ON KORPELEVICH'S EXTRAGRADIENT ALGORITHM

NICHOLAS PISCHKE

Dedicated to Ulrich Kohlenbach on the occasion of his 60th birthday

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

E-mail: pischkenicholas@gmail.com

Abstract. We provide quantitative results on the convergence of the extragradient algorithm of Korpelevich in the form of a computable and highly uniform rate of metastability (in the sense of T. Tao) as well as, under a general metric regularity assumption in the sense of U. Kohlenbach, G. López-Acedo and A. Nicolae, even in the form of a rate of convergence.

Keywords: Extragradient algorithms; rates of convergence; metastability; proof mining
MSC2020 Classification: 47J25, 03F10, 49J40, 47H05

1. Introduction

Let $\mathbb{R}^d$ be the Euclidean space with the usual inner product $\langle \cdot, \cdot \rangle$ and norm $\|\cdot\|$. The variational inequality problem [1, 11] consists of finding a point $u^* \in C$ such that

$$\langle Tu^*, u - u^* \rangle \geq 0 \text{ for all } u \in C,$$

for a given closed and convex set $C \subseteq \mathbb{R}^d$ and a mapping $T : \mathbb{R}^d \to \mathbb{R}^d$. We write $\mathrm{VIP}(T, C)$ for the set of all such solutions $u^* \in C$.

In [19], Korpelevich proposed what is now known as the extragradient algorithm to solve the variational inequality problem which takes the form of a split iteration

$$\begin{cases} \overline{u}^k = P_C(u^k - \alpha T u^k), \\ u^{k+1} = P_C(u^k - \alpha T \overline{u}^k), \end{cases}$$

with a parameter $\alpha > 0$ and where P_C is the projection onto C. Under suitable conditions on T and α, Korpelevich showed convergence of the algorithm. Concretely, we will require T to be monotone, i.e. to satisfy

$$\langle Tx - Ty, x - y \rangle \geq 0 \text{ for all } x, y \in C,$$

as well as to be Lipschitz-continuous with constant L, i.e. to satisfy

$$\|Tx - Ty\| \leq L \|x - y\| \text{ for all } x, y \in C.$$

Then, Korpelevich's convergence result takes the form of the following theorem:

Theorem 1 ([19]). *Let $\mathrm{VIP}(T, C) \neq \emptyset$ and suppose T is monotone and Lipschitz continuous with constant L. If $0 < \alpha < 1/L$, then (u^k) converges to a point in $\mathrm{VIP}(T, C)$.*

This result given by Korpelevich contains no additional quantitative information, like explicit (or possibly effective) rates of convergence or similar. In this note we, in that vein, provide a highly uniform and fully effective rate of so-called metastability in the sense of Tao [28, 29] and even an effective rate of convergence under an additional metric regularity assumption in the sense of [18].

In terms of quantitative information, it is well known since the seminal work of Specker [27] that even for computable sequences of real numbers, in general, there exists no computable rate of convergence and as extensively discussed in [24], these examples can be adapted such that they provide counterexamples for computable rates of convergence for many central iterative methods from nonlinear analysis and optimization. Similarly, it is straightforward to adapt examples from [24] to also rule out computable rates of convergence for Korpelevich's algorithm, even in the simplest case of $d = 1$ and $C = \mathbb{R}$.

Date: June 7, 2023.

However, instead of searching for an upper bound $\Phi(k)$ on the quantifier '$\exists n \in \mathbb{N}$' in the Cauchy property

$$\forall k \in \mathbb{N} \exists n \in \mathbb{N} \forall i, j \geq n \left(d(x_i, x_j) < \frac{1}{k+1} \right)$$

over, say, a metric space (X, d), one can consider the (noneffectively) equivalent reformulation[1]

$$\forall k \in \mathbb{N} \forall g \in \mathbb{N}^{\mathbb{N}} \exists n \in \mathbb{N} \forall i, j \in [n; n+g(n)] \left(d(x_i, x_j) < \frac{1}{k+1} \right).$$

In this formulation, highly uniform and computable bounds $\Phi(k, g)$ on the quantifier '$\exists n \in \mathbb{N}$', known as rates of metastability after Tao [28, 29] as mentioned before, are actually guaranteed to exist in very general situations by general logical metatheorems from the context of the 'proof mining' program where logical machinery is applied to prima facie nonconstructive proofs from ordinary mathematics to extract quantitative information (see [12] for a book treatment and [13] for a recent survey).

