long known to logicians as Kreisel's no-counterexample interpretation and was popularized by Terence Tao under the name of *metastability* [45, 46]. In line with such terminology, a function that given $\varepsilon > 0$ and $f \in \mathbb{N}^{\mathbb{N}}$ outputs a bound on n in ($\ddagger$) is called a rate of metastability. General logical metatheorems (e.g. [16, 22]) guarantee the existence of such rates provided the proof of the convergence statement can be formalized in certain formal systems[1].

We provide a quantitative analysis of the proof in [3] which follows closely the arguments originally given by Boyle and Dykstra [5]. Through our quantitative analysis, it was possible to remove the compactness principles crucial in the original proof and, in this way, obtain effective rates of metastability in the general setting of the convex feasibility problem. Moreover, our quantitative results are a true finitization of Theorem 1.4, in the sense that the infinitary result is fully recovered. In this way, we provide an elementary proof for the convergence of Dykstra's algorithm. Mathematicians naturally prefer rates of convergence to rates of metastability. However, in full generality they are usually unavailable. Furthermore, when such rates are actually available, they are frequently sensitive to the parameters of the problem. In this case, one would expect a rate of convergence to depend on the specifics of the underlying space, of the convex sets and on the initial point on which the iteration is initiated. This should be compared with the uniformity exhibited by the rate of metastability obtain in Theorem 3.11: it only depends on the number of convex sets and on a bound to the distance of x_0 and the intersection set. For the general case augmented with a regularity assumption on the convex sets, we show that it is possible to obtain uniform rates of convergence. From this, we derive rates of convergence for the case of basic semi-algebraic convex sets in $\mathbb{R}^n$ (which in particular covers the polyhedral case in finite dimension). We furthermore argue that the setting with this regularity assumption is actually the only one where uniform rates of convergence are possible, in particular, encompassing the rate of convergence obtained by Hundal and Deutsch for the polyhedral case when $m = 2$.

The quantitative study in this paper is set in the context of the '*proof mining*' program ([23], see also the recent survey [28]) where proof-theoretical techniques are employed to analyze *prima facie* noneffective mathematical proofs with the goal of obtaining additional information. With a broad range of applications, the proof mining program has been in particular very successful in the study of results in convex optimization. A closer connection with the CFP can be found in its applications to splitting methods e.g. [11, 14, 13, 35], in particular in the previous studies of (MAP) in [26]. An important work was the analysis of Bauschke's solution to the zero displacement problem [1] regarding the potentially 'inconsistent feasibility problem' in [27], as well as the quantitative studies of generalizations of Bauschke's result in [41, 42] carried out in the underlying context of the proof mining program.

All these previous works point to the usefulness of proof mining and of logic-based techniques in the study of the convex feasibility problem and connected methods (see e.g.

[1] The class of such admissible proofs is (in practice) very large, encompassing in particular most proofs in classical analysis.

the recent [32]). That being said, we remark that such perspective and techniques only operate in the background and that the central theorems are presented in a way which does not assume any prior knowledge of logic on the reader. Nevertheless, we allow ourselves some simple logical remarks in the final section of the paper, which we think may be of particular interest to logicians and the more inquisitive mathematicians.

2. Preliminaries and Lemmas

2.1. **Quantitative notions.** Let (x_n) be a Cauchy sequence in a normed space $(X, \|\cdot\|)$.

Definition 2.1. *We say that a function $\theta : (0,\infty) \to \mathbb{N}$ is a Cauchy rate if*

$$\forall \varepsilon > 0\ \forall i, j \geq \theta(\varepsilon)\, (\|x_i - x_j\| \leq \varepsilon)\,.$$

If the sequence (x_n) converges to some point $x \in X$ (e.g. if the space is complete), then any Cauchy rate will also be a rate of convergence towards x, i.e.

$$\forall \varepsilon > 0\ \forall i \geq \theta(\varepsilon)\, (\|x_i - x\| \leq \varepsilon)\,.$$

Considerations from computability theory [43] tell us that effective Cauchy rates are in general excluded, and one thus looks at an equivalent but computationally weaker reformulation.

Definition 2.2. *We call a function $\Theta : (0,\infty) \times \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$ a rate of metastability if*

$$\forall \varepsilon > 0\ \forall f : \mathbb{N} \to \mathbb{N}\ \exists n \leq \Theta(\varepsilon, f)\ \forall i, j \in [n; n + f(n)]\, (\|x_i - x_j\| \leq \varepsilon)\,,$$

where $[n; n + f(n)] := \{n, n+1, \cdots, n + f(n)\}$.

The following result is folklore (see e.g. [30]).

Proposition 2.3. *A function $\theta : (0,\infty) \to \mathbb{N}$ is a Cauchy rate if and only if the function $\Theta : (\varepsilon, f) \mapsto \theta(\varepsilon)$ is a rate of metastability.*

2.2. **Lemmas.** Throughout, let $(X, \langle\cdot,\cdot\rangle)$ be a real Hilbert space with inner product $\langle\cdot,\cdot\rangle$ and norm $\|\cdot\| := \sqrt{\langle\cdot,\cdot\rangle}$. We also use the notation $[a; b] := [a, b] \cap \mathbb{N}$ and $\mathbb{N}^* := \mathbb{N} \setminus \{0\}$. In this subsection we collect known results and prove some useful technical lemmas. Recall the following characterization of the metric projection, known as Kolmogorov's criterion.

Proposition 2.4 ([3, Theorem 3.16])**.** *Let C be a nonempty closed convex subset of X. Then, every point $u \in X$ has a unique best approximation on C, which we denote by $P_C(u)$ and call the projection of u onto C. Furthermore, $P_C(u)$ is the unique element of C satisfying*

$$\forall y \in C\ \ (\langle u - P_C(u), y - P_C(u)\rangle \leq 0)\,.$$

We shall require a quantitative version of the proposition above. The first quantitative studies on the metric projection featured in [24] and [25]. The version used here follows the formulation proven in [17, 37]. In [13], it was used in an 'ε/δ-formulation' for the general nonlinear setting of CAT(0) spaces. Let $\overline{B}_r(p)$ denote the closed ball of radius $r \geq 0$ centred at $p \in X$, i.e. $\overline{B}_r(p) := \{x \in X : \|x - p\| \leq r\}$.

Proposition 2.5. *Given $u \in X$, let $b \in \mathbb{N}^*$ be such that $b \geq \|u - p\|$ for some point $p \in \bigcap_{j=1}^m C_j$. For any $\varepsilon > 0$ and function $\delta : (0,\infty) \to (0,\infty)$, there exist $\eta \geq \beta(b, \varepsilon, \delta)$ and $x \in \overline{B}_b(p)$ such that $\bigwedge_{j=1}^m \|x - P_j(x)\| \leq \delta(\eta)$ and*

$$\forall y \in \overline{B}_b(p) \left(\bigwedge_{j=1}^m \|y - P_j(y)\| \leq \eta \to \langle u - x, y - x\rangle \leq \varepsilon \right),$$

where $\beta(b,\varepsilon,\delta) := \dfrac{\varphi^2}{24b}$ *with*

$$\varphi := \min\left\{\tilde{\delta}^{(i)}(1) : i \leq \left\lceil \frac{4b^4}{\varepsilon^2}\right\rceil\right\} \text{ and } \tilde{\delta}(\xi) := \min\left\{\delta\left(\frac{\xi^2}{24b}\right), \frac{\xi^2}{24b}\right\}, \text{ for any } \xi > 0.$$

The proof of this result is an easy modification of the one in [13] regarding common (almost-)fixed points of two nonexpansive maps to common (almost-)fixed points of m projection maps. Indeed, it is clear that the result holds *mutatis mutandis* for any finite number of nonexpansive maps (and so for metric projections in Hilbert spaces). In order to convince the reader, we nevertheless include a proof. First, we require the following two technical lemmas essentially due to Kohlenbach [25].[2]

Lemma 2.6. *Let C be some convex bounded subset of a normed space and $D \in \mathbb{N}^*$ be a bound on the diameter of C. Consider a nonexpansive map $T : C \to C$. Then,*

$$\forall \varepsilon > 0\ \forall x_1, x_2 \in C \left(\bigwedge_{i=1}^{2} \|x_i - T(x_i)\| \leq \frac{\varepsilon^2}{12D} \to \forall t \in [0,1]\,(\|w_t - T(w_t)\| \leq \varepsilon)\right),$$

where $w_t := (1-t)x + ty$.

Lemma 2.7. *Let X be a normed space and for $x, y \in X$ and write $w_t := (1-t)x + ty$, for $t \in [0,1]$. Then, for any $u, x, y \in X$*

$$\forall \varepsilon \in (0, b^2] \left(\forall t \in [0,1]\left(\|u-x\|^2 \leq \|u - w_t\|^2 + \frac{\varepsilon^2}{D^2}\right) \to \langle u - x, y - x\rangle \leq \varepsilon\right),$$

where $D \in \mathbb{N}^*$ is such that $D \geq \|x - y\|$.

