ITERATED LAVRENTIEV REGULARIZATION FOR NONLINEAR ILL-POSED PROBLEMS

P. MAHALE[1] and M. T. NAIR[✉][1]

Abstract

We consider an iterated form of Lavrentiev regularization, using a null sequence (α_k) of positive real numbers to obtain a stable approximate solution for ill-posed nonlinear equations of the form $F(x) = y$, where $F : D(F) \subseteq X \to X$ is a nonlinear operator and X is a Hilbert space. Recently, Bakushinsky and Smirnova ["Iterative regularization and generalized discrepancy principle for monotone operator equations", *Numer. Funct. Anal. Optim.* **28** (2007) 13–25] considered an *a posteriori* strategy to find a stopping index k_δ corresponding to inexact data y^δ with $\|y - y^\delta\| \leq \delta$ resulting in the convergence of the method as $\delta \to 0$. However, they provided no error estimates. We consider an alternate strategy to find a stopping index which not only leads to the convergence of the method, but also provides an order optimal error estimate under a general source condition. Moreover, the condition that we impose on (α_k) is weaker than that considered by Bakushinsky and Smirnova.

1. Introduction

In this paper we are interested in obtaining a stable approximate solution for a nonlinear ill-posed operator equation of the form

$$F(x) = y, \tag{1.1}$$

where $F : D(F) \subset X \to X$ is a monotone operator and X is a Hilbert space. We shall denote the inner product and the corresponding norm on a Hilbert space by $\langle \cdot, \cdot \rangle$ and $\| \cdot \|$, respectively. We recall that F is a monotone operator if it satisfies the relation

$$\langle F(x) - F(y), x - y \rangle \geq 0 \quad \forall x, y \in D(F). \tag{1.2}$$

[1]Department of Mathematics, I.I.T. Madras, Chennai - 600036, India; e-mail: pallavimahale@iitm.ac.in, mtnair@iitm.ac.in.

We assume that (1.1) has a solution, namely $x^\dagger$, and, for $\delta > 0$, $y^\delta \in X$ is the available noisy data in place of y such that

$$\|y - y^\delta\| \le \delta. \tag{1.3}$$

As the given operator equation is ill-posed, its solution need not depend continuously on the data, that is, small perturbations in the data can cause large deviations in the solutions. In order to obtain stable approximate solutions, regularization methods are often used. Tikhonov regularization is one of the classical regularization methods used in the literature (see [4, 6, 7, 11, 12, 15]). Since F is monotone, one can also use the Lavrentiev regularization method (see [16, 17]). In Lavrentiev regularization, the approximate solution is obtained as a solution of the equation

$$F(x) - y^\delta + \alpha(x - x_0) = 0, \tag{1.4}$$

where $\alpha \ge 0$ is called the *regularization parameter* and x_0 is an initial guess for the solution $x^\dagger$. We also assume that F is Fréchet differentiable in an appropriate neighbourhood of $x^\dagger$. From the monotonicity and the Fréchet differentiability of F, it follows that (1.4) has a unique solution (see [5, Theorem 11.2]). Also, from the Fréchet differentiability of F, we note that Equation (1.4) can be reformulated as

$$x = x_0 + (A_x + \alpha I)^{-1}[y^\delta - F(x) + A_x(x - x_0)],$$

where $A_x := F'(x)$ is the Fréchet derivative of F. This motivates us to define the iterations

$$x_{k+1}^\delta = x_0 + (A_k^\delta + \alpha_k I)^{-1}[y^\delta - F(x_k^\delta) + A_x^\delta(x_k^\delta - x_0)] \tag{1.5}$$

or equivalently

$$x_{k+1}^\delta = x_k - (A_k^\delta + \alpha_k I)^{-1}[F(x_k^\delta) - y^\delta + \alpha_k(x_k^\delta - x_0)] \quad \text{for } k = 0, 1, 2, \ldots, \tag{1.6}$$

where $A_k^\delta := F'(x_k^\delta)$ and (α_k) is a sequence of positive real numbers such that $\lim_{k\to\infty} \alpha_k = 0$. It is important to stop the iterations at an appropriate step, say for $k = k_\delta$, and show that x_k is well defined for $0 \le k \le k_\delta$ and $x_{k_\delta}^\delta \to x^\dagger$ as $\delta \to 0$. It is also desirable to see whether the choice k_δ leads to an estimate for $\|x^\dagger - x_{k_\delta}^\delta\|$ which is order optimal with respect to some source set.

The iteration given in (1.6) has been considered by Bakushinsky and Smirnova [3]. They considered a generalized discrepancy principle as in [1, 2] to choose the stopping index k_δ, by requiring it to satisfy

$$\|F(x_{k_\delta}) - y^\delta\|^2 \le \tau\delta < \|F(x_k) - y^\delta\|^2 \quad \text{where } 0 \le k < k_\delta,$$

for some appropriate $\tau > 1$, and they showed that $x_{k_\delta}^\delta \to x^\dagger$ as $\delta \to 0$ under the following assumptions.

(i) There exists $L > 0$ such that $\|F'(x) - F'(y)\| \le L\|x - y\|$ for all $x, y \in D(F)$.
(ii) There exists $p > 0$ such that

$$\frac{\alpha_k - \alpha_{k+1}}{\alpha_k \alpha_{k+1}} \le p \quad \forall k \in \mathbb{N}. \tag{1.7}$$

(iii) $\sqrt{(2 + L\sigma)\|x_0 - x^\dagger\| t d} \leq \sigma - 2\|x_0 - x^\dagger\| t \leq d\alpha_0$, where

$$\sigma := (\sqrt{\tau} - 1)^2, \quad t := p\alpha_0 + 1 \quad \text{and} \quad d = 2(t\,\|x_0 - x^\dagger\| + p\sigma).$$

However, no error estimate was given in [3].

In this paper, motivated by the work of Qi-Nian Jin [10] for an *iteratively regularized Gauss–Newton method*, we consider an alternate stopping criterion which not only ensures the convergence, but also derives an order optimal error estimate under a general source condition on $x^\dagger - x_0$. Moreover, the condition that we impose on (α_k) is weaker than (1.7).

In Section 2, we consider some basic assumptions required throughout the paper. Section 3 deals with the stopping rule and a result that establishes the existence of the stopping index. In Section 4 we prove results for the iterations based on the exact data, and in Section 5 the error analysis for the noisy data case is proved. The main order optimal result using the *a posteriori* stopping rule is proved in Section 6.

2. Basic assumptions and some preliminary results

The results in this paper are derived by making the following assumptions. The first assumption is on the differentiability of F and a certain property that $F'(x)$ must satisfy.

ASSUMPTION 2.1.

(i) There exists $r > 0$ such that $B_r(x^\dagger) \subseteq D(F)$ and F is Fréchet differentiable at all $x \in B_r(x^\dagger)$, where

$$B_r(x^\dagger) := \{x \in X : \|x - x^\dagger\| < r\}.$$

(ii) There exists a constant $k_0 > 0$ such, that for every $x, u \in B_r(x^\dagger)$ and $v \in X$, there exists an element, say $\Phi(x, u, v) \in X$, satisfying

$$[F'(x) - F'(u)]v = F'(u)\Phi(x, u, v), \quad \|\Phi(x, u, v)\| \leq k_0 \|v\| \|x - u\|$$

for all $x, u \in B_r(x^\dagger)$ and $v \in X$.

The following two propositions are based on Assumption 2.1 and will be used in due course.

PROPOSITION 2.2. *For* $x, u \in B_r(x^\dagger)$ *and* $\alpha > 0$,

$$\|(F'(u) + \alpha I)^{-1}[F(x) - F(u) - F'(u)(x - u)]\| \leq \frac{k_0}{2}\|x - u\|^2.$$

PROOF. Using the fundamental theorem of integration, for $x, u \in B_r(x^\dagger)$,

$$F(x) - F(u) = \int_0^1 F'(u + t(x - u))(x - u)\, dt,$$

so that by Assumption 2.1(ii),

$$
\begin{aligned}
F(x) - F(u) - F'(u)(x - u) &= \int_0^1 [F'(u + t(x - u)) - F'(u)](x - u)\, dt \\
&= \int_0^1 F'(u)\Phi(u + t(x - u), u, x - u)\, dt \\
&= F'(u) \int_0^1 \Phi(u + t(x - u), u, x - u)\, dt.
\end{aligned}
$$

Again, by Assumption 2.1(ii) and the inequality $\|(F'(u) + \alpha I)^{-1} F'(u)\| \le 1$,

$$
\begin{aligned}
\|(F'(u) &+ \alpha I)^{-1}[F(x) - F(u) - F'(u)(x - u)]\| \\
&\le \int_0^1 \|\Phi(u + t(x - u), u, x - u)\|\, dt \\
&\le \int_0^1 k_0 \|x - u\|^2 t\, dt \le \frac{k_0}{2} \|x - u\|^2.
\end{aligned}
$$

This completes the proof. □

PROPOSITION 2.3. *For $x, u \in B_r(x^\dagger)$ and $\alpha > 0$,*

$$
\alpha \big\| \big[(F'(x) + \alpha I)^{-1} - (F'(u) + \alpha I)^{-1} \big] \big\| \le k_0 \|x - u\|. \tag{2.1}
$$

PROOF. Denoting $T_{x,u} = \alpha((F'(x) + \alpha I)^{-1} - (\alpha I + F'(u))^{-1})$, for every $v \in X$ we have

$$
T_{x,u} = \alpha(F'(x) + \alpha I)^{-1}(F'(u) - F'(x))(\alpha I + F'(u))^{-1} v.
$$

Hence, using Assumption 2.1(ii) and the relations $\|(F'(x) + \alpha I)^{-1} F'(x)\| \le 1$ and $\|(F'(u) + \alpha I)^{-1}\| \le 1/a$, we obtain

$$
\begin{aligned}
\|T_{x,u} v\| &= \Big\| (F'(x) + \alpha I)^{-1} F'(x) \Phi\Big(u, x, \alpha(\alpha I + F'(u))^{-1} v\Big) \Big\| \\
&\le k_0 \|x - u\| \|v\| \quad \forall v \in X.
\end{aligned}
$$

From this (2.1) follows. □

The next two assumptions concern a source condition based on a *source function φ* and a property that φ must satisfy.

ASSUMPTION 2.4. There exists a continuous, strictly monotonically increasing function $\varphi : (0, a] \to (0, \infty)$ with $a \ge \|F'(x^\dagger)\|$ satisfying $\lim_{\lambda \to 0} \varphi(\lambda) = 0$ and $\rho > 0$, $v \in X$ with $\|v\| \le \rho$ such that $x_0 - x^\dagger = \varphi(F'(x^\dagger))v$.

ASSUMPTION 2.5. The function φ in Assumption 2.4 satisfies

$$
\sup_{\lambda \ge 0} \frac{\alpha \varphi(\lambda)}{\lambda + \alpha} \le c_\varphi \varphi(\alpha) \quad \forall \alpha \in (0, a].
$$

Next we assume a condition on the sequence (α_k) considered in (1.5).

ASSUMPTION 2.6. The sequence (α_k) of positive real numbers is such that $\lim_{k \to \infty} \alpha_k = 0$ and there exists $\mu > 1$ such that

$$1 \leq \frac{\alpha_k}{\alpha_{k+1}} \leq \mu \quad \forall k \in \mathbb{N}. \tag{2.2}$$

We note that the condition (2.2) on (α_k) is weaker than (1.7) considered by Bakushinsky and Smirnova [3]. In fact, if (1.7) is satisfied, then it also satisfies (2.2) with $\mu = p\alpha_0 + 1$, but the converse need not be true. This is seen by taking $\alpha_k = \{1/k^2\}$ or $\alpha_k = \{e^{-k}\}$. Note that, for these choices of (α_k), α_k/α_{k+1} is bounded whereas $(\alpha_k - \alpha_{k+1})/\alpha_k\alpha_{k+1} \to \infty$ as $k \to \infty$.

