

An Inexact Newton Regularization in Banach Spaces based on the Nonstationary Iterated Tikhonov Method

F. Margotti

A. Rieder

Preprint 14/03

INSTITUT FÜR WISSENSCHAFTLICHES RECHNEN
UND MATHEMATISCHE MODELLBILDUNG



Anschriften der Verfasser:

M.Sc. Fabio Margotti
Institut für Angewandte und Numerische Mathematik
Karlsruher Institut für Technologie (KIT)
D-76128 Karlsruhe

Prof. Dr. Andreas Rieder
Institut für Angewandte und Numerische Mathematik
Karlsruher Institut für Technologie (KIT)
D-76128 Karlsruhe

AN INEXACT NEWTON REGULARIZATION IN BANACH SPACES BASED ON THE NONSTATIONARY ITERATED TIKHONOV METHOD

FÁBIO MARGOTTI AND ANDREAS RIEDER

ABSTRACT. A version of the nonstationary iterated Tikhonov method was recently introduced to regularize *linear* inverse problems in Banach spaces [*Inverse Problems* 28, 104011, 2012]. In the present work we employ this method as inner iteration of the inexact Newton regularization method REGINN [*Inverse Problems* 15, 309-327, 1999] which stably solves *nonlinear* ill-posed problems. Further, we propose and analyze a Kaczmarz version of the new scheme which allows fast solution of problems which can be split into smaller subproblems. As special cases we prove strong convergence of Kaczmarz variants of the Levenberg-Marquardt and the iterated Tikhonov methods in Banach spaces.

1. INTRODUCTION

We consider nonlinear ill-posed problems

$$(1) \quad F(x) = y$$

in the abstract framework of Banach spaces, that is, $F: D(F) \subset X \rightarrow Y$ operates between *Banach spaces* X and Y where $D(F)$ denotes the domain of definition of F . Recently, this setting has been attracting and still attracts a lot of research since several real-life applications are naturally modeled with the help of Banach spaces, see, e.g., the first chapter in [15].

The starting point for our investigation is the Newton-type algorithm REGINN (REGularization based on INexact Newton iteration) [13] which improves the current iterate x_n via

$$x_{n+1} = x_n + s_n$$

by a correction step s_n obtained from approximately solving a local linearization of (1):

$$(2) \quad A_n s = b_n$$

where $A_n := F'(x_n)$ is the Fréchet derivative of F in x_n and $b_n := y - F(x_n)$ is the corresponding nonlinear residual. REGINN typically applies an iterative solver to (2), called *inner iteration*, to find a step satisfying

$$(3) \quad \|A_n s_n - b_n\| < \mu \|b_n\|$$

with a pre-defined constant $0 < \mu < 1$.

Now, assume that (1) splits into $d \in \mathbb{N}$ 'smaller' subproblems, that is, Y factorizes into Banach spaces Y_0, \dots, Y_{d-1} : $Y = Y_0 \times Y_1 \times \dots \times Y_{d-1}$. Accordingly, $F =$

Date: April 30, 2014.

2010 Mathematics Subject Classification. 65J20; 65J15.

Key words and phrases. inexact Newton regularization in Banach spaces, nonstationary iterated Tikhonov method, Levenberg-Marquardt scheme, Kaczmarz iteration.

The first author acknowledges support by the German Academic Exchange Service (DAAD).

$(F_0, F_1, \dots, F_{d-1})^\top$, $F_j: D(F) \subset X \rightarrow Y_j$, and $y = (y_0, y_1, \dots, y_{d-1})^\top$. Our task can be recast as: find $x \in D(F)$ such that

$$(4) \quad F_j(x) = y_j, \quad j = 0, \dots, d-1.$$

The Kaczmarz variant of REGINN determines s_n from (3) where, however, $A_n := F'_{n \bmod d}(x_n)$ and $b_n := y_{n \bmod d} - F_{n \bmod d}(x_n)$. Thus, the subsystems are processed cyclically breaking the large-scale system (1) into handy pieces. This kind of cycling strategy was initiated by Kowar and Scherzer [9] and further investigated by several researchers, see, e.g., [2, 4, 8, 11]. We emphasize that systems like (4) arise quite naturally in applications where the data is measured by d individual experiments or observations. For instance, in electrical impedance tomography one wants to find the conductivity of an object by applying, say, d current patterns at the boundary and measuring the resulting voltages at the boundary as well.

In this work we consider the iterated Tikhonov regularization as suggested by Jin and Stals [7] as inner iteration of REGINN. The resulting scheme is called K-REGINN-IT which is short for Kaczmarz version of the REGINN-Iterated-Tikhonov method. As a byproduct we thus generalize the Levenberg-Marquardt regularization (Hanke [5]) to Banach spaces.

On the following pages we present a complete convergence analysis of K-REGINN-IT under the usual assumptions. As our setting is rather abstract and as our arguments are sometimes very technical we try to guide the reader gently through the exposition. Therefore, we collect needed properties and concepts of Banach spaces in Section 2. This material is taken from [3], [14], [16], and [15]. In Section 3 we define K-REGINN-IT, prove its well-definedness and termination. Next we validate strong convergence in the noise-free situation (Section 4) and finally show the regularization property in Section 5.

To keep this exposition lean we restrained from presenting numerical examples. In a forthcoming paper we plan to compare numerically different types of K-REGINN methods. The impatient reader is referred to [12] where we solve the inverse problem of electrical impedance tomography by K-REGINN with an inner iteration of Landweber type.

2. BASIC FACTS ABOUT THE GEOMETRY OF BANACH SPACES

If the context is clear, we always use a generic constant $C > 0$ even if it takes different values at different instances. Sometimes we write $a(x) \lesssim b(x)$ if and only if there exists a positive constant C independent of x such that $a(x) \leq C b(x)$ for all x .

In the following we formulate the assumptions on the Banach space X which we will need later to define and to analyze our method properly.

To cover the lack of an inner product in a general Banach space, we introduce the *duality mapping*. This is the set-valued function $J_p: X \rightarrow 2^{X^*}$ defined by¹

$$J_p(x) := \{x^* \in X^* : \langle x^*, x \rangle = \|x^*\| \|x\| \text{ and } \|x^*\| = \|x\|^{p-1}\}, \quad 1 < p < \infty,$$

where the symbol $\langle \cdot, \cdot \rangle$ denotes the inner product in Hilbert spaces as well as the duality pairing in a general Banach space. For all $x \in X$, $J_p(x) \neq \emptyset$ (see [3]) and the relation

$$(5) \quad J_r(x) = \|x\|^{r-t} J_t(x)$$

holds for each $r, t > 1$. The duality mapping J_2 is called the *normalized* duality mapping and as a consequence of the Riesz representation theorem, it is the identity operator in

¹The duality mapping is often defined in a more general way associated with a so called *gauge function*. We prefer to use here this particular definition, which is actually the duality mapping associated with the gauge function $t \mapsto t^{p-1}$.

any Hilbert space. A *selection* $j_p: X \rightarrow X^*$ of the duality mapping J_p is a single-valued not necessarily continuous function satisfying $j_p(x) \in J_p(x)$ for all $x \in X$. We will often use the estimates $\langle J_p(x), y \rangle \leq \|x\|^{p-1} \|y\|$ and $\langle J_p(x), x \rangle = \|x\|^p$ which are immediate consequences from the definition of J_p . Observe that the inner product also shares these properties for $p = 2$.

We suppose now that the Banach space X has the following geometrical properties where we use the notation $a \vee b := \max\{a, b\}$:

Assumption 1. (a) X is reflexive;

(b) For each $1 < p < \infty$, the duality mapping $J_p: X \rightarrow X^*$ is single-valued, continuous and invertible with a continuous inverse satisfying

$$J_p^{-1} = J_{p^*}^*: X^* \rightarrow X^{**} \cong X,$$

where p and p^* are conjugate numbers, i.e., $\frac{1}{p} + \frac{1}{p^*} = 1$.

(c) There exists a number $1 < s < \infty$ such that, for all $1 < p < \infty$, the conjugate numbers of s and p satisfy

$$\|J_{p^*}^*(x^*) - J_{p^*}^*(y^*)\| \leq C_{p^*, s^*} (\|x^*\| \vee \|y^*\|)^{p^* - s^*} \|x^* - y^*\|^{s^* - 1},$$

for all $x^*, y^* \in X^*$, where $C_{p^*, s^*} > 0$ is a constant which depends only on p and s .

As a substitute for the polarization identity

$$\frac{1}{2} \|x - y\|^2 = \frac{1}{2} \|x\|^2 - \langle x, y \rangle + \frac{1}{2} \|y\|^2,$$

which holds in Hilbert spaces, we introduce the *Bregman distance* $\Delta_p: X \times X \rightarrow \mathbb{R}$,

$$\Delta_p(x, y) := \frac{1}{p} \|x\|^p - \langle J_p(y), x \rangle + \frac{1}{p^*} \|J_p(y)\|^{p^*}.$$

In any real Hilbert space, $\Delta_2(x, y) = \frac{1}{2} \|x - y\|^2$. A straightforward calculation shows the equality²

$$\Delta_p(x, y) = \frac{1}{p} \|x\|^p - \frac{1}{p} \|y\|^p - \langle J_p(y), x - y \rangle$$

and the *three-points identity*

$$(6) \quad \Delta_p(x, y) = \Delta_p(z, y) - \Delta_p(z, x) + \langle J_p(x) - J_p(y), x - z \rangle,$$

for all $x, y, z \in X$. Further,

$$\Delta_p(x, y) \geq \frac{1}{p} \|x\|^p + \frac{1}{p^*} \|y\|^p - \|y\|^{p-1} \|x\|.$$

Using now Young's inequality³, we find that $\Delta_p(\cdot, \cdot) \geq 0$. Moreover, if $(x_n)_{n \in \mathbb{N}} \subset X$ is a sequence and $x \in X$ is a fixed vector, then $\Delta_p(x, x_n) \leq \rho$ implies

$$\|x_n\|^{p-1} \left(\frac{1}{p^*} \|x_n\| - \|x\| \right) \leq \rho.$$

Considering now the cases $\frac{1}{p^*} \|x_n\| - \|x\| \leq \frac{1}{2p^*} \|x_n\|$ and $\frac{1}{p^*} \|x_n\| - \|x\| > \frac{1}{2p^*} \|x_n\|$, we conclude the implication

$$(7) \quad \Delta_p(x, x_n) \leq \rho \implies \|x_n\| \leq 2p^* (\|x\| \vee \rho^{1/p}).$$

²This equivalent form is probably the most common definition of Bregman distance in Banach spaces with single-valued duality mappings.

³ $\forall a, b \geq 0 : ab \leq a^p/p + b^{p^*}/p^*$

Therefore, $(x_n)_{n \in \mathbb{N}}$ is uniformly bounded. A similar result can be proven if $\Delta_p(x_n, x) \leq \rho$. The continuity of the duality mapping (Assumption 1(b)) is handed down to both arguments of the Bregman distance Δ_p . Of course $x = y$ implies $\Delta_p(x, y) = \Delta_p(y, x) = 0$. The reciprocal and other important results are true under further properties of X :

Assumption 2. (a) *The functional $\|\cdot\|^p$ is strictly convex for any $1 < p < \infty$;*
 (b) *Any sequence $(x_n)_{n \in \mathbb{N}} \subset X$ satisfying $x_n \rightharpoonup x$ and $\|x_n\| \rightarrow \|x\|$ as $n \rightarrow \infty$, converges strongly: $\|x_n - x\| \rightarrow 0$ as $n \rightarrow \infty$;*
 (c) *If $1 < p \leq s < \infty$ then*

$$\Delta_p(x, y) \geq C \|x - y\|^s$$

for all $x, y \in B_R(0, \|\cdot\|) := \{z \in X : \|z\| \leq R\}$ where $C > 0$ depends only on p, s and R .