These results and methods obtained in the context of the 'proof mining' program also form the basis for the results presented in this note. Note for this that in particular, on the way to her convergence result, Korpelevich established the following two key lemmas:

Lemma 2 ([19]). *Let $u^* \in \mathrm{VIP}(T, C)$. Then $\left\| u^{k+1} - u^* \right\| \leq \left\| u^k - u^* \right\|$ for all k.*

Lemma 3 ([19]). *Let $\mathrm{VIP}(T, C) \neq \emptyset$. Then $\left\| u^k - \overline{u}^k \right\| \to 0$ for $k \to \infty$.*

In that way, the sequence generated by the algorithm has the property of being Fejér monotone [6] and abstract results established in [16, 18] (similarly in the context of the proof mining program) guarantee the existence for respective rates of metastability and, under additional regularity assumptions as mentioned before, rates of convergence and these can, moreover, be constructed from respective quantitative versions of the properties given in Lemmas 2 and 3. It is the extraction of these quantitative versions of Lemma 2 and Lemma 3, themselves originating from logical considerations, together with the resulting construction of a rate of metastability and (under a metric regularity assumption) a rate of convergence as in [16, 18], that we detail here.

In that way, this note provides a further concrete instance for an exemplary application of the very abstract approach from [16, 18] (which has so far been successfully applied e.g. in the context of the proximal point algorithm [14, 15, 21, 20], subgradient methods for equilibrium problems [26] or regarding the asymptotic regularity of compositions of two mappings [17], among others). Beyond this however, while the original algorithm of Korpelevich is a very concrete case to be analyzed using this proof theoretic approach, we think that the methods used here can be employed, potentially with small modifications, to further provide quantitative results for the various extensions and modifications of Korpelevich's algorithm [3, 4, 5, 8, 9, 23, 22, 31] and other related general extragradient procedures which have received significant attention in the last years. To that end, as common in the context of applications of proof mining, while the present results were obtained using said underlying logical methodology, they are presented in this work without any further reference to logic.

2. Preliminary quantitative results

Throughout, let $M \in \mathbb{N} \setminus \{0\}$ be an upper bound on $\left\| u^k \right\|, \left\| \overline{u}^k \right\|$ for all k as generated by Korpelevich's algorithm for some T which is monotone and Lipschitz continuous with constant L and for $0 < \alpha < 1/L$. Note that the existence of such an M is guaranteed under the assumption that $\mathrm{VIP}(T, C) \neq \emptyset$ as using Lemma 2 on $u^* \in \mathrm{VIP}(T, C)$ yields

$$\left\| u^k \right\| \leq \left\| u^k - u^* \right\| + \left\| u^* \right\| \leq \left\| u^0 - u^* \right\| + \left\| u^* \right\|,$$

i.e. (u^k) is bounded and thus also $(\overline{u}^k)$ is bounded by Lemma 3. Define $X_0 = \overline{B}_M(0) \cap C$ which is compact as C is closed. Adapting from [19], we define

$$\varphi(u) := \left\| u - P_C(u - \alpha T u) \right\|$$

and in general write $\overline{u} := P_C(u - \alpha T u)$ for a given u where P_C is the projection onto C as before which, in Hilbert spaces, is a nonexpansive map.

The rest of this section presents general quantitative versions of Lemmas 2 and 3 as well as of related partial results given in Korpelevich's proof from [19]. These will be combined to a rate of metastability or respectively

[1] Here, we write $[n; n+m] := \{n + i \mid i \in \mathbb{N} \text{ and } 0 \leq i \leq m\}$.

to a rate of convergence in the next section.

To begin with, a proof-theoretic analysis of the proof given in [19] for Lemma 2 yields the following quantitative version.

Lemma 4. *If $u^* \in C$ is such that $\langle Tu^*, u - u^*\rangle \geq -\varepsilon$ for any $u \in X_0$, then*

$$\left\|u^{k+1} - u^*\right\|^2 \leq \left\|u^k - u^*\right\|^2 + 2\alpha\varepsilon$$

for any $k \in \mathbb{N}$.

Proof. From the characterization of the projection P_C in inner product spaces [2, Theorem 3.16], we have $\langle u - P_C u, v - P_C u\rangle \leq 0$ for all $u \in \mathbb{R}^n, v \in C$. Thus in particular $\|u - v\|^2 \geq \|u - P_C u\|^2 + \|v - P_C u\|^2$ for all $u \in \mathbb{R}^n, v \in C$. With $v = u^*$ and $u = u^k - \alpha T\overline{u}^k$, this yields

$$\begin{aligned}
\left\|u^{k+1} - u^*\right\|^2 &\leq \left\|u^k - \alpha T\overline{u}^k - u^*\right\|^2 - \left\|u^k - \alpha T\overline{u}^k - u^{k+1}\right\|^2 \\
&= \left\|u^k - u^*\right\|^2 - 2\alpha\langle T\overline{u}^k, u^k - u^*\rangle + \left\|\alpha T\overline{u}^k\right\|^2 \\
&\quad - \left(\left\|u^{k+1} - u^k\right\|^2 - 2\alpha\langle T\overline{u}^k, u^k - u^{k+1}\rangle + \left\|\alpha T\overline{u}^k\right\|^2\right) \\
&= \left\|u^k - u^*\right\|^2 - \left\|u^{k+1} - u^k\right\|^2 - 2\alpha\langle T\overline{u}^k, u^k - u^*\rangle \\
&\quad + 2\alpha\langle T\overline{u}^k, u^k - u^{k+1}\rangle \\
&= \left\|u^k - u^*\right\|^2 - \left\|u^{k+1} - u^k\right\|^2 + 2\alpha\langle T\overline{u}^k, u^* - u^{k+1}\rangle.
\end{aligned}$$

