MATHEMATICS OF COMPUTATION
Volume 79, Number 272, October 2010, Pages 2191–2211
S 0025-5718(10)02359-8
Article electronically published on April 20, 2010

ON A CLASS OF FROZEN REGULARIZED GAUSS-NEWTON METHODS FOR NONLINEAR INVERSE PROBLEMS

QINIAN JIN

ABSTRACT. In this paper we consider a class of regularized Gauss-Newton methods for solving nonlinear inverse problems for which an a posteriori stopping rule is proposed to terminate the iteration. Such methods have the frozen feature that they require only the computation of the Fréchet derivative at the initial approximation. Thus the computational work is considerably reduced. Under certain mild conditions, we give the convergence analysis and derive various estimates, including the order optimality, on these methods.

1. INTRODUCTION

Nonlinear inverse problems arise from many practical applications that include inverse source problems, inverse scattering problems, tomographies, and parameter identifications in partial differential equations; see [4, 5, 6, 8]. Mathematically, such a problem usually is formulated as the problem of finding a solution $x^\dagger$ of the operator equation

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

where $F : D(F) \subset X \mapsto Y$ is a Fréchet differentiable nonlinear operator between two Hilbert spaces X and Y with domain $D(F)$, Throughout this paper $\|\cdot\|$ and $(\cdot,\cdot)$ will be used to denote the norms and inner products, respectively, for both the spaces X and Y since there is no confusion. The Fréchet derivative of F at $x \in D(F)$ and its adjoint will be denoted as $F'(x)$ and $F'(x)^*$, respectively. It is known that if $F'(x^\dagger) : X \to Y$ is an injective map with a closed range, then (1.1) possesses the local uniqueness and Lipschitz stability; see [15, Theorem 1] for instance. Unfortunately, the closed range condition on $F'(x^\dagger)$ is rarely satisfied since $F'(x^\dagger)$ is compact in general. In fact, a characteristic property of most inverse problems is their ill-posedness in the sense that their solutions do not depend continuously on the data. Since the right-hand side of (1.1) is usually obtained by measurement, thus, instead of y itself, the available data is an approximation y^δ satisfying

$$\|y^\delta - y\| \le \delta \tag{1.2}$$

with a given small noise level $\delta > 0$. Due to the ill-posedness, the computation of a stable approximation to $x^\dagger$ from y^δ becomes an important issue, and the regularization techniques should be taken into account.

Received by the editor September 26, 2008 and, in revised form, June 1, 2009.
2010 *Mathematics Subject Classification.* Primary 65J15, 65J20.
Key words and phrases. Nonlinear inverse problems, frozen regularized Gauss-Newton method, a posteriori stopping rule, convergence, order optimality.

Many regularization methods have been considered to solve (1.1) in the last two decades. Due to the straightforward implementation, iterative methods are attractive for solving nonlinear inverse problems; an overview can be found in the recent book [13]. The general regularized Gauss-Newton method, which defines the iterative solutions $\{x_k^\delta\}$ successively by

$$x_{k+1}^\delta = x_0 - g_{\alpha_k}\left(F'(x_k^\delta)^* F'(x_k^\delta)\right) F'(x_k^\delta)^* \left(F(x_k^\delta) - y^\delta - F'(x_k^\delta)(x_k^\delta - x_0)\right), \tag{1.3}$$

has been considered in several references ([2, 12, 11]), where $x_0^\delta := x_0$ is an initial guess of $x^\dagger$, $\{\alpha_k\}$ is a given sequence of numbers such that

$$\alpha_k > 0, \qquad 1 \le \frac{\alpha_k}{\alpha_{k+1}} \le r \qquad \text{and} \qquad \lim_{k\to\infty} \alpha_k = 0 \tag{1.4}$$

for some constant $r > 1$, and $g_\alpha : [0,\infty) \to (-\infty,\infty)$ is a family of piecewise continuous functions satisfying certain structure conditions.

In order for the method (1.3) to be useful for solving (1.1), the iteration must be terminated properly, that is, a stopping index k_δ must be chosen so that $x_{k_\delta}^\delta$ is indeed a good approximation to $x^\dagger$. Due to the practical applications, a posteriori rules, which use only quantities that arise during computation, should be considered to choose the stopping index of iteration. In our recent paper [11] the discrepancy principle

$$\|F(x_{k_\delta}^\delta) - y^\delta\| \le \tau\delta < \|F(x_k^\delta) - y^\delta\|, \quad 0 \le k < k_\delta, \tag{1.5}$$

with $\tau > 1$, which is widely used in the literature of regularization theory for ill-posed problems, has been considered for the general method (1.3). Several useful results, concerning the approximation of $x_{k_\delta}^\delta$ to $x^\dagger$, were obtained; in particular, it was shown that order optimality can be obtained under merely the Lipschitz condition on F' if $x_0 - x^\dagger$ is smooth enough.

The method (1.3) with $g_\alpha(\lambda) = (\alpha+\lambda)^{-1}$ together with (1.5) has been considered in [3, 7]. Note that when $g_\alpha(\lambda) = (\alpha+\lambda)^{-1}$, the method (1.3) becomes

$$x_{k+1}^\delta = x_k^\delta - \left(\alpha_k I + F'(x_k^\delta)^* F'(x_k^\delta)\right)^{-1} \left(F'(x_k^\delta)^*(F(x_k^\delta) - y^\delta) + \alpha_k(x_k^\delta - x_0)\right) \tag{1.6}$$

which is the iteratively regularized Gauss-Newton method (see [1]). It is known that the best possible rate of convergence for the method defined by (1.6) and (1.5) is $O(\delta^{1/2})$. In order to prevent such saturation, we proposed in [9] an alternative a posteriori stopping rule to choose the stopping index k_δ as the first integer satisfying

$$\alpha_{k_\delta}^{1/2} \left\| \left(\alpha_{k_\delta} I + F'(x_{k_\delta}^\delta) F'(x_{k_\delta}^\delta)^*\right)^{-1/2} \left(F(x_{k_\delta}^\delta) - y^\delta\right) \right\| \le \tau\delta, \tag{1.7}$$

where $\tau > 1$ is a given number. The careful convergence analysis has been given in [9, 10]. It is even shown that the method defined by (1.6) and (1.7) is order optimal under merely the Lipschitz condition on F' if $x_0 - x^\dagger$ is smooth enough.

Note that the method (1.3) requires calculating the Fréchet derivative of F at each iteration, which needs a considerable amount of computational work, and thus makes the method expensive. In order to reduce the computational work, in this paper we will consider the iterative methods which define the iterates $\{x_k^\delta\}$ by

$$x_{k+1}^\delta = x_0 - g_{\alpha_k}(A_0^* A_0) A_0^* \left(F(x_k^\delta) - y^\delta - A_0(x_k^\delta - x_0)\right), \tag{1.8}$$

where $A_0 := F'(x_0)$ is the Fréchet derivative of F at the initial guess x_0. Such methods are called the frozen regularized Gauss-Newton methods since the Fréchet derivative is held at x_0 throughout the iteration process. In order to terminate the iteration (1.8) properly, we need a suitable a posteriori stopping rule. The discrepancy principle (1.5) is certainly a candidate. It turns out, however, that its convergence analysis requires a very restrictive condition on F, thus further investigations are required. The stopping rule (1.7) is also expensive due to the required calculation of the Fréchet derivative at each iteration. Thus, instead of applying it directly, we will consider a frozen version of (1.7) to choose the stopping index of iteration k_δ as the first integer such that

$$\alpha_{k_\delta}^{1/2} \| (\alpha_{k_\delta} I + A_0 A_0^*)^{-1/2} \left(F(x_{k_\delta}^\delta) - y^\delta \right) \| \leq \tau \delta, \tag{1.9}$$

where $\tau > 1$ is a given number.

In this paper we will give a convergence analysis on the method defined by (1.8) and (1.9). Certain conditions should be imposed on $\{g_\alpha\}$, $\{\alpha_k\}$ and F. We start with the assumptions on g_α which is always assumed to be piecewise continuous on $[0, 1/2]$ for each $\alpha > 0$. We set

$$r_\alpha(\lambda) := 1 - \lambda g_\alpha(\lambda), \tag{1.10}$$

which is called the residual function associated with g_α.