Assumption 2(a) implies the strict convexity of Δ_p in its first argument, which in turn ensures that $x = y$ whenever $\Delta_p(x, y) = 0$. In fact, if $x \neq y$ then for $\lambda \in (0, 1)$, we derive the contradiction

$$0 \leq \Delta_p(\lambda x + (1 - \lambda)y, y) < \lambda \Delta_p(x, y) + (1 - \lambda) \Delta_p(y, y) = 0.$$

For our convergence analysis we rely on both Assumptions 1 and 2. We emphasize that the properties of X listed therein are not entirely independent of each other: some of them can be proven assuming the others. Note that both assumptions hold true in case X is s -convex and uniformly smooth (see, e.g., [3], [15] and [14] for details). Important examples are the spaces $L^p(\Omega)$, $l^p(\mathbb{N})$, and $W^{n,p}(\Omega)$ with $1 < p < \infty$. They are uniformly smooth⁴ and $\max\{p, 2\}$ -convex. For the Lebesgue spaces the duality mapping $J_p: L^p(\Omega) \rightarrow L^{p^*}(\Omega)$ is given by

$$(8) \quad J_p(f) = |f|^{p-1} \text{sign}(f).$$

3. THE K-REGINN-IT METHOD

We will need a bunch of standard assumptions about the structure of the nonlinearity F .

Assumption 3. (a) *Equation (1) has a solution $x^+ \in X$ and for a given and fixed number $1 < p < \infty$, there exists a $\rho > 0$ such that*

$$B_\rho(x^+, \Delta_p) := \{v \in X : \Delta_p(x^+, v) < \rho\} \subset D(F).$$

(b) *Suppose that all the functions F_j , $j = 0, \dots, d - 1$, are continuously Fréchet differentiable in $B_\rho(x^+, \Delta_p)$ and that their Fréchet derivatives $F'_j: B_\rho(x^+, \Delta_p) \rightarrow \mathcal{L}(X, Y_j)$ are uniformly bounded by a constant $M > 0$:*

$$\|F'_j(v)\| \leq M \text{ for all } v \in B_\rho(x^+, \Delta_p) \text{ and } j = 0, \dots, d - 1.$$

(c) *(Tangential Cone Condition (TCC)): Suppose that*

$$\|F_j(v) - F_j(w) - F'_j(w)(v - w)\| \leq \eta \|F_j(v) - F_j(w)\|,$$

for all $v, w \in B_\rho(x^+, \Delta_p)$ and $j = 0, \dots, d - 1$, where $0 \leq \eta < 1$ is a constant.

⁴They are actually $\min\{p, 2\}$ -smooth, which is a stronger property.

We suppose to access only noisy versions $y_j^{\delta_j}$ of the exact but unknown data $y_j = F_j(x^+)$ satisfying

$$(9) \quad \|y_j - y_j^{\delta_j}\| \leq \delta_j.$$

The positive *noise levels* δ_j , $j = 0, \dots, d-1$, are assumed to be known. Further, define the maximal noise level

$$(10) \quad \delta := \max \{\delta_j : j = 0, \dots, d-1\} > 0.$$

As the spaces Y_j are arbitrary, the duality mapping J_r does not need to be single-valued (for any $r > 1$). Then, $j_r : Y_j \rightarrow Y_j^*$ represents a selection of J_r .

Now, we define K-REGINN-IT recursively: Let $x_n \in D(F)$ be given. The inner iteration to compute the Newton step s_n starts with setting $z_{n,0} := x_n$ and produces $z_{n,k+1}$ recursively as minimizer in X of the strict convex functional

$$(11) \quad T_{n,k}^\delta(z) := \frac{1}{r} \|b_n^\delta - A_n(z - x_n)\|^r + \alpha_n \Delta_p(z, z_{n,k})$$

with $\alpha_n > 0$. Here,

$$A_n := F'_{[n]}(x_n) \quad \text{and} \quad b_n^\delta := y_{[n]}^{\delta_{[n]}} - F_{[n]}(x_n)$$

where $[n] := n \bmod d$ denotes the remainder of integer division. Observe that the minimizer $z_{n,k+1}$ of $T_{n,k}^\delta : X \rightarrow \mathbb{R}$ exists and is unique due to the strict convexity of Δ_p . Set $s_{n,k} := z_{n,k} - x_n$ and $x_{n+1} := x_n + s_{n,k_n}$ where the final (inner) index k_n is determined as follows: choose $\tau > 1$, $\mu \in (0, 1)$ and $k_{\max} \in \mathbb{N} \cup \{\infty\}$. Define $k_n = 0$ in case of

$$(12) \quad \|b_n^\delta\| \leq \tau \delta_{[n]}.$$

Otherwise set

$$(13) \quad k_{REG} := \min \{k \in \{1, \dots, k_{\max}\} : \|b_n^\delta - A_n s_{n,k}\| < \mu \|b_n^\delta\|\},$$

using $\min \emptyset = \infty$. Finally,

$$k_n = \begin{cases} k_{REG} & : k_{REG} \leq k_{\max}, \\ k_{\max} & : k_{REG} > k_{\max}. \end{cases}$$

Note that $x_{n+1} = z_{n,k_n}$ and $x_{n+1} = x_n$ if and only if (12) holds. The outer iteration stops as soon as the discrepancy principle (12) is satisfied d times in a row. Our approximate solution of (1) is then x_N where $N = N(\delta)$ is the smallest number⁵ which satisfies

$$(14) \quad \|y_j^{\delta_j} - F_j(x_N)\| \leq \tau \delta_j, \quad j = 0, \dots, d-1.$$

See Algorithm 1 for an implementation in pseudocode.

As $z_{n,k+1}$ is the minimizer of $T_{n,k}^\delta$, we have $T_{n,k}^\delta(z_{n,k+1}) \leq T_{n,k}^\delta(z_{n,k})$ yielding

$$(15) \quad \|b_n^\delta - A_n s_{n,k+1}\| \leq \|b_n^\delta - A_n s_{n,k}\|, \quad k = 0, \dots, k_n - 1,$$

where equality holds only if $z_{n,k+1} = z_{n,k}$ as $z_{n,k+1} \neq z_{n,k}$ results in $\alpha_n \Delta_p(z_{n,k+1}, z_{n,k}) > 0$.

Using the *optimality condition* $0 \in \partial T_{n,k}^\delta(z_{n,k+1})$, we arrive at

$$\alpha_n (J_p(z_{n,k+1}) - J_p(z_{n,k})) \in A_n^* J_r(b_n^\delta - A_n s_{n,k+1}).$$

⁵The number N is chosen by a posteriori strategy, it thus depends actually on δ and y^δ : $N = N(\delta, y^\delta)$. But we stick to the simpler notation $N = N(\delta)$.

Algorithm 1 K-REGINN-IT

Input: $x_N; (y^{\delta_j}, \delta_j); F_j; F'_j, j = 0, \dots, d-1; \mu; k_{\max}; \tau;$

Output: x_N with $\|y_j^{\delta_j} - F_j(x_N)\| \leq \tau\delta_j, j = 0, \dots, d-1;$

$\ell := 0; x_0 := x_N; c := 0;$

while $c < d$ **do**

for $j = 0 : d-1$ **do**

$n := \ell d + j;$

$b_n^\delta := y_j^{\delta_j} - F_j(x_n); A_n := F'_j(x_n);$

if $\|b_n^\delta\| \leq \tau\delta_j$ **then**

$x_{n+1} := x_n; c := c + 1;$

else

$k := 0; z_{n,0} := x_n;$

 choose $\alpha_n > 0$ properly;

repeat

$z_{n,k+1} := \arg \min_{z \in X} \left(\frac{1}{r} \|b_n^\delta - A_n(z - x_n)\|^r + \alpha_n \Delta_p(z, z_{n,k}) \right)$

$k := k + 1;$

until $\|b_n^\delta - A_n(z_{n,k} - x_n)\| < \mu \|b_n^\delta\|$ or $k = k_{\max}$

$x_{n+1} := z_{n,k}; c := 0;$

end if

end for

$\ell := \ell + 1;$

end while

$x_N := x_{\ell d - c};$

Hence, there exists some selection j_r such that

$$J_p(z_{n,k+1}) = J_p(z_{n,k}) + \frac{1}{\alpha_n} A_n^* j_r (b_n^\delta - A_n s_{n,k+1}).$$

Remark 4. By definition of $s_{n,k}$ the above equality can be rewritten as the implicit iteration

$$(16) \quad z_{n,k+1} = J_{p^*}^* \left(J_p(z_{n,k}) + \frac{1}{\alpha_n} A_n^* j_r (b_n^\delta - A_n(z_{n,k+1} - x_n)) \right)$$

which can be solved for $z_{n,k+1}$ by a fixed point iteration. Alternatively, one may apply a gradient method like steepest descent (see, e.g., [1]) directly to the nonlinear functional (11) to find its minimizer.

Remark 5. If $k_{\max} = 1$ then $k_n \in \{0, 1\}$ for all $n \in \mathbb{N}$ and x_{n+1} minimizes

$$T_{n,0}^\delta(x) = \frac{1}{r} \|y_{[n]}^{\delta_{[n]}} - F_{[n]}(x_n) - F'_{[n]}(x_n)(x - x_n)\|^r + \alpha_n \Delta_p(x, x_n).$$

In Hilbert spaces this functional reads (with $p = r = 2$)

$$T_{n,0}^\delta(x) = \frac{1}{2} \|y_{[n]}^{\delta_{[n]}} - F_{[n]}(x_n) - F'_{[n]}(x_n)(x - x_n)\|^2 + \frac{\alpha_n}{2} \|x - x_n\|^2$$

revealing K-REGINN-IT as Kaczmarz version of the Levenberg-Marquardt method in Banach spaces. In case of a linear problem ($F_j = A_j$ are linear for all j 's) we have

$$T_{n,0}^\delta(x) = \frac{1}{r} \left\| y_{[n]}^{\delta_{[n]}} - A_{[n]}x \right\|^r + \alpha_n \Delta_p(x, x_n)$$

and the method is now a Kaczmarz version of the iterated Tikhonov method defined in [7] for the particular case $d = 1$.

In the next theorem we prove that K-REGINN-IT is well defined and terminates.

Theorem 6. *Let X and Y be Banach spaces with X satisfying Assumptions 1 and 2 with $1 < p \leq s \leq r$. Let Assumption 3 hold true and start with $x_0 \in B_\rho(x^+, \Delta_p)$. Choose $\bar{\alpha} > 0$ and define the constants*

$$C_0 := 6p^* M \left(\|x^+\| \vee \rho^{\frac{1}{p}} \right) > 0$$

and

$$C_1 := 2p^* \left(\|x^+\| \vee \left(\rho + \frac{C_0^r}{\bar{\alpha} 2^r} \right)^{\frac{1}{p}} \right) > 0.$$

Define $\alpha_{\min} := \min \{\bar{\alpha}, \tilde{\alpha}\} > 0$, where $\tilde{\alpha} := C_2^{\frac{1}{1-s^*}} C_0^{\frac{(r-1)(s^*-1)-1}{s^*-1}} > 0$ with

$$0 < C_2 < \left(2C_{p^*,s^*} C_1^{(p^*-s^*)(p-1)} M^{s^*} \right)^{-1}$$

where C_{p^*,s^*} is the constant from Assumption 1(c). Additionally, assume that the constant of the TCC in Assumption 3(c) satisfies $0 \leq \eta < C_3$ where

$$C_3 := \frac{1}{2^r} - C_4 > 0$$

with

$$C_4 := C_{p^*,s^*} C_1^{(p^*-s^*)(p-1)} M^{s^*} C_2 > 0.$$

Further, let $\alpha_n \in [\alpha_{\min}, \alpha_{\max}]$ for some $\alpha_{\max} > \alpha_{\min}$, $\tau > \frac{\eta+1}{C_3-\eta}$ and $\mu \in \left(\frac{\eta+1}{C_3-\eta}, 1 \right)$. Then, there exists an $N(\delta)$ such that all iterates $\{x_1, \dots, x_{N(\delta)}\}$ of K-REGINN-IT are well defined and remain in $B_\rho(x^+, \Delta_p)$. Moreover, only the final iterate satisfies (14) and

$$(17) \quad \Delta_p(x^+, x_n) \leq \Delta_p(x^+, x_{n-1})$$

for all $1 \leq n \leq N(\delta)$ where equality holds if and only if (12) applies. Furthermore,

$$(18) \quad \|x_n\| \leq C \text{ for all } \delta > 0 \text{ and } n \leq N(\delta),$$

with $C > 0$ being independent of n , $N(\delta)$, and δ .