Using that $\langle Tu^*, u - u^*\rangle \geq -\varepsilon$ for all $u \in X_0$ and using the monotonicity of T, we get

$$\begin{aligned}
0 &\leq \langle Tu - Tu^*, u - u^*\rangle \\
&= \langle Tu, u - u^*\rangle - \langle Tu^*, u - u^*\rangle \\
&\leq \langle Tu, u - u^*\rangle + \varepsilon
\end{aligned}$$

for all $u \in X_0$. For $u = \overline{u}^k \in X_0$, this yields $\langle T\overline{u}^k, u^* - \overline{u}^k\rangle \leq \varepsilon$ and therefore

$$\begin{aligned}
\langle T\overline{u}^k, u^* - u^{k+1}\rangle &= \langle T\overline{u}^k, u^* - \overline{u}^k\rangle + \langle T\overline{u}^k, \overline{u}^k - u^{k+1}\rangle \\
&\leq \langle T\overline{u}^k, \overline{u}^k - u^{k+1}\rangle + \varepsilon.
\end{aligned}$$

Combined, we get

$$\begin{aligned}
\left\|u^{k+1} - u^*\right\|^2 &\leq \left\|u^k - u^*\right\|^2 - \left\|u^{k+1} - u^k\right\|^2 + 2\alpha\langle T\overline{u}^k, \overline{u}^k - u^{k+1}\rangle + 2\alpha\varepsilon \\
&= \left\|u^k - u^*\right\|^2 - \Big(\left\|u^k - \overline{u}^k\right\|^2 + \left\|\overline{u}^k - u^{k+1}\right\|^2 \\
&\quad + 2\langle u^k - \overline{u}^k, \overline{u}^k - u^{k+1}\rangle\Big) + 2\alpha\langle T\overline{u}^k, \overline{u}^k - u^{k+1}\rangle + 2\alpha\varepsilon \\
&= \left\|u^k - u^*\right\|^2 - \left\|u^k - \overline{u}^k\right\|^2 - \left\|\overline{u}^k - u^{k+1}\right\|^2 \\
&\quad + 2\langle u^k - \alpha T\overline{u}^k - \overline{u}^k, u^{k+1} - \overline{u}^k\rangle + 2\alpha\varepsilon.
\end{aligned}$$

Now using the characterization of the projection again, we further have

$$\begin{aligned}
\langle u^k - \alpha T\overline{u}^k - \overline{u}^k, u^{k+1} - \overline{u}^k\rangle &= \langle u^k - \alpha Tu^k - \overline{u}^k, u^{k+1} - \overline{u}^k\rangle \\
&\quad + \langle \alpha Tu^k - \alpha T\overline{u}^k, u^{k+1} - \overline{u}^k\rangle \\
&\leq 0 + \alpha\left\|Tu^k - T\overline{u}^k\right\|\left\|u^{k+1} - \overline{u}^k\right\| \\
&\leq \alpha L\left\|u^k - \overline{u}^k\right\|\left\|u^{k+1} - \overline{u}^k\right\|
\end{aligned}$$

and using

$$\alpha^2L^2\left\|u^k - \overline{u}^k\right\|^2 + \left\|\overline{u}^k - u^{k+1}\right\|^2 \geq 2\alpha L\left\|u^k - \overline{u}^k\right\|\left\|u^{k+1} - \overline{u}^k\right\|$$

we get

$$\begin{aligned}\left\|u^{k+1}-u^*\right\|^2 &\leq \left\|u^k-u^*\right\|^2 - \left\|u^k-\overline{u}^k\right\|^2 - \left\|\overline{u}^k-u^{k+1}\right\|^2 \\ &\qquad + \alpha^2L^2\left\|u^k-\overline{u}^k\right\|^2 + \left\|\overline{u}^k-u^{k+1}\right\|^2 + 2\alpha\varepsilon \\ &\leq \left\|u^k-u^*\right\|^2 + (\alpha^2L^2-1)\left\|u^k-\overline{u}^k\right\|^2 + 2\alpha\varepsilon \\ &\leq \left\|u^k-u^*\right\|^2 + 2\alpha\varepsilon\end{aligned}$$

where we have used that $\alpha L < 1$. □

Implicit in the above proof is the following general bound akin to [19].

Lemma 5. *Let $u^* \in \mathrm{VIP}(T,C)$ be fixed. Then for any $k \in \mathbb{N}$:*

$$\left\|u^{k+1}-u^*\right\|^2 \leq \left\|u^k-u^*\right\|^2 + (\alpha^2L^2-1)\left\|u^k-\overline{u}^k\right\|^2.$$

From this, we can immediately obtain a respective quantitative version for Lemma 3.