Assumption 1.1. (a) There is a positive constant c_1 such that

$$0 \leq r_\alpha(\lambda) \leq 1 \quad \text{and} \quad 0 \leq g_\alpha(\lambda) \leq c_1 \alpha^{-1}$$

for all $\alpha > 0$ and $\lambda \in [0, 1/2]$;

(b) $r_\alpha(\lambda) \leq r_\beta(\lambda)$ for all $0 < \alpha \leq \beta$ and $\lambda \in [0, 1/2]$.

Assumption 1.1 is standard in the analysis of linear regularization methods. It is clear that condition (a) implies, with a constant $c_0 \leq \sqrt{c_1}$,

$$0 \leq g_\alpha(\lambda) \lambda^{1/2} \leq c_0 \alpha^{-1/2} \tag{1.11}$$

for all $\alpha > 0$ and $\lambda \in [0, 1/2]$. In Lemma 2.3 we will give another simple but very useful consequence of Assumption 1.1.

For the sequence of positive numbers $\{\alpha_k\}$, we will always assume that it satisfies (1.4). Moreover, we also need the following condition on $\{\alpha_k\}$ interacting with r_α.

Assumption 1.2. There is a constant $c_2 > 1$ such that

$$r_{\alpha_k}(\lambda) \leq c_2 r_{\alpha_{k+1}}(\lambda)$$

for all k and $\lambda \in [0, 1/2]$.

We remark that for some $\{g_\alpha\}$ Assumption 1.2 is an immediate consequence of (1.4). However, this is not always the case; in some situations, Assumption 1.2 indeed imposes further conditions on $\{\alpha_k\}$. As a rough interpretation, Assumption 1.2 requires for any two successive iterated solutions that the errors do not decrease dramatically. This may be good for the stable numerical implementations of ill-posed problems although it may require more iterations to be performed. It is not yet clear if Assumption 1.2 can be dropped.

Assumption 1.3. (a) $B_\rho(x^\dagger) \subset D(F)$ for some $\rho > 2(1 + c_0/(\tau - 1))\|e_0\|$, where $B_\rho(x^\dagger)$ denotes the ball of radius ρ with center at $x^\dagger$.

(b) The operator A_0 is properly scaled such that $\|A_0\| \leq \beta_0^{1/2}$, where $0 < \beta_0 \leq 1/2$ is such that $r_{\alpha_0}(\lambda) \geq 3/4$ for all $\lambda \in [0, \beta_0]$.

(c) There is a positive constant K_0 such that

$$F'(x) = A_0 R_x \quad \text{and} \quad \|I - R_x\| \leq K_0 \|x - x_0\| \tag{1.12}$$

for all $x \in B_\rho(x^\dagger)$.

Note that $r_{\alpha_0}(0) = 1$, the number β_0 in Assumption 1.3(b) always exists. The scaling condition (b) can always be fulfilled by multiplying the equation (1.1) by a sufficiently small constant.

Theorem 1.1. *Assume that $\tau > 1$ and that $\{\alpha_k\}$, $\{g_\alpha\}$ and F satisfy* (1.4)*, Assumption* 1.1*, Assumption* 1.2 *and Assumption* 1.3*. There exist positive constants η_0 and C depending only on c_1, c_2, τ and r, such that if $K_0\|x_0 - x^\dagger\| \le \eta_0$, then the method given by* (1.8) *and* (1.9) *is well defined and*

$$\|x_{k_\delta}^\delta - x^\dagger\| \le C \inf \left\{ \|r_{\alpha_k}(A_0^* A_0)(x_0 - x^\dagger)\| + \frac{\delta}{\sqrt{\alpha_k}} : k = 0, 1, \cdots \right\}$$

for the integer k_δ determined by the stopping rule (1.9)*.*

Although Theorem 1.1 is an important result on the method defined by (1.8) and (1.9), it does not imply the convergence of $x_{k_\delta}^\delta$ to $x^\dagger$ if there are no further conditions on $\{g_\alpha\}$ (for instance, consider $g_\alpha \equiv 0$ for all $\alpha > 0$). In order to derive the convergence and rate of convergence, we need the following additional but standard condition.

Assumption 1.4. There exists $\bar{\nu} > 0$ such that for every $0 < \nu \le \bar{\nu}$ there is a constant d_ν such that

$$r_\alpha(\lambda)\lambda^\nu \le d_\nu \alpha^\nu$$

for all $\alpha > 0$ and $\lambda \in [0, 1/2]$.

According to [17], the largest number $\bar{\nu} > 0$ such that Assumption 1.4 holds is called the qualification of the linear regularization method defined by $\{g_\alpha\}$.

Corollary 1.1. *Let Assumption* 1.4 *and all the conditions in Theorem* 1.1 *be fulfilled, let $\{x_k^\delta\}$ be defined by* (1.8)*, and let k_δ be the integer defined by the stopping rule* (1.9) *with $\tau > 1$.*

(i) *If $x_0 - x^\dagger \in \mathcal{N}(A_0)^\perp$, then $\lim_{\delta\to 0} x_{k_\delta}^\delta = x^\dagger$.*

(ii) *If $x_0 - x^\dagger = (A_0^* A_0)^\nu \omega$ for some $\omega \in X$ and $0 < \nu \le \bar{\nu}$, then*

$$\|x_{k_\delta}^\delta - x^\dagger\| \le C_\nu \|\omega\|^{1/(1+2\nu)} \delta^{2\nu/(1+2\nu)},$$

where C_ν is a constant depending only on c_1, c_2, r, τ and ν.

(iii) *If $x_0 - x^\dagger = (-\ln(A_0^* A_0))^{-\mu} \omega$ for some $\omega \in X$ and $\mu > 0$, then*

$$\|x_{k_\delta}^\delta - x^\dagger\| \le C_\mu \|\omega\| \left(1 + \left|\ln \frac{\delta}{\|\omega\|}\right|\right)^{-\mu},$$

where C_μ is a constant depending only on c_1, c_2, r, τ and μ.

In the statement of the main results, the smallness condition on $K_0\|x_0 - x^\dagger\|$ is not specified. In Section 2, however, we will spell out all the necessary smallness conditions during the proof. Our proof is based on a simple consequence of Assumption 1.1 given in Lemma 2.3 which enables us to prove the important inequality in Lemma 2.4. The source conditions in (ii) and (iii) of Corollary 1.1 are called the Hölder type source conditions and the logarithmic type source conditions, which are important for dealing with mildly, and respectively, severely ill-posed problems. In Section 3 we will consider some variants of the above method. A numerical example is reported in Section 4 to test the theoretical results given by Theorem 1.1 and Corollary 1.1.

The use of frozen Newton methods is well understood for well-posed problems and its advantages are explored in numerical experiments for a wide variety of inverse problems. The frozen methods reduce the computational work considerably. By choosing x_0 suitably, the computation of A_0 can be easily handled and even an explicit formula can be obtained. Moreover, the convergence analysis of (1.8) and (1.9) can be carried out under quite mild conditions on F. Its obvious disadvantage is that inevitably more iterations are required which, however, can be offset considering the numerous advantages.

The frozen Gauss-Newton method (1.8) was considered previously in [12] under Hölder type source conditions on $x_0 - x^\dagger$ in which the iteration is terminated as long as

$$\max\left\{\|F(x_{k-1}^\delta) - y^\delta\|, \|F(x_{k-1}^\delta) - y^\delta - A_0(x_k^\delta - x_{k-1}^\delta)\|\right\} \le \tau\delta$$

is satisfied for the first time, where τ is required to be sufficiently large. Such a method was reconsidered recently in [14] under general type source conditions. The convergence analysis in [12, 14] is based on the condition that

$$F'(x) = R_x A_0 \quad \text{and} \quad \|I - R_x\| \le C_0\|x - x_0\| \tag{1.13}$$

for all $x \in B_\rho(x^\dagger)$. This condition looks similar to (1.12) in Assumption 1.3, they are, however, essentially different. The validity of (1.13), in many situations, requires the commutativity of A_0 with a family of linear operators which is impossible in general. Therefore, (1.13) is a very restrictive condition. The verification of (1.12), however, turns out to be much easier and indeed it has been checked for a wide variety of nonlinear inverse problems in the literature.