Proof of Proposition 2.5. Let $\varepsilon > 0$ and a function $\delta : (0,\infty) \to (0,\infty)$ be given. We structure the argument in two central claims.
Claim 1. There exist $\eta \geq \varphi$ and $x \in \overline{B}_b(p)$ such that $\bigwedge_{j=1}^m \|x - P_j(x)\| \leq \tilde{\delta}(\eta)$ and

$$\forall y \in \overline{B}_b(p)\left(\bigwedge_{j=1}^{m} \|y - P_j(y)\| \leq \eta \to \|u - x\|^2 \leq \|u - y\|^2 + \frac{\varepsilon^2}{4b^2}\right).$$

Proof of Claim 1. Assume towards a contradiction that for all $\eta \geq \varphi$ and $x \in \overline{B}_b(p)$ such that $\bigwedge_{j=1}^m \|x - P_j(x)\| \leq \tilde{\delta}(\eta)$,

$$\exists y \in \overline{B}_b(p)\left(\bigwedge_{j=1}^{m} \|y - P_j(y)\| \leq \eta \ \wedge\ \|u - y\|^2 < \|u - x\|^2 - \frac{\varepsilon^2}{4b^2}\right). \tag{1}$$

We then define a finite sequence $z_0, \cdots, z_r$ with $r := \left\lceil \frac{4b^4}{\varepsilon^2}\right\rceil$ as follows. Take z_0 to be p. In particular, $z_0 \in \overline{B}_b(p)$ and $\|z_0 - P_j(z_0)\| \leq \tilde{\delta}^{(r)}(1)$ for all $j \in [1;m]$. Consider that for $i \leq r-1$, we have $z_i \in \overline{B}_b(p)$ such that $\|z_i - P_j(z_i)\| \leq \tilde{\delta}^{(r-i)}(1)$ for all $j \in [1;m]$. Then, by (1) there is some $y \in \overline{B}_b(p)$ such that

$$\bigwedge_{j=1}^{m} \|y - P_j(y)\| \leq \tilde{\delta}^{(r-i-1)}(1) \text{ and } \|u - y\|^2 < \|u - z_i\|^2 - \frac{\varepsilon^2}{4b^2}.$$

We define z_{i+1} to be one such y. By construction $\|u - z_{i+1}\|^2 < \|u - z_i\|^2 - \varepsilon/4b^2$ for all $i < r$, and so we obtain

$$\|u - z_r\|^2 < \|u - z_0\|^2 - r\frac{\varepsilon^2}{4b^2} \leq b^2 - \frac{4b^4}{\varepsilon^2}\frac{\varepsilon^2}{4b^2} = 0.$$

[2] Despite some optimization of constants, the results are established as in [25, Lemmas 2.3 and 2.7].

which is a contradiction and concludes the proof of the claim. ■

In the following, consider $\eta_0 \geq \varphi$ and $x \in \overline{B}_b(p)$ as per Claim 1, and define $\eta_1 := \frac{\eta_0^2}{24b}$ which is bounded below by $\beta(b, \varepsilon, \delta)$.

Claim 2. For all $y \in \overline{B}_b(p)$ and $t \in [0, 1]$,

$$\bigwedge_{j=1}^{m} \|y - P_j(y)\| \leq \eta_1 \to \|u - x\|^2 \leq \|u - w_t\|^2 + \frac{\varepsilon^2}{4b^2}.$$

Proof of Claim 2. By the definition of the function $\tilde{\delta}$, we have in particular that

$$\bigwedge_{j=1}^{m} \|x - P_j(x)\| \leq \eta_1.$$

Now, if we consider $y \in \overline{B}_b(p)$ such that $\|y - P_j(y)\| \leq \eta_1$ for all $j \in [1; m]$, then we can apply Lemma 2.6 (with $D = 2b$) to each of the projection maps restricted to $\overline{B}_b(p)$ to conclude that $\bigwedge_{j=1}^{m} \|w_t - P_j(w_t)\| \leq \eta_0$. By convexity, $w_t \in \overline{B}_b(p)$ and the result follows by the assumption on η_0 and x. ■

By definition of $\tilde{\delta}$, we have $\bigwedge_{j=1}^{m} \|x - P_j(x)\| \leq \delta(\eta_1)$, and an application of Lemma 2.7 (with $D = 2b$) concludes the proof. □

We will make use of the following two technical lemmas. The first lemma is a quantitative version of the fact that any summable sequence of nonnegative real numbers must convergence towards zero.

Lemma 2.8. *Let $(a_n) \in \ell^1_+(\mathbb{N})$ and consider $B \in \mathbb{N}$ such that $\sum a_n \leq B$. Then,*

$$\forall \varepsilon > 0\ \forall f : \mathbb{N} \to \mathbb{N}\ \exists n \leq \Psi(B, \varepsilon, f)\ \forall i \in [n; n + f(n)]\, (a_i \leq \varepsilon)\,,$$

where $\Psi(B, \varepsilon, f) := \check{f}^{(R)}(0)$ with $\check{f}(m) := m + f(m) + 1$ and $R := \lfloor \frac{B}{\varepsilon} \rfloor$.

Proof. Let $\varepsilon > 0$ and $f : \mathbb{N} \to \mathbb{N}$ be given. Assume towards a contradiction that

$$\forall n \leq \Psi(B, \varepsilon, f)\ \exists i \in [n; n + f(n)]\, (a_i > \varepsilon)\,.$$

Using this assumption, we recursively construct $(R + 1)$-many distinct indices $i_0, \cdots, i_R$ such that $i_r \in [\check{f}(r)(0); \check{f}(r)(0) + f(\check{f}(r)(0))]$ and $a_{i_r} > \varepsilon$. Then

$$\sum_{i=0}^{\infty} a_i \geq \sum_{r=0}^{R} a_{i_r} > (R + 1)\varepsilon \geq B,$$

which is a contradiction. □

The next lemma corresponds to a quantitative version of [3, Lemma 30.6] (which itself is a variant of a technical lemma proved by Dykstra).

Lemma 2.9. *Let $(a_n) \in \ell^2_+(\mathbb{N})$ and consider $B \in \mathbb{N}$ such that $\sum a_n^2 \leq B$. For all $n \in \mathbb{N}$, set $s_n := \sum_{k=0}^{n} a_k$, and let $m \geq 2$ be given. Then,*

$$\underline{\lim}\, s_n(s_n - s_{n-m-1}) = 0 \text{ with } \liminf\text{-rate } \phi_B(m, \varepsilon, N) := \left\lfloor e^{\left(\frac{(m+1)B}{\varepsilon}\right)^2} \right\rfloor \cdot (N + 1),$$

i.e.

$$\forall \varepsilon > 0\ \forall N \in \mathbb{N}\ \exists n \in [N; N + \phi_B(m, \varepsilon, N)]\, (s_n(s_n - s_{n-m-1}) \leq \varepsilon)\,.$$

Proof. Let $N \in \mathbb{N}$ and $\varepsilon > 0$ be given, and shorten $\phi = \phi_B(m, \varepsilon, N)$. Using the Cauchy-Schwarz inequality, we have for all $n \in \mathbb{N}$

$$s_n \leq \sqrt{n+1}\sqrt{\sum_{k=0}^{n} a_k^2} \leq \sqrt{n+1}\sqrt{B}. \tag{2}$$

Assume towards a contradiction that for all $n \in [\max\{N, m\}; N + \phi]$ we have

$$s_n - s_{n-m-1} = \sum_{i=n-m}^{n} a_i > \frac{\varepsilon}{\sqrt{B}\sqrt{n+1}}. \tag{3}$$

Then, again by the Cauchy-Schwarz inequality, for all $n \in [\max\{N, m\}; N + \phi]$, we have

$$\frac{\varepsilon^2}{B(n+1)} < \left(\sum_{i=n-m}^{n} a_i\right)^2 \leq (m+1) \cdot \sum_{i=n-m}^{n} a_i^2.$$

Note that, by the integral test, the definition of ϕ entails

$$\sum_{n=N}^{N+\phi} \frac{1}{n+1} \geq \log\left(\frac{N+\phi+2}{N+1}\right) \geq \frac{(m+1)^2 B^2}{\varepsilon^2},$$

and so we derive the following contradiction

$$(m+1)^2 B \leq \sum_{n=N}^{N+\phi} \frac{\varepsilon^2}{B(n+1)} < \sum_{n=N}^{N+\phi}\left((m+1)\sum_{i=n-m}^{n} a_i^2\right) = (m+1)\sum_{k=0}^{m}\sum_{i=N}^{N+\phi} a_{i-k}^2 \leq (m+1)^2 B.$$

Hence there is some $n \in [N; N + \phi]$ such that (3) fails, and by (2) the result follows. $\square$

3. Main results

For the remaining sections, let $x_0 \in X$ be given and consider (x_n) to be the iteration generated by (D). We start with some facts that follow easily from the definition of the algorithm.

Lemma 3.1. *For all $n \in \mathbb{N}^*$:*

(*i*) $x_{n-1} - x_n = q_n - q_{n-m}$,
(*ii*) $x_0 - x_n = \sum_{k=n-m+1}^{n} q_k$,
(*iii*) $x_n \in C_n$ *and* $\forall z \in C_n\, (\langle x_n - z, q_n\rangle \geq 0)$,
(*iv*) $\langle x_n - x_{n+m}, q_n\rangle \geq 0$.

Proof. Let $n \in \mathbb{N}^*$ be given. Fact (*i*) follows immediately from the definition of q_n. Now, from (*i*) we easily derive (*ii*). Indeed,

$$x_0 - x_n = \sum_{k=1}^{n} x_{k-1} - x_k = \sum_{k=1}^{n} q_k - q_{k-m} = \sum_{k=1}^{n} q_k - \sum_{k=-(m-1)}^{n-m} q_k$$
$$= \sum_{k=1}^{n} q_k - \sum_{k=1}^{n-m} q_k = \sum_{k=n-m+1}^{n} q_k.$$

The definition of x_n entails (*iii*) using the definition of q_n and the characterization of the metric projection in Proposition 2.4. Finally, point (*iv*) is an immediate consequence of (*iii*) as $x_{n+m} \in C_{n+m} = C_{j_{n+m}} = C_{j_n} = C_n$. $\square$

We also have the following useful inequality.

Lemma 3.2. *For all $n \in \mathbb{N}$,* $\sum_{k=n-m+1}^{n} \|q_k\| \leq \sum_{k=0}^{n-1} \|x_k - x_{k+1}\|$.

Proof. We argue by induction on $n \in \mathbb{N}$. Since $q_{-(m-1)} = \cdots = q_0 = 0$, the base case $n = 0$ is trivial. For the induction step,

$$\begin{aligned}\sum_{k=n-m+2}^{n+1} \|q_k\| &= \sum_{k=n-m+1}^{n} \|q_k\| + \|q_{n+1}\| - \|q_{n-m+1}\| \\ &\overset{\text{IH}}{\leq} \sum_{k=0}^{n-1} \|x_k - x_{k+1}\| + \|q_{n+1} - q_{n+1-m}\| = \sum_{k=0}^{n} \|x_k - x_{k+1}\|,\end{aligned}$$

using Lemma 3.1.(*i*). This concludes the proof. $\square$

We now prove the main equality used throughout Dykstra's proof.