Proof. First, observe that the bounds on C_2 imply $C_4 < 1/2$ which makes C_3 well defined. The lower bound on τ guarantees that we can select μ from an open interval.

We argue inductively: Suppose that x_0, \dots, x_n are in $B_\rho(x^+, \Delta_p)$ and (17) holds. Further assume that x_n is not the final iterate, i.e., (14) is not satisfied for $N = n$.

If x_n satisfies (12), then $x_{n+1} = x_n \in B_\rho(x^+, \Delta_p)$ and (17) becomes an equality. In case x_n violates (12), x_{n+1} is also well defined as we demonstrate now: We define $e_n := x^+ - x_n$ and, by (6), we find that

$$\begin{aligned} \Delta_p(x^+, z_{n,k+1}) - \Delta_p(x^+, z_{n,k}) &= -\Delta_p(z_{n,k+1}, z_{n,k}) + \langle J_p(z_{n,k+1}) - J_p(z_{n,k}), z_{n,k+1} - x^+ \rangle \\ &\leq \langle J_p(z_{n,k+1}) - J_p(z_{n,k}), z_{n,k+1} - x^+ \rangle \end{aligned}$$

$$\begin{aligned}
(19) \quad &= \frac{1}{\alpha_n} \langle j_r (b_n^\delta - A_n s_{n,k+1}), A_n (s_{n,k+1} - e_n) \rangle \\
&= \frac{1}{\alpha_n} \langle j_r (b_n^\delta - A_n s_{n,k+1}), (b_n^\delta - A_n e_n) - (b_n^\delta - A_n s_{n,k+1}) \rangle \\
&\leq \frac{1}{\alpha_n} \left(\|b_n^\delta - A_n s_{n,k+1}\|^{r-1} \|b_n^\delta - A_n e_n\| - \|b_n^\delta - A_n s_{n,k+1}\|^r \right)
\end{aligned}$$

for $k = 0, \dots, k_n - 1$. Now, using Assumption 3(c), we get for all $k \leq k_n - 2$

$$\begin{aligned}
(20) \quad &\|b_n^\delta - A_n e_n\| \leq \|y_{[n]}^{\delta_{[n]}} - y_{[n]}\| + \|F_{[n]}(x^+) - F_{[n]}(x_n) - F'_{[n]}(x_n)(x^+ - x_n)\| \\
&\leq \delta_{[n]} + \eta \|F_{[n]}(x^+) - F_{[n]}(x_n)\| \leq \delta_{[n]} + \eta \left(\|y_{[n]}^{\delta_{[n]}} - y_{[n]}\| + \|b_n^\delta\| \right) \\
&\leq \delta_{[n]} (\eta + 1) + \eta \|b_n^\delta\| < \left(\frac{\eta + 1}{\tau} + \eta \right) \|b_n^\delta\|.
\end{aligned}$$

As $\left(\frac{\eta+1}{\tau} + \eta\right) \|b_n^\delta\| \leq \mu \|b_n^\delta\| < \|b_n^\delta - A_n s_{n,k+1}\|^r$, $k \leq k_n - 2$, and in view of (20) we conclude that the right-hand side of (19) is negative. Then, for all $l \leq k_n$,

$$\begin{aligned}
&\sum_{k=0}^{l-1} (\Delta_p(x^+, z_{n,k+1}) - \Delta_p(x^+, z_{n,k})) \\
&\leq \frac{1}{\alpha_n} \|b_n^\delta - A_n s_{n,l}\|^{r-1} \left(\left(\frac{\eta + 1}{\tau} + \eta \right) \|b_n^\delta\| - \|b_n^\delta - A_n s_{n,l}\| \right) \\
&\leq \frac{\left(\frac{\eta+1}{\tau} + \eta\right)}{\alpha_n} \|b_n^\delta - A_n s_{n,l}\|^{r-1} \|b_n^\delta\|.
\end{aligned}$$

From (15), $\alpha_n \geq \bar{\alpha}$, $\tau > \frac{\eta+1}{C_3-\eta}$ and $C_3 \leq \frac{1}{2^r}$ we deduce that

$$(21) \quad \Delta_p(x^+, z_{n,l}) \leq \Delta_p(x^+, z_{n,0}) + \frac{1}{\bar{\alpha} 2^r} \|b_n^\delta\|^r.$$

From (20) and $\left(\frac{\eta+1}{\tau} + \eta\right) \leq C_3 \leq \frac{1}{2}$,

$$\|b_n^\delta\| - \|A_n e_n\| \leq \|b_n^\delta - A_n e_n\| \leq \left(\frac{\eta + 1}{\tau} + \eta \right) \|b_n^\delta\| \leq \frac{1}{2} \|b_n^\delta\|,$$

yielding $\|b_n^\delta\| \leq 2M \|e_n\|$. By $\|e_n\| \leq \|x_n\| + \|x^+\|$, the induction hypotheses $x_n \in B_\rho(x^+, \Delta_p)$, and by (7), we see that $\|b_n^\delta\| \leq C_0$. As $z_{n,0} = x_n \in B_\rho(x^+, \Delta_p)$, it follows from (21),

$$\Delta_p(x^+, z_{n,l}) \leq \rho + \frac{C_0^r}{\bar{\alpha} 2^r},$$

which in view of (7) implies

$$(22) \quad \|z_{n,l}\| \leq C_1 \text{ for all } l \leq k_n.$$

Rearranging the estimate

$$\begin{aligned}
\|b_n^\delta - A_n s_{n,k}\|^r &\leq \left(\|b_n^\delta - A_n s_{n,k+1}\| + \|A_n (s_{n,k+1} - s_{n,k})\| \right)^r \\
&\leq \left(\|b_n^\delta - A_n s_{n,k+1}\| + M \|z_{n,k+1} - z_{n,k}\| \right)^r \\
&\leq 2^r \left(\|b_n^\delta - A_n s_{n,k+1}\|^r + M^r \|z_{n,k+1} - z_{n,k}\|^r \right)
\end{aligned}$$

we obtain

$$(23) \quad - \|b_n^\delta - A_n s_{n,k+1}\|^r \leq -\frac{1}{2^r} \|b_n^\delta - A_n s_{n,k}\|^r + M^r \|z_{n,k+1} - z_{n,k}\|^r.$$

To bound the rightmost term we use Assumption 1(c) and note that $p^* - s^* \geq 0$ for $p \leq s$:

$$\begin{aligned} \|z_{n,k+1} - z_{n,k}\| &= \|J_{p^*}^*(J_p(z_{n,k+1})) - J_{p^*}^*(J_p(z_{n,k}))\| \\ &\leq C_{p^*,s^*} (\|J_p(z_{n,k+1})\| \vee \|J_p(z_{n,k})\|)^{p^*-s^*} \|J_p(z_{n,k+1}) - J_p(z_{n,k})\|^{s^*-1} \\ &= C_{p^*,s^*} (\|z_{n,k+1}\| \vee \|z_{n,k}\|)^{(p^*-s^*)(p-1)} \left\| \frac{1}{\alpha_n} A_n^* j_r (b_n^\delta - A_n s_{n,k+1}) \right\|^{s^*-1} \\ &\leq C_{p^*,s^*} C_1^{(p^*-s^*)(p-1)} \alpha_n^{1-s^*} M^{s^*-1} \|b_n^\delta - A_n s_{n,k+1}\|^{(r-1)(s^*-1)} \end{aligned}$$

for all $k \leq k_n - 1$. We proceed using (15), $\alpha_n \geq \tilde{\alpha}$, and $(r-1)(s^*-1) - 1 \geq 0$ for $r \geq s$:

$$\begin{aligned} M^r \|z_{n,k+1} - z_{n,k}\|^r &\leq \left(C_{p^*,s^*} C_1^{(p^*-s^*)(p-1)} M^{s^*} \right)^r \frac{\|b_n^\delta\|^{r[(r-1)(s^*-1)-1]}}{\alpha_n^{(s^*-1)r}} \|b_n^\delta - A_n s_{n,k}\|^r \\ &\leq \left(C_{p^*,s^*} C_1^{(p^*-s^*)(p-1)} M^{s^*} \right)^r \left[\frac{C_0^{\frac{(r-1)(s^*-1)-1}{s^*-1}}}{\alpha_n} \right]^{r(s^*-1)} \|b_n^\delta - A_n s_{n,k}\|^r \\ &\leq \left(C_{p^*,s^*} C_1^{(p^*-s^*)(p-1)} M^{s^*} C_2 \right)^r \|b_n^\delta - A_n s_{n,k}\|^r \\ &= C_4^r \|b_n^\delta - A_n s_{n,k}\|^r. \end{aligned}$$

From (23),

$$(24) \quad - \|b_n^\delta - A_n s_{n,k+1}\|^r \leq -\left(\frac{1}{2^r} - C_4^r \right) \|b_n^\delta - A_n s_{n,k}\|^r = -C_3 \|b_n^\delta - A_n s_{n,k}\|^r$$

for all $k \leq k_n - 1$. Inserting (24) into (19) we, in view of (15), arrive at

$$\begin{aligned} \Delta_p(x^+, z_{n,k+1}) - \Delta_p(x^+, z_{n,k}) &\leq \frac{1}{\alpha_n} \|b_n^\delta - A_n s_{n,k}\|^{r-1} \left(\left(\frac{\eta+1}{\tau} + \eta \right) \|b_n^\delta\| - C_3 \|b_n^\delta - A_n s_{n,k}\| \right). \end{aligned}$$

Since $-C_3 \|b_n^\delta - A_n s_{n,k}\| \leq -\mu C_3 \|b_n^\delta\|$ for all $k \leq k_n - 1$ we have that

$$\left(\frac{\eta+1}{\tau} + \eta \right) \|b_n^\delta\| - C_3 \|b_n^\delta - A_n s_{n,k}\| \leq -C_5 \|b_n^\delta\|$$

with $C_5 := \mu C_3 - \left(\frac{\eta+1}{\tau} + \eta \right) > 0$. Further,

$$\sum_{k=0}^{l-1} (\Delta_p(x^+, z_{n,k+1}) - \Delta_p(x^+, z_{n,k})) \leq -C_5 \frac{\|b_n^\delta\|}{\alpha_n} \sum_{k=0}^{l-1} \|b_n^\delta - A_n s_{n,k}\|^{r-1}$$

for all $l \leq k_n$ resulting in

$$(25) \quad C_5 \frac{\|b_n^\delta\|}{\alpha_n} \sum_{k=0}^{l-1} \|b_n^\delta - A_n s_{n,k}\|^{r-1} \leq \Delta_p(x^+, x_n) - \Delta_p(x^+, z_{n,l}) < \infty.$$

As $\|b_n^\delta - A_n s_{n,k}\| \geq \mu \|b_n^\delta\|$ for all $k \leq k_n - 1$,

$$C_5 \mu^{r-1} \frac{\|b_n^\delta\|^r}{\alpha_n} l < \infty,$$

for all $l \leq k_n$, which shows that $l < \infty$ and then $k_n < \infty$. Hence $x_{n+1} = z_{n,k_n}$ is well defined and letting $l = k_n$ in (25) gives

$$(26) \quad C_5 \frac{\|b_n^\delta\|}{\alpha_n} \sum_{k=0}^{k_n-1} \|b_n^\delta - A_n s_{n,k}\|^{r-1} \leq \Delta_p(x^+, x_n) - \Delta_p(x^+, x_{n+1})$$

which finally validates (17): $\Delta_p(x^+, x_{n+1}) < \Delta_p(x^+, x_n) \leq \rho$.