2. Proof of the main result

We first give some simple but useful consequences of Assumption 1.3. From Assumption 1.3 it follows for any $x, z \in B_\rho(x^\dagger)$ that

$$\begin{aligned} F(x) - F(z) - A_0(x-z) &= \int_0^1 \left[F'(tx + (1-t)z) - A_0\right](x-z)dt \\ &= \int_0^1 A_0\left[R_{tx+(1-t)z} - I\right](x-z)dt. \end{aligned}$$

This together with Assumption 1.1(a) then implies

$$\begin{aligned} &\|g_\alpha(A_0^*A_0)A_0^*\left(F(x) - F(z) - A_0(x-z)\right)\| \\ &\qquad \le \int_0^1 \|g_\alpha(A_0^*A_0)A_0^*A_0\left[R_{tx+(1-t)z} - I\right](x-z)\|dt \\ &\qquad \le K_0\int_0^1 \left(t\|x - x_0\| + (1-t)\|z - x_0\|\right)dt\|x - z\|. \end{aligned}$$

Therefore, for any $x, z \in B_\rho(x^\dagger)$ and $\alpha > 0$ there holds

$$\begin{aligned} &\|g_\alpha(A_0^*A_0)A_0^*(F(x) - F(z) - A_0(x-z))\| \\ &\qquad \le \frac{1}{2}K_0\left(\|x - x_0\| + \|z - x_0\|\right)\|x - z\|. \end{aligned} \tag{2.1}$$

Note that $\|(\alpha I + A_0 A_0^*)^{-1/2} A_0\| \le 1$ for $\alpha > 0$, by a similar argument we also have

$$\begin{aligned} &\|(\alpha I + A_0 A_0^*)^{-1/2}(F(x) - F(z) - A_0(x-z))\| \\ &\qquad \le \frac{1}{2} K_0 \left(\|x - x_0\| + \|z - x_0\|\right) \|x - z\| \end{aligned} \tag{2.2}$$

for any $x, z \in B_\rho(x^\dagger)$ and $\alpha > 0$.

Now we are in a position to show that the method given by (1.8) and (1.9) is well defined under the conditions in Theorem 1.1. Without loss of generality, we may assume that $x_0 \ne x^\dagger$. We introduce the integer $\tilde{k}_\delta$ satisfying

$$\alpha_{\tilde{k}_\delta} \le \left(\frac{(\tau - 1)\delta}{2\|x_0 - x^\dagger\|} \right)^2 < \alpha_k, \quad 0 \le k < \tilde{k}_\delta. \tag{2.3}$$

Since $\tau > 1$ and $\{\alpha_k\}$ satisfies (1.4), such $\tilde{k}_\delta$ exists and is finite. In the following we will show that

$$\|x_k^\delta - x^\dagger\| \le 2 \left(1 + \frac{c_0}{\tau - 1} \right) \|x_0 - x^\dagger\| < \rho \tag{2.4}$$

for all $0 \le k \le \tilde{k}_\delta$ and

$$k_\delta \le \tilde{k}_\delta \tag{2.5}$$

for the integer k_δ defined by the stopping rule (1.9).

For the simplicity of presentation, we set $e_0 := x_0 - x^\dagger$ and $e_k^\delta := x_k^\delta - x^\dagger$. In order to show (2.4), we assume that $x_k^\delta \in B_\rho(x^\dagger)$ for some $0 \le k < \tilde{k}_\delta$. Then it follows from (1.8) that

$$e_{k+1}^\delta = r_{\alpha_k}(A_0^* A_0) e_0 - g_{\alpha_k}(A_0^* A_0) A_0^* \left(F(x_k^\delta) - y^\delta - A_0 e_k^\delta \right). \tag{2.6}$$

Applying Assumption 1.1(a), (1.2) and (2.1) we obtain

$$\|e_{k+1}^\delta\| \le \|e_0\| + c_0 \delta \alpha_k^{-1/2} + K_0 \|e_0\| \|e_k^\delta\| + \frac{1}{2} K_0 \|e_k^\delta\|^2.$$

Since $0 \le k < \tilde{k}_\delta$, we have $\delta \alpha_k^{-1/2} \le \frac{2}{\tau - 1} \|e_0\|$. Therefore,

$$\|e_{k+1}^\delta\| \le \left(1 + \frac{2c_0}{\tau - 1} \right) \|e_0\| + K_0 \|e_0\| \|e_k^\delta\| + \frac{1}{2} K_0 \|e_k^\delta\|^2.$$

Thus, if $K_0 \|e_0\|$ is so small that

$$4 \left(1 + \frac{c_0}{\tau - 1} \right) \left(2 + \frac{c_0}{\tau - 1} \right) K_0 \|e_0\| \le 1, \tag{2.7}$$

then we can conclude (2.4) by an induction argument.

Next we show (2.5). It suffices to show that

$$\alpha_{\tilde{k}_\delta}^{1/2} \|(\alpha_{\tilde{k}_\delta} I + A_0 A_0^*)^{-1/2} (F(x_{\tilde{k}_\delta}^\delta) - y^\delta)\| \le \tau \delta. \tag{2.8}$$

We denote by $\tilde{d}(\delta)$ the left-hand side of the above inequality. If $\tilde{k}_\delta = 0$, then its definition implies $\alpha_0^{1/2} \le (\tau - 1)\delta / (2\|e_0\|)$. Thus, by (1.2), (2.2) and the smallness condition (2.7), we have

$$\tilde{d}(\delta) \le \delta + \alpha_0^{1/2} \|e_0\| + \frac{1}{2} K_0 \|e_0\|^2 \alpha_0^{1/2} \le \tau \delta.$$

Therefore, we may assume $\tilde{k}_\delta > 0$. It then follows from (2.2) that

$$\begin{aligned}\tilde{d}(\delta) \le& \|\alpha_{\tilde{k}_\delta}^{1/2}(\alpha_{\tilde{k}_\delta} I + A_0 A_0^*)^{-1/2}(A_0 e_{\tilde{k}_\delta}^\delta - y^\delta + y)\| \\ &+ \frac{1}{2}\alpha_{\tilde{k}_\delta}^{1/2} K_0 \left(2\|e_0\| + \|e_{\tilde{k}_\delta}^\delta\|\right)\|e_{\tilde{k}_\delta}^\delta\|.\end{aligned}$$

Note that (2.6) implies

$$\begin{aligned}A_0 e_{\tilde{k}_\delta}^\delta - y^\delta + y =& A_0 r_{\alpha_{\tilde{k}_\delta - 1}}(A_0^* A_0) e_0 + r_{\alpha_{\tilde{k}_\delta - 1}}(A_0 A_0^*)(y - y^\delta) \\ &- A_0 g_{\alpha_{\tilde{k}_\delta - 1}}(A_0^* A_0) A_0^* \left(F(x_{\tilde{k}_\delta - 1}^\delta) - y - A_0 e_{\tilde{k}_\delta - 1}^\delta\right).\end{aligned}$$

Thus we may use Assumption 1.1(a) and (2.1) to conclude

$$\begin{aligned}\tilde{d}(\delta) \le& \delta + \alpha_{\tilde{k}_\delta}^{1/2}\|e_0\| + \frac{1}{2}\alpha_{\tilde{k}_\delta}^{1/2} K_0 \left(2\|e_0\| + \|e_{\tilde{k}_\delta - 1}^\delta\|\right)\|e_{\tilde{k}_\delta - 1}^\delta\| \\ &+ \frac{1}{2}\alpha_{\tilde{k}_\delta}^{1/2} K_0 \left(2\|e_0\| + \|e_{\tilde{k}_\delta}^\delta\|\right)\|e_{\tilde{k}_\delta}^\delta\|.\end{aligned}$$

By using the definition of $\tilde{k}_\delta$, the estimate (2.4) and the smallness condition (2.7) we obtain

$$\tilde{d}(\delta) \le \delta + \frac{\tau - 1}{2}\delta + 2(\tau - 1)\left(1 + \frac{c_0}{\tau - 1}\right)\left(2 + \frac{c_0}{\tau - 1}\right) K_0 \|e_0\| \delta \le \tau\delta,$$

which is exactly the inequality (2.8).

