Finite Convergence of A Subgradient Projections Method with Expanding Controls

Yair Censor[1], Wei Chen[2*], and Homeira Pajoohesh[3]

[1]Department of Mathematics, University of Haifa
Mt. Carmel, Haifa 31905, Israel
(yair@math.haifa.ac.il)

[2]Department of Computer Science
The Graduate Center, City University of New York
365 Fifth Avenue, New York, NY 10016, USA
(chen.wei@mgh.harvard.edu)

[3]Department of Mathematics
Medgar Evers College, City University of New York
Brooklyn, New York, NY 11225, USA
(hpajoohesh@mec.cuny.edu)

November 21, 2010. Revised: May 10, 2011.

Abstract

We study finite convergence of the modified cyclic subgradient projections (MCSP) algorithm for the convex feasibility problem (CFP) in the Euclidian space. Expanding control sequences allow the indices of the sets of the CFP to re-appear and be used again by the algorithm

*Currently with the Department of Radiation Oncology, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts 02114, USA.

within windows of iteration indices whose lengths are not constant but may increase without bound. Motivated by another development in finitely convergent sequential algorithms that has a significant real-world application in the field of radiation therapy treatment planning, we show that the MCSP algorithm retains its finite convergence when used with an expanding control that is repetitive and fulfills an additional condition.

Keywords. Convex feasibility problem, modified subgradient projections, finite convergence, expanding controls, repetitive control, quasi-cyclic control.

1 Introduction

In this paper we consider the *convex feasibility problem* (CFP) in the finite-dimensional Euclidean space R^n, which is to find a point in the nonempty intersection $Q := \cap_{i=1}^m Q_i$ of a finite family $\{Q_i\}_{i=1}^m$ of subsets $Q_i \subseteq R^n$. It is assumed that the sets are given by level-sets of convex functions g_i, i.e., for all $i \in \mathcal{M} := \{1, 2, \ldots, m\}$,

$$Q_i = \{x \in R^n \mid g_i(x) \leq 0\}. \tag{1}$$

This is a vast field of research in optimization and applied mathematics with many practical consequences. The CFP is a fundamental problem in many areas of mathematics and the physical sciences, see, e.g., Combettes [10, 11] and references therein. It has been used to model significant real-world problems in image reconstruction from projections, see, e.g., Herman [17], in radiation therapy treatment planning, see Censor, Altschuler and Powlis [5] and in crystallography, see Marks, Sinkler and Landree [22], to name but a few. Published works related to the CFP is extensive and includes, among many others, Byrne [2], Cegielski [3], Jiang and Wang [20], Kiwiel [21], and Yamada [25].

Projection methods have been particularly useful in solving such problems and we focus here on one specific projection method, namely, the *modified cyclic subgradient projection* (MCSP) method of De Pierro and Iusem [13]. They have shown that certain perturbations of the *cyclic subgradient projection* (CSP) method of Censor and Lent [6] and Eremin [14, 15] (see also [7, Algorithm 5.3.1]), can make the method *converge finitely* without loosing its

row-action nature, see Censor [4], see also Bauschke and Borwein [1, Example 7.19 and Remark 7.20]. However, the control sequence, which governs the manner by which the sets Q_i are taken up by the algorithm, could be in [13] at most almost cyclic. This means that all indices of $\mathcal{M}$ must re-appear in, i.e., be re-used by, the algorithm as iterations proceed, within iteration index "windows" of bounded lengths.

In this work we make for the MCSP method the leap into the realm of control sequences with *unbounded* window lengths. Inspired by the quasi-cyclic control proposed and studied by Tseng and Bertsekas [24] and Tseng [23], and further used by Combettes [12], we term such control sequences *expanding controls* and show, specifically, that the finite convergence property of the MCSP method is preserved for expanding controls of the repetitive type that fulfill an additional condition (Condition 19 below). To our knowledge, no previous work attempted to show finite convergence with expanding (thus with unbounded windows of indices) controls. Earlier work on finite convergence of iterative projection methods for the CFP appears in Goffin [16].

Condition 19 is algorithmic–dependent and difficult, if not impossible, to verify. Therefore, our efforts to prove finite convergence of the MCSP method with expanding control sequences give limited results by applying only to repetitive controls which fulfill this condition. Further investigation is needed to remove or weaken Condition 19, if possible. The motivation to pursue finite convergence of the MCSP method with expanding control sequences comes from another algorithmic development in finitely convergent sequential algorithms that has a real-world application in the field of radiation therapy treatment planning. We explain this in Section 4.

2 Control sequences for sequential projection methods

We define control sequences for sequential projection methods as follows. Denote by $\mathcal{N} := \{0, 1, 2, \cdots\}$ the set of all positive integers and zero, and denote by $\mathcal{M} := \{1, 2, \ldots, m\}$ a given finite index set.