Now we prove that Algorithm 1 terminates. Define the set $I := \{n \in \mathbb{N} : \|b_n^\delta\| > \tau \delta_{[n]}\}$ and suppose that I has infinitely many elements. Using again $\|b_n^\delta - A_n s_{n,k}\| \geq \mu \|b_n^\delta\|$ for all $k \leq k_n - 1$ and $\alpha_n \leq \alpha_{\max}$, it follows from (26) that

$$C_5 \mu^{r-1} \frac{\|b_n^\delta\|^r}{\alpha_{\max}} k_n \leq \Delta_p(x^+, x_n) - \Delta_p(x^+, x_{n+1})$$

for all $n \in I$. But for $n \notin I$, $k_n = 0$ and the above inequality trivially holds. Therefore,

$$\sum_{n=0}^{\infty} \|b_n^\delta\|^r k_n \lesssim \sum_{n=0}^{\infty} (\Delta_p(x^+, x_n) - \Delta_p(x^+, x_{n+1})) \leq \Delta_p(x^+, x_0) < \infty.$$

Define now $\delta_{\min} := \min_{0 \leq j \leq d-1} \delta_j > 0$ and observe that $k_n \geq 1$ for all $n \in I$. Hence,

$$\sum_{n \in I} (\tau \delta_{\min})^r \leq \sum_{n \in I} \|b_n^\delta\|^r k_n = \sum_{n=0}^{\infty} \|b_n^\delta\|^r k_n < \infty$$

which can hold only if the number of elements in I is finite. Thus, $N(\delta)$ is the largest element in I plus 1. From inequality (17),

$$\Delta_p(x^+, x_n) \leq \Delta_p(x^+, x_0) < \infty$$

for all $\delta > 0$ and $n \leq N(\delta)$. It follows that $\|x_n\| \leq C$ for all $n \leq N(\delta)$ and some $C > 0$ independent on n , N and δ . \square

Weak convergence of K-REGINN-IT is an immediate consequence of (18) and the reflexivity of X , see [10, Corollary 3.5].

Corollary 7. *Let all the assumptions of Theorem 6 hold true. If the operators F_j , $j = 0, \dots, d-1$, are weakly sequentially closed then for any sequence $(y_j^{(\delta_j)_i})_{i \in \mathbb{N}}$ with $\delta^{(i)} = \max\{(\delta_j)_i : j = 0, \dots, d-1\} \rightarrow 0$ as $i \rightarrow \infty$, the sequence $(x_{N(\delta^{(i)})})_{i \in \mathbb{N}}$ contains a subsequence that converges weakly to a solution of (1) in $B_\rho(x^+, \Delta_p)$. If x^+ is the unique solution of (1) in $B_\rho(x^+, \Delta_p)$, then $(x_{N(\delta)})_{\delta > 0}$ converges weakly to x^+ as $\delta = \max\{\delta_j : j = 0, \dots, d-1\} \rightarrow 0$.*

Remark 8. *The constants C_0 and C_1 in Theorem 6 depend on the unknown solution x^+ . But as $x_0 \in B_\rho(x^+, \Delta_p)$ we conclude that both constants are bounded in p , $\|x_0\|$, and ρ .*

Remark 9. *At a first glance the restriction $s \leq r$ in the above theorem might affect the computation of a minimizer of $T_{n,k}^\delta$ (11) via (16). This, however, is not the case due to (5). For instance, if $Y = L^{1,1}$ and $s \geq 2$ we can realize J_r on Y by (5) and (8).*

Remark 10. *The monotonicity estimate (17) actually holds in a more general setting:*

$$(27) \quad \Delta_p(\vartheta_n, x_{n+1}) \leq \Delta_p(\vartheta_n, x_n),$$

whenever ϑ_n is a solution of the $[n]$ -th equation: $y_{[n]} = F_{[n]}(\vartheta_n)$.

Remark 11. *Let $k_{\max} = \infty$ and assume that (12) is violated by x_n . Following [5] we find*

$$\begin{aligned} \|y_{[n]}^{\delta_{[n]}} - F_{[n]}(x_{n+1})\| &\leq \|y_{[n]}^{\delta_{[n]}} - F_{[n]}(x_n) - F'_{[n]}(x_n)(x_{n+1} - x_n)\| \\ &\quad + \|F_{[n]}(x_{n+1}) - F_{[n]}(x_n) - F'_{[n]}(x_n)(x_{n+1} - x_n)\| \\ &\leq \mu \|y_{[n]}^{\delta_{[n]}} - F_{[n]}(x_n)\| + \eta \|F_{[n]}(x_{n+1}) - F_{[n]}(x_n)\| \end{aligned}$$

so that

$$\|y_{[n]}^{\delta_{[n]}} - F_{[n]}(x_{n+1})\| \leq \Lambda \|y_{[n]}^{\delta_{[n]}} - F_{[n]}(x_n)\| \quad \text{where } \Lambda := \frac{\mu + \eta}{1 - \eta}.$$

Now, if $0 \leq \eta < \frac{C_3}{1+2C_3}$ and $\tau > \frac{1+\eta}{C_3(1-2\eta)-\eta}$ then $\frac{1+\eta}{\tau C_3} < 1 - 2\eta$ and restricting μ to $\left(\frac{1+\eta}{\tau C_3}, 1 - 2\eta\right)$ yields $\Lambda < 1$.

4. CONVERGENCE IN THE NOISE-FREE SETTING

From now on, we need to differ clearly between the noisy ($\delta > 0$) and the noise-free ($\delta = 0$) situations. For this reason we exclusively mark quantities by a superscript δ when the data is corrupted by noise: $x_n^\delta, b_n^\delta, A_n^\delta$ etc. Thus, x_n, b_n, A_n etc. originate from exact data. Note that the starting guess is chosen independently of δ : $x_0^\delta = x_0$.

Algorithm 1 is well defined in the noiseless situation when we set $\delta_j = 0$, $\tau = \infty$, and $\tau\delta_j = 0$. Then, the discrepancy principle (12) is replaced by $\|b_n\| = 0$, in which case $x_{n+1} = x_n$. Termination only occurs in the unlikely event that an iterate x_N satisfies $\|y_j - F_j(x_N)\| = 0$ for $j = 0, \dots, d-1$, i.e., x_N solves (4). In general, Algorithm 1 does not stop but produces a sequence which converges strongly to a solution of (1) as we will prove in this section, see Theorem 13 below.

Except for the termination statement, all results of Theorem 6 hold true with an even larger interval for the selection of the tolerances: $\mu \in \left(\frac{\eta}{C_3}, 1\right)$. Accordingly, the constant in (26) is replaced by $C_5 := \mu C_3 - \eta > 0$. Further, $N(\delta) = \infty$ in case we have no premature termination.

With the next lemma we prepare our convergence proof for the exact data case.

Lemma 12. *Assume all the hypotheses from Theorem 6 but with $\mu \in \left(\frac{\eta}{C_3}, 1\right)$. Then,*

$$(28) \quad \Delta_p(x_n, x_{n+1}) \lesssim \Delta_p(x^+, x_n) - \Delta_p(x^+, x_{n+1})$$

for all $n \in \mathbb{N}$.

Proof. From (6),

$$(29) \quad \Delta_p(x_n, x_{n+1}) \leq \Delta_p(x^+, x_{n+1}) - \Delta_p(x^+, x_n) + |\langle J_p(x_{n+1}) - J_p(x_n), x^+ - x_n \rangle|.$$

But from the definition of the scheme and from properties of j_r ,

$$\begin{aligned} |\langle J_p(x_{n+1}) - J_p(x_n), x^+ - x_n \rangle| &= |\langle J_p(z_{n,k_n}) - J_p(z_{n,0}), x^+ - x_n \rangle| \\ &= \left| \sum_{k=0}^{k_n-1} \langle J_p(z_{n,k+1}) - J_p(z_{n,k}), x_n - x^+ \rangle \right| \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{\alpha_n} \left| \sum_{k=0}^{k_n-1} \langle j_r(b_n - A_n s_{n,k+1}), A_n(x_n - x^+) \rangle \right| \\
&\leq \frac{1}{\alpha_n} \sum_{k=0}^{k_n-1} \|b_n - A_n s_{n,k}\|^{r-1} \|A_n(x_n - x^+)\| \\
&\leq (\eta + 1) \frac{\|b_n\|}{\alpha_n} \sum_{k=0}^{k_n-1} \|b_n - A_n s_{n,k}\|^{r-1} \\
&\leq \frac{\eta + 1}{C_5} (\Delta_p(x^+, x_n) - \Delta_p(x^+, x_{n+1})).
\end{aligned}$$

We have used the TCC (Assumption 3(c)), (15) and (26). Inserting this result in (29), we arrive at (28) with constant $\frac{\eta+1}{C_5} - 1 > 0$. \square

In the following convergence proof we adapt ideas from our work [12]. The line of argumentation is similar to [12] but we decided to present all details for the reader's convenience.

Theorem 13. *Let X and Y be Banach spaces with X satisfying Assumptions 1 and 2 with $1 < p \leq s \leq r$. Let Assumption 3 hold true and start with $x_0 \in B_\rho(x^+, \Delta_p)$. Choose the constants α_{\min} , α_{\max} , and C_3 as in Theorem 6 and assume that the constant of the TCC in Assumption 3(c) satisfies $0 \leq \eta < C_3$. Additionally, let $k_{\max} < \infty$ in case $d > 1$.*

If $\alpha_n \in [\alpha_{\min}, \alpha_{\max}]$ and $\mu \in (\eta/C_3, 1)$ then K-REGINN-IT either stops after finitely many iterations with a solution of (1) or the sequence $(x_n)_{n \in \mathbb{N}} \subset B_\rho(x^+, \Delta_p)$ converges strongly in X to a solution of (1). If x^+ is the unique solution in $B_\rho(x^+, \Delta_p)$, then $x_n \rightarrow x^+$ as $n \rightarrow \infty$.

Proof. If Algorithm 1 stops after a finite number of iterations then the current iterate is a solution of (1). Otherwise, $(x_n)_{n \in \mathbb{N}}$ is a Cauchy sequence, as we will prove now.

Let $m, l \in \mathbb{N}$ with $m \leq l$. We consider first the case $d > 1$. Write $m = m_0 d + m_1$ and $l = l_0 d + l_1$ with $m_0, l_0 \in \mathbb{N}$ and $m_1, l_1 \in \{0, \dots, d-1\}$. Of course $m_0 \leq l_0$. Choose $z_0 \in \{m_0, \dots, l_0\}$ such that

$$\begin{aligned}
(30) \quad &\sum_{j=0}^{d-1} (\|y_j - F_j(x_{z_0 d+j})\| + \|x_{z_0 d+j+1} - x_{z_0 d+j}\|) \\
&\leq \sum_{j=0}^{d-1} (\|y_j - F_j(x_{n_0 d+j})\| + \|x_{n_0 d+j+1} - x_{n_0 d+j}\|)
\end{aligned}$$

for all $n_0 \in \{m_0, \dots, l_0\}$. Define $z := z_0 d + z_1$, where $z_1 = l_1$ if $z_0 = l_0$ and $z_1 = d-1$ otherwise. This setting guarantees $m \leq z \leq l$. From Assumption 2(c),

$$\|x_m - x_l\|^s \leq 2^s (\|x_m - x_z\|^s + \|x_z - x_l\|^s) \lesssim \Delta_p(x_z, x_m) + \Delta_p(x_z, x_l).$$

Identity (6) implies now that

$$(31) \quad \|x_m - x_l\|^s \lesssim \beta_{m,z} + \beta_{l,z} + f(z, m, l)$$

with $\beta_{m,z} := \Delta_p(x^+, x_m) - \Delta_p(x^+, x_z)$ and

$$f(z, m, l) := \left| \langle J_p(x_z) - J_p(x_m), x_z - x^+ \rangle \right| + \left| \langle J_p(x_z) - J_p(x_l), x_z - x^+ \rangle \right|.$$

By monotonicity (17), we conclude that $\Delta_p(x^+, x_n) \rightarrow \gamma \geq 0$ as $n \rightarrow \infty$. Thus, $\beta_{m,z}$ and $\beta_{l,z}$ converge to zero as $m \rightarrow \infty$ (which causes $z \rightarrow \infty$ and $l \rightarrow \infty$). Further,

$$(32) \quad f(z, m, l) \leq \sum_{n=m}^{l-1} |\langle J_p(x_{n+1}) - J_p(x_n), x_z - x^+ \rangle|.$$

As in the proof of Lemma 12 we find

$$(33) \quad |\langle J_p(x_{n+1}) - J_p(x_n), x_z - x^+ \rangle| \leq \frac{1}{\alpha_n} \sum_{k=0}^{k_n-1} \|b_n - A_n s_{n,k}\|^{r-1} \|A_n(x_z - x^+)\|.$$