Summarizing the above results we obtain

Lemma 2.1. *Assume that $\tau > 1$, and that $\{\alpha_k\}$, $\{g_\alpha\}$ and F satisfy* (1.4)*, Assumption* 1.1(a) *and Assumption* 1.3*. If $K_0\|e_0\|$ satisfies the smallness condition* (2.7)*, then the method given by* (1.8) *and* (1.9) *is well defined. Moreover, the estimates* (2.4) *and* (2.5) *hold, where $\tilde{k}_\delta$ is defined by* (2.3)*.*

Remark 2.1. When $\{\alpha_k\}$ is chosen as $\alpha_k = \alpha_0 r^{-k}$ for some $r > 1$, it is easy to see that the integer $\tilde{k}_\delta$ defined by (2.3) satisfies $\tilde{k}_\delta \le O(1 + |\log \delta|)$. Consequently, by (2.5), the integer k_δ determined by the stopping rule (1.9) satisfies $k_\delta \le O(1 + |\log \delta|)$. This indicates that, for such a choice of $\{\alpha_k\}$, the method given by (1.8) and (1.9) has the fast convergence feature.

Next we will derive some estimates on the noise-free iterates $\{x_k\}$ defined by

$$x_{k+1} = x_0 - g_{\alpha_k}(A_0^* A_0) A_0^* \left(F(x_k) - y - A_0(x_k - x_0)\right).$$

We set $e_k := x_k - x^\dagger$. It is easy to see that

$$e_{k+1} = r_{\alpha_k}(A_0^* A_0) e_0 - g_{\alpha_k}(A_0^* A_0) A_0^* \left(F(x_k) - y - A_0 e_k\right). \tag{2.9}$$

Thus, if $x_k \in B_\rho(x^\dagger)$, then it follows from (2.1) that

$$\|e_{k+1} - r_{\alpha_k}(A_0^* A_0) e_0\| \le K_0 \|e_0\| \|e_k\| + \frac{1}{2} K_0 \|e_k\|^2. \tag{2.10}$$

Since (2.7) implies $8K_0\|e_0\| \le 1$ and Assumption 1.1(a) implies $\|r_{\alpha_k}(A_0^* A_0) e_0\| \le \|e_0\|$, we can show by induction that $\{x_k\}$ is well defined and

$$\|e_k\| \le (7 - \sqrt{33})\|e_0\| \le \frac{3}{2}\|e_0\| < \rho \qquad \text{for all } k \ge 0. \tag{2.11}$$

Moreover, we have

Lemma 2.2. *Assume that* $\{\alpha_k\}$, $\{g_\alpha\}$ *and* F *satisfy* (1.4), *Assumption* 1.1, *Assumption* 1.2 *and Assumption* 1.3. *If* $7c_2K_0\|x_0 - x^\dagger\| \le 1$, *then*

$$\frac{2}{3}\|r_{\alpha_k}(A_0^*A_0)e_0\| \le \|e_k\| \le \frac{4}{3}c_2\|r_{\alpha_k}(A_0^*A_0)e_0\| \tag{2.12}$$

and

$$\frac{1}{2c_2}\|e_k\| \le \|e_{k+1}\| \le 2\|e_k\| \tag{2.13}$$

for all $k \ge 0$.

Proof. By using (2.11) and the condition $7c_2K_0\|e_0\| \le 1$, we can obtain from (2.10) that

$$\|e_{k+1} - r_{\alpha_k}(A_0^*A_0)e_0\| \le \frac{7}{4}K_0\|e_0\|\|e_k\| \le \frac{1}{4c_2}\|e_k\|. \tag{2.14}$$

Note that Assumption 1.2 implies

$$\|r_{\alpha_k}(A_0^*A_0)e_0\| \le c_2\|r_{\alpha_{k+1}}(A_0^*A_0)e_0\| \tag{2.15}$$

and that Assumption 1.3(b) implies $\|r_{\alpha_0}(A_0^*A_0)e_0\| \ge \frac{3}{4}\|e_0\|$, thus by induction we may conclude from (2.14) that

$$\|e_k\| \le \frac{4}{3}c_2\|r_{\alpha_k}(A_0^*A_0)e_0\| \quad \text{for all } k. \tag{2.16}$$

Note also that Assumption 1.1(b) implies

$$\|r_{\alpha_{k+1}}(A_0^*A_0)e_0\| \le \|r_{\alpha_k}(A_0^*A_0)e_0\|,$$

we can obtain from (2.14) and (2.16) that

$$\|e_{k+1}\| \ge \|r_{\alpha_k}(A_0^*A_0)e_0\| - \frac{1}{4c_2}\|e_k\| \ge \frac{2}{3}\|r_{\alpha_{k+1}}(A_0^*A_0)e_0\|.$$

Thus we obtain (2.12). The inequality (2.13) is an immediate consequence of (2.14) and (2.12). □

From (2.12) and Assumption 1.1(b) it follows that

$$\|e_l\| \le 2c_2\|e_k\| \quad \text{for all } k \le l. \tag{2.17}$$

It is clear that the reverse inequality does not hold for a convergent method. Lemma 2.4 below, however, will show that a reverse inequality could hold if a certain correction term is added. We need the following consequence of Assumption 1.1.

Lemma 2.3. *Under Assumpion* 1.1 *there holds*

$$0 \le r_\beta(\lambda) - r_\alpha(\lambda) \le c_3\frac{\lambda}{\alpha+\lambda}r_\beta(\lambda) \tag{2.18}$$

for all $0 < \alpha \le \beta$ *and* $\lambda \in [0, 1/2]$, *where* $c_3 := \max\{2, 2c_1\}$.

Proof. We first note that $0 \le r_\beta(\lambda) \le 1$ and Assumption 1.1(b) imply

$$0 \le r_\beta(\lambda) - r_\alpha(\lambda) \le r_\beta(\lambda)\left(1 - r_\alpha(\lambda)\right).$$

By using the definition of r_α and the assumption on g_α we have $1 - r_\alpha(\lambda) \le c_1\lambda\alpha^{-1}$. This together with the fact $1 - r_\alpha(\lambda) \le 1$ implies

$$1 - r_\alpha(\lambda) \le c_3\frac{\lambda}{\alpha+\lambda},$$

and (2.18) thus follows. □

The inequality (2.18) plays a significant role in the proof of Lemma 2.4 below. It is surprising that this simple inequality has never been noticed in the literature of regularization theory.

Lemma 2.4. *Let all the conditions in Lemma 2.2 hold. If, in addition,* $2(2c_2 + 2c_2c_3 + c_3)K_0\|e_0\| \le 1$, *then there is a constant* C *depending only on* c_1 *and* c_2 *such that*

$$\|e_k\| \le C\|e_l\| + \frac{C}{\sqrt{\alpha_l}}\|\alpha_k^{1/2}\left(\alpha_k I + A_0A_0^*\right)^{-1/2}\left(F(x_k) - y\right)\| \tag{2.19}$$

for all $0 \le k \le l$.

Proof. It follows from (2.9) that

$$\begin{aligned} x_k - x_l = &\left[r_{\alpha_{k-1}}(A_0^*A_0) - r_{\alpha_{l-1}}(A_0^*A_0)\right] e_0 \\ &- g_{\alpha_{k-1}}(A_0^*A_0)A_0^*\left(F(x_{k-1}) - y - A_0e_{k-1}\right) \\ &+ g_{\alpha_{l-1}}(A_0^*A_0)A_0^*\left(F(x_{l-1}) - y - A_0e_{l-1}\right). \end{aligned}$$

By using the estimates (2.1) and (2.11) we obtain

$$\begin{aligned} \|x_k - x_l\| \le &\frac{1}{2}K_0(2\|e_0\| + \|e_{k-1}\|)\|e_{k-1}\| + \frac{1}{2}K_0\left(2\|e_0\| + \|e_{l-1}\|\right)\|e_{l-1}\| \\ &+ \|\left[r_{\alpha_{k-1}}(A_0^*A_0) - r_{\alpha_{l-1}}(A_0^*A_0)\right] e_0\| \\ \le &\frac{7}{4}K_0\|e_0\|\left(\|e_{k-1}\| + \|e_{l-1}\|\right) + \mathcal{I}_0, \end{aligned} \tag{2.20}$$

where

$$\mathcal{I}_0 := \|\left[r_{\alpha_{k-1}}(A_0^*A_0) - r_{\alpha_{l-1}}(A_0^*A_0)\right] e_0\|.$$