Lemma 6. *Let $u^* \in \mathrm{VIP}(T,C)$ with $B \geq \left\|u^0-u^*\right\|$. For any $\varepsilon > 0$:*

$$\exists k \leq \left\lceil \frac{B^2}{\varepsilon^2(1-\alpha^2L^2)} \right\rceil \left(\left\|u^k-\overline{u}^k\right\| \leq \varepsilon\right).$$

Proof. Lemma 2 yields that for any k:

$$\left\|u^{k+1}-u^*\right\|^2 \leq \left\|u^k-u^*\right\|^2 \leq \left\|u^0-u^*\right\|^2.$$

We show that for any $\varepsilon > 0$:

$$\exists k \leq K(\varepsilon) := \left\lceil \frac{B^2}{\varepsilon^2(1-\alpha^2L^2)} \right\rceil \left(\left\|u^k-u^*\right\| - \left\|u^{k+1}-u^*\right\|^2 \leq \varepsilon^2(1-\alpha^2L^2)\right).$$

For a contradiction, suppose that there exists an $\varepsilon > 0$ such that

$$\forall k \leq K(\varepsilon)\left(\left\|u^k-u^*\right\| - \left\|u^{k+1}-u^*\right\|^2 > \varepsilon^2(1-\alpha^2L^2)\right).$$

Then we get

$$\begin{aligned}\left\|u^0-u^*\right\|^2 &> \left\|u^1-u^*\right\|^2 + \varepsilon^2(1-\alpha^2L^2) \\ &> \cdots > \left\|u^{K(\varepsilon)+1}-u^*\right\|^2 + \left(\left\lceil \frac{B^2}{\varepsilon^2(1-\alpha^2L^2)} \right\rceil + 1\right)\varepsilon^2(1-\alpha^2L^2) \\ &\geq \left\|u^0-u^*\right\|^2\end{aligned}$$

which is a contradiction. Now, using Lemma 5, we get that there exists a $k \leq K(\varepsilon)$ such that

$$\begin{aligned}\left\|u^k-\overline{u}^k\right\| &\leq \frac{\sqrt{\left\|u^k-u^*\right\|^2 - \left\|u^{k+1}-u^*\right\|^2}}{\sqrt{1-\alpha^2L^2}} \\ &\leq \frac{\sqrt{\varepsilon^2(1-\alpha^2L^2)}}{\sqrt{1-\alpha^2L^2}} \\ &\leq \varepsilon.\end{aligned}$$

□

Lemma 7. *The function φ is Lipschitz-continuous with*

$$|\varphi(u)-\varphi(u')| \leq (2+\alpha L)\left\|u-u'\right\|.$$

Proof. Let u, u' be given. Then

$$\begin{aligned}|\varphi(u)-\varphi(u')| &= \left|\,\left\|u-\overline{u}\right\| - \left\|u'-\overline{u}'\right\|\,\right| \\ &\leq \left\|u-u'\right\| + \left\|\overline{u}-\overline{u}'\right\| \\ &\leq \left\|u-u'\right\| + \left\|P_C(u-\alpha Tu) - P_C(u'-\alpha Tu')\right\| \\ &\leq 2\left\|u-u'\right\| + \alpha\left\|Tu-Tu'\right\| \\ &\leq (2+\alpha L)\left\|u-u'\right\|.\end{aligned}$$

□

The following provides a quantitative version of the implicit result given in [19] that $\varphi(u^*) = 0$ implies $u^* \in \mathrm{VIP}(T, C)$.

Lemma 8. *Let $u^* \in X_0$ be given with $N \geq \|u^* - \alpha T u^* - P_C(u^* - \alpha T u^*)\|$. Given $\varepsilon > 0$:*

$$\varphi(u^*) \leq \frac{\alpha\varepsilon}{N + 2M} \text{ implies } \langle Tu^*, u - u^*\rangle \geq -\varepsilon \text{ for any } u \in X_0.$$

Proof. Suppose $\varphi(u^*) \leq \alpha\varepsilon/(N + 2M)$. For $u \in X_0$, we get

$$\langle u^* - \alpha Tu^* - P_C(u^* - \alpha Tu^*), u - P_C(u^* - \alpha Tu^*)\rangle \leq 0$$

by the characterization of P_C. Using this, we compute

$$\begin{aligned}
-\frac{1}{\alpha}&\langle -\alpha Tu^*, u - u^*\rangle \\
&= -\frac{1}{\alpha}\langle u^* - \alpha Tu^* - P_C(u^* - \alpha Tu^*), u - u^*\rangle \\
&\quad + \frac{1}{\alpha}\langle u^* - P_C(u^* - \alpha Tu^*), u - u^*\rangle \\
&= -\frac{1}{\alpha}\langle u^* - \alpha Tu^* - P_C(u^* - \alpha Tu^*), u - P_C(u^* - \alpha Tu^*)\rangle \\
&\quad - \frac{1}{\alpha}\langle u^* - \alpha Tu^* - P_C(u^* - \alpha Tu^*), P_C(u^* - \alpha Tu^*) - u^*\rangle \\
&\quad + \frac{1}{\alpha}\langle u^* - P_C(u^* - \alpha Tu^*), u - u^*\rangle \\
&\geq 0 - \frac{1}{\alpha}N\frac{\alpha\varepsilon}{N + 2M} - \frac{1}{\alpha}\frac{\alpha\varepsilon}{N + 2M}(\|u\| + \|u^*\|) \\
&\geq -\frac{\alpha\varepsilon}{N + 2M}\frac{1}{\alpha}(N + 2M) \\
&\geq -\varepsilon
\end{aligned}$$