From Assumption 3(c),

$$(34) \quad \begin{aligned} \|A_n(x_z - x^+)\| &\leq \|A_n(x_n - x^+)\| + \|A_n(x_z - x_n)\| \\ &\leq \|b_n\| + \|b_n - A_n(x^+ - x_n)\| + \|F_{[n]}(x_z) - F_{[n]}(x_n)\| \\ &\quad + \|F_{[n]}(x_z) - F_{[n]}(x_n) - F'_{[n]}(x_n)(x_z - x_n)\| \\ &\leq (\eta + 1) (\|b_n\| + \|F_{[n]}(x_z) - F_{[n]}(x_n)\|) \\ &\leq (\eta + 1) (2\|b_n\| + \|y_{[n]} - F_{[n]}(x_z)\|). \end{aligned}$$

Observe that in the last norm, the operator $F_{[n]}$ is applied in the "wrong" vector x_z . To estimate this norm, we use Assumption 3. Write $n = n_0d + n_1$ with $n_0 \in \{m_0, \dots, l_0\}$ and $n_1 \in \{0, \dots, d-1\}$. Then,

$$\begin{aligned} \|y_{[n]} - F_{[n]}(x_z)\| &= \|y_{n_1} - F_{n_1}(x_{z_0d+z_1})\| \\ &\leq \|y_{n_1} - F_{n_1}(x_{z_0d+n_1})\| + \sum_{j=0}^{d-1} \|F_{n_1}(x_{z_0d+j+1}) - F_{n_1}(x_{z_0d+j})\| \\ &\leq \|y_{n_1} - F_{n_1}(x_{z_0d+n_1})\| + \frac{1}{1-\eta} \sum_{j=0}^{d-1} \|F'_{n_1}(x_{z_0d+j})(x_{z_0d+j+1} - x_{z_0d+j})\| \\ &\leq \left(1 + \frac{M}{1-\eta}\right) \sum_{j=0}^{d-1} (\|y_j - F_j(x_{z_0d+j})\| + \|x_{z_0d+j+1} - x_{z_0d+j}\|). \end{aligned}$$

From (30),

$$(35) \quad \|y_{[n]} - F_{[n]}(x_z)\| \leq \left(1 + \frac{M}{1-\eta}\right) \sum_{j=0}^{d-1} (\|y_j - F_j(x_{n_0d+j})\| + \|x_{n_0d+j+1} - x_{n_0d+j}\|).$$

Inserting (35) in (34), (34) in (33), and (33) in (32), we arrive at

$$(36) \quad f(z, m, l) \lesssim \sum_{n=m}^{l-1} \frac{\|b_n\|}{\alpha_n} \sum_{k=0}^{k_n-1} \|b_n - A_n s_{n,k}\|^{r-1} + g(z, m, l) + h(z, m, l)$$

where

$$\begin{aligned} g(z, m, l) &:= \sum_{n=m}^{l-1} \frac{1}{\alpha_n} \sum_{k=0}^{k_n-1} \|b_n - A_n s_{n,k}\|^{r-1} \sum_{j=0}^{d-1} \|y_j - F_j(x_{n_0d+j})\|, \\ h(z, m, l) &:= \sum_{n=m}^{l-1} \frac{1}{\alpha_n} \sum_{k=0}^{k_n-1} \|b_n - A_n s_{n,k}\|^{r-1} \sum_{j=0}^{d-1} \|x_{n_0d+j+1} - x_{n_0d+j}\|. \end{aligned}$$

The first term on the right-hand side of (36) can be estimated by (26). It remains to estimate g and h . As

$$\sum_{k=0}^{k_n-1} \|b_n - A_n s_{n,k}\|^{r-1} \leq k_{\max} \|b_n\|^{r-1},$$

and $\|b_n - A_n s_{n,k}\| \geq \mu \|b_n\|$ for all $k \leq k_n - 1$

$$\begin{aligned} (37) \quad g(z, m, l) &\leq \frac{k_{\max}}{\alpha_{\min}} \sum_{n_0=m_0}^{l_0} \left(\sum_{n_1=0}^{d-1} \|y_{n_1} - F_{n_1}(x_{n_0 d+n_1})\|^{r-1} \sum_{j=0}^{d-1} \|y_j - F_j(x_{n_0 d+j})\| \right) \\ &\lesssim \sum_{n_0=m_0}^{l_0} \sum_{n_1=0}^{d-1} \|y_{n_1} - F_{n_1}(x_{n_0 d+n_1})\|^r = \sum_{n=m_0}^{l_0 d+d-1} \|b_n\|^r \\ &\leq \frac{\alpha_{\max}}{\mu^{r-1}} \sum_{n=m_0}^{l_0 d+d-1} \frac{\|b_n\|}{\alpha_n} \sum_{k=0}^{k_n-1} \|b_n - A_n s_{n,k}\|^{r-1}. \end{aligned}$$

Similarly, we estimate the last term in (36),

$$\begin{aligned} (38) \quad h(z, m, l) &\leq \frac{k_{\max}}{\alpha_{\min}} \sum_{n_0=m_0}^{l_0} \left(\sum_{n_1=0}^{d-1} \|y_{n_1} - F_{n_1}(x_{n_0 d+n_1})\|^{r-1} \sum_{j=0}^{d-1} \|x_{n_0 d+j+1} - x_{n_0 d+j}\| \right) \\ &\lesssim \sum_{n_0=m_0}^{l_0} \left(\sum_{n_1=0}^{d-1} \|y_{n_1} - F_{n_1}(x_{n_0 d+n_1})\|^r + \sum_{n_1=0}^{d-1} \|x_{n_0 d+n_1+1} - x_{n_0 d+n_1}\|^r \right) \\ &= \sum_{n=m_0}^{l_0 d+d-1} \|b_n\|^r + \sum_{n=m_0}^{l_0 d+d-1} \|x_{n+1} - x_n\|^r \\ &\leq \frac{\alpha_{\max}}{\mu^{r-1}} \sum_{n=m_0}^{l_0 d+d-1} \frac{\|b_n\|}{\alpha_n} \sum_{k=0}^{k_n-1} \|b_n - A_n s_{n,k}\|^{r-1} + \sum_{n=m_0}^{l_0 d+d-1} \|x_{n+1} - x_n\|^r. \end{aligned}$$

Relying on Assumption 2(c) once again, we obtain

$$\|x_{n+1} - x_n\|^s \lesssim \Delta_p(x_n, x_{n+1}) \stackrel{(28)}{\lesssim} \Delta_p(x^+, x_n) - \Delta_p(x^+, x_{n+1}).$$

As $r \geq s$, we have for m, l large enough that

$$\sum_{n=m_0}^{l_0 d+d-1} \|x_{n+1} - x_n\|^r \leq \sum_{n=m_0}^{l_0 d+d-1} \|x_{n+1} - x_n\|^s \lesssim \Delta_p(x^+, x_{m_0}) - \Delta_p(x^+, x_{l_0 d+d}) = \beta_{m_0, l_0 d+d}.$$

Plugging this bound into (38), inserting then inequalities (38) and (37) in (36), (36) in (31), and using (26), we end up with

$$(39) \quad \|x_m - x_l\|^s \lesssim \beta_{m,z} + \beta_{l,z} + \beta_{m_0, l_0 d+d}.$$

Now we consider the case $d = 1$ (where $k_{\max} = \infty$ is allowed). This situation is easier because we only need to change the definition x_z in (30) to the vector with the smallest residuum in the outer iteration, i.e., choose $z \in \{m, \dots, l\}$ such that $\|b_z\| \leq \|b_n\|$, for all $n \in \{m, \dots, l\}$. Then, from (34),

$$\|A_n(x_z - x^+)\| \leq (\eta + 1)(2\|b_n\| + \|b_z\|) \leq 3(\eta + 1)\|b_n\|$$

which, together with (33) and (32), leads to

$$f(z, m, l) \leq 3(\eta + 1) \sum_{n=m}^{l-1} \frac{\|b_n\|}{\alpha_n} \sum_{k=0}^{k_n-1} \|b_n - A_n s_{n,k}\|^{r-1}.$$

Plugging now this result into (31) and using again (26), we arrive at (39) with m_0 and $l_0 d + d$ replaced by m and l , respectively.

In any case the right-hand side of inequality (39) converges to zero as $m \rightarrow \infty$ revealing $(x_n)_{n \in \mathbb{N}}$ to be a Cauchy sequence. As X is complete, it converges to some $x_\infty \in X$. Observe that $k_n \geq 1$ if $\|b_n\| \neq 0$ and as $\|b_n - A_n s_{n,k}\| \geq \mu \|b_n\|$ for all $k \leq k_n - 1$,

$$\frac{\mu^{r-1}}{\alpha_{\max}} \sum_{n=0}^{\infty} \|b_n\|^r \leq \sum_{n=0}^{\infty} \frac{\|b_n\|}{\alpha_n} k_n (\mu \|b_n\|)^{r-1} \leq \sum_{n=0}^{\infty} \frac{\|b_n\|}{\alpha_n} \sum_{k=0}^{k_n-1} \|b_n - A_n s_{n,k}\|^{r-1} \stackrel{(26)}{<} \infty.$$

Then, $\|y_{[n]} - F_{[n]}(x_n)\| = \|b_n\| \rightarrow 0$ as $n \rightarrow \infty$ and as the F_j 's are continuous for all $j = 0, \dots, d-1$, we have $y_j = F_j(x_\infty)$. If (1) has only one solution in $B_\rho(x^+, \Delta_p)$ then $x_\infty = x^+$. \square

5. REGULARIZATION PROPERTY

In this section we validate that K-REGINN-IT is a regularization scheme for solving (1) with noisy data y^δ . Indeed, we show that the family $(x_{N(\delta)}^\delta)_{\delta > 0}$ of outputs of Algorithm 1 relative to the inputs $(y^\delta)_{\delta > 0}$ converges strongly to solutions of (1) with exact data y .

To avoid possible wrong interpretations, we will not use the notation δ_j , $j = 0, \dots, d-1$, as in (9) any more. Instead, when we write δ_i , we mean a positive number in a sequence of δ 's as defined in (10), i.e., $\delta_i := \max\{(\delta_j)_i : j = 0, \dots, d-1\} > 0$.

We follow ideas from [10] and [12]. In a first step we investigate the stability of the scheme, i.e., we study the behavior of the n -th iterate x_n^δ as δ approaches zero. The sets \mathcal{X}_n defined below play an important role.

Definition 14. Let $\mathcal{X}_0 := \{x_0\}$ and define \mathcal{X}_{n+1} from \mathcal{X}_n by the following procedure: for each $\xi \in \mathcal{X}_n$, define $\sigma_{n,0}(\xi) := \xi$ and $\sigma_{n,k+1}(\xi)$ as the minimizer of

$$(40) \quad W_{n,k}(z) := \frac{1}{r} \|\tilde{b}_n - F'_{[n]}(\xi)(z - \xi)\|^r + \alpha_n \Delta_p(z, \sigma_{n,k}(\xi)),$$

where $\tilde{b}_n := y_{[n]} - F_{[n]}(\xi)$. Define

$$(41) \quad k_{REG}(\xi) := \min \left\{ k \in \{1, \dots, k_{\max}\} : \|\tilde{b}_n - F'_{[n]}(\xi)(\sigma_{n,k}(\xi) - \xi)\| < \mu \|\tilde{b}_n\| \right\},$$

and

$$k_n(\xi) := \begin{cases} 0 & : \tilde{b}_n = 0, \\ k_{REG}(\xi) & : k_{REG}(\xi) \leq k_{\max}, \\ k_{\max} & : k_{REG}(\xi) > k_{\max}. \end{cases}$$

Then $\sigma_{n,k}(\xi) \in \mathcal{X}_{n+1}$ for $k = 1, \dots, k_n(\xi)$ in case $k_n(\xi) \geq 1$ and only for $k = 0$ in case $k_n(\xi) = 0$. We call $\xi \in \mathcal{X}_n$ the **predecessor** of the vectors $\sigma_{n,k}(\xi) \in \mathcal{X}_{n+1}$ and these ones **successors** of ξ

Of course $x_n \in \mathcal{X}_n$ and \mathcal{X}_n is finite for all $n \in \mathbb{N}$. Moreover, from (17) we get

$$\Delta_p(x^+, \xi_{n+1}) \leq \Delta_p(x^+, \xi_n)$$

whenever $\xi_{n+1} \in \mathcal{X}_{n+1}$ is as successor of $\xi_n \in \mathcal{X}_n$. We emphasize that the sets \mathcal{X}_n , $n \in \mathbb{N}_0$, are defined with respect to exact data y .