Assumption 2.1(ii) is used in the literature for regularization of many nonlinear ill-posed problems (see [10, 11, 16–18]).

Assumption 2.4 is known as a general source condition, and it is similar to that considered in [14] by Nair and Tautenhahn for the linear case and it is also analogous to the source condition considered recently by the current authors [13]. It includes both the well-known source conditions, namely, the Hölder-type source condition, that is, $\varphi(\lambda) = \lambda^\mu$, and the logarithmic source condition, that is, $\varphi(\lambda) = [\log(1/\lambda)]^{-\mu}$. Assumption 2.5 shows how the source condition should be related to the Lavrentiev regularization.

3. Stopping rule

Let $\delta > 0$ be given and let $y^\delta \in Y$ be such that (1.3) is satisfied.

THE STOPPING RULE. Given $c_0 > 4$, choose k_δ to be the first nonnegative integer such that x_k^δ in (1.5) is defined for each $k \in \{0, 1, \ldots, k_\delta\}$ and

$$\|\alpha_{k_\delta}(A_{k_\delta}^\delta + \alpha_{k_\delta}I)^{-1}(F(x_{k_\delta}^\delta) - y^\delta)\| \leq c_0\delta \quad \text{for } c_0 > 4. \tag{3.1}$$

In the following we establish the existence of such a k_δ. First, we consider the positive integer $N \in \mathbb{N}$ satisfying

$$\alpha_N \leq \frac{(c-1)\delta}{\|x_0 - x^\dagger\|} < \alpha_k \quad \forall k \in \{0, 1, \ldots, N-1\}, \tag{3.2}$$

where $c > 1$ and $\alpha_0 > (c-1)\delta/\|x_0 - x^\dagger\|$.

The following technical lemma will be used to prove some of the results of subsequent sections.

LEMMA 3.1. Let $a > 0$ and let $b \geq 0$ be such that $4ab \leq 1$, and let $\theta := (1 - \sqrt{1 - 4ab})/2a$. Let $\theta_1, \ldots, \theta_n$ be nonnegative real numbers such that $\theta_{k+1} \leq a\theta_k^2 + b$ and $\theta_0 \leq \theta$. Then $\theta_k \leq \theta$ for all $k = 1, \ldots, n$.

PROOF. We note that $\theta := (1 - \sqrt{1 - 4ab})/2a$ is one of the two (real) roots of the equation $at^2 + b = t$. Given that $\theta_0 \leq \theta$, now let us assume that $\theta_k \leq \theta$ for some $k \in \mathbb{N} \cup \{0\}$. Hence,

$$\theta_{k+1} \leq a\theta_k^2 + b \leq a\theta^2 + b = \theta.$$

Thus, $\theta_k \leq \theta, k = 1, \ldots, n$. □

The following theorem establishes the existence of the stopping index k_δ as suggested in (3.1).

THEOREM 3.2. *Let* (1.2), (1.3), (2.2) *and Assumption* 2.1 *be satisfied. Let N be as in* (3.2) *for some $c > 1$ and $2ck_0\|x_0 - x^\dagger\|/(c - 1) \le 1$. Then x_k^δ is defined iteratively for each $k \in \{0, 1, \ldots, N\}$, and*

$$\|x_k^\delta - x^\dagger\| \le \frac{2c\|x_0 - x^\dagger\|}{c - 1} \quad \forall k \in \{0, 1, \ldots, N\}. \tag{3.3}$$

In particular, if $r > 2c\|x_0 - x^\dagger\|/(c - 1)$, then $x_k^\delta \in B_r(x^\dagger)$ for $k \in \{0, 1, \ldots, N\}$. Moreover,

$$\|\alpha_N(A_N^\delta + \alpha_N I)^{-1}(F(x_N^\delta) - y^\delta)\| \le c_0\delta \quad for\ c_0 := 3c + 1. \tag{3.4}$$

PROOF. We will show (3.3) by induction. It is obvious that (3.3) holds for $k = 0$. Now, assume that (3.3) holds for some $k \in \{0, 1, \ldots, N\}$. From (1.6),

$$
\begin{aligned}
x_{k+1}^\delta - x^\dagger &= x_k^\delta - x^\dagger - (A_k^\delta + \alpha_k I)^{-1}\{F(x_k^\delta) - y^\delta + \alpha_k(x_k^\delta - x_0)\} \\
&= (A_k^\delta + \alpha_k I)^{-1}\Big((A_k^\delta + \alpha_k I)(x_k^\delta - x^\dagger) - \{F(x_k^\delta) - y^\delta + \alpha_k(x_k^\delta - x_0)\}\Big) \\
&= (A_k^\delta + \alpha_k I)^{-1}\Big(A_k^\delta(x_k^\delta - x^\dagger) + y^\delta - F(x_k^\delta) + \alpha_k(x_0 - x^\dagger)\Big) \\
&= \alpha_k(A_k^\delta + \alpha_k I)^{-1}(x_0 - x^\dagger) + (A_k^\delta + \alpha_k)^{-1}(y^\delta - y) \\
&\quad + (A_k^\delta + \alpha_k I)^{-1}\{F(x^\dagger) - F(x_k^\delta) + A_k^\delta(x_k^\delta - x^\dagger)\}. \tag{3.5}
\end{aligned}
$$

Using (1.3), the estimates $\|(A_k^\delta + \alpha_k)^{-1}\| \le 1/\alpha_k$ and $\|(A_k^\delta + \alpha_k)^{-1}A_k^\delta\| \le 1$, and Proposition 2.2,

$$\|\alpha_k(A_k^\delta + \alpha_k I)^{-1}(x_0 - x^\dagger) + (A_k^\delta + \alpha_k)^{-1}(y^\delta - y)\| \le \|x_0 - x^\dagger\| + \frac{\delta}{\alpha_k}$$

and

$$\|(A_k^\delta + \alpha_k I)^{-1}\{F(x^\dagger) - F(x_k^\delta) + A_k^\delta(x_k^\delta - x^\dagger)\}\| \le \frac{k_0}{2}\|x_k^\delta - x^\dagger\|^2.$$

Thus,

$$\|x_{k+1}^\delta - x^\dagger\| \le \|x_0 - x^\dagger\| + \frac{\delta}{\alpha_k} + \frac{k_0}{2}\|x_k^\delta - x^\dagger\|^2.$$

But, by (3.2), $\delta/\alpha_k \le \|x_0 - x^\dagger\|/(c - 1)$. Hence,

$$\|x_{k+1}^\delta - x^\dagger\| \le \frac{c\|x_0 - x^\dagger\|}{c - 1} + \frac{k_0}{2}\|x_k^\delta - x^\dagger\|^2$$

which leads to the recurrence relation

$$\theta_{k+1} \le a\theta_k^2 + b,$$

where

$$\theta_k = \|x_k^\delta - x^\dagger\|, \quad a = \frac{k_0}{2} \quad \text{and} \quad b = \frac{c\|x_0 - x^\dagger\|}{c - 1}.$$

From the hypothesis of the theorem, we have $4ab = 2ck_0\|x_0 - x^\dagger\|/(c - 1) < 1$. It is obvious that

$$\theta_0 \le \|x_0 - x^\dagger\| \le \theta := \frac{1 - \sqrt{1 - 4ab}}{2a} = \frac{2b}{1 + \sqrt{1 - 4ab}} \le 2b = \frac{2c\|x_0 - x^\dagger\|}{c - 1}.$$

Hence, by Lemma 3.1, we get

$$\|x_k^\delta - x^\dagger\| \le \theta \le \frac{2c\|x_0 - x^\dagger\|}{c - 1} \quad \text{for } k \in \{0, 1, \dots, N\}. \tag{3.6}$$

In particular, if $r > 2c\|x_0 - x^\dagger\|/(c - 1)$, then we have that $x_k^\delta \in B_r(x^\dagger)$ for all $k \in \{0, 1, 2, \dots, N\}$.

Next, let $\gamma = \|\alpha_N(A_N^\delta + \alpha_N I)^{-1}(F(x_N^\delta) - y^\delta)\|$. Then, using the estimates

$$\|\alpha_N(A_N^\delta + \alpha_N I)^{-1}\| \le 1, \quad \|\alpha_N(A_N^\delta + \alpha_N I)^{-1}A_N^\delta\| \le \alpha_k$$

and Proposition 2.2,

$$\gamma \le \delta + \left\| \alpha_N(A_N^\delta + \alpha_N I)^{-1}\left(F(x_N^\delta) - y + A_N^\delta(x_N^\delta - x^\dagger) - A_N^\delta(x_N^\delta - x^\dagger)\right) \right\|$$

$$= \delta + \left\| \alpha_N(A_N^\delta + \alpha_N I)^{-1}[F(x_N^\delta) - F(x^\dagger) - A_N^\delta(x_N^\delta - x^\dagger) + A_N^\delta(x_N^\delta - x^\dagger)] \right\|$$

$$\le \delta + \alpha_N\left\{ k_0\frac{\|x_N^\delta - x^\dagger\|^2}{2} + \|x_N^\delta - x^\dagger\| \right\}$$

$$\le \delta + \alpha_N\|x_N^\delta - x^\dagger\|\left\{ 1 + k_0\frac{\|x_N^\delta - x^\dagger\|}{2} \right\}$$

$$\le \delta + \frac{2\alpha_N\|x_0 - x^\dagger\|c}{c - 1}\left[1 + \frac{k_0\|x_0 - x^\dagger\|c}{c - 1} \right] \le \delta + 2c\delta\left[1 + \frac{1}{2} \right] \le (3c + 1)\delta.$$

So, we have $\|\alpha_N(A_N^\delta + \alpha_N I)^{-1}(F(x_N^\delta) - y^\delta)\| \le c_0\delta$, where $c_0 = 3c + 1$. $\qquad\square$

4. Error bound for the case of noise-free data

To carry out the error analysis concerning the iterates x_k^δ in Sections 5 and 6, in this section we shall make use of some results proved for the iterates x_k, obtained from (1.6) by replacing y^δ by y, that is,

$$x_{k+1} = x_k - (A_k + \alpha_k I)^{-1}[F(x_k) - y + \alpha_k(x_k - x_0)] \quad \text{for } k = 0, 1, 2, \dots. \tag{4.1}$$

We show that each x_k is well defined and belongs to $B_r(x^\dagger)$ for $r > 2\|x_0 - x^\dagger\|$. To do so, we make use of the following lemma.

LEMMA 4.1. *Let Assumption* 2.1 *hold. Suppose that, for each* $k \in \{0, 1, \ldots, n\}$, x_k *in* (4.1) *is well defined, and* $\rho_k := \|\alpha_k(A_k + \alpha_k I)^{-1}(x_0 - x^\dagger)\|$ *for some* $n \in \mathbb{N}$. *Then*

$$\rho_k - k_0 \frac{\|x_k - x^\dagger\|^2}{2} \le \|x_{k+1} - x^\dagger\| \le \rho_k + k_0 \frac{\|x_k - x^\dagger\|^2}{2}, \qquad (4.2)$$

for every $k \in \{0, 1, \ldots, n\}$.