Lemma 3.3. *For all $z \in X$ and $i, n \in \mathbb{N}$ with $i \geq n$,*[3]

$$\begin{aligned}\|x_n - z\|^2 =& \|x_i - z\|^2 + \sum_{k=n}^{i-1} \left(\|x_k - x_{k+1}\|^2 + 2\langle x_{k-m+1} - x_{k+1}, q_{k-m+1}\rangle\right) \\ &+ 2 \sum_{k=i-m+1}^{i} \langle x_k - z, q_k\rangle - 2 \sum_{k=n-m+1}^{n} \langle x_k - z, q_k\rangle,\end{aligned} \tag{4}$$

and in particular

$$\|x_i - z\|^2 \leq \|x_n - z\|^2 + 2 \sum_{k=n-m+1}^{n} \langle x_k - z, q_k\rangle - 2 \sum_{k=i-m+1}^{i} \langle x_k - z, q_k\rangle. \tag{5}$$

Proof. The proof of identity (4) is by induction on i. The base case $i = n$ is trivial. For $i + 1$, we first see that

$$\begin{aligned}\langle x_{i+1} - z, q_{i+1} - q_{i+1-m}\rangle &= \langle x_{i+1} - z, q_{i+1}\rangle \\ &\quad - \left(\langle x_{i+1} - x_{i-m+1}, q_{i-m+1}\rangle + \langle x_{i-m+1} - z, q_{i-m+1}\rangle\right) \\ &= \langle x_{i-m+1} - x_{i+1}, q_{i-m+1}\rangle \\ &\quad + \langle x_{i+1} - z, q_{i+1}\rangle - \langle x_{i-m+1} - z, q_{i-m+1}\rangle,\end{aligned}$$

and so, using Lemma 3.1.(*i*)

$$\begin{aligned}\|x_i - z\|^2 &= \langle (x_{i+1} - z) + (x_i - x_{i+1}), (x_{i+1} - z) + (x_i - x_{i+1})\rangle \\ &= \|x_{i+1} - z\|^2 + \|x_i - x_{i+1}\|^2 + 2\langle x_{i+1} - z, x_i - x_{i+1}\rangle \\ &= \|x_{i+1} - z\|^2 + \|x_i - x_{i+1}\|^2 + 2\langle x_{i+1} - z, q_{i+1} - q_{i+1-m}\rangle \\ &= \|x_{i+1} - z\|^2 + \|x_i - x_{i+1}\|^2 + 2\langle x_{i-m+1} - x_{i+1}, q_{i-m+1}\rangle \\ &\quad + 2\langle x_{i+1} - z, q_{i+1}\rangle - 2\langle x_{i-m+1} - z, q_{i-m+1}\rangle.\end{aligned}$$

[3]Here one considers $x_{-(m-1)}, \cdots, x_{-1}$ arbitrary points in X.

We can now verify the induction step,

$$\begin{aligned}
\|x_n - z\|^2 &\overset{\text{IH}}{=} \|x_i - z\|^2 + \sum_{k=n}^{i-1} \left(\|x_k - x_{k+1}\|^2 + 2\langle x_{k-m+1} - x_{k+1}, q_{k-m+1}\rangle\right) \\
&\quad + 2\sum_{k=i-m+1}^{i} \langle x_k - z, q_k\rangle - 2\sum_{k=n-m+1}^{n} \langle x_k - z, q_k\rangle \\
&= \Big(\|x_{i+1} - z\|^2 + \|x_i - x_{i+1}\|^2 + 2\langle x_{i-m+1} - x_{i+1}, q_{i-m+1}\rangle \\
&\quad + 2\langle x_{i+1} - z, q_{i+1}\rangle - 2\langle x_{i-m+1} - z, q_{i-m+1}\rangle\Big) \\
&\quad + \sum_{k=n}^{i-1} \left(\|x_k - x_{k+1}\|^2 + 2\langle x_{k-m+1} - x_{k+1}, q_{k-m+1}\rangle\right) \\
&\quad + 2\sum_{k=i-m+1}^{i} \langle x_k - z, q_k\rangle - 2\sum_{k=n-m+1}^{n} \langle x_k - z, q_k\rangle \\
&= \|x_{i+1} - z\|^2 + \sum_{k=n}^{i} \left(\|x_k - x_{k+1}\|^2 + 2\langle x_{k-m+1} - x_{k+1}, q_{k-m+1}\rangle\right) \\
&\quad + 2\sum_{k=i-m+2}^{i+1} \langle x_k - z, q_k\rangle - 2\sum_{k=n-m+1}^{n} \langle x_k - z, q_k\rangle,
\end{aligned}$$

which concludes the induction and proves (4). To verify (5), just note that the terms in the first sum are all nonnegative. □

The assumption of feasibility, $\bigcap_{j=1}^m C_j \neq \emptyset$, now entails that the iteration (x_n) is bounded and $\sum \|x_k - x_{k+1}\|^2 < \infty$. For the sequel, we fix some point $p \in \bigcap_{j=1}^m C_j$ and a natural number $b \in \mathbb{N}^*$ such that $b \geq \|x_0 - p\|$.

Lemma 3.4. *For all $n \in \mathbb{N}$,*

$$\|x_n - p\| \leq b \text{ and } \sum_{k=0}^{n} \|x_k - x_{k+1}\|^2 \leq b^2.$$

Proof. The result follows immediately from (4) and (5) with $z = p$ and $n = 0$ using Lemma 3.1.*(iii)* and the fact that $\sum_{k=-(m-1)}^{0} \langle x_k - p, q_k\rangle = 0$. □

Using Lemma 2.9, we derive a lim inf-rate for the first step in the convergence proof of Dykstra's algorithm.

Proposition 3.5. *We have $\underline{\lim}_n \sum_{k=n-m+1}^{n} |\langle x_k - x_n, q_k\rangle| = 0$, and moreover, for all $\varepsilon > 0$ and $N \in \mathbb{N}$,*

$$\exists n \in [N; N + \Phi(b, m, \varepsilon, N)] \left(\sum_{k=n-m+1}^{n} |\langle x_k - x_n, q_k\rangle| \leq \varepsilon\right), \tag{6}$$

where $\Phi(b, m, \varepsilon, N) := \phi_{b^2}(m, \varepsilon, N)$, with ϕ as defined in Lemma 2.9.

Proof. Let $\varepsilon > 0$ and $N \in \mathbb{N}$ be given. As we have seen $\sum \|x_k - x_{k+1}\|^2 \leq b^2$ and so we can apply Lemma 2.9 (instantiated with $a_n = \|x_n - x_{n+1}\|$ and $B = b^2$) to conclude that there is $n \in [N; N + \Phi(b, m, \varepsilon, N)]$ such that

$$\left(\sum_{k=n-m+1}^{n} \|x_k - x_{k+1}\|\right) \cdot \left(\sum_{k=0}^{n} \|x_k - x_{k+1}\|\right) = (s_n - s_{n-m})s_n \leq \varepsilon.$$

By triangle inequality, for all $k \in [n-m+1; n]$,

$$\|x_k - x_n\| \leq \sum_{\ell=k}^{n-1} \|x_\ell - x_{\ell+1}\| \leq \sum_{\ell=n-m+1}^{n-1} \|x_\ell - x_{\ell+1}\|,$$

and thus using Cauchy-Schwarz and Lemma 3.2, we get

$$\begin{aligned}
\sum_{k=n-m+1}^{n} |\langle x_k - x_n, q_k\rangle| &\leq \sum_{k=n-m+1}^{n} \|x_k - x_n\| \cdot \|q_k\| \\
&\leq \left(\sum_{k=n-m+1}^{n} \|q_k\|\right)\left(\sum_{\ell=n-m+1}^{n-1} \|x_\ell - x_{\ell+1}\|\right) \\
&\leq \left(\sum_{k=0}^{n} \|x_k - x_{k+1}\|\right)\left(\sum_{k=n-m+1}^{n} \|x_k - x_{k+1}\|\right) \leq \varepsilon,
\end{aligned}$$

which, in particular, means that $\underline{\lim}_n \sum_{k=n-m+1}^{n} |\langle x_k - x_n, q_k\rangle| = 0$. □

3.1. **Asymptotic regularity.** Here we discuss the asymptotic regularity of (x_n). In the previous subsection, we saw that $\sum_{k=0}^{n} \|x_k - x_{k+1}\|^2 \leq b^2$. Hence by Lemma 2.8 we immediately derive the following result.

Proposition 3.6. *We have* $\lim \|x_n - x_{n+1}\| = 0$*, and moreover*

$$\forall \varepsilon > 0\ \forall f \in \mathbb{N}^{\mathbb{N}}\ \exists n \leq \Psi(b^2, \varepsilon^2, f)\ \forall k \in [n; n+f(n)]\ (\|x_k - x_{k+1}\| \leq \varepsilon),$$

where Ψ *is as defined in Lemma 2.8.*

One can now show the asymptotic regularity with respect to the individual projections.