as $u, u^* \in X_0$, so $\|u\|, \|u^*\| \leq M$. □

3. Quantitative versions of the convergence of Korpelevich's algorithm

In this section, we apply the abstract results of [16] as well as of [18] to the particular instance of Korpelevich's algorithm. Concretely, we will appropriately translate the previous preliminary quantitative versions of the main ingredients of the convergence proof presented in [19] to fit the respective formal setups given in [16, 18] and then derive from this a rate of metastability or a rate of convergence, respectively.

3.1. A rate of metastability.

The results given in [16] rely on uniform reformulations of the respective properties like Fejér monotonicity in terms of approximations instead of full solutions (see [16] for details). In that way, we define

$$F = X_0 \cap \mathrm{zer}\varphi$$

as the solution set as well as

$$AF_n = \left\{ u \in X_0 \mid \varphi(u) \leq \frac{1}{n+1} \right\}$$

as the set of approximate solutions in terms of a specific error.

Then, we can translate the results of the previous section into the relevant bounds and moduli required in the general abstract setup presented in [16] in the context of these uniform reformulations.

Lemma 9. *(u^k) is uniformly $((\cdot)^2, (\cdot)^2)$-Fejér monotone w.r.t. F and AF_n in the sense of* [16] *with modulus*

$$\chi(k, m, r) = \lceil 2(r + 2)m(1 + \alpha(LM + D) + 2M)\rceil \dot{-} 1$$

where $D \geq \|T0\|$, that is for any $k, m, r \in \mathbb{N}$, any $u^* \in AF_{\chi(k,m,r)}$ and any $l \leq m$:

$$\left\|u^{k+l} - u^*\right\|^2 < \left\|u^k - u^*\right\|^2 + \frac{1}{r+1}.$$

Proof. Let $k, m, r \in \mathbb{N}$ be given and $u^* \in AF_\chi(k,m,r)$, i.e. $u^* \in X_0$ and

$$\varphi(u^*) \leq \frac{1}{\chi(k,m,r)+1} \leq \frac{\alpha}{2(r+2)m\alpha(1+\alpha(LM+D)+2M)}$$

By Lemma 8, as

$$\begin{aligned}\|u^* - \alpha Tu^* - P_C(u^* - \alpha Tu^*)\| &\leq \|u^* - \overline{u}^*\| + \alpha \|Tu^*\| \\ &\leq 1 + \alpha(\|Tu^* - T0\| + \|T0\|) \\ &\leq 1 + \alpha(LM + D),\end{aligned}$$

we get $\langle Tu^*, u - u^* \rangle \geq -1/2(r+2)m\alpha$ for any $u \in X_0$. By iterating Lemma 4, we get

$$\left\|u^{k+l} - u^*\right\|^2 \leq \left\|u^k - u^*\right\|^2 + 2m\alpha\frac{1}{2(r+2)m\alpha} < \left\|u^k - u^*\right\|^2 + \frac{1}{r+1}.$$

□

Lemma 10. *Let $u^* \in \mathrm{VIP}(T,C)$ with $B \geq \left\|u^0 - u^*\right\|$. Then (u^k) has approximate F-points in the sense of* [16] *with an approximate F-point bound*

$$\Phi(n) = \left\lceil \frac{(n+1)^2 B^2}{1-\alpha^2 L^2} \right\rceil,$$

that is

$$\forall n \in \mathbb{N} \exists k \leq \Phi(n) \left(u^k \in AF_n\right).$$

Proof. By Lemma 6 with $\varepsilon = 1/(n+1)$:

$$\exists k \leq \left\lceil \frac{(n+1)^2 B^2}{1-\alpha^2 L^2} \right\rceil \left(\left\|u^k - \overline{u}^k\right\| \leq \frac{1}{n+1} \right).$$

□

Lemma 11. *F is uniformly closed w.r.t. AF_n in the sense of* [16] *with moduli*

$$\begin{cases} \delta_F(n) = 2n+1 \\ \omega_F(n) = \lceil (2+\alpha L)(2n+2) \rceil \dot{-} 1, \end{cases}$$

that is for any $n \in \mathbb{N}$, any $q \in AF_{\delta_F(n)}$ and any p with $\|p - q\| \leq 1/(\omega_F(n)+1)$, we have $p \in AF_n$.