Definition 1 *Given a monotonically increasing sequence* $\{\tau_k\}_{k=0}^{\infty} \subset \mathcal{N}$, *a mapping* $i : \mathcal{N} \to \mathcal{M}$ *is called a* `control with respect to the sequence`

$\{\tau_k\}_{k=0}^{\infty}$ *if it defines a* `control sequence` $\{i(t)\}_{t=0}^{\infty}$, *such that, for all* $k \geq 0$,

$$\mathcal{M} \subseteq \{i(\tau_k), i(\tau_k + 1), \ldots, i(\tau_{k+1} - 1)\}. \tag{2}$$

Call the set $\{\tau_k, \tau_k + 1, \ldots, \tau_{k+1} - 1\}$ *the* k`-th window` *(with respect to the given sequence* $\{\tau_k\}_{k=0}^{\infty}$*) and define its* `length` *by* $C_k := \tau_{k+1} - \tau_k$. *A control with respect to a sequence* $\{\tau_k\}_{k=0}^{\infty}$ *for which* $\{C_k\}_{k=0}^{\infty}$ *is an unbounded sequence (in contrast with bounded) will be called an* `expanding control`.

Different choices of the sequence $\{\tau_k\}_{k=0}^{\infty}$ and different conditions on the window lengths C_k give rise to the following specific control sequences. The first three are controls with bounded window lengths.

Example 2 ***Almost cyclic control****: This is any control sequence* $\{i(t)\}_{t=0}^{\infty}$ *for which there exists some monotonically increasing sequence* $\{\tau_k\}_{k=0}^{\infty}$, *such that (2) is satisfied and* $\tau_k = Ck$ *where* $C \geq m$, *i.e., every window must contain the set* $\mathcal{M}$ *and the windows are all of the same fixed length* C.

Example 3 m***-window control****: This is any almost cyclic control for which* $\tau_k = mk$, *for all* $k \geq 0$.

Example 4 ***Cyclic control****: This is any* m*-window control for which the order of the indices in the first window repeats in all subsequent windows, namely,* $i(t) = t(\bmod m) + 1$, *for all* $t \in \mathcal{N}$.

The next three controls allow the window lengths to increase without bound, as k increases, and may, therefore, be *expanding controls*.

Example 5 ***Repetitive control****: This is a control sequence* $\{i(t)\}_{t=0}^{\infty}$ *for which there exists some monotonically increasing sequence* $\{\tau_k\}_{k=0}^{\infty}$, *such that (2) is satisfied and* $\lim_{k\to\infty} C_k = +\infty$ *without any additional condition imposed. This means that the only thing that matters is that every index* $i \in \mathcal{M}$ *should appear again, no matter how large* C_k *is. Alternatively, a control sequence* $\{i(t)\}_{t=0}^{\infty}$ *is repetitive in* $\mathcal{N}$ *if for every* $t \in \mathcal{N}$ *there exists a positive integer* Δ_t *such that* $\mathcal{M} \subseteq \{t, t+1, \ldots, t+\Delta_t - 1\}$. *Elsewhere this control is called also* `chaotic`.

Example 6 ***Quasi-cyclic control****: This is any repetitive control such that the window lengths* $C_k \geq m$ *and* $\sum_{k=0}^{\infty} 1/C_k = +\infty$. *This means that the lengths of the windows may (or may not) grow without bound but cannot grow too fast with* k.

Example 7 ***Linearly increasing windows control****: This is any quasi-cyclic control with $\tau_{k+1} = \tau_k + mk$ which implies $C_k = mk$, i.e., the lengths of windows increase linearly with k.*

The quasi-cyclic control was proposed and studied by Tseng and Bertsekas [24] and by Tseng [23] who showed that the method of successive orthogonal projections for the convex feasibility problem still converges under such a control. The quasi-cyclic control was further used by Combettes [12] along with several other controls to pursue convergence properties of his Extrapolated Method of Parallel Projections (EMOPP). By our definitions, a cyclic control is a m-window control, which is an almost cyclic control, which is a linearly increasing windows control, which is a quasi-cyclic control, which is a repetitive control. Additional controls, not mentioned here, such as the *remotest set control*, the *approximately remotest set control*, and the *most violated constraint control*, see, e.g., [7, Section 5.1], and the *admissible*, *coercive*, and *chaotically coercive* from Combettes [12], have also been used.

3 Finite convergence of the modified cyclic subgradient projection method with repetitive control

The modified cyclic subgradient projection (MCSP) method with repetitive control is described as follows.

Algorithm 8 ***The modified cyclic subgradient projection (MCSP) method with repetitive control.***
Initialization*: Pick an arbitrary $x^0 \in R^n$, let $\{\varepsilon_t\}_{t=0}^{\infty}$ be a monotonically decreasing sequence of positive numbers such that $\lim_{t\to\infty} \varepsilon_t = 0$ and $\sum_{t=0}^{\infty} \varepsilon_t = \infty$, and choose any, arbitrarily small, $0 < \beta_1, \beta_2 < 1$.*
Iterative Step*: Given the current iterate x^t, pick a control index $i(t)$ and compute the next iterate by*

$$x^{t+1} = \begin{cases} x^t, & \text{if } g_{i(t)}(x^t) \leq 0, \\ x^t - \alpha_t \dfrac{g_{i(t)}(x^t) + \varepsilon_t}{\left\|s^t\right\|^2} s^t, & \text{otherwise,} \end{cases} \tag{3}$$

where $s^t \in \partial g_{i(t)}(x^t)$ is a subgradient of $g_{i(t)}$ at x^t and $\{\alpha_t\}_{t=0}^{\infty}$ are relaxation parameters.