Proposition 3.7. *For all* $j \in [1; m]$*, we have* $\lim \|x_n - P_j(x_n)\| = 0$*, and*

$$\forall \varepsilon > 0\ \forall f \in \mathbb{N}^{\mathbb{N}}\ \exists n \leq \alpha(b, m, \varepsilon, f)\ \forall k \in [n; n+f(n)] \left(\bigwedge_{j=1}^{m} \|x_k - P_j(x_k)\| \leq \varepsilon\right)$$

where $\alpha(b, m, \varepsilon, f) := \Psi(b^2, \left(\frac{\varepsilon}{m-1}\right)^2, f + m - 2)$*, with* Ψ *as defined in Lemma 2.8.*

Proof. For given $\varepsilon > 0$ and $f : \mathbb{N} \to \mathbb{N}$, by Proposition 3.6 there is $n \leq \alpha(b, m, \varepsilon, f)$ such that

$$\forall k \in [n; n+f(n)+m-2] \left(\|x_k - x_{k+1}\| \leq \frac{\varepsilon}{m-1}\right). \tag{7}$$

Consider $k \in [n; n+f(n)]$. By the definition, we have $x_k \in C_{j_k}$ with $j_k := [k-1]+1$. For all $i \in [0; m-1]$, as $x_{k+i} \in C_{j_{k+i}}$ from the definition of the projection map $P_{j_{k+i}}$, we have

$$\begin{aligned}
\|x_k - P_{j_{k+i}}(x_k)\| \leq \|x_k - x_{k+i}\| &\leq \sum_{\ell=k}^{k+i-1} \|x_\ell - x_{\ell+1}\| \\
&\leq \sum_{\ell=k}^{k+m-2} \|x_\ell - x_{\ell+1}\| \leq (m-1) \cdot \frac{\varepsilon}{m-1} = \varepsilon,
\end{aligned}$$

where in the last step we use the fact that $[k; k+m-2] \subset [n; n+f(n)+m-2]$ and (7). The conclusion now follows from observing that for any $k \in \mathbb{N}$ the cyclic definition of $C_{(\cdot)}$ entails $\{P_{j_{k+i}} : i \in [0; m-1]\} = \{P_1, \cdots, P_m\}$. □

Remark 3.8. *Note that the previous argument is clearly constructive, and the reason we only obtain a rate of metastability is due to having only a metastability statement in Proposition 3.6. This is already clear by Proposition 2.3 and the fact that the counterfunction f only appears in the definition of α when the bounding information from Proposition 3.6 depends on the functional parameter. Indeed, if the conclusion of Proposition 3.6 would hold with a rate of asymptotic regularity ψ,*

$$\forall \varepsilon > 0\ \forall k \geq \psi(\varepsilon)\left(\|x_k - x_{k+1}\| \leq \varepsilon\right),$$

then the same argument would entail a rate satisfying

$$\forall \varepsilon > 0\ \forall k \geq \widetilde{\psi}(\varepsilon)\left(\bigwedge_{j=1}^{m} \|x_k - P_j(x_k)\| \leq \varepsilon\right),$$

with $\widetilde{\psi}(\varepsilon) := \psi(\varepsilon/(m-1))$.

3.2. **Metastability.** Here we show the full quantitative version of Theorem 1.4. Moreover, our proof bypasses all of the compactness principles used in the original proof. We start with an easy remark regarding some of the data obtained so far.

Remark 3.9. *The function α (Proposition 3.7) is monotone in ε,*

$$\varepsilon \leq \varepsilon' \to \alpha(b, m, \varepsilon, f) \geq \alpha(b, m, \varepsilon', f).$$

The function Φ (Proposition 3.5) is monotone in N,

$$N \leq N' \to \Phi(b, m, \varepsilon, N) \leq \Phi(b, m, \varepsilon, N').$$

We have the following result which plays a central role in bypassing the compactness principles used in the original argument.

Proposition 3.10. *Let $\varepsilon > 0$ and a function $\Delta : \mathbb{N} \to (0, \infty)$ be given. Then,*

$$\exists n \leq \gamma(b, m, \varepsilon, \Delta)\ \exists x \in \overline{B}_b(p)$$
$$\left(\bigwedge_{j=1}^{m} \|x - P_j(x)\| \leq \Delta(n) \wedge \|x - x_n\| \leq \varepsilon \wedge \sum_{k=n-m+1}^{n} \langle x_k - x_n, q_k\rangle \leq \varepsilon\right),$$

where $\gamma(b, m, \varepsilon, \Delta) := \overline{\alpha}(\overline{\beta}) + \Phi_\varepsilon\big(\overline{\alpha}(\overline{\beta})\big)$ with

$$\overline{\beta} := \beta\left(b, \frac{\varepsilon^2}{2}, \delta\right),$$

$$\delta(\eta) := \min\left\{\frac{\varepsilon^2}{8b\big(\overline{\alpha}(\eta) + \Phi_\varepsilon\big(\overline{\alpha}(\eta)\big)\big)}, \widetilde{\Delta}\Big(\overline{\alpha}(\eta) + \Phi_\varepsilon\big(\overline{\alpha}(\eta)\big)\Big)\right\},\ \textit{for all } \eta > 0,$$

$$\overline{\alpha}(\eta) := \alpha(b, m, \eta, \Phi_\varepsilon),\ \textit{for all } \eta > 0,$$

$$\Phi_\varepsilon(N) := \Phi\left(b, m, \frac{\varepsilon^2}{4}, N\right),\ \textit{for all } N \in \mathbb{N},$$

$$\widetilde{\Delta}(k) := \min\{\Delta(k') : k' \leq k\},\ \textit{for all } k \in \mathbb{N},$$

α, β, Φ are as in Propositions 3.7, 2.5 and 3.5, respectively.

Proof. By Proposition 2.5 with $u = x_0$, there are $\eta_0 \geq \overline{\beta}$ and $x \in \overline{B}_b(p)$ such that

$$\bigwedge_{j=1}^{m} \|x - P_j(x)\| \leq \delta(\eta_0) \tag{8}$$

and

$$\forall y \in \overline{B}_b(p) \left(\bigwedge_{j=1}^{m} \|y - P_j(y)\| \leq \eta_0 \to \langle x_0 - x, y - x \rangle \leq \frac{\varepsilon^2}{2} \right). \tag{9}$$

Considering Proposition 3.7 with $\varepsilon = \eta_0$ and $f = \Phi_\varepsilon$, we obtain

$$\exists N_0 \leq \overline{\alpha}(\eta_0)\ \forall i \in [N_0; N_0 + \Phi_\varepsilon(N_0)] \left(\bigwedge_{j=1}^{m} \|x_i - P_j(x_i)\| \leq \eta_0 \right).$$

Since $(x_n) \subseteq \overline{B}_b(p)$, by (9) we have $\forall i \in [N_0; N_0 + \Phi_\varepsilon(N_0)] \left(\langle x_0 - x, x_i - x \rangle \leq \varepsilon^2/2\right)$.

On the other hand, from Proposition 3.5 (with $\varepsilon = \frac{\varepsilon^2}{4}$ and $N = N_0$) and the definition of the function Φ_ε, there exists $n_0 \in [N_0; N_0 + \Phi_\varepsilon(N_0)]$ such that

$$\sum_{k=n_0-m+1}^{n_0} \langle x_k - x_{n_0}, q_k \rangle \leq \frac{\varepsilon^2}{4}.$$

At this point, we remark that $n_0 \leq \gamma(b, m, \varepsilon, \Delta)$. Indeed, by Remark 3.9 and as $\eta_0 \geq \overline{\beta}$,

$$n_0 \leq N_0 + \Phi_\varepsilon(N_0) \leq \overline{\alpha}(\eta_0) + \Phi_\varepsilon\big(\overline{\alpha}(\eta_0)\big) \leq \overline{\alpha}(\overline{\beta}) + \Phi_\varepsilon\big(\overline{\alpha}(\overline{\beta})\big) = \gamma(b, m, \varepsilon, \Delta).$$

The definition of the function δ, then entails

$$\delta(\eta_0) \leq \widetilde{\Delta}(\overline{\alpha}(\eta_0) + \Phi_\varepsilon\left(\overline{\alpha}(\eta_0)\right)) \leq \Delta(n_0).$$

Hence, the first and the last term of the conjunction in the result hold true *a fortiori*, and it remains to verify that $\|x - x_{n_0}\| \leq \varepsilon$. Note that the definition of δ also entails

$$\delta(\eta_0) \leq \frac{\varepsilon^2}{8b(\overline{\alpha}(\eta_0) + \Phi_\varepsilon(\overline{\alpha}(\eta_0))} \leq \frac{\varepsilon^2}{8b(N_0 + \Phi_\varepsilon(N_0))} \leq \frac{\varepsilon^2}{8bn_0}.$$

Thus,
$$\begin{aligned}
\|x - x_{n_0}\|^2 &= \langle x - x_{n_0}, x - x_0 \rangle + \langle x - x_{n_0}, x_0 - x_{n_0} \rangle \\
&\leq \frac{\varepsilon^2}{2} + \langle x - x_{n_0}, x_0 - x_{n_0} \rangle \\
&= \frac{\varepsilon^2}{2} + \sum_{k=n_0-m+1}^{n_0} \langle x - x_{n_0}, q_k \rangle \qquad \text{by Lemma 3.1.}(ii) \\
&= \frac{\varepsilon^2}{2} + \sum_{k=n_0-m+1}^{n_0} \langle x_k - x_{n_0}, q_k \rangle + \sum_{k=n_0-m+1}^{n_0} \langle x - x_k, q_k \rangle \\
&\leq \frac{\varepsilon^2}{2} + \frac{\varepsilon^2}{4} + \sum_{k=n_0-m+1}^{n_0} \langle \underbrace{P_k(x)}_{\in C_k} - x_k, q_k \rangle + \sum_{k=n_0-m+1}^{n_0} \langle x - P_k(x), q_k \rangle \\
&\leq \frac{3\varepsilon^2}{4} + \sum_{k=n_0-m+1}^{n_0} \langle x - P_k(x), q_k \rangle \qquad \text{by Lemma 3.1.}(iii) \\
&\leq \frac{3\varepsilon^2}{4} + \sum_{k=n_0-m+1}^{n_0} \|x - P_k(x)\| \cdot \|q_k\| \\
&\leq \frac{3\varepsilon^2}{4} + \delta(\eta_0) \sum_{k=0}^{n_0-1} \|x_k - x_{k+1}\| \qquad \text{by Lemma 3.2} \\
&\leq \frac{3\varepsilon^2}{4} + \delta(\eta_0) \cdot n_0 \cdot 2b \leq \frac{3\varepsilon^2}{4} + \frac{\varepsilon^2}{8bn_0} \cdot n_0 \cdot 2b = \varepsilon^2,
\end{aligned}$$

which gives $\|x - x_{n_0}\| \leq \varepsilon$ and concludes the proof. □

We are now ready to prove the quantitative version of Theorem 1.4.