PROOF. By definition,

$$\begin{aligned}
x_{k+1} - x^\dagger &= x_k - x^\dagger - (A_k + \alpha_k I)^{-1}[F(x_k) - y + \alpha_k(x_k - x_0)] \\
&= (A_k + \alpha_k I)^{-1}\{(A_k + \alpha_k I)(x_k - x^\dagger) - [F(x_k) - y + \alpha_k(x_k - x_0)]\} \\
&= (A_k + \alpha_k I)^{-1}\{F(x^\dagger) - F(x_k) - A_k(x^\dagger - x_k) + \alpha_k(x_0 - x^\dagger)\} \\
&= \alpha_k(A_k + \alpha_k I)^{-1}(x_0 - x^\dagger) \\
&\quad + (A_k + \alpha_k I)^{-1}\{F(x^\dagger) - F(x_k) - A_k(x^\dagger - x_k)\}.
\end{aligned}$$

Note that $\rho_k = \|\alpha_k(A_k + \alpha_k I)^{-1}(x_0 - x^\dagger)\|$ and, by Proposition 2.2,

$$\|(A_k + \alpha_k I)^{-1}\{F(x^\dagger) - F(x_k) - A_k(x^\dagger - x_k)\}\| \le \frac{k_0}{2}\|x^\dagger - x_k\|^2.$$

Thus, it follows that

$$\rho_k - k_0 \frac{\|x_k - x^\dagger\|^2}{2} \le \|x_{k+1} - x^\dagger\| \le \rho_k + k_0 \frac{\|x_k - x^\dagger\|^2}{2},$$

completing the proof. □

THEOREM 4.2. *Let Assumption* 2.1 *hold. If* $2k_0\|x_0 - x^\dagger\| \le 1$ *and* $r > 2\|x_0 - x^\dagger\|$, *then, for each* $k \in \mathbb{N}$, *the iterates* x_k *in* (4.1) *are well-defined and*

$$\|x_k - x^\dagger\| \le \frac{2\|x_0 - x^\dagger\|}{1 + \sqrt{1 - 2k_0\|x_0 - x^\dagger\|}} \le 2\|x_0 - x^\dagger\| \quad \forall k \in \mathbb{N}. \qquad (4.3)$$

PROOF. Assume that $2k_0\|x_0 - x^\dagger\| \le 1$ and $r > 2\|x_0 - x^\dagger\|$. Note that

$$\frac{2\|x_0 - x^\dagger\|}{1 + \sqrt{1 - 2k_0\|x_0 - x^\dagger\|}} \le 2\|x_0 - x^\dagger\| < r.$$

We prove (4.3) by induction on k so that $x_k \in B_r(x^\dagger)$ for each $k \in \mathbb{N} \cup \{0\}$. Clearly (4.3) holds for $k = 0$ and $x_0 \in B_r(x^\dagger)$. Suppose that (4.3) holds for some $k \in \mathbb{N} \cup \{0\}$. Then, by taking

$$\theta_k = \|x_k - x^\dagger\|, \quad a = \frac{k_0}{2} \quad \text{and} \quad b = \|x_0 - x^\dagger\|,$$

it follows from (4.2) by noting $\|\alpha_k(A_k + \alpha_k I)^{-1}(x_0 - x^\dagger)\| \le \|x_0 - x^\dagger\|$ that

$$\theta_{k+1} \le a\theta_k^2 + b.$$

Now, by the hypothesis of the theorem we have $4ab = 2k_0\|x_0 - x^\dagger\| \le 1$. Hence, by Lemma 3.1, we get

$$\|x_{k+1} - x^\dagger\| = \theta_{k+1} \le \frac{2\|x_0 - x^\dagger\|}{1 + \sqrt{1 - 2k_0\|x_0 - x^\dagger\|}} \le 2\|x_0 - x^\dagger\|.$$

Thus, we have proved that $x_k \in B_r(x^\dagger)$ and (4.3) hold for all $k \in \mathbb{N}$.   □

The next lemma is crucial for proving subsequent results of this paper.

LEMMA 4.3. *Let Assumptions 2.1 and 2.6 hold and let $r > 2\|x_0 - x^\dagger\|$. Assume that $\|A\| \le \eta\alpha_0$ and $2\mu(1 + \eta^{-1})k_0\|x_0 - x^\dagger\| \le 1$ for some η with $0 < \eta < 1$. Then, for each $k \in \mathbb{N}$,*

$$\frac{1}{(1+\eta)\mu}\|x_k - x^\dagger\| \le \|\alpha_k(A + \alpha_k I)^{-1}(x_0 - x^\dagger)\| \le \frac{1}{1-\eta}\|x_k - x^\dagger\|, \quad (4.4)$$

$$\frac{1-\eta}{(1+\eta)\mu}\|x_k - x^\dagger\| \le \|x_{k+1} - x^\dagger\| \le \left(\frac{1}{1-\eta} + \frac{\eta}{(1+\eta)\mu}\right)\|x_k - x^\dagger\|. \quad (4.5)$$

PROOF. For $k \in \mathbb{N} \cup \{0\}$, let

$$\rho_k := \|\alpha_k(A_k + \alpha_k I)^{-1}(x_0 - x^\dagger)\|, \quad \tilde{\rho}_k := \|\alpha_k(A + \alpha_k I)^{-1}(x_0 - x^\dagger)\|,$$

where $A_k = F'(x_k)$ and $A = F'(x^\dagger)$. Then

$$|\tilde{\rho}_k - \rho_k| = \left| \|\alpha_k(\alpha_k I + A)^{-1}(x_0 - x^\dagger)\| - \|\alpha_k(\alpha_k I + A_k)^{-1}(x_0 - x^\dagger)\| \right|$$

$$\le \|\alpha_k\{(A + \alpha_k I)^{-1} - (\alpha_k I + A_k)^{-1}\}(x_0 - x^\dagger)\|.$$

Hence, by Proposition 2.3, we have $|\tilde{\rho}_k - \rho_k| \le k_0\|x_k - x^\dagger\|\|x_0 - x^\dagger\|$, which also implies that

$$-k_0\|x_k - x^\dagger\|\|x_0 - x^\dagger\| \le \rho_k - \tilde{\rho}_k \le k_0\|x_k - x^\dagger\|\|x_0 - x^\dagger\|. \quad (4.6)$$

From Lemma 4.1,

$$\rho_k - \frac{k_0\|x_k - x^\dagger\|^2}{2} \le \|x_{k+1} - x^\dagger\| \le \rho_k + \frac{k_0\|x_k - x^\dagger\|^2}{2}. \quad (4.7)$$

Using (4.6) in (4.7), we get

$$\tilde{\rho}_k - k_0\|x_k - x^\dagger\|\|x_0 - x^\dagger\| - \frac{k_0\|x_k - x^\dagger\|^2}{2} \le \|x_{k+1} - x^\dagger\| \quad (4.8)$$

and

$$\|x_{k+1} - x^\dagger\| \le \tilde{\rho}_k + k_0\|x_k - x^\dagger\|\|x_0 - x^\dagger\| + \frac{k_0\|x_k - x^\dagger\|^2}{2}. \quad (4.9)$$

From Theorem 4.2, we know that $\|x_k - x^\dagger\| \leq 2\|x_0 - x^\dagger\|$. Using this inequality in (4.8) and (4.9), we get

$$\tilde{\rho}_k - 2k_0\|x_k - x^\dagger\|\|x_0 - x^\dagger\| \leq \|x_{k+1} - x^\dagger\| \leq \tilde{\rho}_k + 2k_0\|x_k - x^\dagger\|\|x_0 - x^\dagger\|.$$

Thus, denoting $\beta := 2k_0\|x_0 - x^\dagger\|$,

$$\tilde{\rho}_k - \beta\|x_k - x^\dagger\| \leq \|x_{k+1} - x^\dagger\| \leq \tilde{\rho}_k + \beta\|x_k - x^\dagger\|. \tag{4.10}$$

Now, from the right-hand inequality in (4.10),

$$\frac{\|x_{k+1} - x^\dagger\|}{\tilde{\rho}_{k+1}} \leq \frac{\tilde{\rho}_k}{\tilde{\rho}_{k+1}} + \beta\frac{\|x_k - x^\dagger\|\tilde{\rho}_k}{\tilde{\rho}_{k+1}\tilde{\rho}_k}. \tag{4.11}$$

Let E_λ be the spectral family of $A := F'(x^\dagger)$. Then, using the relations

$$\frac{\alpha_{k+1}^2}{(\lambda + \alpha_{k+1})^2} \leq \frac{\alpha_k^2}{\alpha_{k+1}^2} \leq \frac{\alpha_k^2}{\alpha_{k+1}^2}\frac{\alpha_{k+1}^2}{(\lambda + \alpha_{k+1})^2}$$

and the spectral representation

$$\tilde{\rho}_k^2 = \int_0^{\|A\|} \frac{\alpha_k^2}{(\lambda + \alpha_k)^2}\, d\|E_\lambda(x_0 - x^\dagger)\|^2 \quad \forall k \in \mathbb{N} \cup \{0\},$$

it follows that

$$\tilde{\rho}_{k+1} \leq \tilde{\rho}_k \leq \mu\tilde{\rho}_{k+1}. \tag{4.12}$$

Denoting $\sigma_k := \|x_k - x^\dagger\|/\tilde{\rho}_k$, from (4.11) and (4.12) we obtain

$$\sigma_{k+1} \leq \mu + \beta\mu\sigma_k. \tag{4.13}$$

Using the relation $\beta\mu \leq \eta/(\eta + 1)$ in (4.13), we obtain

$$\sigma_{k+1} \leq \mu + \frac{\eta}{1 + \eta}\sigma_k. \tag{4.14}$$

We note that

$$\tilde{\rho}_0 = \|\alpha_0(A + \alpha_0 I)^{-1}(x_0 - x^\dagger)\| \geq \frac{\alpha_0\|x_0 - x^\dagger\|}{\|A + \alpha_0 I\|} \geq \frac{\alpha_0\|x_0 - x^\dagger\|}{\eta + 1}.$$

Thus,

$$\sigma_0 = \frac{\|x_0 - x^\dagger\|}{\tilde{\rho}_0} \leq \eta + 1.$$

Now, from (4.14), it follows by induction that $\sigma_k \leq \mu(1 + \eta)$ for all $k \in \mathbb{N} \cup \{0\}$. Thus,

$$\|x_k - x^\dagger\| \leq \tilde{\rho}_k\mu(\eta + 1) \quad \forall k \in \mathbb{N} \cup \{0\}. \tag{4.15}$$

Now from the left-hand inequality of (4.10),

$$\|x_{k+1} - x^\dagger\| \geq \tilde{\rho}_k - \beta\|x_k - x^\dagger\| \geq \tilde{\rho}_k - \frac{\eta}{(\eta + 1)\mu}\|x_k - x^\dagger\| \quad \forall k \in \mathbb{N} \cup \{0\}.$$

This, together with (4.12), implies that

$$\|x_{k+1} - x^\dagger\| \geq (1 - \eta)\tilde{\rho}_k \geq (1 - \eta)\tilde{\rho}_{k+1}.$$

We also note that

$$\tilde{\rho}_0 \leq \|x_0 - x^\dagger\| \leq \frac{1}{1 - \eta}\|x_0 - x^\dagger\|$$

which implies $\|x_0 - x^\dagger\| \geq (1 - \eta)\tilde{\rho}_0$. Thus,

$$\|x_k - x^\dagger\| \geq (1 - \eta)\tilde{\rho}_k \quad \forall k \in \mathbb{N} \cup \{0\}. \tag{4.16}$$

Now, from (4.15) and (4.16), we get the following relation:

$$\frac{1}{(1 + \eta)\mu}\|x_k - x^\dagger\| \leq \tilde{\rho}_k \leq \frac{1}{1 - \eta}\|x_k - x^\dagger\|. \tag{4.17}$$

Using the relation (4.17) in (4.10), we find that

$$\left(\frac{1}{\mu(1 + \eta)} - \beta\right)\|x_k - x^\dagger\| \leq \|x_{k+1} - x^\dagger\| \leq \left(\frac{1}{1 - \eta} + \beta\right)\|x_k - x^\dagger\|. \tag{4.18}$$

Using the relation $\beta\mu \leq \eta/(\eta + 1)$ in (4.18), we get

$$\frac{1 - \eta}{(\eta + 1)\mu}\|x_k - x^\dagger\| \leq \|x_{k+1} - x^\dagger\| \leq \left(\frac{1}{1 - \eta} + \frac{\eta}{(\eta + 1)\mu}\right)\|x_k - x^\dagger\|.$$

This completes the proof. □

The following corollary follows from Lemma 4.3 by taking $\eta = 1/3$. We shall see that this particular case of Lemma 4.3 is better suited for our later results.