Relaxation Parameters: *The relaxation parameters* $\{\alpha_t\}_{t=0}^{\infty}$ *must all lie in the interval* $\beta_1 \leq \alpha_t \leq 2 - \beta_2$.
Control Sequence: *The control sequence* $\{i(t)\}_{t=0}^{\infty}$ *is repetitive.*

The following string of lemmas leads to the required finite convergence result. Our analysis follows closely that of [13]. We need a Slater condition on the CFP to guarantee the existence of an interior point of Q. For any $\varepsilon > 0$ define

$$Q_\varepsilon := \{x \in R^n \mid g_i(x) + \varepsilon \leq 0, \text{ for all } i \in \mathcal{M}\}. \tag{4}$$

Condition 9 *Slater condition: There exists a positive* $\hat{\varepsilon}$ *such that* $Q_{\hat{\varepsilon}} \neq \emptyset$.

Remark 10 *Observe, for later use, that Condition 9 implies that* $Q_\varepsilon \neq \emptyset$ *for all* $\varepsilon \in [0, \hat{\varepsilon}]$. *Therefore, for any sequence* $\{\varepsilon_t\}_{t=0}^{\infty}$, *as defined in Algorithm 8, there exists an integer* T *such that* $Q_{\varepsilon_t} \neq \emptyset$ *for all* $t \geq T$, *and such that the following sets are nested as* $Q_{\varepsilon_t} \subset Q_{\varepsilon_{\bar{t}}} \subset Q$ *for all* $\bar{t} \geq t \geq T$.

A sequence $\{x^t\}_{t=0}^{\infty}$ is called `Fejér-monotone with respect to a set` $\Omega \subset R^n$ if for any $\omega \in \Omega$ we have $\|x^{t+1} - \omega\| \leq \|x^t - \omega\|$ for all $t \geq 0$.

Lemma 11 *[13, Lemma 1] Let* $Q_{\hat{\varepsilon}}$ *be as in Condition 9. Then for any sequence* $\{x^t\}_{t=0}^{\infty}$, *generated by Algorithm 8 with* $\{\varepsilon_t\}_{t=0}^{\infty}$, *and with the index* T *as defined in Remark 10, the tail sequence* $\{x^t\}_{t=T}^{\infty}$ *is Fejér-monotone with respect to* $Q_{\hat{\varepsilon}}$.

Lemma 12 *[13, Lemma 2] Let* $Q_{\hat{\varepsilon}}$ *be as in Condition 9 and let* Δ *be any (fixed) positive integer. Then for any sequence* $\{x^t\}_{t=0}^{\infty}$, *generated by Algorithm 8 with* $\{\varepsilon_t\}_{t=0}^{\infty}$, *and with the index* T *as defined in Remark 10, we have for all* $t \geq T$, *and any* $\hat{x} \in Q_{\hat{\varepsilon}}$

$$\left\|x^{t+\Delta} - \hat{x}\right\|^2 \leq \left\|x^t - \hat{x}\right\|^2 - \frac{\beta_2}{2\Delta}\left(\sum_{j=t}^{t+\Delta-1} \left\|x^{j+1} - x^j\right\|\right)^2, \tag{5}$$

where β_2 *is defined in Algorithm 8.*

We comment that in [13, Lemma 2] this lemma is proven for the "almost cyclicality constant" C instead of the Δ here, but the proof there does not depend on the nature of the constant Δ. With the definition $d(x, S) := \inf\{\|x - y\| \mid y \in S\}$ of the distance from a point x to a set S, the following corollary follows.

Corollary 13 *If Condition 9 holds, and if Δ is any (fixed) positive integer, then for any sequence $\{x^t\}_{t=0}^{\infty}$, generated by Algorithm 8 with $\{\varepsilon_t\}_{t=0}^{\infty}$, and with the index T as defined in Remark 10, we have for all $t \geq T$,*

$$d^2(x^{t+\Delta}, Q_{\varepsilon_t}) \leq d^2(x^t, Q_{\varepsilon_t}) - \frac{\beta_2}{2\Delta} \left(\sum_{j=t}^{t+\Delta-1} \left\| x^{j+1} - x^j \right\| \right)^2 . \tag{6}$$

Proof. Condition 9 and Remark 10 assure us that $Q_{\varepsilon_t} \neq \emptyset$ for all $t \geq T$, so that Lemma 12 applies, with $\tilde{Q} = Q_{\varepsilon_t}$. By taking the infimum of both sides of (5) over all $\hat{x} \in Q_{\varepsilon_t}$ the result is obtained. ■

For any sequence $\{x^t\}_{t=0}^{\infty}$, generated by Algorithm 8 with $\{\varepsilon_t\}_{t=0}^{\infty}$, and with the index T as defined in Remark 10 and any $\hat{x} \in Q_{\hat{\varepsilon}}$, define

$$M := \max \left\{ \frac{-\left\| x^T - \hat{x} \right\|}{g_i(\hat{x})} \mid i \in \mathcal{M} \right\} . \tag{7}$$

$M > 0$ since $g_i(\hat{x}) \leq -\hat{\varepsilon}$ for all $i \in \mathcal{M}$.