The proof of the next lemma basically adapts ideas of [7] and [10].

Lemma 15. *Let all the assumptions of Theorem 6 hold true. If $\delta_i \rightarrow 0$ as $i \rightarrow \infty$ then for $n \leq N(\delta_i)$ with $\delta_i > 0$ sufficiently small, the sequence $(x_n^{\delta_i})_{i \in \mathbb{N}}$ splits into convergent subsequences, all of which converge to elements of \mathcal{X}_n .*

Proof. We prove the statement by induction. For $n = 0$, $x_0^{\delta_i} = x_0 \in \mathcal{X}_0$. Now, suppose that for some $n \in \mathbb{N}$ with $n+1 \leq N(\delta_i)$ for i large enough, $(x_n^{\delta_i})_{i \in \mathbb{N}}$ splits into convergent subsequences, all of which converge to elements of \mathcal{X}_n . To simplify the notation, let $(x_n^{\delta_i})_{i \in \mathbb{N}}$ itself be a subsequence which converges to an element of \mathcal{X}_n , say,

$$(42) \quad \lim_{i \rightarrow \infty} x_n^{\delta_i} = \xi \in \mathcal{X}_n.$$

We must prove that the sequence $(x_{n+1}^{\delta_i})_{i \in \mathbb{N}}$ splits in convergent subsequences, each one converging to a point of \mathcal{X}_{n+1} . Let us prove beforehand by induction over k that

$$(43) \quad z_{n,k}^{\delta_i} \rightarrow \sigma_{n,k}(\xi) \text{ as } i \rightarrow \infty \text{ for all } k \leq k_n(\xi).$$

In the remainder of this proof we suppress the dependence of the $\sigma_{n,k}$'s on ξ .

For $k = 0$, $z_{n,0}^{\delta_i} = x_n^{\delta_i} \rightarrow \xi = \sigma_{n,0}$ as $i \rightarrow \infty$. Suppose for some $k \leq k_n(\xi) - 1$ that $z_{n,k}^{\delta_i} \rightarrow \sigma_{n,k}$ as $i \rightarrow \infty$. As the family $(z_{n,k+1}^{\delta_i})_{\delta_i > 0}$ is uniformly bounded (see (22)) and X is reflexive (Assumption 1(a)), there exists, by picking a subsequence if necessary, some $\bar{z} \in X$ such that $z_{n,k+1}^{\delta_i} \rightharpoonup \bar{z}$ as $i \rightarrow \infty$. Now, for all $g \in Y_{[n]}^*$,

$$\langle g, F'_{[n]}(x_n^{\delta_i}) s_{n,k+1}^{\delta_i} \rangle = \langle g, (F'_{[n]}(x_n^{\delta_i}) - F'_{[n]}(\xi)) s_{n,k+1}^{\delta_i} \rangle + \langle g, F'_{[n]}(\xi) s_{n,k+1}^{\delta_i} \rangle.$$

But as $s_{n,k+1}^{\delta_i} = z_{n,k+1}^{\delta_i} - x_n^{\delta_i} \rightharpoonup \bar{z} - \xi =: \bar{s}$ as $i \rightarrow \infty$ and $F'_{[n]}(\xi)^* g \in X^*$,

$$\langle g, F'_{[n]}(\xi) s_{n,k+1}^{\delta_i} \rangle = \langle F'_{[n]}(\xi)^* g, s_{n,k+1}^{\delta_i} \rangle \rightarrow \langle F'_{[n]}(\xi)^* g, \bar{s} \rangle = \langle g, F'_{[n]}(\xi) \bar{s} \rangle.$$

Now, as $F'_{[n]}$ is continuous and $x_n^{\delta_i} \rightarrow \xi$,

$$|\langle g, (F'_{[n]}(x_n^{\delta_i}) - F'_{[n]}(\xi)) s_{n,k+1}^{\delta_i} \rangle| \leq \|g\|_{Y_{[n]}^*} \|F'_{[n]}(x_n^{\delta_i}) - F'_{[n]}(\xi)\|_{\mathcal{L}(X, Y_{[n]})} \|s_{n,k+1}^{\delta_i}\|_X \rightarrow 0$$

as $i \rightarrow \infty$ because $\|s_{n,k+1}^{\delta_i}\| \leq \|z_{n,k+1}^{\delta_i}\| + \|x_n^{\delta_i}\|$ is uniformly bounded. Then,

$$(44) \quad \langle g, F'_{[n]}(x_n^{\delta_i}) s_{n,k+1}^{\delta_i} \rangle \rightarrow \langle g, F'_{[n]}(\xi) \bar{s} \rangle$$

and as $g \in Y_{[n]}^*$ is arbitrary,

$$F'_{[n]}(x_n^{\delta_i}) s_{n,k+1}^{\delta_i} \rightharpoonup F'_{[n]}(\xi) \bar{s}.$$

From (9) we conclude that

$$b_n^{\delta_i} - F'_{[n]}(x_n^{\delta_i}) s_{n,k+1}^{\delta_i} \rightharpoonup \tilde{b}_n - F'_{[n]}(\xi) \bar{s}$$

and then

$$(45) \quad \|\tilde{b}_n - F'_{[n]}(\xi) \bar{s}\| \leq \liminf \|b_n^{\delta_i} - F'_{[n]}(x_n^{\delta_i}) s_{n,k+1}^{\delta_i}\|.$$

Now, Assumption 1(b) guarantees that J_p is continuous and similarly to (44) we get

$$\langle J_p(z_{n,k}^{\delta_i}), z_{n,k+1}^{\delta_i} \rangle = \langle J_p(z_{n,k}^{\delta_i}) - J_p(\sigma_{n,k}), z_{n,k+1}^{\delta_i} \rangle + \langle J_p(\sigma_{n,k}), z_{n,k+1}^{\delta_i} \rangle \rightarrow \langle J_p(\sigma_{n,k}), \bar{z} \rangle$$

which in turn implies

$$\begin{aligned}
(46) \quad \Delta_p(\bar{z}, \sigma_{n,k}) &= \frac{1}{p} \|\bar{z}\|^p + \frac{1}{p^*} \|\sigma_{n,k}\|^p - \langle J_p(\sigma_{n,k}), \bar{z} \rangle \\
&\leq \liminf \left(\frac{1}{p} \|z_{n,k+1}^{\delta_i}\|^p + \frac{1}{p^*} \|z_{n,k}^{\delta_i}\|^p - \langle J_p(z_{n,k}^{\delta_i}), z_{n,k+1}^{\delta_i} \rangle \right) \\
&= \liminf \Delta_p(z_{n,k+1}^{\delta_i}, z_{n,k}^{\delta_i}).
\end{aligned}$$

From (42), (45), (46) and due to the minimality property of $z_{n,k+1}^{\delta_i}$,

$$W_{n,k}(\bar{z}) \leq \liminf T_{n,k}^{\delta_i}(z_{n,k+1}^{\delta_i}) \leq \liminf T_{n,k}^{\delta_i}(\sigma_{n,k+1}) = \lim_{i \rightarrow \infty} T_{n,k}^{\delta_i}(\sigma_{n,k+1}) = W_{n,k}(\sigma_{n,k+1}),$$

where $W_{n,k}$ and $T_{n,k}^{\delta_i}$ are defined in (40) and (11), respectively. Using minimality and uniqueness of $\sigma_{n,k+1}$, we conclude that $\sigma_{n,k+1} = \bar{z}$ and then $z_{n,k+1}^{\delta_i} \rightarrow \sigma_{n,k+1}$. Accordingly, $s_{n,k+1}^{\delta_i} \rightarrow \sigma_{n,k+1} - \xi$ which implies that $\bar{s} = \sigma_{n,k+1} - \xi$. We prove now that

$$(47) \quad \Delta_p(z_{n,k+1}^{\delta_i}, z_{n,k}^{\delta_i}) \rightarrow \Delta_p(\sigma_{n,k+1}, \sigma_{n,k}) \text{ as } i \rightarrow \infty.$$

Define

$$\begin{aligned}
a_i &:= \Delta_p(z_{n,k+1}^{\delta_i}, z_{n,k}^{\delta_i}), \quad a := \limsup a_i, \quad c := \Delta_p(\sigma_{n,k+1}, \sigma_{n,k}), \\
re_i &:= \frac{1}{r} \|b_n^{\delta_i} - F'_{[n]}(x_n^{\delta_i}) s_{n,k+1}^{\delta_i}\|^r, \text{ and } re := \liminf re_i.
\end{aligned}$$

In view of (46), it is enough to prove that $a \leq c$. Suppose that $a > c$. From the definition of lim sup there exists, for all $M \in \mathbb{N}$, some index $i > M$ such that

$$(48) \quad a_i > a - \frac{a-c}{4}.$$

From the definition of lim inf, there exists $N_1 \in \mathbb{N}$ such that

$$(49) \quad re_i \geq re - \frac{\alpha_n(a-c)}{4},$$

for all $i \geq N_1$. As above, $\lim_{i \rightarrow \infty} T_{n,k}^{\delta_i}(\sigma_{n,k+1}) = W_{n,k}(\sigma_{n,k+1})$ and then there is an $N_2 \in \mathbb{N}$ such that

$$(50) \quad T_{n,k}^{\delta_i}(\sigma_{n,k+1}) < W_{n,k}(\sigma_{n,k+1}) + \frac{\alpha_n(a-c)}{2}$$

for all $i \geq N_2$. Using (45) and setting $M = N_1 \vee N_2$, there exists some index $i > M$ such that

$$\begin{aligned}
W_{n,k}(\sigma_{n,k+1}) &\leq re + \alpha_n c = re + \alpha_n a - \alpha_n(a-c) \\
&\leq re_i + \frac{\alpha_n(a-c)}{4} + \alpha_n a_i + \frac{\alpha_n(a-c)}{4} - \alpha_n(a-c) \\
&= re_i + \alpha_n a_i - \frac{\alpha_n(a-c)}{2} = T_{n,k}^{\delta_i}(z_{n,k+1}^{\delta_i}) - \frac{\alpha_n(a-c)}{2} \\
&\leq T_{n,k}^{\delta_i}(\sigma_{n,k+1}) - \frac{\alpha_n(a-c)}{2}
\end{aligned}$$

where the second inequality comes from (49) and (48) and the last one, from the minimality of $z_{n,k+1}^{\delta_i}$. From (50) we obtain the contradiction $W_{n,k}(\sigma_{n,k+1}) < W_{n,k}(\sigma_{n,k+1})$. Thus, $a \leq c$ and (47) holds. From the definition of a Bregman distance we have $\|z_{n,k+1}^{\delta_i}\| \rightarrow \|\sigma_{n,k+1}\|$. As $z_{n,k+1}^{\delta_i} \rightarrow \sigma_{n,k+1}$ we conclude that $z_{n,k+1}^{\delta_i} \rightarrow \sigma_{n,k+1}$ as $i \rightarrow \infty$,

see Assumption 2(b). So far, we have shown that each positive zero sequence $(\delta_i)_{i \in \mathbb{N}}$ contains a subsequence $(\delta_{i_j})_{j \in \mathbb{N}}$ such that $z_{n,k+1}^{\delta_{i_j}} \rightarrow \sigma_{n,k+1}$ as $j \rightarrow \infty$ which proves (43). Consequently,

$$s_{n,k}^{\delta_i} \rightarrow \sigma_{n,k} - \xi$$

and

$$(51) \quad b_n^{\delta_i} - F'_{[n]}(x_n^{\delta_i}) s_{n,k}^{\delta_i} \rightarrow \tilde{b}_n - F'_{[n]}(\xi) (\sigma_{n,k} - \xi)$$

as $i \rightarrow \infty$ for all $k \leq k_n(\xi)$. Now we have to differ three cases:

Case 1: $1 \leq k_n(\xi) = k_{REG}(\xi)$. From definition (41),

$$\|\tilde{b}_n - F'_{[n]}(\xi) (\sigma_{n,k_n(\xi)} - \xi)\| < \mu \|\tilde{b}_n\|.$$

It follows from (51) that for i large enough

$$\|b_n^{\delta_i} - F'_{[n]}(x_n^{\delta_i}) s_{n,k_n(\xi)}^{\delta_i}\| < \mu \|b_n^{\delta_i}\|$$

which in view of (13) implies $k_n^{\delta_i} \leq k_n(\xi)$. Then $k_n^{\delta_i} \in \{0, \dots, k_n(\xi)\}$ and we conclude using (43) that $x_{n+1}^{\delta_i} = z_{n,k_n^{\delta_i}}^{\delta_i}$ splits in at most $k_n(\xi) + 1$ convergent subsequences, each one converging to an element of \mathcal{X}_{n+1} .