COROLLARY 4.4. *Let Assumptions 2.1 and 2.6 hold and let $r > 2\|x_0 - x^\dagger\|$. Assume that $\|A\| \leq \alpha_0/3$ and $8\mu k_0\|x_0 - x^\dagger\| \leq 1$. Then, for each $k \in \mathbb{N}$,*

$$\frac{3}{4\mu}\|x_k - x^\dagger\| \leq \|\alpha_k(A + \alpha_k I)^{-1}(x_0 - x^\dagger)\| \leq \frac{3}{2}\|x_k - x^\dagger\|, \tag{4.19}$$

$$\frac{\|x_k - x^\dagger\|}{2\mu} \leq \|x_{k+1} - x^\dagger\| \leq 2\|x_k - x^\dagger\|.$$

The next theorem shows that under the appropriate conditions on $x^\dagger - x_0$ and F the iterates defined by (4.1) converge to the solution $x^\dagger$ as $k \to \infty$.

THEOREM 4.5. *Let the assumptions of Lemma 4.3 hold. If x_0 is chosen such that $x_0 - x^\dagger \in N(F'(x^\dagger))^\perp$, then $\lim_{k\to\infty} x_k = x^\dagger$.*

PROOF. From the left-hand inequality of (4.4),

$$\|x_k - x^\dagger\| \leq (1 + \eta)\mu\|\alpha_k(A + \alpha_k I)^{-1}(x_0 - x^\dagger)\| \quad \text{with } A = F'(x^\dagger).$$

Hence it is enough to show that $\|\alpha_k(A + \alpha_k I)^{-1}(x_0 - x^\dagger)\| \to 0$ as $k \to \infty$. We show that $\|\alpha_k(A + \alpha_k I)^{-1}w\| \to 0$ for every $w \in R(A)$. Let $w \in R(A)$ and $u \in D(F)$ such that $w = Au$. Then we have

$$\|\alpha_k(A + \alpha_k I)^{-1}w\| = \alpha_k \|(A + \alpha_k I)^{-1}Au\| \leq \alpha_k \|u\| \to 0 \quad \text{as } k \to \infty.$$

Now using the fact that $\|\alpha_k(A + \alpha_k I)^{-1}\| \leq 1$ and for $R(A)$ a dense subspace of $N(A)^\perp$, we get

$$\|\alpha_k(A + \alpha_k I)^{-1}(x_0 - x^\dagger)\| \to 0 \quad \text{as } k \to \infty,$$

which gives $x_k \to x^\dagger$ as $k \to \infty$. $\square$

LEMMA 4.6. *Let the assumptions of Lemma* 4.3 *hold for* η *satisfying*

$$\left(1 - \sqrt{1 - \frac{\eta}{(1 + \eta)\mu}}\right)[1 + (2\mu - 1)\eta + 2\mu] + 4\eta < 2. \tag{4.20}$$

Then, for $k, l \in \mathbb{N} \cup \{0\}$ *with* $k \geq l$,

$$\|x_l - x^\dagger\| \leq c_\eta \left\{\|x_k - x^\dagger\| + \frac{\|\alpha_l(\alpha_l + A)^{-1}(F(x_l) - y)\|}{\alpha_k}\right\},$$

where

$$c_\eta := (1 - b_\eta)^{-1} \max\left\{\mu, 1 + \frac{(\epsilon + 1)\eta}{2(1 - \eta)}\right\},$$

$$b_\eta := \frac{(1 + \epsilon)\eta}{1 - \eta} + \frac{\epsilon a}{2}, \quad \epsilon := \frac{1 - \sqrt{1 - a}}{a} \quad \text{and} \quad a := \frac{\eta}{(1 + \eta)\mu}.$$

PROOF. Let $k, l \in \mathbb{N} \cup \{0\}$ with $k \geq l$. We observe that

$$x_k - x_l = u_1 + u_2 + u_3,$$

where

$$u_1 := [\alpha_{k-1}(A_{k-1} + \alpha_{k-1}I)^{-1} - \alpha_{l-1}(A_{l-1} + \alpha_{l-1}I)^{-1}](x_0 - x^\dagger),$$
$$u_2 := (A_{l-1} + \alpha_{l-1}I)^{-1}(F(x_{l-1}) - y - A_{l-1}(x_{l-1} - x^\dagger)),$$
$$u_3 := (A_{k-1} + \alpha_{k-1}I)^{-1}(y - F(x_{k-1}) - A_{k-1}(x^\dagger - x_{k-1})).$$

By Proposition 2.2,

$$\|u_2\| \leq \frac{k_0}{2}\|x_{l-1} - x^\dagger\|^2, \quad \|u_3\| \leq \frac{k_0}{2}\|x_{k-1} - x^\dagger\|^2. \tag{4.21}$$

We note that $k_0\|x_0 - x^\dagger\| \leq \eta/2(1 + \eta)\mu$. Using this inequality in (4.3), we get

$$\|x_{l-1} - x^\dagger\| \leq 2\frac{\sqrt{(\eta + 1)\mu}\|x_0 - x^\dagger\|}{\sqrt{(\eta + 1)\mu} + \sqrt{(\eta + 1)\mu - \eta}} = 2\epsilon\|x_0 - x^\dagger\|. \tag{4.22}$$

Using (4.22) in (4.21), we get

$$\|u_2\| \leq k_0\epsilon\|x_0 - x^\dagger\|\|x_{l-1} - x^\dagger\|, \quad \|u_3\| \leq k_0\epsilon\|x_0 - x^\dagger\|\|x_{k-1} - x^\dagger\|. \tag{4.23}$$

Now, we find an estimate for $\|u_1\|$:

$$
\begin{aligned}
u_1 &:= [\alpha_{k-1}(A_{k-1} + \alpha_{k-1}I)^{-1} - \alpha_{l-1}(A_{l-1} + \alpha_{l-1}I)^{-1}](x_0 - x^\dagger) \\
&= \alpha_{k-1}(A_{k-1} + \alpha_{k-1}I)^{-1}(x_0 - x^\dagger) - \alpha_{k-1}(A + \alpha_{k-1}I)^{-1}(x_0 - x^\dagger) \\
&\quad + \alpha_{k-1}(A + \alpha_{k-1}I)^{-1}(x_0 - x^\dagger) - \alpha_{l-1}(A + \alpha_{l-1}I)^{-1}(x_0 - x^\dagger) \\
&\quad + \alpha_{l-1}(A + \alpha_{l-1}I)^{-1}(x_0 - x^\dagger) - \alpha_{l-1}(A_{l-1} + \alpha_{l-1}I)^{-1}(x_0 - x^\dagger),
\end{aligned}
$$

$$
\begin{aligned}
\|u_1\| &\leq \left\| \alpha_{k-1}\big((A_{k-1} + \alpha_{k-1}I)^{-1} - (A + \alpha_{k-1}I)^{-1}\big)(x_0 - x^\dagger) \right\| \\
&\quad + \|[\alpha_{l-1}(A + \alpha_{l-1}I)^{-1}(x_0 - x^\dagger) - \alpha_{k-1}(A + \alpha_{k-1}I)^{-1}](x_0 - x^\dagger)\| \\
&\quad + \|\alpha_{l-1}[(A_{l-1} + \alpha_{l-1}I)^{-1} - (A + \alpha_{l-1}I)^{-1}](x_0 - x^\dagger)\|. \tag{4.24}
\end{aligned}
$$

By Proposition 2.3,

$$
\|\alpha_{k-1}[(A_{k-1} + \alpha_{k-1}I)^{-1} - (A + \alpha_{k-1}I)^{-1}](x_0 - x^\dagger)\| \leq k_0 \|x_{k-1} - x^\dagger\|\|x_0 - x^\dagger\| \tag{4.25}
$$

and

$$
\|\alpha_{l-1}[(A_{l-1} + \alpha_{l-1}I)^{-1} - (A + \alpha_{l-1}I)^{-1}](x_0 - x^\dagger)\| \leq k_0 \|x_{l-1} - x^\dagger\|\|x_0 - x^\dagger\|. \tag{4.26}
$$

Using (4.25) and (4.26) in (4.24), we get

$$
\|u_1\| = \|v\| + k_0\|x_0 - x^\dagger\|\big(\|x_{k-1} - x^\dagger\| + \|x_{l-1} - x^\dagger\|\big), \tag{4.27}
$$

where

$$
v := \Big(\alpha_{l-1}(A + \alpha_{l-1}I)^{-1} - \alpha_{k-1}(A + \alpha_{k-1}I)^{-1}\Big)(x_0 - x^\dagger) = v_1 + v_2 + v_3
$$

with

$$
v_1 := \left(1 - \frac{\alpha_{k-1}}{\alpha_{l-1}}\right)(A + \alpha_{k-1}I)^{-1}(F(x_l) - y),
$$

$$
v_2 := \left(1 - \frac{\alpha_{k-1}}{\alpha_{l-1}}\right)(A + \alpha_{k-1}I)^{-1}(A(x_l - x^\dagger) - F(x_l) + y),
$$

$$
v_3 := \left(1 - \frac{\alpha_{k-1}}{\alpha_{l-1}}\right)(A + \alpha_{k-1}I)^{-1}A\Big(\alpha_{l-1}(A + \alpha_{l-1}I)^{-1}(x_0 - x^\dagger) - (x_l - x^\dagger)\Big).
$$

By Proposition 2.2 and the relation (4.22),

$$
\|v_2\| \leq k_0 \frac{\|x_l - x^\dagger\|^2}{2} \leq k_0\epsilon\|x_0 - x^\dagger\|\|x_l - x^\dagger\|. \tag{4.28}
$$

It can be seen that

$$
\begin{aligned}
v_3 &= \alpha_{l-1}\big((A + \alpha_{l-1}I)^{-1} - (\alpha_{l-1}I + A_{l-1})^{-1}\big)(x_0 - x^\dagger) \\
&\quad + (A_{l-1} + \alpha_{l-1}I)^{-1}\big(F(x_{l-1}) - y - A_{l-1}(x_{l-1} - x^\dagger)\big).
\end{aligned}
$$

Hence, by Propositions 2.2 and 2.3 and the relation (4.22), we obtain

$$
\|v_3\| \leq k_0\|x_{l-1} - x^\dagger\|\|x_0 - x^\dagger\| + \epsilon k_0\|x_{l-1} - x^\dagger\|\|x_0 - x^\dagger\|,
$$

or equivalently

$$\|v_3\| \le (1 + \epsilon)k_0\|x_{l-1} - x^{\dagger}\|\|x_0 - x^{\dagger}\|. \tag{4.29}$$

Now, we estimate $\|v_1\|$. We have that

$$\|v_1\| := \left\| \left(1 - \frac{\alpha_{k-1}}{\alpha_{l-1}}\right)(A + \alpha_{k-1}I)^{-1}(F(x_l) - y) \right\|$$

$$\le \|(A + \alpha_{k-1}I)^{-1}(F(x_l) - y)\|.$$

Therefore, using the spectral representation of $(A + \alpha_{k-1}I)^{-1}$ and the fact that $\alpha_k \le \alpha_{k-1} \le \alpha_{l-1}$,

$$\|(A + \alpha_{k-1}I)^{-1}x\|^2 = \int_0^{\|A\|} \frac{d\|E_\lambda x\|^2}{(\alpha_{k-1} + \lambda)^2}$$

$$\le \frac{1}{\alpha_{k-1}^2} \int_0^{\|A\|} \left(\frac{\alpha_{l-1}}{\alpha_{l-1} + \lambda}\right)^2 d\|E_\lambda x\|^2$$

$$\le \left(\frac{\alpha_{l-1}}{\alpha_{k-1}\alpha_l}\right)^2 \int_0^{\|A\|} \left(\frac{\alpha_l}{\alpha_l + \lambda}\right)^2 d\|E_\lambda x\|^2$$

$$\le \left(\frac{\mu\alpha_l}{\alpha_{k-1}}\right)^2 \int_0^{\|A\|} \frac{1}{(\alpha_l + \lambda)^2} d\|E_\lambda x\|^2$$

$$= \left(\frac{\mu}{\alpha_k}\right)^2 \|\alpha_l(A + \alpha_l I)^{-1}x\|^2.$$

Thus,

$$\|v_1\| \le \mu\frac{\|\alpha_l(A + \alpha_l I)^{-1}(F(x_l) - y)\|}{\alpha_k}. \tag{4.30}$$