Theorem 3.11. *Let $C_1, \cdots, C_m$ be $m \geq 2$ convex subsets of a Hilbert space X such that $\bigcap_{j=1}^m C_j \neq \emptyset$. Let $x_0 \in X$ and $b \in \mathbb{N}^*$ be given such that $b \geq \|x_0 - p\|$ for some $p \in \bigcap_{j=1}^m C_j$. Then, the sequence (x_n) generated by (D) is a Cauchy sequence and for all $\varepsilon \in (0,1]$ and $f : \mathbb{N} \to \mathbb{N}$,*

$$\exists n \leq \Omega(m, b, \varepsilon, f)\ \forall i, j \in [n; n + f(n)]\ (\|x_i - x_j\| \leq \varepsilon)\,,$$

where $\Omega(b, m, \varepsilon, f) := \gamma(b, m, \tilde{\varepsilon}, \Delta_{\varepsilon,f})$ with γ as defined in Proposition 3.10, $\tilde{\varepsilon} := \dfrac{\varepsilon^2}{96b}$ and

$$\Delta_{\varepsilon,f}(k) := \frac{\varepsilon^2}{48b \cdot \max\{k + f(k), 1\}}, \text{ for all } k \in \mathbb{N}.$$

Proof. Let $\varepsilon \in (0,1]$ and a function $f : \mathbb{N} \to \mathbb{N}$ be given. By Proposition 3.10, there are $n_0 \leq \Omega(b, m, \varepsilon, f)$ and $x \in \overline{B}_b(p)$ such that

(*a*) $\bigwedge_{j=1}^m \|x - P_j(x)\| \leq \Delta_{\varepsilon,f}(n_0)$,
(*b*) $\|x_{n_0} - x\| \leq \tilde{\varepsilon} \leq \frac{\varepsilon}{\sqrt{12}}$,[4]
(*c*) $\sum_{k=n_0-m+1}^{n_0} \langle x_k - x_{n_0}, q_k\rangle \leq \tilde{\varepsilon} \leq \frac{\varepsilon^2}{48}$.

In order to verify that the result holds for such n_0, we consider $i \in [n_0; n_0 + f(n_0)]$. We assume that $f(n_0) \geq 1$, and thus $\max\{n_0 + f(n_0), 1\} = n_0 + f(n_0)$, otherwise the result trivially holds. Since $i \geq n_0$, by (5) and using (*b*), we have

$$\begin{aligned}
\|x_i - x\|^2 &\leq \|x_{n_0} - x\|^2 + 2\sum_{k=n_0-m+1}^{n_0} \langle x_k - x, q_k\rangle - 2\sum_{k=i-m+1}^{i} \langle x_k - x, q_k\rangle \\
&\leq \frac{\varepsilon^2}{12} + 2\underbrace{\sum_{k=n_0-m+1}^{n_0} \langle x_k - x, q_k\rangle}_{t_1} + 2\underbrace{\sum_{k=i-m+1}^{i} \langle x - x_k, q_k\rangle}_{t_2}.
\end{aligned}$$

Using (*b*), (*c*) and Lemma 3.1.(*ii*), we get

$$\begin{aligned}
t_1 &= \sum_{k=n_0-m+1}^{n_0} \langle x_k - x_{n_0}, q_k\rangle + \sum_{k=n_0-m+1}^{n_0} \langle x_{n_0} - x, q_k\rangle \\
&\leq \frac{\varepsilon^2}{48} + \langle x_{n_0} - x, \sum_{k=n_0-m+1}^{n_0} q_k\rangle = \frac{\varepsilon^2}{48} + \langle x_{n_0} - x, x_0 - x_{n_0}\rangle \\
&\leq \frac{\varepsilon^2}{48} + 2b \cdot \|x_{n_0} - x\| \leq \frac{\varepsilon^2}{48} + 2b \cdot \tilde{\varepsilon} = \frac{\varepsilon^2}{48} + 2b \cdot \frac{\varepsilon^2}{96b} = \frac{\varepsilon^2}{24}.
\end{aligned}$$

[4]For simplicity, we assumed that $\varepsilon \in (0,1]$ to avoid replacing ε with $\min\{\varepsilon, 1\}$ in the definition of $\tilde{\varepsilon}$ and still have the second inequality in (b).

and, using (a) and Kolmogorov's criterion, we get

$$\begin{aligned}
t_2 &= \sum_{k=i-m+1}^{i} \langle x - P_k(x), q_k\rangle + \underbrace{\sum_{k=i-m+1}^{i} \langle \underbrace{P_k(x)}_{\in C_k} - x_k, q_k\rangle}_{\leq 0} \\
&\leq \sum_{k=i-m+1}^{i} \langle x - P_k(x), q_k\rangle \leq \sum_{k=i-m+1}^{i} \|x - P_k(x)\| \|q_k\| \\
&\leq \Delta_{\varepsilon,f}(n_0) \sum_{k=i-m+1}^{i} \|q_k\| \leq \Delta_{\varepsilon,f}(n_0) \cdot \sum_{k=0}^{i-1} \|x_k - x_{k+1}\| \quad \text{by Lemma 3.2} \\
&\leq 2b \cdot i \cdot \Delta_{\varepsilon,f}(n_0) = 2b \cdot i \cdot \frac{\varepsilon^2}{48b(n_0 + f(n_0))} \leq \frac{\varepsilon^2}{24},
\end{aligned}$$

using in the last inequality the fact that $i \leq n_0 + f(n_0)$. Overall, we conclude that

$$\|x_i - x\|^2 \leq \frac{\varepsilon^2}{12} + \frac{\varepsilon^2}{12} + \frac{\varepsilon^2}{12} = \frac{\varepsilon^2}{4},$$

and thus $\|x_i - x\| \leq \varepsilon/2$, which entails the result by triangle inequality. □

In contrast with the lack of a full rate of convergence, the reader should note the high uniformity of the rate of metastability obtained. Our function does not depend on specifics of the underlying space nor on any additional geometric properties of the convex sets. The rate only depends on the parameters $m \geq 2$ for the number of sets, and $b \in \mathbb{N}^*$ for a bound on the distance between the initial point and the feasibility set.

Remark 3.12. *Theorem 3.11 is a true finitization of Dykstra's convergence result in the sense that, besides only discussing properties for a finite number of terms, it implies back the original statement in a mathematically simple way. Indeed, if the sets are closed, as the sequence (x_n) satisfies the metastability property it is a Cauchy sequence, and by completeness it converges to some point of the space, say $z = \lim x_n$. By Proposition 3.7 and continuity of the projection maps P_j, we conclude that z must be a common fixed point for all the projection, i.e. $z \in \bigcap_{j=1}^m C_j$. It only remains to argue that the limit point is actually the feasible point closest to x_0. Let $C := \bigcap_{j=1}^m C_j$ and write $P_C(x_0)$ for the projection of x_0 onto the intersection. Consider $\varepsilon > 0$ to be arbitrarily given and $N_0 \in \mathbb{N}$ such that $\|x_n - z\| \leq \min\{\frac{\varepsilon^2}{8b}, \frac{\varepsilon}{2}\}$ for all $n \geq N_0$ (with b as before). As per Proposition 3.5, we may consider some $n_0 \geq N_0$ such that*

$$\sum_{k=n_0-m+1}^{n_0} \langle x_k - x_{n_0}, q_k\rangle \leq \frac{\varepsilon^2}{8}.$$

Since $z \in C$, by Proposition 2.4, we also have

$$\langle P_C(x_0) - x_{n_0}, P_C(x_0) - x_0\rangle \leq \langle z - x_{n_0}, P_C(x_0) - x_0\rangle \leq b \cdot \|z - x_{n_0}\| \leq \frac{\varepsilon^2}{8}.$$

It is now easy to see that

$$
\begin{aligned}
\|P_C(x_0) - x_{n_0}\|^2 &\leq \frac{\varepsilon^2}{8} + \langle P_C(x_0) - x_{n_0}, x_0 - x_{n_0}\rangle \\
&= \frac{\varepsilon^2}{8} + \sum_{k=n_0-m+1}^{n_0} \langle P_C(x_0) - x_{n_0}, q_k\rangle \qquad \text{by Lemma 3.1.(ii)} \\
&= \frac{\varepsilon^2}{8} + \sum_{k=n_0-m+1}^{n_0} \langle x_k - x_{n_0}, q_k\rangle + \underbrace{\sum_{k=n_0-m+1}^{n_0} \langle P_C(x_0) - x_k, q_k\rangle}_{\leq 0, \text{ by Lemma 3.1.(iii)}} \\
&\leq \frac{\varepsilon^2}{8} + \frac{\varepsilon^2}{8} = \left(\frac{\varepsilon}{2}\right)^2,
\end{aligned}
$$

which entails $\|P_C(x_0) - z\| \leq \varepsilon$ *and so, as* ε *is arbitrary,* $z = P_C(x_0)$.

4. Rates of convergence and regularity

In this section, we study the rate of convergence for Dykstra's algorithm under a regularity assumption on the structure of the convex sets. We remark that a regularity assumption on the convex sets $C_1, \cdots, C_m$ is known to allow for rates of convergence already for simpler iterative methods. As explained in [29, pp 291–292], a solution to the CFP can be obtained via a Mann-type iteration and, based on the work in [20], rates of convergence are available under a regularity condition. Moreover, in [29, pp 288] the authors obtained rates of convergence for (MAP) even in a general nonlinear setting and an extremely fast rate was given in [29, Corollary 4.17]. However, these studies were concerned with the interplay between regularity and iterative methods which are Fejér monotone with respect to the feasibility set. In this context, the study of Dykstra's algorithm is of particular interest as the iteration fails to be Fejér monotone and yet it was possible to obtain rates of convergence.