Case 2: $k_n(\xi) = k_{\max}$. In this case, $k_n^{\delta_i} \leq k_{\max} = k_n(\xi)$ and we proceed as in Case 1.

Case 3: $k_n(\xi) = 0$. Then $\tilde{b}_n = 0$, that is, $y_{[n]} = F_{[n]}(\xi)$ and $\xi \in \mathcal{X}_{n+1}$. We will prove that $x_{n+1}^{\delta_i} \rightarrow \xi$ as $i \rightarrow \infty$. Assume the contrary, then there exist an $\epsilon > 0$ and a subsequence $(\delta_{i_m})_{m \in \mathbb{N}}$ such that $\epsilon < \|\xi - x_{n+1}^{\delta_{i_m}}\|^s$ and using Assumption 2(c),

$$\epsilon \leq \frac{1}{C} \Delta_p(\xi, x_{n+1}^{\delta_{i_m}}) \stackrel{(27)}{\leq} \frac{1}{C} \Delta_p(\xi, x_n^{\delta_{i_m}}) \xrightarrow{i \rightarrow \infty} \frac{1}{C} \Delta_p(\xi, \xi) = 0$$

contradicting $\epsilon > 0$. □

In the second step towards establishing the regularization property we provide a kind of uniform convergence of the set sequence $(\mathcal{X}_n)_n$ to solutions of (1). For the rigorous formulation in Lemma 16 below we need to introduce further notation: Let $l \in \mathbb{N}$ and set $\xi_0^{(l)} := x_0$. Now define $\xi_{n+1}^{(l)} := \sigma_{n,k_n^{(l)}}(\xi_n^{(l)})$ by choosing $k_n^{(l)} \in \{1, \dots, k_n(\xi_n^{(l)})\}$ in case of $k_n(\xi_n^{(l)}) \geq 1$ and $k_n^{(l)} = 0$ otherwise. Then $\xi_{n+1}^{(l)}$ is a successor of $\xi_n^{(l)}$. Of course $\xi_n^{(l)} \in \mathcal{X}_n$ for all $n \in \mathbb{N}$ and reciprocally, each element in \mathcal{X}_n can be written as $\xi_n^{(l)}$ for some $l \in \mathbb{N}$.

Observe that $(\xi_n^{(l)})_{n \in \mathbb{N}}$ represents a sequence generated by K-REGINN-IT with the inner iteration stopped with an arbitrary stop index $k_n^{(l)}$ less than or equal $k_n(\xi_n^{(l)})$. Due to this fact, we call the sequence $(k_n^{(l)})_{n \in \mathbb{N}}$ a *stop rule*. Then each element of $(\xi_n^{(l)})_{n \in \mathbb{N}}$ satisfies

$$\|y_{[n]} - F_{[n]}(\xi_n^{(l)}) - F'_{[n]}(\xi_n^{(l)}) (\sigma_{n,k}(\xi_n^{(l)}) - \xi_n^{(l)})\| \geq \mu \|y_{[n]} - F_{[n]}(\xi_n^{(l)})\|, \quad k \leq k_n^{(l)} - 1,$$

see (41). Hence, Theorem 13 applies to $(\xi_n^{(l)})_{n \in \mathbb{N}}$, that is, the limit

$$(52) \quad x_\infty^{(l)} := \lim_{n \rightarrow \infty} \xi_n^{(l)}$$

exists and is a solution of (1) in $B_\rho(x^+, \Delta_p)$.

The following result is adapted from [6] and generalizes Proposition 19 in [12].

Lemma 16. *Let all assumptions of Theorem 13 hold true and let $(\xi_n^{(l)})_{n \in \mathbb{N}}$ denote the sequence generated by the stop rule $(k_n^{(l)})_{n \in \mathbb{N}}$. Then, for each $\epsilon > 0$ there exists an $M = M(\epsilon) \in \mathbb{N}$ such that*

$$\|\xi_n^{(l)} - x_\infty^{(l)}\| < \epsilon \text{ for all } n \geq M \text{ and all } l \in \mathbb{N}.$$

In particular, if x^+ is the unique solution of (1) in $B_\rho(x^+, \Delta_p)$ then $\|\xi_n^{(l)} - x^+\| < \epsilon$ for all $n \geq M$ and all $l \in \mathbb{N}$.

Proof. Assume the statement is not true. Then, there exist an $\epsilon > 0$ and sequences $(n_j)_j, (l_j)_j \subset \mathbb{N}$ with $(n_j)_j$ strictly increasing such that

$$\|\xi_{n_j}^{(l_j)} - x_\infty^{(l_j)}\| > \epsilon \text{ for all } j \in \mathbb{N}$$

where $(\xi_n^{(l_j)})_n$ represents the sequence generated by the stop rule $(k_n^{(l_j)})_n$. We stress the fact that the iterates $\xi_{n_j}^{(l_j)}$ must be generated by infinitely many different sequences of stop rules (otherwise, as $\xi_{n_j}^{(l)} \rightarrow x_\infty^{(l)}$ as $j \rightarrow \infty$ for each l and as the l_j 's attain only a finite number of values, we would have $\|\xi_{n_j}^{(l_j)} - x_\infty^{(l_j)}\| < \epsilon$ for n_j large enough). Next we reorder the numbers l_j (excluding some iterates if necessary) such that

$$(53) \quad \|\xi_{n_l}^{(l)} - x_\infty^{(l)}\| > \epsilon \text{ for all } l \in \mathbb{N}.$$

Now we construct inductively an auxiliary sequence $(\widehat{\xi}_n)_n$, which is generated by a stop rule $(\widehat{k}_n)_n$, as well as a sequence of unbounded sets $(\mathcal{L}_n)_n$ such that $\mathcal{L}_n \subset \mathbb{N} \setminus \{1, \dots, n\}$ with $\mathcal{L}_{n+1} \subset \mathcal{L}_n$, $n \in \mathbb{N}_0$, and

$$(54) \quad \widehat{k}_n = k_n^{(l)} \text{ for all } l \in \mathcal{L}_n, n \in \mathbb{N}_0.$$

Set $\widehat{\xi}_0 := x_0$. As $k_0(\widehat{\xi}_0) < \infty$ there exists $\widehat{k}_0 \in \{0, \dots, k_0(\widehat{\xi}_0)\}$ such that $\widehat{k}_0 = k_0^{(l)}$ for infinitely many $l \in \mathbb{N}$. Let $\mathcal{L}_0 \subset \mathbb{N}$ be the set of those indices l . Assume that $\widehat{\xi}_0, \dots, \widehat{\xi}_n; \widehat{k}_0, \dots, \widehat{k}_n$, and $\mathcal{L}_0, \dots, \mathcal{L}_n$ have been constructed with the requested properties. Then, define $\widehat{\xi}_{n+1} := \sigma_{n, \widehat{k}_n}(\widehat{\xi}_n)$, see Definition 14. Since $k_{n+1}(\widehat{\xi}_{n+1}) < \infty$ we find a $\widehat{k}_{n+1} \in \{0, \dots, k_{n+1}(\widehat{\xi}_{n+1})\}$ such that $\widehat{k}_{n+1} = k_{n+1}^{(l)}$ for infinitely many $l \in \mathcal{L}_n \setminus \{1, \dots, n+1\}$. Those l 's are collected in $\mathcal{L}_{n+1} \subset \mathcal{L}_n$ and the inductive construction of $(\widehat{\xi}_n)_n$ is complete.

In view of (52) the limit $\widehat{x}_\infty := \lim_{n \rightarrow \infty} \widehat{\xi}_n$ exists in $B_\rho(x^+, \Delta_p)$ and solves (1). Observe that, if $l \in \mathcal{L}_0$, then $\xi_1^{(l)} = \sigma_{0, k_0^{(l)}}(\xi_0^{(l)}) \stackrel{(54)}{=} \sigma_{0, \widehat{k}_0}(\xi_0^{(l)}) = \sigma_{0, \widehat{k}_0}(\widehat{\xi}_0) = \widehat{\xi}_1$. By induction,

$$(55) \quad l \in \mathcal{L}_n \implies \xi_{n+1}^{(l)} = \widehat{\xi}_{n+1} \text{ for all } n \in \mathbb{N}_0.$$

Since the "diagonal" sequence $\widehat{\xi}_n$ converges to \widehat{x}_∞ as $n \rightarrow \infty$, there exists $M = M(\epsilon) \in \mathbb{N}$ such that

$$(56) \quad \Delta_p(\widehat{x}_\infty, \widehat{\xi}_n) < \frac{C\epsilon^s}{2^{s+1}} \text{ for all } n > M,$$

where $C > 0$ is the constant in Assumption 2(c). We can additionally suppose that $\widehat{\xi}_n \in B_\rho(x^+, \Delta_p)$ for all $n > M$. In fact, as $\lim_{n \rightarrow \infty} \widehat{\xi}_n = \widehat{x}_\infty$ and the mappings J_p and $\Delta_p(\widehat{x}_\infty, \cdot)$ are continuous, we have that $\Delta_p(\widehat{x}_\infty, \widehat{\xi}_n)$ and $\langle J_p(\widehat{\xi}_n) - J_p(\widehat{x}_\infty), \widehat{x}_\infty - x^+ \rangle$ converge to zero as $n \rightarrow \infty$. From (6),

$$\Delta_p(x^+, \widehat{\xi}_n) = \Delta_p(\widehat{x}_\infty, \widehat{\xi}_n) + \Delta_p(x^+, \widehat{x}_\infty) + \langle J_p(\widehat{\xi}_n) - J_p(\widehat{x}_\infty), \widehat{x}_\infty - x^+ \rangle$$

and as $\Delta_p(x^+, \widehat{x}_\infty) < \rho$, we conclude that $\Delta_p(x^+, \widehat{\xi}_n) < \rho$ for n large enough.

Now, for $l_0 \in \mathcal{L}_M$ fixed,

$$\Delta_p(\widehat{x}_\infty, \xi_{M+1}^{(l_0)}) \stackrel{(55)}{=} \Delta_p(\widehat{x}_\infty, \widehat{\xi}_{M+1}) \stackrel{(56)}{<} \frac{C\epsilon^s}{2^{s+1}}.$$

As \widehat{x}_∞ is a solution of (1) and $\xi_{M+1}^{(l_0)} = \widehat{\xi}_{M+1} \in B_\rho(x^+, \Delta_p)$, inequality (17) applies and the errors $\Delta_p(\widehat{x}_\infty, \xi_n^{(l_0)})$ are monotonically decreasing in n for all $n \geq M+1$. In particular, $n_{l_0} \geq l_0 \geq M+1$ (because $l_0 \in \mathcal{L}_M \subset \mathbb{N} \setminus \{1, \dots, M\}$). Then

$$\Delta_p(\widehat{x}_\infty, \xi_{n_{l_0}}^{(l_0)}) \leq \Delta_p(\widehat{x}_\infty, \xi_{M+1}^{(l_0)}) < \frac{C\epsilon^s}{2^{s+1}}.$$

As $\xi_n^{(l_0)} \rightarrow x_\infty^{(l_0)}$ as $n \rightarrow \infty$, we conclude that

$$\Delta_p(\widehat{x}_\infty, x_\infty^{(l_0)}) = \lim_{n \rightarrow \infty} \Delta_p(\widehat{x}_\infty, \xi_n^{(l_0)}) \leq \frac{C\epsilon^s}{2^{s+1}}.$$

From the inequality in Assumption 2(c),

$$\begin{aligned} \|\xi_{n_{l_0}}^{(l_0)} - x_\infty^{(l_0)}\|^s &\leq 2^s \left(\|\xi_{n_{l_0}}^{(l_0)} - \widehat{x}_\infty\|^s + \|\widehat{x}_\infty - x_\infty^{(l_0)}\|^s \right) \\ &\leq \frac{2^s}{C} \left(\Delta_p(\widehat{x}_\infty, \xi_{n_{l_0}}^{(l_0)}) + \Delta_p(\widehat{x}_\infty, x_\infty^{(l_0)}) \right) < \epsilon^s, \end{aligned}$$

contradicting (53). \square

We are now well prepared to prove our main result.