Hence, from (4.28), (4.29) and (4.30),

$$\|v\| \le \mu\frac{\|\alpha_l(A + \alpha_l I)^{-1}(F(x_l) - y)\|}{\alpha_k} + k_0\epsilon\|x_0 - x^{\dagger}\|\|x_l - x^{\dagger}\|$$

$$+ (1 + \epsilon)k_0\|x_{l-1} - x^{\dagger}\|\|x_0 - x^{\dagger}\|. \tag{4.31}$$

Using (4.31) in (4.27),

$$\|u_1\| = \mu\frac{\|\alpha_l(A + \alpha_l I)^{-1}(F(x_l) - y)\|}{\alpha_k} + k_0\epsilon\|x_0 - x^{\dagger}\|\|x_l - x^{\dagger}\|$$

$$+ (2 + \epsilon)k_0\|x_{l-1} - x^{\dagger}\|\|x_0 - x^{\dagger}\| + k_0\|x_{k-1} - x^{\dagger}\|\|x_0 - x^{\dagger}\|. \tag{4.32}$$

From (4.23) and (4.32), we get

$$\|x_k - x_l\| \le \mu\frac{\|\alpha_l(A + \alpha_l I)^{-1}(F(x_l) - y)\|}{\alpha_k} + k_0\epsilon\|x_0 - x^{\dagger}\|\|x_l - x^{\dagger}\|$$

$$+ 2(1 + \epsilon)k_0\|x_{l-1} - x^{\dagger}\|\|x_0 - x^{\dagger}\|$$

$$+ (\epsilon + 1)k_0\|x_{k-1} - x^{\dagger}\|\|x_0 - x^{\dagger}\|. \tag{4.33}$$

From the left-hand side of (4.5), we get the inequality

$$\|x_{k-1} - x^\dagger\| \le \frac{(1+\eta)\mu}{1-\eta}\|x_k - x^\dagger\|.$$

Using this inequality in (4.33), we get

$$
\begin{aligned}
\|x_k - x_l\| &\le \mu \frac{\|\alpha_l(A + \alpha_l I)^{-1}(F(x_l) - y)\|}{\alpha_k} \\
&+ \left(\epsilon + \frac{2(1+\epsilon)(\eta+1)\mu}{1-\eta}\right)k_0\|x_0 - x^\dagger\|\|x_l - x^\dagger\| \\
&+ \frac{(1+\eta)\mu}{1-\eta}(\epsilon + 1)k_0\|x_0 - x^\dagger\|\|x_k - x^\dagger\| \\
&\le \mu \frac{\|\alpha_l(A + \alpha_l I)^{-1}(F(x_l) - y)\|}{\alpha_k} \\
&+ \left[\frac{\eta}{1-\eta} + \epsilon\left(\frac{\eta}{2(1+\eta)\mu} + \frac{\eta}{1-\eta}\right)\right]\|x_l - x^\dagger\| \\
&+ \frac{(\epsilon+1)\eta}{2(1-\eta)}\|x_k - x^\dagger\|.
\end{aligned}
\tag{4.34}
$$

Using (4.34) with $\|x_l - x^\dagger\| \le \|x_l - x_k\| + \|x_k - x^\dagger\|$ gives

$$
\begin{aligned}
\|x_l - x^\dagger\| &\le \mu \frac{\|\alpha_l(A + \alpha_l I)^{-1}(F(x_l) - y)\|}{\alpha_k} \\
&+ \left[\frac{\eta}{1-\eta} + \epsilon\left(\frac{\eta}{2(1+\eta)\mu} + \frac{\eta}{1-\eta}\right)\right]\|x_l - x^\dagger\| \\
&+ \left[1 + \frac{(\epsilon+1)\eta}{2(1-\eta)}\right]\|x_k - x^\dagger\|.
\end{aligned}
$$

Let

$$b_\eta := \frac{(1+\epsilon)\eta}{1-\eta} + \frac{\eta\epsilon}{2(1+\eta)\mu}.$$

It can be seen that $b_\eta < 1$ if and only if

$$\left(1 - \sqrt{1 - \frac{\eta}{(1+\eta)\mu}}\right)[1 + (2\mu - 1)\eta + 2\mu] + 4\eta < 2.$$

Thus, for η satisfying (4.20),

$$
\begin{aligned}
&\|x_l - x^\dagger\| \\
&\le \frac{(1 - b_\eta)^{-1}}{\alpha_k}\left[\mu\|\alpha_l(A + \alpha_l I)^{-1}(F(x_l) - y)\| + \left(1 + \frac{(\epsilon+1)\eta}{2(1-\eta)}\right)\|x_k - x^\dagger\|\right].
\end{aligned}
$$

Hence, we get

$$\|x_l - x^\dagger\| \leq c_\eta \left\{ \|x_k - x^\dagger\| + \frac{\|\alpha_l (A + \alpha_l)^{-1}(F(x_l) - y)\|}{\alpha_k} \right\},$$

where $c_\eta = (1 - b_\eta)^{-1} \max\{\mu, 1 + (\epsilon + 1)\eta/2(1 - \eta)\}$. □

REMARK 4.7. It can be seen that (4.20) is satisfied if $\eta \leq 1/3 + 1/24$.

Now, if we take $\eta = 1/3$, that is, $k_0 \|x_0 - x^\dagger\| \mu \leq 1/8$ in Lemma 4.6, then it takes the following form.

LEMMA 4.8. *Let the assumptions of Lemma* 4.3 *hold with* $\eta = 1/3$. *Then, for* $k \geq l \geq 0$,

$$\|x_l - x^\dagger\| \leq c_{1/3} \left\{ \|x_k - x^\dagger\| + \frac{\|\alpha_l (A + \alpha_l)^{-1}(F(x_l) - y)\|}{\alpha_k} \right\},$$

where

$$c_{1/3} := \left\{ 1 - \frac{4\mu + (4\mu + 1)\epsilon}{8\mu} \right\}^{-1} \max \left\{ \mu, 1 + \frac{\epsilon + 1}{4} \right\} \quad and$$

$$\epsilon := \frac{\sqrt{4\mu}}{\sqrt{4\mu} + \sqrt{4\mu - 1}}.$$

5. Error analysis with noisy data

The first result in this section gives an error estimate for $\|x_k^\delta - x_k\|$ where $k = 0, 1, 2, \ldots, N$.

LEMMA 5.1. *Let Assumption* 2.1 *hold and let* $k_0 \|x_0 - x^\dagger\| \leq 1/m$ *where we have* $m > (5 + \sqrt{33})/2$ *and let N be the integer satisfying* (3.2) *with*

$$c > \frac{m^2 - 4m - 2}{m^2 - 5m - 2}.$$

Then, for all $k \in \{0, 1, \ldots, N\}$,

$$\|x_k^\delta - x_k\| \leq \frac{\delta}{(1 - \kappa)\alpha_k} \quad where \quad \kappa := \frac{1}{m} \left(5 + \frac{1}{c - 1} + \frac{2}{m} \right). \quad (5.1)$$

PROOF. Inequality (5.1) is trivial for $k = 0$. From the definition of x_k^δ and x_k,

$$x_{k+1}^\delta - x_{k+1} = w_1 + w_2 + w_3,$$

where

$$w_1 = (A_k + \alpha_k I)^{-1} \left(F(x_k) - y - A_k(x_k - x^\dagger) \right)$$

$$- (A_k^\delta + \alpha_k I)^{-1} \left(F(x_k^\delta) - y - A_k^\delta(x_k^\delta - x^\dagger) \right),$$

$$w_2 = \alpha_k \Big((A_k^\delta + \alpha_k I)^{-1} - (A_k + \alpha_k I)^{-1} \Big)(x_0 - x^\dagger),$$

$$w_3 = (A_k^\delta + \alpha_k I)^{-1}(y^\delta - y).$$

Using Proposition 2.3,

$$\|w_2\| \le k_0 \|x_k^\delta - x_k\| \|x_0 - x^\dagger\|.$$

From (1.3) and the estimate $\|(A_k^\delta + \alpha_k I)^{-1}\| \le 1/\alpha_k$, we have $\|w_3\| \le \delta/\alpha_k$. So,

$$\|x_{k+1}^\delta - x_{k+1}\| \le \|w_1\| + k_0 \|x_k^\delta - x_k\| \|x_0 - x^\dagger\| + \frac{\delta}{\alpha_k}.$$

Now, let us write $w_1 = w_1^1 + w_1^2$ with

$$w_1^1 = \Big((A_k + \alpha_k I)^{-1} - (A_k^\delta + \alpha_k I)^{-1} \Big)\Big(F(x_k) - y - A_k(x_k - x^\dagger) \Big),$$

$$w_1^2 = (A_k^\delta + \alpha_k I)^{-1}\Big(F(x_k) - F(x_k^\delta) - A_k(x_k - x^\dagger) + A_k^\delta(x_k^\delta - x^\dagger) \Big).$$

Using Assumption 2.1(ii), we get

$$w_1^1 = [(A_k^\delta + \alpha_k I)^{-1} - (A_k + \alpha_k I)^{-1}][y - F(x_k) - A_k(x^\dagger - x_k)]$$
$$= (A_k^\delta + \alpha_k I)^{-1}(A_k - A_k^\delta)(A_k + \alpha_k I)^{-1}[F(x^\dagger) - F(x_k) - A_k(x^\dagger - x_k)]$$
$$= (A_k^\delta + \alpha_k I)^{-1} A_k^\delta \Phi(x_k, x_k^\delta, z_k),$$

where $z_k := (A_k + \alpha_k I)^{-1}(F(x^\dagger) - F(x_k) - A_k(x^\dagger - x_k))$. Hence, by the relation $\|(A_k^\delta + \alpha_k I)^{-1} A_k^\delta\| \le 1$, the Assumption 2.1(ii) and Proposition 2.2, we have

$$\|w_1^1\| \le \|(A_k^\delta + \alpha_k I)^{-1} A_k^\delta \Phi(x_k, x_k^\delta, z_k)\|$$
$$\le k_0 \|x_k - x_k^\delta\| \|z_k\| \le \frac{k_0^2}{2} \|x_k - x_k^\delta\| \|x_k - x^\dagger\|^2.$$

Next, we observe that

$$w_1^2 = (A_k^\delta + \alpha_k I)^{-1}\Big(F(x_k) - F(x_k^\delta) - A_k(x_k - x^\dagger) + A_k^\delta(x_k^\delta - x^\dagger) \Big)$$
$$= (A_k^\delta + \alpha_k I)^{-1}[F(x_k) - F(x_k^\delta) - A_k^\delta(x_k - x_k^\delta) + (A_k^\delta - A_k)(x_k - x^\dagger)].$$

By Proposition 2.2 and Assumption 2.1(ii),

$$\|(A_k^\delta + \alpha_k I)^{-1}[F(x_k) - F(x_k^\delta) - A_k^\delta(x_k - x_k^\delta)]\| \le \frac{k_0}{2} \|x_k - x_k^\delta\|^2$$

and

$$\|(A_k^\delta + \alpha_k I)^{-1}(A_k^\delta - A_k)(x_k - x^\dagger)\| \le \|(A_k^\delta + \alpha_k I)^{-1} A_k^\delta \Phi(x_k^\delta, x_k, x_k - x^\dagger)\|$$
$$\le \|\Phi(x_k^\delta, x_k, x_k - x^\dagger)\|$$
$$\le k_0 \|x_k^\delta - x_k\| \|x_k - x^\dagger\|.$$

Thus,

$$\|w_1^2\| \le k_0(\|x_k - x_k^\delta\| + 2\|x_k - x^\dagger\|)\frac{\|x_k^\delta - x_k\|}{2}.$$