4.1. **The rate of convergence.** Denote $C := \bigcap_{j=1}^m C_j$ and let p be some point of X. We call a function $\mu : \mathbb{N} \times (0, \infty) \to (0, \infty)$ satisfying for all $\varepsilon > 0$ and $r \in \mathbb{N}$,

$$
(\star) \qquad \forall x \in \overline{B}_r(p) \left(\bigwedge_{j=1}^m \|x - P_j(x)\| \leq \mu_r(\varepsilon) \to \exists z \in C\, \|x - z\| \leq \varepsilon \right)
$$

a *modulus of regularity* for the sets $C_1, \cdots, C_m$ (centred at p). Thus, a modulus of regularity lets us know how close to the individual sets (i.e. $\mu_r(\varepsilon)$-almost P_j fixed point) must a point be so that we are sure that it is close to the intersection set (i.e. an ε-almost P_C fixed point). We refer the reader to [29] where this notion was developed and shown to be an effective tool for a unified discussion of several concepts in convex optimization.

Remark 4.1. *Clearly the conclusion of* $(\star)$ *is equivalent to* $\mathrm{dist}(x, C) \leq \varepsilon$. *Furthermore, the existence of a modulus of regularity centred at* p *(say* μ^p*), obviously entails the existence of a modulus of regularity centred at any other* $q \in X$ *(say* μ^q*) – it is easy to verify that for any* $q \in X$, $\mu^q_r : \varepsilon \mapsto \mu^p_{r+\lceil\|p-q\|\rceil}(\varepsilon)$ *works. The choice of the point* p *is always clear by the context and so we just write* μ.

When a modulus of regularity is available, we can actually give rates of convergence for Dykstra's iteration.

Theorem 4.2. *Consider* $x_0 \in X$ *and a natural number* $b \in \mathbb{N}^*$ *such that* $b \geq \|x_0 - p\|$ *for some* $p \in C$. *Let* μ *be a function satisfying* $(\star)$. *Then,*

$$
\forall \varepsilon > 0\ \forall i, j \geq \Theta(b, m, \varepsilon)\, (\|x_i - x_j\| \leq \varepsilon),
$$

where $\Theta(b,m,\varepsilon) := \alpha(b,m,\mu_b(\tilde{\varepsilon}),\Phi_\varepsilon) + \Phi_\varepsilon(\alpha(b,m,\mu_b(\tilde{\varepsilon}),\Phi_\varepsilon))$ *with*

$$\tilde{\varepsilon} := \frac{\varepsilon^2}{32b},\ \Phi_\varepsilon(N) := \Phi\left(b,m,\frac{\varepsilon^2}{16},N\right) \text{ for all } N\in\mathbb{N},$$
$$\alpha,\Phi \text{ are as in Propositions 3.7 and 3.5, respectively.}$$

In particular, (x_n) *converges with rate* Θ.

Proof. By Proposition 3.7, there is $N_0 \leq \alpha(b,m,\mu_b(\tilde{\varepsilon}),\Phi_\varepsilon)$ such that

$$\forall n \in [N_0; N_0+\Phi_\varepsilon(N_0)] \left(\bigwedge_{j=1}^{m} \|x_n - P_j(x_n)\| \leq \mu_b(\tilde{\varepsilon})\right).$$

Since $(x_n) \subseteq \overline{B}_b(p)$, by the assumption ($\star$) on μ it follows that

$$\forall n \in [N_0; N_0+\Phi_\varepsilon(N_0)]\ \exists z \in C \left(\|x_n - z\| \leq \frac{\varepsilon^2}{32b}\right). \tag{10}$$

Applying Proposition 3.5 with $\varepsilon = \dfrac{\varepsilon^2}{16}$ and $N = N_0$, we have $n_0 \in [N_0; N_0+\Phi_\varepsilon(N_0)]$ such that

$$\sum_{k=n_0-m+1}^{n_0} \langle x_k - x_{n_0}, q_k\rangle \leq \frac{\varepsilon^2}{16}.$$

By (10), let $z_0 \in C$ be such that $\|z_0 - x_{n_0}\| \leq \dfrac{\varepsilon^2}{32b}$. Thus, for any $i \geq n_0$,

$$\begin{aligned}
\sum_{k=i-m+1}^{i} \langle x_k - x_{n_0}, q_k\rangle &= \underbrace{\sum_{k=i-m+1}^{i} \langle x_k - z_0, q_k\rangle}_{\geq 0,\ \text{by Lemma 3.1.}(iii)} + \sum_{k=i-m+1}^{i} \langle z_0 - x_{n_0}, q_k\rangle \\
&\geq \langle z_0 - x_{n_0}, \sum_{k=i-m+1}^{i} q_k\rangle \\
&= \langle z_0 - x_{n_0}, x_0 - x_i\rangle \qquad \text{by Lemma 3.1.}(ii) \\
&\geq -\|z_0 - x_{n_0}\| \cdot \|x_0 - x_i\| \geq -\frac{2b\varepsilon^2}{32b} = -\frac{\varepsilon^2}{16}.
\end{aligned}$$

Now by (5) with $n = n_0$ and $z = x_{n_0}$,

$$\|x_i - x_{n_0}\|^2 \leq 2\sum_{k=n_0-m+1}^{n_0} \langle x_k - x_{n_0}, q_k\rangle - 2\sum_{k=i-m+1}^{i} \langle x_k - x_{n_0}, q_k\rangle \leq \frac{\varepsilon^2}{4},$$

which entails that $\|x_i - x_{n_0}\| \leq \dfrac{\varepsilon}{2}$ and the result follows by triangle inequality. $\square$

In particular, Θ is also a rate of asymptotic regularity for the sequence (x_n) and, by Remark 3.8, the function $\Theta' : \varepsilon \mapsto \Theta(b,m,\frac{\varepsilon}{m-1})$ is a rate of asymptotic regularity with respect to the individual projections.

We now recall the following class of convex sets in $\mathbb{R}^n$.

Definition 4.3. *A set* $C \subseteq \mathbb{R}^n$ *is called a basic semi-algebraic convex set in* $\mathbb{R}^n$ *if there exist* $\gamma \geq 1$ *convex polynomial functions* g_i *on* $\mathbb{R}^n$ *such that*

$$C := \{x \in \mathbb{N}^n : g_i(x) \leq 0, i \in [1;\gamma]\}.$$

We remark that the class of basic semi-algebraic convex sets is a broad class of convex sets which includes in particular the polyhedral case and the class of convex sets described by convex quadratic functions. It was observed in [29] that the study of Hölderian regularity in [4] entails the existence of a modulus of regularity for basic semi-algebraic convex sets with respect to compact sets. As such, we immediately have the following example of application of Theorem 4.2.

Example 4.4. *Let $C_1, \cdots, C_m \subseteq \mathbb{R}^n$ be basic semi-algebraic sets described by convex polynomials $g_{i,j}$ with degree at most $d \in \mathbb{N}$, and such that $\bigcap_{j=1}^m C_j \neq \emptyset$. Consider some $p \in \mathbb{R}^n$. Then, for any $r \in \mathbb{N}$ there exists $c > 0$ such that*

$$\mu_r(\varepsilon) := \frac{(\varepsilon/c)^\sigma}{m}, \text{ with } \sigma := \min\left\{\frac{(2d-1)^n+1}{2}, B(n-1)d^n\right\},$$

where $B(n) := \binom{n}{\lfloor n/2 \rfloor}$ is the central binomial coefficient with respect to n, is a modulus of regularity for $C_1, \cdots, C_m$ centred at p.[5] *Hence, by Theorem 4.2 one has a uniform rate of convergence for Dykstra's cyclic projections algorithm for basic semi-algebraic convex sets in $\mathbb{R}^n$.*

4.2. **Regularity.** We now argue that a modulus of regularity is a necessary condition for the existence of uniform convergence rates.

Proposition 4.5. *Let $x_0 \in X$ and $b \in \mathbb{N}^*$ be such that $b \geq \|x_0 - p\|$ for some $p \in \bigcap_{j=1} C_j$. Consider (x_n) the iteration generated by (D) with initial point x_0. Then,*

$$\forall \varepsilon > 0\ \forall n \in \mathbb{N} \left(\bigwedge_{j=1}^m \|x_0 - P_j(x_0)\| \leq \frac{\varepsilon^2}{4bn} \to \|x_n - x_0\| \leq \varepsilon \right).$$

Proof. Let $\varepsilon > 0$ and $n \in \mathbb{N}$ be given. Assuming the premise of the implication, by (5) (with $i = n$, $n = 0$ and $z = x_0$), we have

$$\begin{aligned}
\|x_n - x_0\|^2 &\leq 2 \underbrace{\sum_{k=-(m-1)}^{0} \langle x_k - x_0, q_k \rangle}_{=0} + 2 \sum_{k=n-m+1}^{n} \langle x_0 - x_k, q_k \rangle \\
&= \sum_{k=n-m+1}^{n} \langle x_0 - P_k(x_0), q_k \rangle + 2 \sum_{k=n-m+1}^{n} \underbrace{\langle \underbrace{P_k(x_0)}_{\in C_k} - x_k, q_k \rangle}_{\leq 0 \text{ by Lemma 3.1.}(iii)} \\
&\leq 2 \sum_{k=n-m+1}^{n} \|x_0 - P_k(x_0)\| \cdot \|q_k\| \leq \frac{\varepsilon^2}{2bn} \sum_{k=n-m+1}^{n} \|q_k\| \\
&\leq \frac{\varepsilon^2}{2bn} \sum_{k=0}^{n-1} \|x_k - x_{k+1}\| \leq \varepsilon^2 \qquad \text{using Lemma 3.2,}
\end{aligned}$$

which entails $\|x_n - x_0\| \leq \varepsilon$ and concludes the proof. ∎

Note that the natural proof that the scheme (D) satisfies

$$x_0 \in \bigcap_{j=1}^m C_j \to \forall n \in \mathbb{N}\, (x_n = x_0)$$

[5] Which is to a modulus of regularity for $C_1 \cap \cdots \cap C_m$ with respect to the compact set $\overline{B}_r(p) \subseteq \mathbb{R}^n$, in the terminology used in [29].

doesn't require the knowledge that the functions P_j are the metric projections, but only that they are extensional. Since the functions P_j, being nonexpansive, have a trivial modulus of uniform continuity independent of any majorizability assumption, logical considerations make it clear that there must exist a bound which does not depend on the additional constant $b \in \mathbb{N}$. With such perspective, we give an alternative version of Proposition 4.5.