Lemma 14 *Let Condition 9 hold and let $\{x^t\}_{t=0}^{\infty}$ be any sequence, generated by Algorithm 8. If $x^t \neq x^{t+1}$ then there exists an index $\ell(t) \in \mathcal{M}$ such that*

$$d^2(x^t, Q_{\varepsilon_t}) \leq M^2 \left(g_{\ell(t)}(x^t) + \varepsilon_t \right)^2 , \tag{8}$$

where M is given by (7).

Proof. Since the control of Algorithm 8 does not enter here, the proof follows verbatim from that of [13, Lemma 3]. ■

In the next lemma we make use of the uniform boundedness on bounded sets of the subgradients of a family of convex functions, defined next.

Definition 15 *Given a family $\{g_i\}_{i \in \mathcal{M}}$ of convex functions and any bounded set $U \subseteq R^n$, if there exists a constant R, called* `a uniform bound`*, such that $\|s\| \leq R$ for all subgradients $s \in \partial g_i(x)$ for all $i \in \mathcal{M}$ and all $x \in U$, then we say that the family of convex functions has the uniform boundedness on bounded sets of the subgradients property.*

Remark 16 *Uniform boundedness on bounded sets of the subgradients is a standard property, frequently used in theorems on subgradient projection methods. It holds if the effective domain of all functions is the whole space R^n, as is the case here, see, e.g., [1, Proposition 7.8 and Corollary 7.9]. Therefore, we use this property in the sequel.*

Lemma 17 *Let Condition 9 hold and let $\{x^t\}_{t=0}^{\infty}$ be any sequence generated by Algorithm 8. If $x^t \neq x^{t+1}$ for some t, then*

$$d^2(x^t, Q_{\varepsilon_t}) \leq \left(\frac{2MR}{\beta_1}\right)^2 \left(\sum_{j=t}^{t+\Delta_t-1} \left\|x^{j+1} - x^j\right\|\right)^2, \tag{9}$$

where the real constant R is a uniform bound of the subgradients as in Definition 15, M is defined by (7) and Δ_t is as in Example 5.

Proof. The proof follows closely that of [13, Lemma 4]. We give the details for completeness. From Lemma 14 there exists an index $\ell(t) \in \mathcal{M}$ such that (8) holds. By repetitiveness there exists an index r such that

$$t \leq r \leq t + \Delta_t - 1 \quad \text{and} \quad i(r) = \ell(t). \tag{10}$$

For any sequence $\{x^t\}_{t=0}^{\infty}$, generated by Algorithm 8, it follows from Lemma 11 with $\tilde{Q} = Q_{\varepsilon_t}$, that $\|x^t - \hat{x}\| \leq \left\|x^T - \hat{x}\right\|$ for all $t \geq T$, where the index T is defined in Remark 10. Therefore, the compact set

$$U := \{x \in R^n \mid \|x - \hat{x}\| \leq \left\|x^T - \hat{x}\right\|\} \tag{11}$$

contains the sequence $\{x^t\}_{t=T}^{\infty}$. So, from the subgradient inequality and from uniform boundedness of the subgradients over U of (11) with bound R,

$$\begin{aligned} g_{\ell(t)}(x^t) &\leq g_{\ell(t)}(x^r) + \left\langle s^t, x^t - x^r \right\rangle \leq g_{\ell(t)}(x^r) + R\left\|x^t - x^r\right\| \\ &\leq g_{i(r)}(x^r) + R\sum_{j=t}^{r-1} \left\|x^{j+1} - x^j\right\|. \end{aligned} \tag{12}$$

We claim now that

$$g_{\ell(t)}(x^t) \leq \frac{R}{\beta_1} \sum_{j=t}^{t+\Delta_t-1} \left\|x^{j+1} - x^j\right\|. \tag{13}$$

Indeed, if $g_{i(r)}(x^r) \leq 0$ then, from (12) and the fact that $0 < \beta_1 < 1$,

$$g_{\ell(t)}(x^t) \leq R\sum_{j=t}^{r-1} \left\|x^{j+1} - x^j\right\| \leq \frac{R}{\beta_1} \sum_{j=t}^{t+\Delta_t-1} \left\|x^{j+1} - x^j\right\|. \tag{14}$$