Hence,

$$\|w_1\| \le \|w_1^1\| + \|w_1^2\| \le \frac{k_0}{2}\|x_k - x_k^\delta\|\|x_k - x^\dagger\|(2 + k_0\|x_k - x^\dagger\|) + \frac{k_0}{2}\|x_k - x_k^\delta\|^2.$$

Now, denoting $\varepsilon_k := \|x_k - x_k^\delta\|$ and using the estimate $\|x_k - x^\dagger\| \le 2\|x_0 - x^\dagger\|$, we get

$$\|w_1\| \le 2k_0\varepsilon_k\|x_0 - x^\dagger\|(1 + k_0\|x_0 - x^\dagger\|) + \frac{k_0}{2}\varepsilon_k^2 \quad \text{and} \quad \|w_2\| \le k_0\varepsilon_k\|x_0 - x^\dagger\|.$$

Thus,

$$\begin{aligned}
\varepsilon_{k+1} &\le k_0\varepsilon_k\|x_0 - x^\dagger\|(2 + 2k_0\|x_0 - x^\dagger\|) + \frac{k_0}{2}\varepsilon_k^2 + k_0\varepsilon_k\|x_0 - x^\dagger\| + \frac{\delta}{\alpha_k} \\
&\le k_0\varepsilon_k\|x_0 - x^\dagger\|(3 + 2k_0\|x_0 - x^\dagger\|) + \frac{k_0}{2}\varepsilon_k^2 + \frac{\delta}{\alpha_k} \\
&\le k_0\varepsilon_k\|x_0 - x^\dagger\|(3 + 2k_0\|x_0 - x^\dagger\|) + \frac{k_0}{2}\varepsilon_k(\|x_k^\delta - x^\dagger\| + \|x_k - x^\dagger\|) + \frac{\delta}{\alpha_k}.
\end{aligned}$$

Now using the estimates

$$\|x_k - x^\dagger\| \le 2\|x_0 - x^\dagger\| \quad \text{and} \quad \|x_k^\delta - x^\dagger\| \le \frac{2c\|x_0 - x^\dagger\|}{c - 1},$$

we get

$$\begin{aligned}
\varepsilon_{k+1} &\le k_0\varepsilon_k\|x_0 - x^\dagger\|(3 + 2k_0\|x_0 - x^\dagger\|) + k_0\|x_0 - x^\dagger\|\varepsilon_k\left(\frac{c}{c - 1} + 1\right) + \frac{\delta}{\alpha_k} \\
&\le k_0\|x_0 - x^\dagger\|\varepsilon_k\left(4 + \frac{c}{c - 1} + 2k_0\|x_0 - x^\dagger\|\right) + \frac{\delta}{\alpha_k}. \tag{5.2}
\end{aligned}$$

Using the assumption that $k_0\|x_0 - x^\dagger\| \le 1/m$ in (5.2), we obtain

$$\varepsilon_{k+1} \le \frac{1}{m}\left(5 + \frac{1}{c - 1} + \frac{2}{m}\right)\varepsilon_k + \frac{\delta}{\alpha_k}. \tag{5.3}$$

Denoting $\kappa = (5 + 1/(c - 1) + 2/m)/m$ and multiplying both sides of (5.3) by α_{k+1},

$$\alpha_{k+1}\varepsilon_{k+1} \le \frac{\kappa\alpha_{k+1}\alpha_k\varepsilon_k}{\alpha_k} + \frac{\alpha_{k+1}\delta}{\alpha_k}.$$

Denoting $\ell_{k+1} := \alpha_{k+1}\varepsilon_{k+1}$,

$$\ell_{k+1} \le \kappa\ell_k + \delta. \tag{5.4}$$

We note that $\kappa < 1$ for $c > (m^2 - 4m - 2)/(m^2 - 5m - 2)$. Now, we claim that

$$\ell_k \le \frac{\delta}{(1 - \kappa)} \quad \text{for } k \in \{0, 1, \dots, N\}.$$

The inequality is trivial for $k = 0$. Now, let us assume that $\ell_k \leq \delta/(1 - \kappa)$ for some $k \in \{0, 1, \ldots, N - 1\}$. From (5.4),

$$\ell_{k+1} \leq \kappa \ell_k + \delta \leq \frac{\delta}{(1 - \kappa)}.$$

Hence we get

$$\|x_k^\delta - x_k\| \leq \frac{\delta}{(1 - \kappa)\alpha_k} \quad \text{for } k \in \{0, 1, \ldots, N\}. \qquad \square$$

If we take $m = 8$ in Lemma 5.1, then we get the following corollary as a particular case of Lemma 5.1. We make use of it in the following error analysis.

COROLLARY 5.2. *Let Assumption* 2.1 *hold and take* $8k_0\|x_0 - x^\dagger\| \leq 1$. *Let N be the integer defined by* (3.2) *with $c > 15/11$. Then, for all $\kappa \in \{0, 1, \ldots, N\}$,*

$$\|x_k^\delta - x_k\| \leq \frac{\delta}{(1 - \kappa)\alpha_k} \quad \text{where } \kappa = \frac{21c - 17}{32(c - 1)}.$$

LEMMA 5.3. *Let the assumptions of Lemma* 5.1 *hold. Then*

$$\|\alpha_{k_\delta}(A + \alpha_{k_\delta}I)^{-1}(F(x_{k_\delta}) - y)\| \leq c_1\delta.$$

Moreover, if $k_\delta > 0$, then for all integers $0 \leq k < k_\delta$,

$$\|\alpha_k(A + \alpha_k I)^{-1}(F(x_k) - y)\| \geq c_2\delta,$$

where

$$c_1 = \left(1 + \frac{2ck_0\|x_0 - x^\dagger\|}{c - 1}\right)\left(c_0 + \frac{2 - \kappa}{1 - \kappa} + \frac{k_0\mu\|x_0 - x^\dagger\|}{2(1 - \kappa)^2(c - 1)}\right),$$

$$c_2 = \frac{c_0 - ((2 - \kappa)/(1 - \kappa)) - (k_0\|x_0 - x^\dagger\|/2(1 - \kappa)^2(c - 1))}{1 + 2(ck_0\|x_0 - x^\dagger\|/(c - 1))},$$

with $c_0 = 3c + 1$ and κ as in Lemma 5.1.

PROOF. Let k_δ be as in (3.4) and $0 \leq k \leq k_\delta$. Since

$$\left\|\alpha_k(A_k^\delta + \alpha_k I)^{-1}(F(x_k^\delta) - F(x_k))\right\|$$

can be written as

$$\left\|\alpha_k(A_k^\delta + \alpha_k I)^{-1}\big(F(x_k) - F(x_k^\delta) - A_k^\delta(x_k - x_k^\delta) + A_k^\delta(x_k - x_k^\delta)\big)\right\|,$$

it follows from the relation $\|\alpha_k(A_k^\delta + \alpha_k I)^{-1}A_k^\delta\| \leq 1$ and Proposition 2.2 that

$$\|\alpha_k(A_k^\delta + \alpha_k I)^{-1}(F(x_k^\delta) - F(x_k))\|$$
$$\leq \alpha_k\left\{\frac{k_0\|x_k - x_k^\delta\|^2}{2} + \|x_k - x_k^\delta\|\right\} = \alpha_k\left(\frac{k_0\|x_k - x_k^\delta\|}{2} + 1\right)\|x_k - x_k^\delta\|.$$

Using the estimate $\|x_k - x_k^\delta\| \le \delta/(1-\kappa)\alpha_k$ from Lemma 5.1, we get

$$\|\alpha_k(A_k^\delta + \alpha_k I)^{-1}(F(x_k^\delta) - F(x_k))\| \le \left(1 + \frac{\delta k_0}{2(1-\kappa)\alpha_k}\right)\frac{\delta}{(1-\kappa)}. \qquad (5.5)$$

Similarly

$$\|\alpha_{k_\delta}(A_{k_\delta}^\delta + \alpha_{k_\delta} I)^{-1}(F(x_{k_\delta}^\delta) - F(x_{k_\delta}))\| \le \left(1 + \frac{\delta k_0}{2(1-\kappa)\alpha_{k_\delta}}\right)\frac{\delta}{(1-\kappa)}. \qquad (5.6)$$

By (2.2) and the definition of N,

$$\frac{\delta}{\alpha_{k_\delta}} \le \frac{\mu\delta}{\alpha_{k_\delta-1}} \le \frac{\mu\|x_0 - x^\dagger\|}{c-1} \qquad (5.7)$$

and

$$\frac{\delta}{\alpha_k} \le \frac{\|x_0 - x^\dagger\|}{(c-1)}. \qquad (5.8)$$

Using (5.7) in (5.6) and (5.8) in (5.5), we get

$$\|\alpha_{k_\delta}(A_{k_\delta}^\delta + \alpha_{k_\delta} I)^{-1}(F(x_{k_\delta}^\delta) - F(x_{k_\delta}))\| \le \left(1 + k_0\frac{\mu\|x_0 - x^\dagger\|}{2(1-\kappa)(c-1)}\right)\frac{\delta}{(1-\kappa)} \qquad (5.9)$$

and

$$\|\alpha_k(A_k^\delta + \alpha_k I)^{-1}(F(x_k^\delta) - F(x_k))\| \le \left(1 + k_0\frac{\|x_0 - x^\dagger\|}{2(1-\kappa)(c-1)}\right)\frac{\delta}{(1-\kappa)}. \qquad (5.10)$$

Now using (1.3), (3.1) and (5.9), we get

$$\begin{aligned}
&\|\alpha_{k_\delta}(A_{k_\delta}^\delta + \alpha_{k_\delta} I)^{-1}(F(x_{k_\delta}) - y)\| \\
&\le \|\alpha_{k_\delta}(A_{k_\delta}^\delta + \alpha_{k_\delta} I)^{-1}(F(x_{k_\delta}^\delta) - F(x_{k_\delta}))\| \\
&\quad + \|\alpha_{k_\delta}(A_{k_\delta}^\delta + \alpha_{k_\delta} I)^{-1}(F(x_{k_\delta}^\delta) - y^\delta)\| + \|y^\delta - y\| \\
&\le \left(1 + k_0\frac{\mu\|x_0 - x^\dagger\|}{2(1-\kappa)(c-1)}\right)\frac{\delta}{1-\kappa} + c_0\delta + \delta \\
&\le \left(c_0 + \frac{2-\kappa}{1-\kappa} + \frac{k_0\mu\|x_0 - x^\dagger\|}{2(1-\kappa)^2(c-1)}\right)\delta.
\end{aligned}$$

Again using (1.3), (3.1) and (5.10), we obtain

$$\begin{aligned}
&\|\alpha_k(A_k^\delta + \alpha_k I)^{-1}(F(x_k) - y)\| \\
&= \left\|\alpha_k(A_k^\delta + \alpha_k I)^{-1}\left(F(x_k^\delta) - y^\delta - F(x_k^\delta) + F(x_k) + y^\delta - y\right)\right\| \\
&\ge \|\alpha_k(A_k^\delta + \alpha_k I)^{-1}(F(x_k^\delta) - y^\delta)\| \\
&\quad - \|\alpha_k(A_k^\delta + \alpha_k I)^{-1}(F(x_k^\delta) - F(x_k))\| - \|y^\delta - y\|
\end{aligned}$$

$$\geq c_0 \delta - \left(1 + k_0 \frac{\|x_0 - x^\dagger\|}{2(1-\kappa)(c-1)}\right) \frac{\delta}{(1-\kappa)} - \delta$$

$$\geq \left(c_0 - \frac{2-\kappa}{1-\kappa} - \frac{k_0\|x_0 - x^\dagger\|}{2(1-\kappa)^2(c-1)}\right)\delta. \tag{5.11}$$