Proposition 4.6. *Consider (x_n) to be the iteration generated by (D) with some initial point $x_0 \in X$. Then,*

$$\forall \varepsilon > 0 \ \forall n \in \mathbb{N}^* \left(\bigwedge_{j=1}^{m} \|x_0 - P_j(x_0)\| \leq \frac{\varepsilon}{5^{n-1}} \to \|x_n - x_0\| \leq \varepsilon \right).$$

Proof. By induction we show the stronger assertion that for all $n \in \mathbb{N}^*$,

$$\forall \varepsilon > 0 \left(\bigwedge_{j=1}^{m} \|x_0 - P_j(x_0)\| \leq \frac{\varepsilon}{5^{n-1}} \to \forall n' \in [1;n] \left(\|x_{n'} - x_0\| \leq \varepsilon \wedge \|q_{n'-m}\| \leq \varepsilon \right) \right).$$

For $n = 1$, we have $q_{1-m} = 0$ and

$$\|x_1 - x_0\| = \|P_1(x_0) - x_0\| \leq \frac{\varepsilon}{5^0} = \varepsilon.$$

Assuming that the claim holds for some $n \in \mathbb{N}^*$, suppose that

$$\bigwedge_{j=1}^{m} \|x_0 - P_j(x_0)\| \leq \frac{\varepsilon}{5^n}.$$

By the induction hypothesis, we have

$$\forall n' \in [1;n] \left(\|x_{n'} - x_0\| \leq \frac{\varepsilon}{5} \wedge \|q_{n'-m}\| \leq \frac{\varepsilon}{5} \right), \tag{11}$$

and, in particular, we just need to verify the consequent for $n' = n + 1$. Let us first focus on q_{n+1-m}. If $n < m$, then $q_{n+1-m} = 0$ and so we assume $n \geq m$. We have,

$$\|q_{n+1-m}\| = \|x_{n-m} + q_{n+1-2m} - x_{n+1-m}\| \leq \|x_{n-m} - x_{n+1-m}\| + \|q_{n+1-2m}\|$$

Since $n + 1 - m \in [1;n]$, by (11) we have $\|q_{n+1-2m}\| \leq \varepsilon/5$. On the other hand, we have $\|x_{n-m} - x_{n+1-m}\| \leq 2\varepsilon/5$. Indeed, if $n = m$ then

$$\|x_{n-m} - x_{n+1-m}\| = \|x_0 - x_1\| = \|x_0 - P_1(x_0)\| \leq \frac{\varepsilon}{5^n} \leq \frac{2\varepsilon}{5}.$$

If $n > m$, then $n - m$, $n + 1 - m \in [1;n]$ and by (11),

$$\|x_{n-m} - x_{n+1-m}\| \leq \|x_{n-m} - x_0\| + \|x_{n+1-m} - x_0\| \leq \frac{2\varepsilon}{5}.$$

Overall, we conclude that $\|q_{n+1-m}\| \leq \frac{3\varepsilon}{5}$, which in particular gives the second conjunct. It is now possible to verify that

$$\begin{aligned}
\|x_{n+1} - x_0\| &\leq \|P_{n+1}(x_n + q_{n+1-m}) - P_{n+1}(x_0)\| + \|P_{n+1}(x_0) - x_0\| \\
&\leq \|x_n - x_0\| + \|q_{n+1-m}\| + \|P_{n+1}(x_0) - x_0\| \\
&\leq \frac{\varepsilon}{5} + \frac{3\varepsilon}{5} + \frac{\varepsilon}{5^n} \leq \varepsilon,
\end{aligned}$$

concluding the proof. □

The next result states that the existence of rates of convergence that are uniform for initial points in $\overline{B}_b(p)$, entails the existence of a modulus of regularity for the convex sets.

Proposition 4.7. *Let $p \in \bigcap_{j=1}^m C_j =: C$. For any $b \in \mathbb{N}$, assume the existence of a common rate of convergence towards the limit $P_C(x_0)$ for any iteration generated by* (D) *with initial point $x_0 \in \overline{B}_b(p)$, i.e.*

$$\|x_0 - p\| \leq b \ \to\ \forall \varepsilon > 0\ \forall n \geq \rho(b, \varepsilon)\, (\|x_n - P_C(x_0)\| \leq \varepsilon)\,.$$

Then, the function

$$\mu(b, \varepsilon) := \max\left\{\frac{\varepsilon^2}{16b \cdot \rho(b, \varepsilon/2)}, \frac{\varepsilon}{2 \cdot 5^{\rho(b,\varepsilon/2)}}\right\}$$

is a modulus of regularity for the sets $C_1, \cdots, C_m$ centred at p.

Proof. Consider $\varepsilon > 0$, $b \in \mathbb{N}$ and $x \in \overline{B}_b(p)$, and assume that

$$\bigwedge_{j=1}^{m} \|x - P_j(x)\| \leq \max\left\{\frac{\varepsilon^2}{16b \cdot \rho(b, \varepsilon/2)}, \frac{\varepsilon}{2 \cdot 5^{\rho(b,\varepsilon/2)}}\right\}.$$

Let (x_n) be the iteration generated by (D) with initial point $x_0 := x$. Then by both Proposition 4.5 and 4.6, $\|x_{\rho(b,\varepsilon/2)} - x\| \leq \frac{\varepsilon}{2}$. On the other hand, by the assumption on ρ, we have $\|x_{\rho(b,\varepsilon/2)} - P_C(x)\| \leq \frac{\varepsilon}{2}$. Hence,

$$\|x - P_C(x)\| \leq \|x - x_{\rho(b,\varepsilon/2)}\| + \|x_{\rho(b,\varepsilon/2)} - P_C(x)\| \leq \varepsilon. \qquad \square$$

Note that the requirement $p \in C$ is tied with Proposition 4.5 and can we can take p an arbitrary point in X with suitable changes. The key idea of the previous argument is that a rate of convergence for an iteration which remains constant whenever the initial point is already in the target set, will entail the existence of a modulus of regularity. An analogous argument was used in [29, Proposition 4.4] for the case of the Picard iteration in a metric setting. Indeed, this reasoning can be stated in a general framework, but we refrain from doing it here and direct the reader to [31, 39].

If a rate of convergence is available but it is sensitive to the initial point, then we obtain a weaker result (with unclear usefulness).

Proposition 4.8. *For any $x_0 \in X$, assume the existence of rate of convergence towards $P_C(x_0)$ for the iteration generated by* (D) *with initial point x_0, i.e.*

$$\forall \varepsilon > 0\ \forall n \geq \rho(x_0, \varepsilon)\, (\|x_n - P_C(x_0)\| \leq \varepsilon)\,.$$

Then,

$$\forall \varepsilon > 0\ \forall x \in X \left(\bigwedge_{j=1}^{m} \|x - P_j(x)\| \leq \frac{\varepsilon}{2 \cdot 5^{\rho(x,\varepsilon/2)}} \to \|x - P_C(x)\| \leq \varepsilon\right).$$

For the particular case of the intersection of two half-spaces, Deutsch and Hundal [10] obtained a rate of convergence for Dykstra's algorithm which is uniform on the choice of the initial point depending only on a bound to its distance to the intersection set. By Propositions 4.7, such situation entails a modulus of regularity for the two half-spaces. In full generality, but provided there exists a modulus of regularity, Theorem 4.2 guarantees the existence of uniform rates of convergence. This still leaves open the possibility that no modulus of regularity exists and yet rates of convergence are available. Such rates would necessarily be sensitive to the initial point of the iteration – such is the case with the rates of convergence obtained by Deutsch and Hundal for the general polyhedral case. In contrast, we obtained rates of metastability in full generality which are uniform in all the parameters of the convex feasibility problem.

5. Final Remarks

This quantitative study analyzes the proof of strong convergence of Dykstra's cyclic projection algorithm. Although the original proof relies on several strong mathematical principles, in the end we obtain simple computable metastability rates (primitive recursive in f in the sense of Kleene [21]) which are highly uniform in the parameters of the convex feasibility problem. Indeed, our rates only require information on the number of convex sets m, and an upper bound b on the distance between the initial point and the feasibility set. Furthermore, under a regularity assumption, we adapt the argument to actually derive uniform rates of convergence towards the feasible point closest to the starting term of the iteration. We show that the regularity condition is actually necessary for the existence of such uniform rates. The regularity assumption comes in the form of a modulus of regularity μ which (informally) guarantees that the point is ε-close to the intersection whenever it is $\mu(\varepsilon)$-close to all the individual convex sets. In the general case, the finitary version follows through the crucial observation that the role of the weak limit can actually be replaced by that of a weak version of the projection of x_0 onto the intersection set. We show that our main result (Theorem 3.11) is a true finitary version of Theorem 1.4 in the sense that it only regards a finite number of iteration terms and the full statement is fully recovered in an elementary way from the quantitative version (cf. Remark 3.12).

This kind of argument is in line with the macro developed in [17]. The ability to establish the Cauchy property of the iteration without the use of sequential weak compactness is of paramount relevance as it ensures that the final quantitative bound information will be of a simple nature (namely, it can be described without the need of Spector's bar-recursive functionals [44]). This technique has been applied several times in proof mining (e.g. in [38, 12, 11]). Moreover, even if one is not concerned with quantitative information, a simpler proof which bypasses complex comprehension principles (in this case the arithmetical comprehension required to justify weak compactness) allows for easier generalizations of the original result (see e.g. the recent [13] where a quantitative approach allowed to establish a fully new result in a geodesic setting in which weak compactness arguments, common in Hilbert spaces, are harder to employ). For the case at hand, the simple quantitative perspective on the convergence of Dykstra's algorithm allowed for the recent work in [40] establishing the convergence of Dykstra's method with Bregman projections (first introduced in [7]) in general (reflexive) Banach spaces.