If, on the other hand, $g_{i(r)}(x^r) > 0$ then, from (12) and Remark 16 we obtain

$$\begin{aligned} g_{\ell(t)}(x^t) &\leq g_{i(r)}(x^r) + \varepsilon_r + R\sum_{j=t}^{r-1} \left\|x^{j+1} - x^j\right\| \\ &\leq \frac{R}{\alpha_r}\left\|x^{r+1} - x^r\right\| + R\sum_{j=t}^{r-1}\left\|x^{j+1} - x^j\right\| \leq \frac{R}{\beta_1}\sum_{j=t}^{t+\Delta_t-1}\left\|x^{j+1} - x^j\right\| \end{aligned} \tag{15}$$

because, by Algorithm 8,

$$g_{i(r)}(x^r) + \varepsilon_r \leq \frac{\|s^r\|}{\alpha_r}\left\|x^{r+1} - x^r\right\| \leq \frac{R}{\alpha_r}\left\|x^{r+1} - x^r\right\|. \tag{16}$$

Thus (13) is proven. Now, $g_{i(t)}(x^t) > 0$ since $x^t \neq x^{t+1}$, thus

$$\begin{aligned} \varepsilon_t \leq g_{i(t)}(x^t) + \varepsilon_t &\leq \frac{R}{\alpha_t}\left\|x^{t+1} - x^t\right\| \leq \frac{R}{\beta_1}\left\|x^{t+1} - x^t\right\| \\ &\leq \frac{R}{\beta_1}\sum_{j=t}^{t+\Delta_t-1}\left\|x^{j+1} - x^j\right\|. \end{aligned} \tag{17}$$

From (13) and (17) we get

$$\left(g_{\ell(t)}(x^t) + \varepsilon_t\right)^2 \leq \left(\frac{2R}{\beta_1}\sum_{j=t}^{t+\Delta_t-1}\left\|x^{j+1} - x^j\right\|\right)^2, \tag{18}$$

which, together with (8), yields (9). ■

Corollary 18 *Under the assumptions of Lemma 17, if $x^t \neq x^{t+1}$ for some t then*

$$d(x^{t+\Delta_t}, Q_{\varepsilon_t}) \leq \sigma_t d(x^t, Q_{\varepsilon_t}), \tag{19}$$

where

$$\sigma_t = \left(1 - \frac{\beta_2}{8\Delta_t}\left(\frac{\beta_1}{MR}\right)^2\right)^{1/2}. \tag{20}$$

Proof. The proof follows from Lemma 17 and Corollary 13. ■

Observe that, due to control repetitiveness, $\lim_{t\to\infty}\Delta_t = \infty$ is true and it implies that $\lim_{t\to\infty}\sigma_t = 1$. Since this prevents us from bounding the

sequence $\{\sigma_t\}_{t=0}^{\infty}$ from above by any upper bound smaller than 1, we need to use in the finite convergence theorem a refined argument not present in [13, Theorem 1]. Denote

$$\gamma := \frac{\beta_2}{8}\left(\frac{\beta_1}{MR}\right)^2 \tag{21}$$

so that (20) becomes

$$\sigma_t = \left(1 - \frac{\gamma}{\Delta_t}\right)^{1/2}. \tag{22}$$

Consider Algorithm 8 with $\{\varepsilon_t\}_{t=0}^{\infty}$, and with the index T as defined in Remark 10. For any integer r define

$$h_T(r) := \tau_T + \sum_{\mu=T}^{\mu=T+r} C_\mu \tag{23}$$

where $\{\tau_k\}_{k=0}^{\infty}$ and C_μ are as in Definition 1. Also define the distance

$$\delta_t := d(x^t, Q_{\varepsilon_t}) \tag{24}$$

between an iteration vector x^t and a set Q_{ε_t}.

Condition 19 *Assume that the following limit exists*

$$\liminf_{r\to\infty}\left(\left(\frac{\varepsilon_{h_T(r)}}{R\delta_T}\right)^{4/r} + \frac{2\gamma}{r}\sum_{\substack{\ell=0\\ \ell\ even}}^{r-2}\frac{1}{C_{\ell+T}}\right) > 1, \tag{25}$$

where R is the uniform bound of the subgradients used in Lemma 17.

We present now our finite convergence theorem. Finite convergence means that from a certain iteration index onward the algorithm does not create further changes of the iteration vectors.

Theorem 20 *Under the assumptions of Lemma 17 and Condition 19, any sequence $\{x^t\}_{t=0}^{\infty}$, generated by Algorithm 8, converges finitely to a point in Q.*

Proof. Using $\delta_t := d(x^t, Q_{\varepsilon_t})$, Corollary 13, and Remark 10, we obtain, for all $t \geq T$,

$$\delta_{t+1} = d(x^{t+1}, Q_{\varepsilon_{t+1}}) \leq d(x^{t+1}, Q_{\varepsilon_t}) \leq d(x^t, Q_{\varepsilon_t}) = \delta_t, \tag{26}$$

proving that $\{\delta_t\}_{t=T}^{\infty}$ is monotonically decreasing. If the sequence $\{x^t\}_{t=0}^{\infty}$ does not converge finitely then, for each even r, there exists an iteration index t_r such that

$$t_r \in \left[\tau_T + \sum_{\mu=T}^{\mu=T+r-1} C_\mu, \tau_T + \sum_{\mu=T}^{\mu=T+r} C_\mu\right) = [h_T(r-1), h_T(r)) \tag{27}$$

for which $x^{t_r} \neq x^{t_r+1}$, see the definition of $h_T(r)$ in (23). This is so because this is the $(T+r)$-th window and if there was not such a t_r there then finite convergence would have been established because all indices of $\mathcal{M}$ are included in the window.