Denoting

$$a_k := \|\alpha_k(A + \alpha_k I)^{-1}(F(x_k) - y)\| \quad \text{and} \quad b_k := \|\alpha_k(A_k^\delta + \alpha_k I)^{-1}(F(x_k) - y)\|$$

we have

$$\begin{aligned}
a_{k_\delta} &= \|\alpha_{k_\delta}(A + \alpha_{k_\delta}I)^{-1}(F(x_{k_\delta}) - y)\| \\
&\leq \|\alpha_{k_\delta}(A_{k_\delta}^\delta + \alpha_{k_\delta}I)^{-1}(F(x_{k_\delta}) - y)\| \\
&\quad + \|\alpha_{k_\delta}(\alpha_{k_\delta}I + A)^{-1} - (A_{k_\delta}^\delta + \alpha_{k_\delta}I)^{-1}(F(x_{k_\delta}) - y)\| \\
&\leq b_{k_\delta} + \|(A + \alpha_{k_\delta}I)^{-1}(A_{k_\delta}^\delta - A)\alpha_{k_\delta}(A_{k_\delta}^\delta + \alpha_{k_\delta}I)^{-1}(F(x_{k_\delta}) - y)\| \\
&\leq b_{k_\delta} + k_0\|x_{k_\delta}^\delta - x^\dagger\|\|\alpha_{k_\delta}(\alpha_{k_\delta}I + A_{k_\delta}^\delta)^{-1}(F(x_{k_\delta}) - y)\| \\
&\leq \left(1 + 2\frac{c}{c-1}k_0\|x_0 - x^\dagger\|\right)b_{k_\delta} \\
&\leq \left(1 + 2\frac{c}{c-1}k_0\|x_0 - x^\dagger\|\right)\left(c_0 + \frac{2-\kappa}{1-\kappa} + \frac{k_0\mu\|x_0 - x^\dagger\|}{2(1-\kappa)^2(c-1)}\right)\delta.
\end{aligned}$$

Hence,

$$\|\alpha_{k_\delta}(A + \alpha_{k_\delta})^{-1}(F(x_{k_\delta}) - y)\| \leq c_1\delta,$$

where

$$c_1 = \left(1 + \frac{2ck_0\|x_0 - x^\dagger\|}{c-1}\right)\left(c_0 + \frac{2-\kappa}{1-\kappa} + \frac{k_0\mu\|x_0 - x^\dagger\|}{2(1-\kappa)^2(c-1)}\right).$$

Now we have

$$\begin{aligned}
b_k &= \|\alpha_k(A_k^\delta + \alpha_k I)^{-1}(F(x_k) - y)\| \\
&\leq \|\alpha_k(A + \alpha_k I)^{-1}(F(x_k) - y)\| \\
&\quad + \|\alpha_k(A_k^\delta + \alpha_k I)^{-1} - (A + \alpha_k I)^{-1}(F(x_k) - y)\| \\
&\leq a_k + \|(A_k^\delta + \alpha_k I)^{-1}(A - A_k^\delta)\alpha_k(A + \alpha_k I)^{-1}(F(x_k) - y)\| \\
&\leq a_k + k_0\|x_k^\delta - x^\dagger\|\|\alpha_k(A + \alpha_k I)^{-1}(F(x_k) - y)\| \leq \left(1 + \frac{2ck_0\|x_0 - x^\dagger\|}{c-1}\right)a_k.
\end{aligned}$$

So, we get

$$\begin{aligned}
a_k &\geq \frac{b_k}{1 + (2ck_0\|x_0 - x^\dagger\|/(c-1))} \\
&\geq \frac{c_0 - ((2-\kappa)/(1-\kappa)) - (k_0\|x_0 - x^\dagger\|/2(1-\kappa)^2(c-1))}{1 + (2ck_0\|x_0 - x^\dagger\|/(c-1))}.
\end{aligned}$$

Thus, $\|\alpha_k(A + \alpha_k I)^{-1}(F(x_k) - y)\| \geq c_2 \delta$, where

$$c_2 = \frac{c_0 - ((2 - \kappa)/(1 - \kappa)) - (k_0\|x_0 - x^\dagger\|/2(1 - \kappa)^2(c - 1))}{1 + (2ck_0\|x_0 - x^\dagger\|/(c - 1))}. \qquad \square$$

THEOREM 5.4. *Let Assumptions* 2.1 *and* 2.6 *hold. If* $8\mu k_0\|x_0 - x^\dagger\| \leq 1$ *and the integer* k_δ *is chosen according to stopping rule* (3.1) *with* $c_0 > 56/11$, *then*

$$\|x_{k_\delta}^\delta - x^\dagger\| \leq \xi \inf\left\{\|x_k - x^\dagger\| + \frac{\delta}{\alpha_k} : k = 0, 1, \ldots\right\}, \qquad (5.12)$$

where

$$\xi = \max\left\{2\mu\varrho, \frac{c_{1/3}c_1 + 1}{1 - \kappa}, c\right\} \quad and \quad \varrho := 1 + \frac{\mu}{c_2(1 - \kappa)}(1 + k_0\|x_0 - x^\dagger\|) \qquad (5.13)$$

with $c_{1/3}$ *and* κ *as in Lemma* 4.8 *and Corollary* 5.2, *respectively and* c_1 *and* c_2 *as in Lemma* 5.3.

PROOF. We consider two cases.

Case (i) $k > N$. From the definition of N and the estimate

$$\|x_k^\delta - x^\dagger\| \leq \frac{2c\|x_0 - x^\dagger\|}{c - 1}$$

from Theorem 3.2,

$$\frac{\delta}{\alpha_k} \geq \frac{2\|x_0 - x^\dagger\|}{c - 1} \geq \frac{\|x_{k_\delta}^\delta - x^\dagger\|}{c},$$

which implies that

$$\|x_{k_\delta}^\delta - x^\dagger\| \leq \frac{c\delta}{\alpha_k} \quad \forall k > N.$$

Hence, we get

$$\|x_{k_\delta}^\delta - x^\dagger\| \leq \frac{c\delta}{\alpha_k} + \|x_k - x^\dagger\| \quad \forall k > N.$$

So, we get

$$\|x_{k_\delta}^\delta - x^\dagger\| \leq c\left\{\frac{\delta}{\alpha_k} + \|x_k - x^\dagger\|\right\} \quad \forall k > N. \qquad (5.14)$$

Case (ii) $0 \leq k \leq N$. To prove the result in this case we will show that, for all integers $k \leq N$,

$$\|x_{k_\delta} - x^\dagger\| + \frac{\delta}{\alpha_{k_\delta}} \leq \tilde{c}\left\{\|x_k - x^\dagger\| + \frac{\delta}{\alpha_k}\right\},$$

for some constant $c > 0$. We further divide this case into two subcases.

Case (iia) $k_\delta \leq k \leq N$. From Lemma 5.3, we get

$$\|\alpha_{k_\delta}(A + \alpha_{k_\delta}I)^{-1}(F(x_{k_\delta}) - y)\| \leq c_1\delta.$$

Now from Lemma 4.8,

$$\|x_{k_\delta} - x^\dagger\| \leq c_{1/3} \left\{ \|x_k - x^\dagger\| + \frac{\|\alpha_{k_\delta}(A + \alpha_{k_\delta}I)^{-1}(F(x_{k_\delta}) - y)\|}{\alpha_k} \right\}$$

$$\leq c_{1/3} \left\{ \|x_k - x^\dagger\| + c_1 \frac{\delta}{\alpha_k} \right\} \leq c_{1/3}\, c_1 \left\{ \|x_k - x^\dagger\| + \frac{\delta}{\alpha_k} \right\}.$$

So, we get

$$\|x_{k_\delta} - x^\dagger\| + \frac{\delta}{\alpha_{k_\delta}(1 - \kappa)} \leq c_{1/3}\, c_1 \left\{ \|x_k - x^\dagger\| + \frac{\delta}{\alpha_k} \right\} + \frac{\delta}{\alpha_{k_\delta}(1 - \kappa)}.$$

We note that $\delta/\alpha_{k_\delta}(1 - \kappa) \leq \delta/\alpha_k(1 - \kappa)$. Hence we have

$$\|x_{k_\delta} - x^\dagger\| + \frac{\delta}{\alpha_{k_\delta}(1 - \kappa)} \leq \frac{1 + c_{1/3}\, c_1}{(1 - \kappa)} \left\{ \|x_k - x^\dagger\| + \frac{\delta}{\alpha_k} \right\}. \quad (5.15)$$

Case (iib) $k \leq k_\delta$. First we show that the map $\alpha \to \|\alpha(A + \alpha I)^{-1}(x_0 - x^\dagger)\|$ is monotonically increasing on $[0, \infty)$.

Let $\alpha \leq \beta$ and E_λ be the spectral family generated by $A := F'(x^\dagger)$. Then,

$$\|\alpha(A + \alpha I)^{-1}(x_0 - x^\dagger)\|^2 = \int_0^{\|A\|} \frac{\alpha^2}{(\alpha + \lambda)^2}\, d\|E_\lambda(x_0 - x^\dagger)\|^2$$

$$\leq \int_0^{\|A\|} \frac{\beta^2}{(\beta + \lambda)^2}\, d\|E_\lambda(x_0 - x^\dagger)\|^2$$

$$= \|\beta(A + \beta I)^{-1}(x_0 - x^\dagger)\|^2,$$

which shows that the map $\alpha \to \|\alpha(\alpha I + A)^{-1}(x_0 - x^\dagger)\|$ is monotonically increasing on $(0, \infty)$. Hence by (4.19), for all $l \geq m$,

$$\|x_l - x^\dagger\| \leq \tfrac{4}{3}\mu\|\alpha_l(A + \alpha_l I)^{-1}(x_0 - x^\dagger)\|$$

$$\leq \tfrac{4}{3}\mu\|\alpha_m(A + \alpha_m I)^{-1}(x_0 - x^\dagger)\| \leq 2\mu\|x_m - x^\dagger\|,$$

which in particular implies that

$$\|x_{k_\delta} - x^\dagger\| \leq 2\mu\|x_k - x^\dagger\|. \quad (5.16)$$

Now from Lemma 5.3, Proposition 2.3 and the estimate $\|x_k^\delta - x^\dagger\| \leq 2\|x_0 - x^\dagger\|$, we get

$$c_2\delta \leq \left\| \alpha_{k_\delta - 1}(A + \alpha_{k_\delta - 1}I)^{-1}\big(F(x_{k_\delta - 1}) - y\big) \right\|$$

$$\leq \left\| \alpha_{k_\delta - 1}(A + \alpha_{k_\delta - 1}I)^{-1}\big(F(x_{k_\delta - 1}) - F(x^\dagger) - F'(x^\dagger)(x_{\kappa_\delta - 1} - x^\dagger) \right.$$

$$\left. + F'(x^\dagger)(x_{k_\delta - 1} - x^\dagger)\big) \right\|$$

$$\leq \alpha_{k_\delta - 1}\left(\frac{1}{2}k_0\|x_{k_\delta - 1} - x^\dagger\|^2 + \|x_{k_\delta - 1} - x^\dagger\| \right)$$

$$\leq \alpha_{k_\delta - 1}\left(1 + k_0 \|x_0 - x^\dagger\|\right)\|x_{k_\delta - 1} - x^\dagger\|$$

$$\leq \mu\alpha_{k_\delta}\left(1 + k_0 \|x_0 - x^\dagger\|\right)\|x_{k_\delta - 1} - x^\dagger\|$$

$$\leq 2\mu^2\alpha_{k_\delta}\left(1 + k_0 \|x_0 - x^\dagger\|\right)\|x_k - x^\dagger\|.$$

So, we get

$$\frac{\delta}{\alpha_{k_\delta}} \leq 2\mu^2\left(1 + k_0\|x_0 - x^\dagger\|\right)\frac{\|x_k - x^\dagger\|}{c_2}.$$

Using (5.16), we get

$$\|x_{k_\delta} - x^\dagger\| + \frac{\delta}{(1-\kappa)\alpha_{k_\delta}}$$

$$\leq 2\mu\|x_k - x^\dagger\|\left(1 + \frac{\mu}{c_2(1-\kappa)}\left(1 + k_0\|x_0 - x^\dagger\|\right)\right) + \frac{\delta}{\alpha_k}.$$

So,

$$\|x_{k_\delta} - x^\dagger\| + \frac{\delta}{\alpha_{k_\delta}(1-\kappa)}$$

$$\leq 2\mu\left(1 + \frac{\mu}{c_2(1-\kappa)}\left(1 + k_0\|x_0 - x^\dagger\|\right)\right)\left\{\|x_k - x^\dagger\| + \frac{\delta}{\alpha_k}\right\}. \quad (5.17)$$

From (5.14), (5.15) and (5.17), we get

$$\|x_{k_\delta}^\delta - x^\dagger\| \leq \xi \inf\left\{\|x_k - x^\dagger\| + \frac{\delta}{\alpha_k} : k = 0, 1, \ldots\right\},$$

where

$$\xi = \max\left\{2\mu\left(1 + \frac{\mu}{c_2(1-\kappa)}\left(1 + k_0\|x_0 - x^\dagger\|\right)\right), \frac{c_{1/3}\,c_1 + 1}{1-\kappa}, c\right\}. \qquad \square$$

6. Order optimal result with an a posterior stopping rule

Now we are in a position to give the main result of this paper. In the following theorem we show the convergence $x_{k_\delta} \to x^\dagger$ as $\delta \to 0$ and also give an optimal error estimate for $\|x_{k_\delta} - x^\dagger\|$.