Previous eliminations of weak compactness principles were applied to Halpern-type iterations and to convergence proofs following a similar common proof structure. However, Dykstra's algorithm doesn't appear to have any connection with the Halpern iteration and the proof follows a completely different argument. Thus, it was not a priori known if it would be possible to bypass the compactness arguments crucial in the original proof (see the extensive proof-theoretical discussion in [39]). Furthermore, regarding the discussion under a regularity assumption, as is explained in [29], it is known that for Fejér monotone iterations a modulus of regularity allows one to obtain rates of convergence. Note that in this case however, Dykstra's method fails to be Fejér monotone, and still it was possible to extract uniform rates of convergence. This phenomenon is explained in terms of a generalized notion of Fejér monotonicity in the recent [31].

Acknowledgment. This work benefited from discussions with Ulrich Kohlenbach and Nicholas Pischke. The proof mining study of Dykstra's algorithm is one of the objectives in Ulrich Kohlenbach's DFG Project KO 1737/6-2.
Funding. The research was supported by the DFG Project KO 1737/6-2.
Conflicts of interests. The author has no conflicts of interests to declare.

References

[1] H. H. Bauschke. The composition of projections onto closed convex sets in Hilbert space is asymptotically regular. *Proc. Amer. Math. Soc.*, 131(1):141–146, 2003.

[2] H. H. Bauschke and J. M. Borwein. On projection algorithms for solving convex feasibility problems. *SIAM Rev.*, 38(3):367–426, 1996.

[3] H. H. Bauschke and P. L. Combettes. *Convex analysis and monotone operator theory in Hilbert spaces.* CMS Books in Mathematics. Springer, 2nd edition, 2017.

[4] J. M. Borwein, G. Li, and L. Yao. Analysis of the convergence rate for the cyclic projection algorithm applied to basic semialgebraic convex sets. *SIAM J. Optim.*, 24(1):498–527, 2014.

[5] J. P. Boyle and R. L. Dykstra. A method for finding projections onto the intersection of convex sets in Hilbert spaces. In *Advances in order restricted statistical inference (Iowa City, Iowa, 1985)*, volume 37 of *Lect. Notes Stat.*, pages 28–47. Springer, Berlin, 1986.

[6] L. M. Bregman. The method of successive projections for finding a common point of convex sets. In *Soviet Math. Dokl.*, volume 6, pages 688–692, 1965.

[7] Y. Censor and S. Reich. The Dykstra algorithm with Bregman projections. *Commun. Appl. Anal.*, 2(3):407–419, 1998.

[8] F. Deutsch. The method of alternating orthogonal projections. In *Approximation theory, spline functions and applications (Maratea, 1991)*, volume 356 of *NATO Adv. Sci. Inst. Ser. C: Math. Phys. Sci.*, pages 105–121. Kluwer Acad. Publ., Dordrecht, 1992.

[9] F. Deutsch. Dykstra's cyclic projections algorithm: the rate of convergence. In *Approximation theory, wavelets and applications (Maratea, 1994)*, volume 454 of *NATO Adv. Sci. Inst. Ser. C: Math. Phys. Sci.*, pages 87–94. Kluwer Acad. Publ., Dordrecht, 1995.

[10] F. Deutsch and H. Hundal. The rate of convergence of Dykstra's cyclic projections algorithm: the polyhedral case. *Numer. Funct. Anal. Optim.*, 15(5-6):537–565, 1994.

[11] B. Dinis and P. Pinto. On the convergence of algorithms with Tikhonov regularization terms. *Optim. Lett.*, 15(4):1263–1276, 2021.

[12] B. Dinis and P. Pinto. Quantitative results on the multi-parameters proximal point algorithm. *J. Convex Anal.*, 28(3):729–750, 2021.

[13] B. Dinis and P. Pinto. Strong convergence for the alternating Halpern-Mann iteration in CAT(0) spaces. *SIAM Journal on Optimization*, 33(2):785–815, 2023.

[14] B. Dinis and P. Pinto. Effective metastability for a method of alternating resolvents. *Fixed Point Theory*, 25(1):61–98, 2024.

[15] R. L. Dykstra. An algorithm for restricted least squares regression. *J. Amer. Statist. Assoc.*, 78(384):837–842, 1983.

[16] P. Engrácia and F. Ferreira. Bounded functional interpretation with an abstract type. In *Contemporary logic and computing*, volume 1 of *Landsc. Log*, pages 87–112. Coll. Publ., London, 2020.

[17] F. Ferreira, L. Leuştean, and P. Pinto. On the removal of weak compactness arguments in proof mining. *Adv. Math.*, 354:106728, 55pp, 2019.

[18] I. Halperin. The product of projection operators. *Acta Sci. Math. (Szeged)*, 23:96–99, 1962.

[19] H. Hundal. An alternating projection that does not converge in norm. *Nonlinear Anal.*, 57(1):35–61, 2004.

[20] M. A. A. Khan and U. Kohlenbach. Quantitative image recovery theorems. *Nonlinear Anal.*, 106:138–150, 2014.

[21] S. C. Kleene. *Introduction to metamathematics.* D. Van Nostrand Co., Inc., New York, 1952.

[22] U. Kohlenbach. Some logical metatheorems with applications in functional analysis. *Transactions of the American Mathematical Society*, 357(1):89–128, 2005.

[23] U. Kohlenbach. *Applied Proof Theory: Proof Interpretations and their Use in Mathematics*. Springer Monographs in Mathematics. Springer-Verlag Berlin Heidelber, 2008.

[24] U. Kohlenbach. On the logical analysis of proofs based on nonseparable Hilbert space theory. In *Proofs, categories and computations*, volume 13 of *Tributes*, pages 131–143. Coll. Publ., London, 2010.

[25] U. Kohlenbach. On quantitative versions of theorems due to F.E. Browder and R. Wittmann. *Adv. Math.*, 226(3):2764–2795, 2011.

[26] U. Kohlenbach. On the quantitative asymptotic behavior of strongly nonexpansive mappings in Banach and geodesic spaces. *Israel J. Math.*, 216(1):215–246, 2016.

[27] U. Kohlenbach. A polynomial rate of asymptotic regularity for compositions of projections in Hilbert space. *Found. Comput. Math.*, 19(1):83–99, 2019.

[28] U. Kohlenbach. Proof-theoretic Methods in Nonlinear Analysis. In Boyan Sirakov, Paulo Ney de Souza, and Marcelo Viana, editors, *Proceedings of the International Congress of Mathematicians 2018*, volume 2, pages 62–82. World Scientific, 2019.

[29] U. Kohlenbach, G. López-Acedo, and A. Nicolae. Moduli of regularity and rates of convergence for Fejér monotone sequences. *Israel J. Math.*, 232(1):261–297, 2019.

[30] U. Kohlenbach and P. Pinto. Quantitative translations for viscosity approximation methods in hyperbolic spaces. *J. Math. Anal. Appl.*, 507(2):125823, 33pp, 2022.

[31] U. Kohlenbach and P. Pinto. Fejér monotone sequences revisited. submitted, 2024.

[32] U. Kohlenbach and N. Pischke. Proof theory and non-smooth analysis. *Philos. Trans. Roy. Soc. A*, 381(2248):Paper No. 20220015 , 21, 2023.

[33] E. Kopecká and S. Reich. A note on the von Neumann alternating projections algorithm. *J. Nonlinear Convex Anal.*, 5(3):379–386, 2004.

[34] E. Kopecká and S. Reich. Another note on the von Neumann alternating projections algorithm. *J. Nonlinear Convex Anal.*, 11(3):455–460, 2010.

[35] L. Leuştean and P. Pinto. Rates of asymptotic regularity for the alternating Halpern-Mann iteration. To appear in *Optimization Letters*, 2023. arXiv:2206.02226.

[36] E. Matoušková and S. Reich. The Hundal example revisited. *J. Nonlinear Convex Anal.*, 4(3):411–427, 2003.

[37] P. Pinto. *Proof mining with the bounded functional interpretation.* PhD thesis, Universidade de Lisboa, 2019.

[38] P. Pinto. A rate of metastability for the Halpern type proximal point algorithm. *Numer. Funct. Anal. Optim.*, 42(3):320–343, 2021.

[39] P. Pinto. Proof mining and the convex feasibility problem: The curious case of Dykstra's algorithm. submitted, 2024.

[40] P. Pinto and N. Pischke. On Dykstra's algorithm with Bregman projections. submitted, Oberwolfach Preprint OWP-2024-04, 2024.

[41] A. Sipoş. Quantitative inconsistent feasibility for averaged mappings. *Optim. Lett.*, 16(6):1915–1925, 2022.

[42] A. Sipoş. The computational content of super strongly nonexpansive mappings and uniformly monotone operators. *To appear in: Israel Journal of Mathematics*, 2023. ArXiv Preprint 2303.02768.

[43] E. Specker. Nicht konstruktiv beweisbare Sätze der Analysis. *J. Symbolic Logic*, 14:145–158, 1949.

[44] C. Spector. Provably recursive functionals of analysis: a consistency proof of analysis by an extension of principles formulated in current intuitionistic mathematics. In *Proc. Sympos. Pure Math., Vol. V*, pages 1–27. Amer. Math. Soc., Providence, RI, 1962.

[45] T. Tao. Norm convergence of multiple ergodic averages for commuting transformations. *Ergodic Theory Dynam. Systems*, 28(2):657–688, 2008.

[46] T. Tao. *Structure and Randomness: Pages from Year One of a Mathematical Blog.* American Mathematical Society Providence, RI, 2008.

[47] J. von Neumann. *Functional Operators. II. The Geometry of Orthogonal Spaces.* Annals of Mathematics Studies, No. 22. Princeton University Press, Princeton, N. J., 1950. Reprint of mimeographed lecture notes first distributed in 1933.