Proof. Let $q \in AF_{\delta_F(n)}$, i.e. $\varphi(q) \leq 1/(\delta_F(n)+1)$ and p with $\|p-q\| \leq 1/(\omega_F(n)+1)$. Using Lemma 7, we get

$$\begin{aligned}\varphi(p) &\leq \varphi(q) + |\varphi(p) - \varphi(q)| \\ &\leq \varphi(q) + (2+\alpha L)\|p-q\| \\ &\leq \frac{1}{\delta_F(n)+1} + (2+\alpha L)\frac{1}{\omega_F(n)+1} \\ &\leq \frac{1}{2n+2} + (2+\alpha L)\frac{1}{(2+\alpha L)(2n+2)} \\ &\leq \frac{1}{n+1}.\end{aligned}$$

□

Theorem 12. *Let $u^* \in \mathrm{VIP}(T,C)$ with $B \geq \left\|u^0 - u^*\right\|$, $D \geq \|T0\|$ and define $\gamma(n) := \left\lceil 2(n+1)\sqrt{d}M \right\rceil^d$. Let T be monotone and Lipschitz continuous with constant L and assume that $0 < \alpha < 1/L$. Then (u^k) is Cauchy and, moreover, for any $n \in \mathbb{N}$ and any $g : \mathbb{N} \to \mathbb{N}$*

$$\exists N \leq \Psi(n,g) \forall i,j \in [N, N+g(N)] \left(\left\|u^i - u^j\right\| \leq \frac{1}{n+1} \wedge u^i \in AF_n \right),$$

where $\Psi(n,g) = \Psi_0(P, n_0, g)$ with $P = \gamma\left(\left\lceil \sqrt{8(n+1)^2+1} - 1 \right\rceil\right)$ and

$$n_0 = \max\left\{ n, \left\lceil \frac{\lceil (2+\alpha L)(2n+2) \rceil - 2}{2} \right\rceil \right\}$$

and where $\Psi_0(m,n,g)$ *is defined by recursion via*

$$\begin{cases}\Psi_0(0,n,g) := 0,\\ \Psi_0(m+1,n,g) := \Phi(\chi^M_{n,g}(\Psi_0(m,n,g), 8(n+1)^2)),\end{cases}$$

with

$$\chi_n(k,m,r) = \max\{2n+1, \lceil (r+2)m2(1+\alpha(LM+D)+2M)\rceil \dot{-} 1\},$$
$$\chi_{n,g}(k,r) = \chi_n(k,g(k),r),\ \chi^M_{n,g}(k,r) = \max\{\chi_{n,g}(i,r) \mid i \leq k\},$$

as well as

$$\Phi(n) = \left\lceil \frac{(n+1)^2B^2}{1-\alpha^2L^2} \right\rceil.$$

Proof. The theorem is an instance of Theorem 5.3 given in [16]. Lemma 10 established that Φ is an approximate F-point bound, Lemma 9 established that χ is a modulus of uniform (G,H)-Fejér monotonicity for $G = H = (\cdot)^2$, Lemma 11 established that δ_F and ω_F are moduli of uniform closedness for F and AF_n and, lastly,

$$\gamma(n) := \left\lceil 2(n+1)\sqrt{d}M \right\rceil^d$$

is a II-modulus of total boundedness for X_0 in the sense of [7, 16] which follows from Example 2.8 in [16]. The result follows (after some obvious simplifications in the bounds) by noting that $\alpha_G(n) = \lceil \sqrt{n+1} - 1\rceil$ is a G-modulus for $(\cdot)^2$ and $\beta_H(n) = (n+1)^2 - 1$ is an H-modulus for $(\cdot)^2$ in the sense of [16]. □

The above Theorem 12 is a finitization in the sense of Tao of Korpelevich's convergence result given in Theorem 1 since the above theorem only talks about finite segments of (u^k) but trivially implies back Theorem 1 (which can be shown in a similar way as outlined in Remark 5.5 in [16]).

3.2. **A rate of convergence.** As mentioned before, while in general even for Fejér monotone sequences computable rates of convergence are ruled out (see again the discussions in [24]), we can provide such rates under additional assumptions. A large class of such assumptions in the context of Fejér monotone sequences, generalizing various concepts known from nonlinear analysis and optimization such as error bounds and metric subregularity, among others, was introduced and studied in [18] under the name of moduli of regularity. In our context, we consider the following instantiation of this notion from [18]:

Definition 13. Assume $X_0 \cap \mathrm{zer}\varphi \neq \emptyset$ and let $u^* \in X_0 \cap \mathrm{zer}\varphi \neq \emptyset$ and $r > 0$. We say that $\rho : (0,\infty) \to (0,\infty)$ is a *modulus of regularity* for φ w.r.t. $\mathrm{zer}\varphi$ and $\overline{B}_r(u^*)$ if for all $\varepsilon > 0$ and all $x \in \overline{B}_r(u^*)$, we have that

$$\varphi(x) < \rho(\varepsilon) \text{ implies } \mathrm{dist}(x, \mathrm{zer}\varphi) < \varepsilon.$$

Abstract results given in [18] (see Theorem 4.1) then allow for the construction of a rate of convergence for our sequence at hand under the assumption that φ possesses a modulus of regularity in the above sense:

Theorem 14. *Assume* $X_0 \cap \mathrm{zer}\varphi \neq \emptyset$ *and let* $u^* \in X_0 \cap \mathrm{zer}\varphi$ *and* $B \geq \|u^0 - u^*\|$, *let* T *be monotone and Lipschitz-continuous with constant* L *and assume that* $0 < \alpha < 1/L$. *If* ρ *is a modulus of regularity for* φ *w.r.t.* $\mathrm{zer}\varphi$ *and* $\overline{B}_B(u^*)$, *then* (u^k) *is Cauchy and*

$$\forall \varepsilon > 0 \forall k, j \geq \Phi(\rho(\varepsilon/2)) \left(\|u^k - u^j\| < \varepsilon\right)$$

where

$$\Phi(\varepsilon) = \left\lceil \frac{B^2}{(\varepsilon/2)^2(1-\alpha^2L^2)} \right\rceil.$$

Proof. The theorem is an instance of Theorem 4.1 given in [18] which follows immediately by noting Lemma 6, from which we obtain

$$\exists k \leq \left\lceil \frac{B^2}{(\varepsilon/2)^2(1-\alpha^2L^2)} \right\rceil \left(\|u^k - \overline{u}^k\| \leq \frac{\varepsilon}{2} < \varepsilon\right)$$

for any $\varepsilon > 0$. □

Based on the general discussions given in [18], it is to be expected that this notion of a modulus of regularity for φ encompasses many of the usual regularity-type assumptions made in the context of extragradient algorithms to derive rates of convergence with specific (low) complexities. In particular, it is rather immediate to see that this notion especially generalizes the influential regularity notion due to Tseng [30] which continues to be prevalent in low complexity estimates on the rates associated with various types of extragradient methods (see e.g. [10]).

Definition 15. The problem $\mathrm{VIP}(T, C)$ satisfies Tseng's regularity assumption if it has a solution and there exist $\delta, \eta > 0$ such that for any $u \in C$:

$$\varphi(u) \leq \delta \text{ implies } \mathrm{dist}(u, \mathrm{zer}\varphi) \leq \eta\varphi(u).$$

As shown by Tseng in [30], this regularity assumption in particular holds true if

(1) T is strongly monotone and Lipschitz continuous,
(2) C is a polyhedral set and T is affine,
(3) C is a polyhedral set and T is of the form $T(x) = E^\top G(Ex) + q$ where E is a $m \times d$ matrix with no zero column, q is a vector in $\mathbb{R}^d$ and $G : \mathbb{R}^m \to \mathbb{R}^m$ is a strongly monotone and Lipschitz continuous function.

E.g., in the first case, if T is strongly monotone with constant $\alpha > 0$, i.e.

$$\langle x - y, Tx - Ty \rangle \geq \alpha \|x - y\|^2$$

as well as Lipschitz continuous with constant L, then the constants δ and η can be chosen specifically with an *arbitrary* δ and $\eta = (L + 1)/\alpha$ since, as shown in Theorem 3.1 in [25], the inequality

$$\mathrm{dist}(u, \mathrm{zer}\varphi) \leq \frac{(L+1)}{\alpha}\varphi(u)$$

holds already unconditionally for every $u \in C$.

It is clear that if δ, η are values witnessing that Tseng's regularity assumption holds, then

$$\rho(\varepsilon) = \min\left\{\delta, \frac{\varepsilon}{\eta}\right\}$$

is a modulus of regularity for φ w.r.t. $\mathrm{zer}\varphi$ and $\overline{B}_r(u^*)$ for any r and any $u^* \in X_0 \cap \mathrm{zer}\varphi$. In that case, the rate obtained in Theorem 14 correspondingly simplifies to

$$\Phi(\rho(\varepsilon/2)) = \left\lceil \frac{B^2}{\left(\min\left\{\delta, \frac{\varepsilon/2}{\eta}\right\}/2\right)^2 (1 - \alpha^2 L^2)} \right\rceil$$

which is quadratic in the error. So, in regard to the previously discussed special case, if T is e.g. strongly monotone with constant α and Lipschitz continuous with constant L, then we can further simplify the rate to

$$\left\lceil \frac{16B^2(L+1)^2}{\alpha^2\varepsilon^2(1 - \alpha^2 L^2)} \right\rceil.$$

Acknowledgments: The author was supported by the 'Deutsche Forschungsgemeinschaft' Project DFG KO 1737/6-2.