Now we describe how this non-finite convergence assumption leads to a contradiction. It follows, from (27) and from the fact that $\{C_k\}_{k=0}^{\infty}$ is a non-decreasing sequence, that

$$\begin{aligned} t_r &\geq \tau_T + C_T + C_{T+1} + \cdots + C_{T+r-2} + C_{T+r-1} \\ &= (\tau_T + C_T + C_{T+1} + \cdots + C_{T+r-2}) + C_{T+r-1} \\ &\geq t_{r-2} + C_{T+r-1} \geq t_{r-2} + C_{T+r-2}, \end{aligned} \tag{28}$$

thus,

$$t_r \geq t_{r-2} + C_{T+r-2} \tag{29}$$

and we apply Corollary 18 together with (26) to obtain:

$$\begin{aligned} \delta_{t_r} &\leq \delta_{t_{r-2}+C_{T+r-2}} \text{ [by (29) and (26)]} \qquad (30) \\ &= d(x^{t_{r-2}+C_{T+r-2}}, Q_{\varepsilon_{t_{r-2}+C_{T+r-2}}}) \text{ [by definition]} \\ &\leq d(x^{t_{r-2}+C_{T+r-2}}, Q_{\varepsilon_{t_{r-2}}}) \text{ [by nested sets: } Q_{\varepsilon_{t_{r-2}}} \subset Q_{\varepsilon_{t_{r-2}+C_{T+r-2}}}] \\ &\leq \sigma_{t_{r-2}}\delta_{t_{r-2}} \text{ [by Corollary 18 with } t_{r-2} \text{ and with } \Delta_{t_{r-2}} = C_{T+r-2}] \\ &\leq \sigma_{t_{r-2}}\delta_{t_{r-4}+C_{T+r-4}} \text{ [applying again the argument of (30)]} \\ &\leq \sigma_{t_{r-2}}\sigma_{t_{r-4}}\delta_{t_{r-4}} \\ &\leq \cdots \leq \sigma_{t_{r-2}}\sigma_{t_{r-4}}\cdots\sigma_{\tau_T}\delta_{\tau_T} \\ &\leq \sigma_{t_{r-2}}\sigma_{t_{r-4}}\cdots\sigma_{\tau_T}\delta_T, \qquad (31) \end{aligned}$$

where the last inequality follows since $\{\delta_t\}_{t=T}^{\infty}$ is monotonically decreasing (see (26)) and $T \leq \tau_T$. Now let z^k be the closest point to x^k in Q_{ε_k}, so that $\delta_{t_r} = \left\|x^{t_r} - z^{t_r}\right\|$ and let $u = i(t_r)$. Then we get from the subgradient inequality

$$g_u(z^{t_r}) \geq g_u(x^{t_r}) + \left\langle s^{t_r}, z^{t_r} - x^{t_r} \right\rangle. \tag{32}$$

Since $g_u(x^{t_r}) > 0$ and $g_u(z^{t_r}) + \varepsilon_{t_r} = 0$, adding ε_{t_r} to both sides of (32) gives

$$\left\langle s^{t_r}, z^{t_r} - x^{t_r} \right\rangle \leq -\varepsilon_{t_r}, \tag{33}$$

which, together with (31), yields

$$\varepsilon_{t_r} \leq \left|\left\langle s^{t_r}, z^{t_r} - x^{t_r} \right\rangle\right| \leq \left\|s^{t_r}\right\| \delta_{t_r} \leq R\delta_{t_r} \leq R\sigma_{t_{r-2}}\sigma_{t_{r-4}} \cdots \sigma_{\tau_T}\delta_T. \tag{34}$$

Setting $t_0 := \tau_T$, denoting $\Delta_{\tau_T} = \Delta_{t_0}$ and using (22), this means that

$$\left(\frac{\varepsilon_{t_r}}{R\delta_T}\right)^2 \leq (\sigma_{t_{r-2}}\sigma_{t_{r-4}} \cdots \sigma_{\tau_T})^2 = \prod_{\substack{\ell=0 \\ \ell \text{ even}}}^{r-2} \left(1 - \frac{\gamma}{\Delta_{t_\ell}}\right). \tag{35}$$

By the standard inequality between the arithmetic and geometric means of a finite set of real numbers, we get

$$\prod_{\substack{\ell=0 \\ \ell \text{ even}}}^{r-2} \left(1 - \frac{\gamma}{\Delta_{t_\ell}}\right) \leq \left((2/r) \sum_{\substack{\ell=0 \\ \ell \text{ even}}}^{r-2} \left(1 - \frac{\gamma}{\Delta_{t_\ell}}\right)\right)^{r/2}, \tag{36}$$

which, in turn, leads to

$$\left(\frac{\varepsilon_{t_r}}{R\delta_T}\right)^{4/r} \leq 1 - \frac{2\gamma}{r} \sum_{\substack{\ell=0 \\ \ell \text{ even}}}^{r-2} \frac{1}{\Delta_{t_\ell}}. \tag{37}$$