In order to estimate the term $\mathcal{I}_0$, let $\{E_\lambda\}$ denote the spectral family generated by the self-adjoint operator $A_0^*A_0$. It then follows from (2.18) in Lemma 2.3 that

$$\begin{aligned} \mathcal{I}_0 &= \left(\int_0^{1/2}\left[r_{\alpha_{k-1}}(\lambda) - r_{\alpha_{l-1}}(\lambda)\right]^2 d\|E_\lambda e_0\|^2\right)^{1/2} \\ &\le c_3\left(\int_0^{1/2}(\alpha_{l-1} + \lambda)^{-2}\lambda^2 r_{\alpha_{k-1}}(\lambda)^2 d\|E_\lambda e_0\|^2\right)^{1/2} \\ &= c_3\|(\alpha_{l-1}I + A_0^*A_0)^{-1}A_0^*A_0 r_{\alpha_{k-1}}(A_0^*A_0)e_0\|. \end{aligned}$$

Therefore, by using the inequality $\|(\alpha_{l-1}I + A_0^*A_0)^{-1}A_0^*A_0\| \le 1$, we have

$$\mathcal{I}_0 \le c_3\|e_k - r_{\alpha_{k-1}}(A_0^*A_0)e_0\| + c_3\|\left(\alpha_{l-1}I + A_0^*A_0\right)^{-1}A_0^*A_0e_k\|. \tag{2.21}$$

The combination of (2.20), (2.21) and (2.14) gives

$$\begin{aligned} \|x_k - x_l\| \le &\frac{7}{4}K_0\|e_0\|\left(\|e_{k-1}\| + \|e_{l-1}\|\right) + \frac{7}{4}c_3K_0\|e_0\|\|e_{k-1}\| \\ &+ c_3\|\left(\alpha_{l-1}I + A_0^*A_0\right)^{-1}A_0^*A_0e_k\|. \end{aligned}$$

Thus, by using (2.2) and (2.11), we have

$$
\begin{aligned}
\|x_k - x_l\| \le & \frac{7}{4} K_0 \|e_0\| \left(\|e_{k-1}\| + \|e_{l-1}\|\right) + \frac{7}{4} c_3 K_0 \|e_0\| \|e_{k-1}\| \\
& + c_3 \|A_0^* \left(\alpha_{l-1} I + A_0 A_0^*\right)^{-1} \left(F(x_k) - y - A_0 e_k\right)\| \\
& + c_3 \| \left(\alpha_{l-1} I + A_0^* A_0\right)^{-1} A_0^* (F(x_k) - y)\| \\
\le & \frac{7}{4} K_0 \|e_0\| \left(\|e_{k-1}\| + \|e_{l-1}\|\right) + \frac{7}{4} c_3 K_0 \|e_0\| \|e_{k-1}\| + \frac{7}{4} c_3 K_0 \|e_0\| \|e_k\| \\
& + c_3 \| \left(\alpha_{l-1} I + A_0^* A_0\right)^{-1} A_0^* (F(x_k) - y)\|. \qquad (2.22)
\end{aligned}
$$

In order to proceed further, we consider the bounded linear operator

$$
\mathcal{L} := \alpha_{l-1}^{1/2} \alpha_k^{-1/2} \left(\alpha_{l-1} I + A_0^* A_0\right)^{-1} A_0^* \left(\alpha_k I + A_0 A_0^*\right)^{1/2}.
$$

In order to estimate $\|\mathcal{L}\|$, we note that $k < l$ and (1.4) imply $\alpha_{l-1}(\alpha_k + \lambda) \le \alpha_k(\alpha_{l-1} + \lambda)$, by using $\{F_\lambda\}$ to denote the spectral family generated by the self-adjoint operator $A_0 A_0^*$, we have for any $u \in Y$ that

$$
\begin{aligned}
\|\mathcal{L} u\|^2 &= \|\alpha_{l-1}^{1/2} \alpha_k^{-1/2} (A_0 A_0^*)^{1/2} \left(\alpha_{l-1} I + A_0 A_0^*\right)^{-1} \left(\alpha_k I + A_0 A_0^*\right)^{1/2} u\|^2 \\
&= \int_0^{1/2} \alpha_{l-1} \alpha_k^{-1} \lambda (\alpha_{l-1} + \lambda)^{-2} (\alpha_k + \lambda) d\|F_\lambda u\|^2 \\
&= \int_0^{1/2} \frac{\lambda}{\alpha_{l-1} + \lambda} \cdot \frac{\alpha_{l-1}(\alpha_k + \lambda)}{\alpha_k(\alpha_{l-1} + \lambda)} d\|F_\lambda u\|^2 \\
&\le \int_0^{1/2} d\|F_\lambda u\|^2 = \|u\|^2.
\end{aligned}
$$

This implies that $\|\mathcal{L}\| \le 1$. Therefore,

$$
\begin{aligned}
\| & \left(\alpha_{l-1} I + A_0^* A_0\right)^{-1} A_0^* (F(x_k) - y)\| \\
& = \alpha_{l-1}^{-1/2} \|\mathcal{L} \cdot \alpha_k^{1/2} \left(\alpha_k I + A_0 A_0^*\right)^{-1/2} \left(F(x_k) - y\right)\| \\
& \le \alpha_l^{-1/2} \|\alpha_k^{1/2} \left(\alpha_k I + A_0 A_0^*\right)^{-1/2} \left(F(x_k) - y\right)\|.
\end{aligned}
$$

This together with (2.22) and (2.13) implies

$$
\begin{aligned}
\|x_k - x_l\| \le & \frac{7}{4} \left(2c_2 + 2c_2 c_3 + c_3\right) K_0 \|e_0\| \|e_k\| + \frac{7}{2} c_2 K_0 \|e_0\| \|e_l\| \\
& + \frac{c_3}{\sqrt{\alpha_l}} \|\alpha_k^{1/2} \left(\alpha_k I + A_0 A_0^*\right)^{-1/2} \left(F(x_k) - y\right)\|.
\end{aligned}
$$

Since $2(2c_2 + 2c_2 c_3 + c_3) K_0 \|e_0\| \le 1$, we immediately obtain (2.19). □

Lemma 2.5. *Let all the conditions in Lemma* 2.1 *hold. If, in addition,* $(11 + 4c_0/(\tau - 1)) K_0 \|e_0\| \le 2$, *then for all* $0 \le k \le \tilde{k}_\delta$ *there hold*

$$
\|x_k^\delta - x_k\| \le 2 c_0 \delta \alpha_k^{-1/2} \qquad (2.23)
$$

and

$$
\|\alpha_k^{1/2} \left(\alpha_k I + A_0 A_0^*\right)^{-1/2} \left(F(x_k^\delta) - F(x_k) - y^\delta + y\right)\| \le (1 + \varepsilon)\, \delta, \qquad (2.24)
$$

where $\varepsilon := \left(11 + 4c_0/(\tau - 1)\right) c_0 K_0 \|e_0\|$.

Proof. From (2.6) and (2.9) it follows for $0 \le k < \tilde{k}_\delta$ that

$$x_{k+1}^\delta - x_{k+1} = g_{\alpha_k}(A_0^* A_0) A_0^* \left(F(x_k) - F(x_k^\delta) - A_0(x_k - x_k^\delta) \right) + g_{\alpha_k}(A_0^* A_0) A_0^* (y^\delta - y). \tag{2.25}$$

By using (2.1), (1.2) and Assumption 1.1(a) we obtain

$$\|x_{k+1}^\delta - x_{k+1}\| \le c_0 \delta \alpha_k^{-1/2} + \frac{1}{2} K_0 \left(2\|e_0\| + \|e_k\| + \|e_k^\delta\| \right) \|x_k^\delta - x_k\|.$$

Using the estimates (2.4) and (2.11) on $\|e_k\|$ and $\|e_k^\delta\|$ yields

$$\begin{aligned} \|x_{k+1}^\delta - x_{k+1}\| &\le c_0 \delta \alpha_k^{-1/2} + \left(\frac{11}{4} + \frac{c_0}{\tau - 1} \right) K_0 \|e_0\| \|x_k^\delta - x_k\| \\ &\le c_0 \delta \alpha_k^{-1/2} + \frac{1}{2} \|x_k^\delta - x_k\|. \end{aligned}$$

Thus, we can obtain the desired estimate (2.23) by an induction argument. In order to show (2.24), we denote by $d_k(\delta)$ the left-hand side. Then it follows from (2.2) that

$$\begin{aligned} d_k(\delta) \le & \|\alpha_k^{1/2} (\alpha_k I + A_0 A_0^*)^{-1/2} \left(F(x_k^\delta) - F(x_k) - A_0(x_k^\delta - x_k) \right) \| \\ & + \|\alpha_k^{1/2} (\alpha_k I + A_0 A_0^*)^{-1/2} \left(A_0(x_k^\delta - x_k) - y^\delta + y \right) \| \\ \le & \frac{1}{2} \alpha_k^{1/2} K_0 \left(2\|e_0\| + \|e_k\| + \|e_k^\delta\| \right) \|x_k^\delta - x_k\| \\ & + \|\alpha_k^{1/2} (\alpha_k I + A_0 A_0^*)^{-1/2} \left(A_0(x_k^\delta - x_k) - y^\delta + y \right) \|. \end{aligned}$$