References

[1] G. Allen. Variational inequalities, complementarity problems, and duality theorems. *Journal of Mathematical Analysis and Applications*, 58(1):1–10, 1977.
[2] H.H. Bauschke and P.L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. CMS Books in Mathematics. Springer Cham, 2017.
[3] Y. Censor, A. Gibali, and S. Reich. Strong convergence of subgradient extragradient methods for the variational inequality problem in Hilbert space. *Optimization Methods and Software*, 26(4–5):827–845, 2011.
[4] Y. Censor, A. Gibali, and S. Reich. The Subgradient Extragradient Method for Solving Variational Inequalities in Hilbert Space. *Journal of Optimization Theory and Applications*, 148:318–335, 2011.
[5] Y. Censor, A. Gibali, and S. Reich. Extensions of Korpelevich's extragradient method for the variational inequality problem in Euclidean space. *Optimization*, 61(9):1119–1132, 2012.
[6] P.L. Combettes. Fejér monotonicity in convex optimization. In C.A. Floudas and P.M Pardalos, editors, *Encyclopedia of Optimization*, pages 1016–1024. Springer New York, 2009.
[7] P. Gerhardy. Proof mining in topological dynamics. *Notre Dame Journal of Formal Logics*, 49:431–446, 2008.
[8] A.N. Iusem and L.R.L. Pérez. An extragradient-type algorithm for non-smooth variational inequalities. *Optimization*, 48(3):309–332, 2000.
[9] P.D. Khanh. A New Extragradient Method for Strongly Pseudomonotone Variational Inequalities. *Numerical Functional Analysis and Optimization*, 37(9):1131–1143, 2016.
[10] P.D. Khanh. On the convergence rate of a modified extragradient method for pseudomonotone variational inequalities. *Vietnam Journal of Mathematics*, 45:397–408, 2017.
[11] D. Kinderlehrer and G. Stampacchia. *An Introduction to Variational Inequalities and Their Applications*. Academic Press, 1980.

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

[13] U. Kohlenbach. Proof-theoretic Methods in Nonlinear Analysis. In B. Sirakov, P. Ney de Souza, and M. Viana, editors, *Proc. ICM 2018*, volume 2, pages 61–82. World Scientific, 2019.

[14] U. Kohlenbach. Quantitative results on the Proximal Point Algorithm in uniformly convex Banach spaces. *Journal of Convex Analysis*, 28(1):11–18, 2021.

[15] U. Kohlenbach. On the Proximal Point Algorithm and its Halpern-type variant for generalized monotone operators in Hilbert space. *Optimization Letters*, 16:611–621, 2022.

[16] U. Kohlenbach, L. Leuştean, and A. Nicolae. Quantitative Results on Fejér Monotone Sequences. *Communications in Contemporary Mathematics*, 20(2), 2018.

[17] U. Kohlenbach, G. López-Acedo, and A. Nicolae. Quantitative Asymptotic Regularity Results for the Composition of Two Mappings. *Optimization*, 66:1291–1299, 2017.

[18] U. Kohlenbach, G. López-Acedo, and A. Nicolae. Moduli of regularity and rates of convergence for Fejér monotone sequences. *Israel Journal of Mathematics*, 232:261–297, 2019.

[19] G.M. Korpelevich. The extragradient method for finding saddle points and other problems. *Ekonomika i matematicheskie metody*, 12(4):747–756, 1976.

[20] L. Leuştean and A. Sipoş. An application of proof mining to the proximal point algorithm in CAT(0) spaces. In A. Bellow, C. Calude, and T. Zamfirescu, editors, *Mathematics Almost Everywhere. In Memory of Solomon Marcus*, pages 153–168. World Scientific, 2018.

[21] L. Leuştean and A. Sipoş. Effective strong convergence of the proximal point algorithm in CAT(0) spaces. *Journal of Nonlinear and Variational Analysis*, 2:219–228, 2018.

[22] Y.V. Malitsky and V.V. Semenov. An Extragradient Algorithm for Monotone Variational Inequalities. *Cybernetics and Systems Analysis*, 50:271–277, 2014.

[23] R.D.C. Monteiro and B.F. Svaiter. On the Complexity of the Hybrid Proximal Extragradient Method for the Iterates and the Ergodic Mean. *SIAM Journal on Optimization*, 20(6):2755–2787, 2010.

[24] Eike Neumann. Computational Problems in Metric Fixed Point Theory and their Weihrauch Degrees. *Logical Methods in Computer Science*, 11(4), 2015.

[25] J.-S. Pang. A posteriori error bounds for the linearly-constrained variational inequality problem. *Mathematics of Operations Research*, 12(3):474–484, 1987.

[26] N. Pischke and U. Kohlenbach. Quantitative Analysis of a Subgradient-Type Method for Equilibrium Problems. *Numerical Algorithms*, 90(1):197–219, 2022.

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

[28] T. Tao. Norm Convergence of Multiple Ergodic Averages for Commuting Transformations. *Ergodic Theory and Dynamical Systems*, 28:657–688, 2008.

[29] T. Tao. *Structure and Randomness: Pages from Year One of a Mathematical Blog*, chapter Soft analysis, hard analysis, and the finite convergence principle. American Mathematical Society, Providence, RI, 2008.

[30] P. Tseng. On linear convergence of iterative methods for the variational inequality problem. *Journal of Computational and Applied Mathematics*, 60:237–252, 1995.

[31] L.-C. Zeng and J.-C. Yao. Strong convergence theorem by an extragradient method for fixed point problems and variational inequality problems. *Taiwanese Journal of Mathematics*, 10(5):1293–1303, 2006.