But letting $\Delta_{t_\ell} = t_{\ell+1} - t_\ell$ be $C_{\ell+T}$ (see the 3rd line of (30)), and since $t_r \leq h_T(r) = \tau_T + \sum_{\mu=T}^{\mu=T+r} C_\mu$ (by (27)), we obtain $\varepsilon_{h_T(r)} \leq \varepsilon_{t_r}$, thus, for even number r we have

$$\left(\left(\frac{\varepsilon_{h_T(r)}}{R\delta_T}\right)^{4/r} + \frac{2\gamma}{r} \sum_{\substack{\ell=0 \\ \ell \text{ even}}}^{r-2} \frac{1}{C_{\ell+T}}\right) \leq 1. \tag{38}$$

which contradicts Condition 19. ■

We note that Condition 19 is algorithmic-dependent and difficult, if not impossible, to verify. Thus, further investigation is called for to remove or weaken it, if possible. However, our result still partially is a logical generalization of [13, Theorem 1] due to the next lemma.

Lemma 21 *Any almost cyclic control for which*

$$\lim_{r\to\infty} \left(\varepsilon_{h_T(r)}\right)^{4/r} > 1 - \frac{\gamma}{C}, \tag{39}$$

where γ and C are as in (21) and in Example 2, respectively, is a repetitive control that fulfills Condition 19.

Proof. By the definition of almost cyclic control, given in Example 2, $C_k = C$ for all $k \geq 0$. Then the second summand in (25) is

$$\lim_{r\to\infty} \frac{2\gamma}{r} \sum_{\substack{\ell=0 \\ \ell \text{ even}}}^{r-2} \frac{1}{C_{\ell+T}} = \lim_{r\to\infty} \frac{2\gamma}{r} \frac{r}{2C} = \frac{\gamma}{C}. \tag{40}$$

For the first summand in (25)

$$\lim_{r\to\infty} \left(\frac{\varepsilon_{h_T(r)}}{R\delta_T}\right)^{4/r} = \lim_{r\to\infty} \left(\varepsilon_{h_T(r)}\right)^{4/r}. \tag{41}$$

Both $\{\varepsilon_{h_T(r)}\}_{r=0}^{\infty}$ and $\{4/r\}_{r=0}^{\infty}$ tend to zero as $r \to \infty$. (Because $h_T(r) \to \infty$, as $r \to \infty$.) Therefore, and since R and δ_T are constants, the limit in (41) depends on which of the two sequences tends to zero faster. The condition in (39) is sufficient to guarantee Condition 19. ■

4 Motivation from an algorithmic problem

Even though some sequential row-action algorithms [4] have been proved to converge, finitely or asymptotically, under more general controls than a cyclic control (e.g., almost cyclic control, repetitive control, quasi-cyclic control, etc.), there did not exist, to the best of our knowledge, an implementation of a more complicated non-cyclic control for any good reason, until our recent work in [19]. We explain below the main thrust of [19], which is the motivation for our present study.

The algorithm *Algebraic Reconstruction Technique 3* ($ART3$) [18] is a cyclic projection algorithm for solving a system of linear interval inequalities. It is proved to be finitely convergent if the feasible region is full-dimensional. In [19] we proposed an improved version of $ART3$, called $ART3+$. It differs from its predecessor $ART3$ only in that $ART3+$ picks the linear interval constraint to be used in each iteration in a specially designed non-cyclic manner. We proved there that $ART3+$ is still finitely convergent, and demonstrated by numerical experiments that $ART3+$ finds a solution faster than $ART3$. In this way we demonstrated the possibility of improving the performance of a sequential algorithm by resorting to specially designed non-cyclic controls. This speeded-up $ART3+$ algorithm has been successfully applied to radiation therapy treatment planning in [9].

Since the speed-up strategy of [19] is actually independent of the algorithm $ART3$, this speed-up strategy has been generalized to any finitely convergent algorithm in [8]. Instead of proving the finite convergence of $ART3+$ directly as in [19], the logic of the proof of the more general theorem in [8] is as follows:

(1) Prove that the generalized $ART3$ is finitely convergent when its control is repetitive.

(2) Prove that the control sequences that are used by $ART3+$ are repetitive.

(3) Derive the finite convergence of $ART3+$ from (1) and (2) above.

Along these lines, we proposed in [8] a formal transformation from any sequential algorithm ALG to its improved version $ALG+$, and proved that for any sequential algorithm ALG, under some reasonable conditions, if ALG is finitely convergent under repetitive control, then $ALG+$ is finitely convergent. The efficiency of $ALG+$ can be illustrated by comparing the speed of ALG and $ALG+$ in real world applications.