By the estimates (2.4), (2.11) and (2.23) we obtain

$$d_k(\delta) \le \|\alpha_k^{1/2} (\alpha_k I + A_0 A_0^*)^{-1/2} \left(A_0(x_k^\delta - x_k) - y^\delta + y \right) \| + \left(\frac{11}{2} + \frac{2c_0}{\tau - 1} \right) c_0 K_0 \|e_0\| \delta. \tag{2.26}$$

Note that (2.25) implies

$$\begin{aligned} & A_0(x_k^\delta - x_k) - y^\delta + y \\ & \quad = A_0 g_{\alpha_{k-1}}(A_0^* A_0) A_0^* \left(F(x_{k-1}) - F(x_{k-1}^\delta) - A_0(x_{k-1} - x_{k-1}^\delta) \right) \\ & \quad + r_{\alpha_{k-1}}(A_0 A_0^*)(y - y^\delta). \end{aligned}$$

Thus, by using (1.2), (2.1), Assumption 1.1(a), (2.4), (2.11), and (2.23) we have

$$\begin{aligned} & \|\alpha_k^{1/2} (\alpha_k I + A_0 A_0^*)^{-1/2} \left(A_0(x_k^\delta - x_k) - y^\delta + y \right) \| \\ & \qquad \le \frac{1}{2} \alpha_k^{1/2} K_0 \left(2\|e_0\| + \|e_{k-1}\| + \|e_{k-1}^\delta\| \right) \|x_{k-1}^\delta - x_{k-1}\| + \delta \\ & \qquad \le \delta + \left(\frac{11}{2} + \frac{2c_0}{\tau - 1} \right) c_0 K_0 \|e_0\| \delta. \end{aligned}$$

This together with (2.26) implies the estimate (2.24). □

Now we are ready to complete the proof of the main results. For ease of exposition, in the following we will use the convention $\Phi \lesssim \Psi$ to mean that $\Phi \le C\Psi$ for some generic constant C depending only on c_1, c_2, r and τ.

Proof of Theorem 1.1. We first consider the case $k \ge k_\delta$. It follows from (2.23) in Lemma 2.5, (2.19) in Lemma 2.4, and the fact $\alpha_k \le \alpha_{k_\delta}$ that

$$\begin{aligned}\|e_{k_\delta}^\delta\| &\lesssim \|e_{k_\delta}\| + \frac{\delta}{\sqrt{\alpha_{k_\delta}}}\\ &\lesssim \|e_k\| + \frac{1}{\sqrt{\alpha_k}}\|\alpha_{k_\delta}^{1/2}\left(\alpha_{k_\delta} I + A_0A_0^*\right)^{-1/2}\left(F(x_{k_\delta}) - y\right)\| + \frac{\delta}{\sqrt{\alpha_k}}.\end{aligned}$$

By the estimate (2.24) in Lemma 2.5, the fact $k_\delta \le \tilde{k}_\delta$ given by (2.5), and the definition of k_δ, we have

$$\begin{aligned}&\|\alpha_{k_\delta}^{1/2}\left(\alpha_{k_\delta} I + A_0A_0^*\right)^{-1/2}\left(F(x_{k_\delta}) - y\right)\|\\ &\qquad \lesssim \|\alpha_{k_\delta}^{1/2}\left(\alpha_{k_\delta} I + A_0A_0^*\right)^{-1/2}\left(F(x_{k_\delta}^\delta) - y^\delta\right)\| + \delta \lesssim \delta.\end{aligned}$$

Therefore, by using (2.12) in Lemma 2.2 we have for $k \ge k_\delta$ that

$$\|e_{k_\delta}^\delta\| \lesssim \|e_k\| + \frac{\delta}{\sqrt{\alpha_k}} \lesssim \|r_{\alpha_k}(A_0A_0^*)e_0\| + \frac{\delta}{\sqrt{\alpha_k}}.$$

Next we consider the case $0 \le k < k_\delta$. We first obtain from (2.23) and (2.17) that

$$\|e_{k_\delta}^\delta\| \lesssim \|e_{k_\delta}\| + \frac{\delta}{\sqrt{\alpha_{k_\delta}}} \lesssim \|e_k\| + \frac{\delta}{\sqrt{\alpha_{k_\delta}}}. \tag{2.27}$$

By the definition of k_δ and (2.24) we have

$$\begin{aligned}\tau\delta \le& \|\alpha_{k_\delta-1}^{1/2}\left(\alpha_{k_\delta-1} I + A_0A_0^*\right)^{-1/2}\left(F(x_{k_\delta-1}^\delta) - y^\delta\right)\|\\ \le& \|\alpha_{k_\delta-1}^{1/2}\left(\alpha_{k_\delta-1} I + A_0A_0^*\right)^{-1/2}\left(F(x_{k_\delta-1}) - y\right)\| + (1+\varepsilon)\delta.\end{aligned}$$

Thus, if $K_0\|e_0\|$ is so small that $\varepsilon \le (\tau - 1)/2$, then

$$\delta \lesssim \|\alpha_{k_\delta-1}^{1/2}\left(\alpha_{k_\delta-1} I + A_0A_0^*\right)^{-1/2}\left(F(x_{k_\delta-1}) - y\right)\|.$$

With the help of (2.2) and (2.11) we obtain

$$\delta \lesssim \alpha_{k_\delta-1}^{1/2}\|e_{k_\delta-1}\| + \alpha_{k_\delta-1}^{1/2}K_0\left(\|e_0\| + \|e_{k_\delta-1}\|\right)\|e_{k_\delta-1}\| \lesssim \alpha_{k_\delta-1}^{1/2}\|e_{k_\delta-1}\|.$$

Using (1.4) and (2.17) yields

$$\frac{\delta}{\sqrt{\alpha_{k_\delta}}} \lesssim \|e_{k_\delta-1}\| \lesssim \|e_k\|.$$

This together with (2.27) and Lemma 2.2 implies for all $0 \le k < k_\delta$ that

$$\|e_{k_\delta}^\delta\| \lesssim \|e_k\| \lesssim \|r_{\alpha_k}(A_0A_0^*)e_0\|.$$

The proof is therefore complete. □

In order to complete the proof of Corollary 1.1, in particular, the assertion (iii), we need the simple consequence of Assumption 1.4 which says for every $\mu > 0$ there is a positive constant b_μ such that

$$r_\alpha(\lambda)(-\ln\lambda)^{-\mu} \le b_\mu\left(-\ln(\alpha/(2\alpha_0))\right)^{-\mu} \tag{2.28}$$

for all $0 < \alpha \le \alpha_0$ and $\lambda \in [0, 1/2]$. To see this, we pick a ν_0 with $0 < \nu_0 < \bar{\nu}$. Then $0 \le r_\alpha(\lambda) \le 1$ in Assumption 1.1 and Assumption 1.4 implies for every $\mu > 0$ that

$$r_\alpha(\lambda)(-\ln\lambda)^{-\mu} \le \min\left\{(-\ln\lambda)^{-\mu}, d_{\nu_0}\alpha^{\nu_0}\lambda^{-\nu_0}(-\ln\lambda)^{-\mu}\right\}$$

for all $0 < \alpha \le \alpha_0$ and $\lambda \in [0, 1/2]$. It is clear that $(-\ln\lambda)^{-\mu} \le (-\ln(\alpha/(2\alpha_0)))^{-\mu}$ for $0 \le \lambda \le \alpha/(2\alpha_0)$. By using the fact that the function $\lambda \to \lambda^{-\nu_0}(-\ln\lambda)^{-\mu}$ is decreasing on the interval $(0, e^{-\mu/\nu_0}]$ and is increasing on the interval $[e^{-\mu/\nu_0}, 1)$, it is easy to show that there is a positive constant a_μ such that $d_{\nu_0}\alpha^{\nu_0}\lambda^{-\nu_0}(-\ln\lambda)^{-\mu} \le a_\mu(-\ln(\alpha/(2\alpha_0)))^{-\mu}$ for $\alpha/(2\alpha_0) \le \lambda \le 1/2$. We thus obtain (2.28).