THEOREM 6.1. *Let the assumptions of Theorem* 5.4 *hold and let* k_δ *be the integer chosen by* (3.1). *If* x_0 *is chosen such that* $x_0 - x^\dagger \in N(F'(x^\dagger))^\perp$, *then we have that* $\lim_{\delta \to 0} x_{k_\delta}^\delta = x^\dagger$. *Moreover, if Assumptions* 2.4 *and* 2.5 *are satisfied, then*

$$\|x_{k_\delta}^\delta - x^\dagger\| \leq \xi'\mu c_\varphi\|v\|\psi^{-1}\left(\frac{\delta}{\|v\|c_\varphi}\right),$$

where $\xi' := 8\mu\xi/3$ *with* ξ *as in Theorem* 5.4.

PROOF. From (5.12),

$$\|x_{k_\delta}^\delta - x^\dagger\| \leq \xi \inf\left\{\|x_k - x^\dagger\| + \frac{\delta}{\alpha_k} : k = 0, 1, \ldots\right\}. \tag{6.1}$$

Now, using the estimate $\|x_k - x^\dagger\| \leq 4\mu\|\alpha_k(A + \alpha_k I)^{-1}(x_0 - x^\dagger)\|/3$ from (4.19) in (6.1), we get

$$\|x_{k_\delta}^\delta - x^\dagger\| \leq \xi' \inf\left\{\|\alpha_k(A + \alpha_k I)^{-1}(x_0 - x^\dagger)\| + \frac{\delta}{\alpha_k} : k = 0, 1, \ldots\right\}, \tag{6.2}$$

where

$$\xi' = \frac{4\mu}{3} \max\left\{2\mu\left(1 + \frac{\mu}{c_2(1 - \kappa)}(1 + k_0\|x_0 - x^\dagger\|)\right), \frac{c_{1/3}\, c_1 + 1}{(1 - \kappa)}, c\right\}.$$

We choose an integer m_δ such that $m_\delta = \max\{k : \alpha_k \geq \sqrt{\delta}\}$. Then we have

$$\|x_{k_\delta}^\delta - x^\dagger\| \leq \xi'\left\{\|\alpha_{m_\delta}(A + \alpha_{m_\delta}I)^{-1}(x_0 - x^\dagger)\| + \frac{\delta}{\alpha_{m_\delta}}\right\}. \tag{6.3}$$

We note that

$$\frac{\delta}{\alpha_{m_\delta}} \leq \sqrt{\delta} \quad \text{and} \quad \frac{\delta}{\alpha_{m_\delta}} \to 0 \text{ as } \delta \to 0.$$

To show that

$$\|\alpha_{m_\delta}(A + \alpha_{m_\delta}I)^{-1}(x_0 - x^\dagger)\| \to 0 \quad \text{as } \delta \to 0$$

it is enough to prove that $\|\alpha_{m_\delta}(A + \alpha_{m_\delta}I)^{-1}w\| \to 0$ for every $w \in R(F'(x^\dagger))$. We observe that, for every $w \in R(F'(x^\dagger))$, we have $w = F'(x^\dagger)u$ for some $u \in D(F)$,

$$\|\alpha_{m_\delta}(A + \alpha_{m_\delta}I)^{-1}(x_0 - x^\dagger)\| \leq \alpha_{m_\delta}\|u\| \to 0 \quad \text{as } \delta \to 0.$$

Now using the fact that $\|\alpha_{m_\delta}(A + \alpha_{m_\delta}I)^{-1}\| \leq 1$ and for $R(F'(x^\dagger))$ a dense subspace of $N(F'(x^\dagger))^\perp$, we get

$$\|\alpha_{m_\delta}(A + \alpha_{m_\delta}I)^{-1}(x_0 - x^\dagger)\| \to 0 \quad \text{as } \delta \to 0.$$

Now from (6.3),

$$\|x_{k_\delta}^\delta - x^\dagger\| \leq \xi'\left\{\|\alpha_{m_\delta}(A + \alpha_{m_\delta}I)^{-1}(x_0 - x^\dagger)\| + \frac{\delta}{\alpha_{m_\delta}}\right\} \to 0 \quad \text{as } \delta \to 0.$$

Hence, we get $\lim_{\delta \to 0} x_{k_\delta}^\delta = x^\dagger$. Now using Assumptions 2.4 and 2.5,

$$\|\alpha_k(A + \alpha_k I)^{-1}(x_0 - x^\dagger)\| \leq c_\varphi \varphi(\alpha_k)\|v\|. \tag{6.4}$$

Using (6.4) in (6.2), we get

$$\|x_{k_\delta}^\delta - x^\dagger\| \leq \xi' \inf\left\{c_\varphi \varphi(\alpha_k)\|v\| + \frac{\delta}{\alpha_k} : k = 0, 1, \ldots\right\}. \tag{6.5}$$

We choose $\hat{k}_\delta$ such that

$$\varphi(\alpha_{\hat{k}_\delta})\alpha_{\hat{k}_\delta} \leq \frac{\delta}{\|v\|c_\varphi} < \varphi(\alpha_k)\alpha_k \quad \text{for } k = 0, 1, \ldots, \hat{k}_\delta - 1 \qquad (6.6)$$

which also implies that

$$\psi\varphi(\hat{k}_\delta) \leq \frac{\delta}{\|v\|c_\varphi} < \psi\varphi(\alpha_k) \quad \text{for } k = 0, 1, \ldots, \hat{k}_\delta - 1, \qquad (6.7)$$

where $\psi : (0, \varphi(a)] \to (0, a\varphi(a)]$ is defined as

$$\psi(\lambda) := \lambda\varphi^{-1}(\lambda), \quad \lambda \in (0, \varphi(a)].$$

From (6.5),

$$\|x_{k_\delta}^\delta - x^\dagger\| \leq \xi' \left\{ c_\varphi \varphi(\alpha_{\hat{k}_\delta})\|v\| + \frac{\delta}{\alpha_{\hat{k}_\delta}} \right\}.$$

Using (6.6), (6.7) and the definition of ψ, we get

$$\|x_{k_\delta}^\delta - x^\dagger\| \leq 2\xi' \frac{\delta}{\alpha_{\hat{k}_\delta}} \leq 2\xi'\mu \frac{\delta}{\alpha_{\hat{k}_\delta-1}} \leq 2\xi'\mu \frac{\delta}{\varphi^{-1}\psi^{-1}(\delta/c_\varphi\|v\|)}$$

$$= 2\xi'\mu c_\varphi \|v\| \psi^{-1}\left(\frac{\delta}{c_\varphi\|v\|}\right),$$

which completes the proof. $\qquad\qquad\square$

7. Concluding remarks

In this paper, we have considered an iterative regularization procedure for nonlinear ill-posed monotone operator equations. We have introduced a stopping criterion analogous to the one considered by Qi-Nian Jin in [10] for the *iteratively regularized Gauss–Newton method*. We have shown that this stopping criterion not only ensures convergence, but also gives an order optimal error estimate. Here, an optimal error estimate has been obtained under a general source condition. Hence, the results of this paper can be applied to the larger class of ill-posed problems including both severely as well as mildly ill-posed problems (see [8, 9]).

References

[1] A. Bakushinsky and A. Smirnova, "On application of generalized discrepancy principle to iterative methods for nonlinear ill-posed problems", *Numer. Funct. Anal. Optim.* **26** (2005) 35–48.

[2] A. Bakushinsky and A. Smirnova, "A posteriori stopping rule for regularized fixed point iterations", *Nonlinear Anal.* **64** (2006) 1255–1261.

[3] A. Bakushinsky and A. Smirnova, "Iterative regularization and generalized discrepancy principle for monotone operator equations", *Numer. Funct. Anal. Optim.* **28** (2007) 13–25.

[4] A. Binder, H. W. Engl, C. W. Groetsch, A. Neubauer and O. Scherzer, "Weakly closed nonlinear operators and parameter identification in parabolic equations by Tikhonov regularization", *Appl. Anal.* **55** (1994) 215–235.

[5] K. Deimling, *Nonlinear functional analysis* (Springer-Verlag, New York, 1985).

[6] H. W. Engl, M. Hanke and A. Neubauer, *Regularization of inverse problems,* Volume 375 of *Mathematics and its Applications* (Kluwer, Dordrecht, 1996).

[7] H. W. Engl, K. Kunish and A. Neubauer, "Convergent rates for Tikhonov regularisation of nonlinear ill-posed problems", *Inverse Problems* **5** (1989) 523–540.

[8] T. Hohage, "Iterative methods in inverse obstacle scattering: regularization theory of linear and nonlinear exponentially ill-posed problems" Ph. D. Thesis, University of Linz, 1999.

[9] T. Hohage, "Regularization of exponentially ill-posed problem", *Numer. Funct. Anal. Optim.* **21** (2000) 439–464.

[10] Q. Jin, "On the iteratively regularized Gauss–Newton method for solving nonlinear ill-posed problems", *Math. Comp.* **69** (2000) 1603–1623.

[11] Q. Jin and Zong-Yi Hou, "On the choice of the regularization parameter for ordinary and iterated Tikhonov regularization of nonlinear illposed problems", *Inverse Problems* **13** (1997) 815–827.

[12] Q. Jin and Zong-Yi Hou, "On an a posteriori parameter choice strategy for Tikhonov regularization of nonlinear ill-posed problems", *Numer. Math.* **83** (1999) 139–159.

[13] P. Mahale and M. T. Nair, "General source conditions for nonlinear ill-posed equations", *Numer. Funct. Anal. Optim.* **28** (2007) 111–126.

[14] M. T. Nair and U. Tautenhahn, "Lavrentiev regularization for linear ill-posed problems under general source conditions", *Z. Anal. Anwendungen* **23** (2004) 167–185.

[15] O. Scherzer, H. W. Engl and K. Kunisch, "Optimal a posteriori parameter choice for Tikhonov regularization for solving nonlinear ill-posed problems", *SIAM J. Numer. Anal.* **30** (1993) 1796–1838.

[16] U. Tautenhahn, "On the method of Lavrentiev regularization for nonlinear ill-posed problems", *Inverse Problems* **18** (2002) 191–207.

[17] U. Tautenhahn, "Lavrentiev regularization of nonlinear ill-posed problems", *Vietnam J. Math.* **32** (2004) 29–41.

[18] U. Tautenhahn and Q. Jin, "Tikhonov regularization and a posteriori rules for solving nonlinear ill posed problems", *Inverse Problems* **19** (2003) 1–21.