The $MCSP$ algorithm of [13] is a candidate of such an ALG for solving a, not necessarily linear, CFP. If we could prove that the $MCSP$ is finitely convergent under repetitive control, then automatically we would have a proof for the claim that the new algorithm $MCSP+$, whose relative efficiency over $MCSP$ can be verified, is finitely convergent. However, [13] supplies a finite convergence theorem for $MCSP$ only under almost cyclic control. This is why we need a stronger theorem which states that the $MCSP$ is finitely convergent under some expanding control and this is the rational for the investigation present here. We are unable to produce a counter-example to the claim that the MCSP algorithm with repetitive control is finitely

convergent. Our Theorem 20 gives a partial affirmative answer to the claim.

Condition 19, which contains parameters ε and β, is built specifically for MCSP. ART3+ has no such parameters so Condition 19 does not simply apply to it. We hope that this work will be continued and better understanding of finite convergence in the presence of expanding controls will emerge.

Acknowledgments. We heartily thank Gabor Herman for his many insightful comments on this study. We appreciate the advice of two anonymous referees whose comments helped us to improve an earlier version of the paper. This work was partially supported by the United States-Israel Binational Science Foundation (BSF) Grant number 200912 and by Award Number R01HL070472 from the National Heart, Lung, And Blood Institute.

References

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

[2] C. Byrne, Iterative oblique projection onto convex sets and the split feasibility problem, *Inverse Problems,* **18** (2002), 441–453.

[3] A. Cegielski, Generalized relaxations of nonexpansive operators and convex feasibility problems, *Contemporary Mathematics*, **513** (2010), 111–123.

[4] Y. Censor, Row-action methods for huge and sparse systems and their applications, *SIAM Review,* **23** (1981), 444–466.

[5] Y. Censor, M.D. Altschuler, and W.D. Powlis, On the use of Cimmino's simultaneous projections method for computing a solution of the inverse problem in radiation therapy treatment planning, *Inverse Problems,* **4** (1988), 607–623.

[6] Y. Censor and A. Lent, Cyclic subgradient projections, *Mathematical Programming,* **24** (1982), 233–235.

[7] Y. Censor and S.A. Zenios, *Parallel Optimization: Theory, Algorithms, and Applications*, Oxford University Press, New York, NY, USA, 1997.

[8] W. Chen and G.T. Herman, Efficient controls for finitely convergent sequential algorithms, *ACM Transactions on Mathematical Software*, **37** (2010), Article No. 14.

[9] W. Chen, D. Craft, T.M. Madden, K. Zhang, H.M. Kooy and G.T. Herman, A fast optimization algorithm for multi-criteria intensity modulated proton therapy planning, *Medical Physics,* **37** (2010), 4938–4945.

[10] P.L. Combettes, The foundations of set-theoretic estimation, *Proceedings of the IEEE,* **81** (1993), 182–208.

[11] P.L. Combettes, The convex feasibility problem in image recovery, *Advances in Imaging and Electron Physics,* **95** (1996), 155–270.

[12] P.L. Combettes, Hilbertian convex feasibility problem: Convergence of projection methods, *Applied Mathematics and Optimization*, **35** (1997), 311–330.

[13] A.R. De Pierro and A. Iusem, A finitely convergent "row-action" method for the convex feasibility problem, *Applied Mathematics and Optimization*, **17** (1988), 225–235.

[14] I.I. Eremin, The relaxation method for solving systems of inequalities with convex functions on the left-hand side, *Soviet. Math. Doklady,* **6** (1965), 219–222.

[15] I.I. Eremin, On some iterative methods in convex programming, *Ekonomika i Matematichesky Methody,* **2** (1966), 870–886 [In Russian].

[16] J.L. Goffin, The relaxation method for solving systems of linear inequalities, *Mathematics of Operations Research,* **5** (1980), 388–414.

[17] G.T. Herman, *Fundamentals of Computerized Tomography: Image Reconstruction from Projections*, 2nd ed., Springer, 2009.

[18] G.T. Herman, A relaxation method for reconstructing objects from noisy x-rays, *Mathematical Programming*, **8** (1975), 1–19.

[19] G.T. Herman and W. Chen, A fast algorithm for solving a linear feasibility problem with application to intensity-modulated radiation therapy, *Linear Algebra and its Application*, **428** (2008), 1207–1217.

[20] M. Jiang and G. Wang, Convergence studies on iterative algorithms for image reconstruction, *IEEE Transactions on Medical Imaging,* **22** (2003), 569–579.

[21] K.C. Kiwiel, Block-iterative surrogate projection methods for convex feasibility problems, *Linear Algebra and Applications*, **215** (1995), 225–259.

[22] L.D. Marks, W. Sinkler, and E. Landree, A feasible set approach to the crystallographic phase problem, *Acta Crystallographica,* **A55** (1999), 601–612.

[23] P. Tseng, On the convergence of the products of firmly nonexpansive mappings, *SIAM Journal on Optimization*, **2** (1992), 425–434.

[24] P. Tseng and D.P. Bertsekas, Relaxation methods for problems with strictly convex separable costs and linear constraints, *Mathematical Programming*, **38** (1987), 303–321.

[25] I. Yamada, Hybrid steepest descent method for variational inequality problem over the fixed point set of certain quasi-nonexpansive mappings, *Numerical Functional Analysis and Optimization,* **25** (2004), 619–655.