Proof of Corollary 1.1. We first prove (i), we may assume $x_0 \ne x^\dagger$. Let $\bar{k}_\delta$ be the first integer such that $\alpha_{\bar{k}_\delta} \le \delta$. By (1.4) such $\bar{k}_\delta$ exists, $\alpha_{\bar{k}_\delta} \to 0$ and $\delta\alpha_{\bar{k}_\delta}^{-1/2} \to 0$ as $\delta \to 0$. By Theorem 1.1 we have

$$\|x_{k_\delta}^\delta - x^\dagger\| \le C\left(\|r_{\alpha_{\bar{k}_\delta}}(A_0^*A_0)e_0\| + \delta\alpha_{\bar{k}_\delta}^{-1/2}\right).$$

Since Assumption 1.4 and $e_0 \in \mathcal{N}(A_0)^\perp$ imply $\|r_\alpha(A_0^*A_0)e_0\| \to 0$ as $\alpha \to 0$, we therefore obtain the convergence.

In order to show (ii), we note that Assumption 1.4 and the source condition $e_0 = (A_0^*A_0)^\nu\omega$ imply $\|r_\alpha(A_0^*A_0)e_0\| \le d_\nu\|\omega\|\alpha^\nu$ for $\alpha > 0$. By choosing the integer m_δ satisfying

$$\alpha_{m_\delta}^{\nu+1/2} \le \frac{\delta}{\|\omega\|} < \alpha_k^{\nu+1/2}, \quad 0 \le k < m_\delta.$$

then from Theorem 1.1 and (1.4) we obtain

$$\|x_{k_\delta}^\delta - x^\dagger\| \le C\left(d_\nu\alpha_{m_\delta}^\nu\|\omega\| + \delta\alpha_{m_\delta}^{-1/2}\right) \le C_\nu\|\omega\|^{1/(1+2\nu)}\delta^{2\nu/(1+2\nu)}.$$

Finally, we prove (iii). We define n_δ to be the integer satisfying

$$\alpha_{n_\delta}^{1/2}\left(-\ln\left(\frac{\alpha_{n_\delta}}{2\alpha_0}\right)\right)^{-\mu} \le \frac{\delta}{\|\omega\|} < \alpha_k^{1/2}\left(-\ln\left(\frac{\alpha_k}{2\alpha_0}\right)\right)^{-\mu}, \quad 0 \le k < n_\delta.$$

By an elementary argument we can show from (1.4) that there is a constant $c_\mu > 0$ such that

$$\alpha_{n_\delta} \ge r^{-1}\alpha_{n_\delta-1} \ge c_\mu\left(\frac{\delta}{\|\omega\|}\right)^2\left(1 + \left|\ln\frac{\delta}{\|\omega\|}\right|\right)^{2\mu}.$$

Thus, by using Theorem 1.1, the source condition $e_0 = (-\ln(A_0^*A_0))^{-\mu}\omega$, (2.28) and the definition of n_δ we obtain

$$\begin{aligned}\|x_{k_\delta}^\delta - x^\dagger\| &\le C_\mu\left(\left(-\ln\left(\frac{\alpha_{n_\delta}}{2\alpha_0}\right)\right)^{-\mu}\|\omega\| + \frac{\delta}{\sqrt{\alpha_{n_\delta}}}\right) \le C_\mu\frac{\delta}{\sqrt{\alpha_{n_\delta}}}\\ &\le C_\mu\|\omega\|\left(1 + \left|\ln\frac{\delta}{\|\omega\|}\right|\right)^{-\mu}.\end{aligned}$$

The proof is therefore complete. ☐

Remark 2.2. By checking the proof of the main result carefully, one can see that Assumption 1.2 is superfluous if F is linear. Thus, we provide some new insights on the regularization methods for linear ill-posed problems.

Remark 2.3. Here is a minor remark on Assumption 1.3(c). From the proof of Theorem 1.1, one can see that the full strength of Assumption 1.3(c) is not used;

what we used are (2.1) and (2.2). Let $\{F_\lambda\}$ denote the spectral family generated by $A_0A_0^*$, then by Assumption 1.1(a) we have for all $u \in Y$,

$$\begin{aligned}\|g_\alpha(A_0^*A_0)A_0^*(\alpha I + A_0A_0^*)^{1/2}u\|^2 &= \int_0^{1/2} g_\alpha(\lambda)^2\lambda(\alpha+\lambda)d\|F_\lambda u\|^2 \\ &\le (1+c_1)\|u\|^2.\end{aligned}$$

Therefore, for all $\alpha > 0$

$$\|g_\alpha(A_0^*A_0)A_0^*(\alpha I + A_0A_0^*)^{1/2}\| \le \sqrt{1+c_1}.$$

Consequently, (2.2) implies (2.1) but with the number $1/2$ on the right-hand side replaced by $\sqrt{1+c_1}/2$. Thus, Assumption 1.3(c) becomes unnecessary if one can check (2.2) directly.

Example 2.1. (a) In the *iterated Tikhonov regularization of order m*,

$$g_\alpha(\lambda) = \frac{(\alpha+\lambda)^m - \alpha^m}{\lambda(\alpha+\lambda)^m} \quad \text{and} \quad r_\alpha(\lambda) = \frac{\alpha^m}{(\alpha+\lambda)^m}.$$

The *ordinary Tikhonov regularization* corresponds to $m = 1$. It is well known that Assumption 1.1 and Assumption 1.4 hold with $c_1 = m$ and $\bar{\nu} = m$. For any sequence $\{\alpha_k\}$ satisfying (1.4), Assumption 1.2 is satisfied with $c_2 = r^m$.

(b) In *Landweber iteration* we have

$$g_\alpha(\lambda) = \sum_{i=0}^{[1/\alpha]} (1-\lambda)^i \quad \text{and} \quad r_\alpha(\lambda) = (1-\lambda)^{[1/\alpha]+1}.$$

Then Assumption 1.1 and Assumption 1.4 hold with $c_1 = 2$ and $\bar{\nu} = \infty$. If the sequence $\{\alpha_k\}$ is chosen as $\alpha_k = 1/n_k$, where $\{n_k\}$ is a sequence of positive integers such that $\lim_{k\to\infty} n_k = \infty$ and $0 \le n_{k+1} - n_k \le q$ for some $q \ge 1$, then both (1.4) and Assumption 1.2 are satisfied with $r = 1 + q/n_0$ and $c_2 = 2^q$.

(c) In the *asymptotic regularization*, we have

$$g_\alpha(\lambda) = (1 - e^{-\lambda/\alpha})/\lambda \quad \text{and} \quad r_\alpha(\lambda) = e^{-\lambda/\alpha}.$$

Assumption 1.1 and Assumption 1.4 hold with $c_1 = 1$ and $\bar{\nu} = \infty$. If $\{\alpha_k\}$ is a sequence of positive numbers satisfying $\lim_{k\to\infty} \alpha_k = 0$ and $0 \le 1/\alpha_{k+1} - 1/\alpha_k \le \theta_0$ for some $\theta_0 > 0$, then both (1.4) and Assumption 1.2 are satisfied with $r = 1 + \theta_0\alpha_0$ and $c_2 = e^{\theta_0}$.

3. Miscellaneous variants

3.1. **A continuous analog.** We consider the regularization scheme in which the regularized solution x_α^δ is defined implicitly by the equation

$$x_\alpha^\delta = x_0 - g_\alpha(A_0^*A_0)A_0^*\left(F(x_\alpha^\delta) - y^\delta - A_0(x_\alpha^\delta - x_0)\right), \tag{3.1}$$

where $\alpha > 0$ is the regularization parameter which is determined by the following rule: Let $\tau > 1$ be a given number. If $\|F(x_0) - y^\delta\| \le \tau\delta$, we choose $\alpha(\delta) := \infty$, i.e., we choose x_0 as an approximation of $x^\dagger$; otherwise we choose $\alpha(\delta)$ as the root of the equation

$$\|\alpha^{1/2}\left(\alpha I + A_0A_0^*\right)^{-1/2}\left(F(x_\alpha^\delta) - y^\delta\right)\| = \tau\delta. \tag{3.2}$$

Such a method can be viewed as a continuous analog of the method defined by (1.8) and (1.9).