Theorem 17 (Regularization property). *Let all assumptions of Theorem 6 hold true but with $k_{\max} < \infty$ for $d > 1$. If $\delta_i \rightarrow 0$ as $i \rightarrow \infty$, then the sequence $(x_{N(\delta_i)}^{\delta_i})_{i \in \mathbb{N}}$ splits in convergent subsequences, all of which converge strongly to solutions of (1) as $i \rightarrow \infty$. If x^+ is the unique solution of (1) in $B_\rho(x^+, \Delta_p)$ then*

$$\lim_{i \rightarrow \infty} \|x_{N(\delta_i)}^{\delta_i} - x^+\| = 0.$$

Proof. If $N(\delta_i) \leq I$ for some $I \in \mathbb{N}$ as $i \rightarrow \infty$ then the sequence $(x_{N(\delta_i)}^{\delta_i})_{i \in \mathbb{N}}$ splits in subsequences of the form $(x_n^{\delta_i})_{j \in \mathbb{N}}$ where n is an iteration index less than or equal to I . According to Lemma 15, each of these subsequences splits into convergent subsequences. Hence each limit of such a subsequence must be a solution of (1) due to the discrepancy principle (14). In fact, if $x_n^{\delta_i} \rightarrow a$ as $i \rightarrow \infty$, then using (9),

$$\|y_j - F_j(a)\| = \lim_{i \rightarrow \infty} \|y_j - F_j(x_n^{\delta_i})\| \leq \lim_{i \rightarrow \infty} (\tau + 1) \delta_i = 0, \quad j = 0, \dots, d-1.$$

Suppose now that $N(\delta_i) \rightarrow \infty$ as $i \rightarrow \infty$ and let $\epsilon > 0$ be given. As the Bregman distance is a continuous function in both arguments, there exists $\gamma = \gamma(\epsilon) > 0$ such that

$$(57) \quad \Delta_p(x, x_n^{\delta_i}) < C\epsilon^s \text{ whenever } \|x - x_n^{\delta_i}\| \leq \gamma$$

where $C > 0$ is the constant appearing in Assumption 2(c). From Lemma 16, there is an $M \in \mathbb{N}$ such that, for each $\xi_M^{(l)} \in \mathcal{X}_M$, there exists a solution $x_\infty^{(l)}$ of (1) satisfying

$$\|x_\infty^{(l)} - \xi_M^{(l)}\| < \frac{\gamma}{2}.$$

According to Lemma 15, the sequence $x_M^{\delta_i}$ splits into convergent subsequences, each one converging to an element of \mathcal{X}_M . Let $(x_M^{\delta_{i_j}})_{j \in \mathbb{N}}$ be a generic convergent subsequence, which converges to an element of \mathcal{X}_M , say

$$\lim_{j \rightarrow \infty} x_M^{\delta_{i_j}} = \xi_M^{(l_0)} \in \mathcal{X}_M.$$

We will prove that the subsequence $(x_{N(\delta_{i_j})}^{\delta_{i_j}})_{j \in \mathbb{N}}$ converges to the solution $x_\infty^{(l_0)}$. In fact, since $x_M^{\delta_{i_j}} \rightarrow \xi_M^{(l_0)}$ as $j \rightarrow \infty$, there exists a $J_1 = J_1(\epsilon)$ such that

$$\|\xi_M^{(l_0)} - x_M^{\delta_{i_j}}\| < \frac{\gamma}{2} \text{ for all } j \geq J_1.$$

As $N(\delta_{i_j}) \rightarrow \infty$ as $j \rightarrow \infty$, we have $N(\delta_{i_j}) \geq M$ for all $j \geq J$ where $J \geq J_1$ is a sufficiently large number. Then, for all $j \geq J$,

$$\|x_\infty^{(l_0)} - x_M^{\delta_{i_j}}\| \leq \|x_\infty^{(l_0)} - \xi_M^{(l_0)}\| + \|\xi_M^{(l_0)} - x_M^{\delta_{i_j}}\| \leq \gamma.$$

Finally, (57) and Assumption 2(c) lead to

$$\|x_{N(\delta_{i_j})}^{\delta_{i_j}} - x_\infty^{(l_0)}\|^s \leq \frac{1}{C} \Delta_p \left(x_\infty^{(l_0)}, x_{N(\delta_{i_j})}^{\delta_{i_j}} \right) \stackrel{(17)}{\leq} \frac{1}{C} \Delta_p \left(x_\infty^{(l_0)}, x_M^{\delta_{i_j}} \right) \leq \epsilon^s.$$

□

We like to emphasize that the regularization property of K-REGINN-IT holds without any additional assumption on Y other than it is a general Banach space.

REFERENCES

- [1] Thomas Bonesky, Kamil S. Kazimierski, Peter Maass, Frank Schöpfer, and Thomas Schuster. Minimization of Tikhonov functionals in Banach spaces. *Abstr. Appl. Anal.*, pages Art. ID 192679, 19, 2008.
- [2] Martin Burger and Barbara Kaltenbacher. Regularizing Newton-Kaczmarz methods for nonlinear ill-posed problems. *SIAM J. Numer. Anal.*, 44(1):153–182 (electronic), 2006.
- [3] Ioana Cioranescu. *Geometry of Banach spaces, duality mappings and nonlinear problems*, volume 62 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1990.
- [4] Markus Haltmeier, Antonio Leitão, and Otmar Scherzer. Kaczmarz methods for regularizing nonlinear ill-posed equations. I. Convergence analysis. *Inverse Probl. Imaging*, 1(2):289–298, 2007.
- [5] Martin Hanke. A regularizing Levenberg-Marquardt scheme, with applications to inverse ground-water filtration problems. *Inverse Problems*, 13(1):79–95, 1997.
- [6] Martin Hanke. Regularizing properties of a truncated Newton-CG algorithm for nonlinear inverse problems. *Numer. Funct. Anal. Optim.*, 18(9-10):971–993, 1997.
- [7] Qinian Jin and Linda Stals. Nonstationary iterated Tikhonov regularization for ill-posed problems in Banach spaces. *Inverse Problems*, 28:104011, 15, 2012.
- [8] Qinian Jin and Wei Wang. Landweber iteration of Kaczmarz type with general non-smooth convex penalty functionals. *Inverse Problems*, 29(8):085011, 22, 2013.
- [9] Richard Kowar and Otmar Scherzer. Convergence analysis of a Landweber-Kaczmarz method for solving nonlinear ill-posed problems. In *Ill-posed and inverse problems*, pages 253–270. VSP, Zeist, 2002.
- [10] Armin Lechleiter and Andreas Rieder. Towards a general convergence theory for inexact Newton regularizations. *Numer. Math.*, 114(3):521–548, 2010.
- [11] Antonio Leitão and M. Marques Alves. On Landweber-Kaczmarz methods for regularizing systems of ill-posed equations in Banach spaces. *Inverse Problems*, 28(10):104008, 15, 2012.
- [12] Fabio Margotti, Andreas Rieder, and Antonio Leitao. A Kaczmarz version of the inexact Newton regularization REGINN-Landweber in Banach spaces. *SIAM J. Numer. Anal.*, 2014. To appear.

- [13] Andreas Rieder. On the regularization of nonlinear ill-posed problems via inexact Newton iterations. *Inverse Problems*, 15(1):309–327, 1999.
- [14] Frank Schöpfer, Alfred K. Louis, and Thomas Schuster. Nonlinear iterative methods for linear ill-posed problems in Banach spaces. *Inverse Problems*, 22(1):311–329, 2006.
- [15] Thomas Schuster, Barbara Kaltenbacher, Bernd Hofmann, and Kamil S. Kazimierski. *Regularization methods in Banach spaces*, volume 10 of *Radon Series on Computational and Applied Mathematics*. Walter de Gruyter GmbH & Co. KG, Berlin, 2012.
- [16] Zong Ben Xu and G. F. Roach. Characteristic inequalities of uniformly convex and uniformly smooth Banach spaces. *J. Math. Anal. Appl.*, 157(1):189–210, 1991.

DEPARTMENT OF MATHEMATICS, KARLSRUHE INSTITUTE OF TECHNOLOGY (KIT), D-76128 KARLSRUHE, GERMANY

E-mail address: `fabio.margotti@partner.kit.edu`

E-mail address: `andreas.rieder@kit.edu`

IWRMM-Preprints seit 2012

- Nr. 12/01 Branimir Anic, Christopher A. Beattie, Serkan Gugercin, Athanasios C. Antoulas: Interpolatory Weighted-H2 Model Reduction
- Nr. 12/02 Christian Wieners, Jiping Xin: Boundary Element Approximation for Maxwell's Eigenvalue Problem
- Nr. 12/03 Thomas Schuster, Andreas Rieder, Frank Schöpfer: The Approximate Inverse in Action IV: Semi-Discrete Equations in a Banach Space Setting
- Nr. 12/04 Markus Bürg: Convergence of an hp-Adaptive Finite Element Strategy for Maxwell's Equations
- Nr. 12/05 David Cohen, Stig Larsson, Magdalena Sigg: A Trigonometric Method for the Linear Stochastic Wave Equation
- Nr. 12/06 Tim Kreuzmann, Andreas Rieder: Geometric Reconstruction in Bioluminescence Tomography
- Nr. 12/07 Tobias Jahnke, Michael Kreim: Error bound for piecewise deterministic processes modeling stochastic reaction systems
- Nr. 12/08 Haojun Li, Kirankumar Hiremath, Andreas Rieder, Wolfgang Freude: Adaptive Wavelet Collocation Method for Simulation of Time Dependent Maxwell's Equations
- Nr. 12/09 Andreas Arnold, Tobias Jahnke: On the approximation of high-dimensional differential equations in the hierarchical Tucker format
- Nr. 12/10 Mike A. Botchev, Volker Grimm, Marlis Hochbruck: Residual, Restarting and Richardson Iteration for the Matrix Exponential
- Nr. 13/01 Willy Dörfler, Stefan Findeisen: Numerical Optimization of a Waveguide Transition Using Finite Element Beam Propagation
- Nr. 13/02 Fabio Margotti, Andreas Rieder, Antonio Leitao: A Kaczmarz Version of the Reginn-Landweber Iteration for Ill-Posed Problems in Banach Spaces
- Nr. 13/03 Andreas Kirsch, Andreas Rieder: On the Linearization of Operators Related to the Full Waveform Inversion in Seismology
- Nr. 14/01 Robert Winkler, Andreas Rieder: Resolution-Controlled Conductivity Discretization in Electrical Impedance Tomography
- Nr. 14/02 Andreas Kirsch, Andreas Rieder: Seismic Tomography is Locally Ill-Posed
- Nr. 14/03 Fabio Margotti, Andreas Rieder: An Inexact Newton Regularization in Banach Spaces based on the Nonstationary Iterated Tikhonov Method

Eine aktuelle Liste aller IWRMM-Preprints finden Sie auf:

www.math.kit.edu/iwrmm/seite/preprints

Kontakt

Karlsruher Institut für Technologie (KIT)
Institut für Wissenschaftliches Rechnen
und Mathematische Modellbildung

Prof. Dr. Christian Wieners
Geschäftsführender Direktor

Campus Süd
Engesserstr. 6
76131 Karlsruhe

E-Mail: Bettina.Haindl@kit.edu

www.math.kit.edu/iwrmm/

Herausgeber

Karlsruher Institut für Technologie (KIT)
Kaiserstraße 12 | 76131 Karlsruhe

Mai 2014

www.kit.edu