In the following we will show that the above method is well defined under Assumption 1.1 and Assumption 1.3. We set

$$\alpha_0(\delta) := \left(\frac{(\tau-1)\delta}{2\|x_0 - x^\dagger\|}\right)^2. \tag{3.3}$$

We will show that for each $\alpha \geq \alpha_0(\delta)$ the equation (3.1) has a unique solution x_α^δ satisfying

$$\|x_\alpha^\delta - x^\dagger\| \leq r := 2\left(1 + \frac{c_0}{\tau - 1}\right)\|x_0 - x^\dagger\|, \tag{3.4}$$

and if $\|F(x_0) - y^\delta\| > \tau\delta$, then the equation (3.2) has a root $\alpha(\delta)$ satisfying $\alpha(\delta) \geq \alpha_0(\delta)$.

To this end, we set $\mathcal{D}_r := \overline{B_r(x^\dagger)}$ as the closed ball of radius r with center at $x^\dagger$. For each $\alpha \geq \alpha_0(\delta)$ we consider the function

$$\Phi_\alpha(x) := x_0 - g_\alpha(A_0^* A_0) A_0^* \left(F(x) - y^\delta - A_0(x - x_0)\right)$$

which can be written as

$$\begin{aligned} \Phi_\alpha(x) &= x^\dagger + r_\alpha(A_0^* A_0)(x_0 - x^\dagger) + g_\alpha(A_0^* A_0) A_0^* (y^\delta - y) \\ &\quad - g_\alpha(A_0^* A_0) A_0^* \left(F(x) - y - A_0(x - x^\dagger)\right). \end{aligned} \tag{3.5}$$

This together with Assumption 1.1, (1.2) and (2.1) implies, with $e_0 := x_0 - x^\dagger$,

$$\begin{aligned} \|\Phi_\alpha(x) - x^\dagger\| &\leq \|e_0\| + c_0\delta\alpha^{-1/2} + K_0\|e_0\|\|x - x^\dagger\| + \frac{1}{2}K_0\|x - x^\dagger\|^2 \\ &\leq \left(1 + \frac{2c_0}{\tau - 1}\right)\|e_0\| + K_0\|e_0\|\|x - x^\dagger\| + \frac{1}{2}K_0\|x - x^\dagger\|^2. \end{aligned}$$

Thus, if $K_0\|e_0\|$ satisfies the smallness condition (2.7), then for any $x \in \mathcal{D}_r$ we have $\|\Phi(x) - x^\dagger\| \leq r$. Thus Φ_α maps $\mathcal{D}_r$ into itself.

Next we will show that Φ_α is a contractive mapping. By using (3.5) we have for any $x, z \in \mathcal{D}_r$ that

$$\Phi_\alpha(x) - \Phi_\alpha(z) = g_\alpha(A_0^* A_0) A_0^* \left(F(z) - F(x) - A_0(z - x)\right).$$

With the help of (2.1) and (2.7), we obtain

$$\|\Phi_\alpha(x) - \Phi_\alpha(z)\| \leq \frac{1}{2}K_0\left(2\|e_0\| + \|x - x^\dagger\| + \|z - x^\dagger\|\right)\|x - z\| \leq \frac{1}{2}\|x - z\|.$$

Thus $\Phi_\alpha : \mathcal{D}_r \to \mathcal{D}_r$ is a contractive mapping. It follows from the contractive mapping theorem that Φ_α has a unique fixed point x_α^δ in $\mathcal{D}_r$ for each $\alpha \geq \alpha_0(\delta)$. We thus obtain (3.4).

Now we assume that $\|F(x_0) - y^\delta\| > \tau\delta$. We denote by $d^\delta(\alpha)$ the left-hand side of (3.2). By using (3.1), the estimate (3.4), (1.11) and the continuity of F, it is easy to see that $\lim_{\alpha\to\infty} x_\alpha^\delta = x_0$. Thus,

$$\lim_{\alpha\to\infty} d^\delta(\alpha) = \|F(x_0) - y^\delta\| > \tau\delta.$$

By using a similar manner in deriving (2.8) we can show that for $\alpha_0(\delta)$ defined by (3.3) there holds

$$d^\delta(\alpha_0(\delta)) \leq \tau\delta.$$

Thus, in order to show that (3.2) has a root, it suffices to show $\alpha \to x_\alpha^\delta$ is continuous on $[\alpha_0(\delta), \infty)$. From (3.1) it follows for any $\alpha, \beta \ge \alpha_0(\delta)$ that

$$\begin{aligned} x_\alpha^\delta - x_\beta^\delta =& g_\alpha(A_0^* A_0) A_0^* \left(F(x_\beta^\delta) - F(x_\alpha^\delta) - A_0(x_\beta^\delta - x_\alpha^\delta) \right) \\ &+ \left[g_\beta(A_0^* A_0) - g_\alpha(A_0^* A_0) \right] A_0^* \left(F(x_\beta^\delta) - y^\delta - A_0(x_\beta^\delta - x_0) \right). \end{aligned}$$

Therefore, by using (2.1) we have

$$\begin{aligned} \|x_\alpha^\delta - x_\beta^\delta\| =& \frac{1}{2} K_0 \left(2\|e_0\| + \|x_\alpha^\delta - x^\dagger\| + \|x_\beta^\delta - x^\dagger\| \right) \|x_\alpha^\delta - x_\beta^\delta\| \\ &+ \| \left[g_\beta(A_0^* A_0) - g_\alpha(A_0^* A_0) \right] A_0^* \left(F(x_\beta^\delta) - y^\delta - A_0(x_\beta^\delta - x_0) \right) \|. \end{aligned}$$

With the help of (3.4) and the smallness condition (2.7) we obtain

$$\|x_\alpha^\delta - x_\beta^\delta\| \le 2 \| \left[g_\beta(A_0^* A_0) - g_\alpha(A_0^* A_0) \right] A_0^* \left(F(x_\beta^\delta) - y^\delta - A_0(x_\beta^\delta - x_0) \right) \|.$$

Thus, we can conclude the continuity of $\alpha \to x_\alpha^\delta$ if $\{g_\alpha\}$ satisfies the following additional condition.

Assumption 3.1. The function $\alpha \to g_\alpha(\lambda)$ is continuous on $(0, \infty)$ uniformly with respect to $\lambda \in [0, 1/2]$.

Now we turn to consider the approximation property of $x_{\alpha(\delta)}^\delta$ to $x^\dagger$. We will use x_α to denote the solution of the noise-free equation

$$x_\alpha = x_0 - g_\alpha(A_0^* A_0) A_0^* \left(F(x_\alpha) - y - A_0(x_\alpha - x_0) \right). \tag{3.6}$$

By using a similar argument as above we can show that for each $\alpha > 0$ the equation (3.6) has a unique solution x_α satisfying

$$\|x_\alpha - x^\dagger\| \le \frac{3}{2} \|x_0 - x^\dagger\| \quad \text{and} \quad \|x_\alpha - x^\dagger\| \lesssim \|r_\alpha(A_0^* A_0)(x_0 - x^\dagger)\|.$$

By checking the proof of Lemma 2.4 one can easily see, without using any estimates as in Lemma 2.2, that for any $\alpha \ge \beta > 0$ there holds

$$\|x_\alpha - x^\dagger\| \lesssim \|x_\beta - x^\dagger\| + \frac{1}{\sqrt{\beta}} \|\alpha^{1/2} (\alpha I + A_0 A_0^*)^{-1/2} \left(F(x_\alpha) - y \right) \|. \tag{3.7}$$

Similar estimates as in Lemma 2.5 still hold and read as

$$\|x_\alpha^\delta - x_\alpha\| \le 2 c_0 \delta \alpha^{-1/2}$$

and

$$\|\alpha^{1/2} (\alpha I + A_0 A_0^*)^{-1/2} \left(F(x_\alpha^\delta) - F(x_\alpha) - y^\delta + y \right) \| \le (1 + \varepsilon) \delta$$

for all $\alpha \ge \alpha_0(\delta)$.

With the above ingredients, one can follow the proofs of Theorem 1.1 and Corollary 1.1 to obtain the following result.

Theorem 3.1. *Assume that $\tau > 1$ and that $\{g_\alpha\}$ and F satisfy Assumption* 1.1*, Assumption* 3.1 *and Assumption* 1.3*. There exist positive constants η_1 and C depending only on c_1 and τ such that if $K_0 \|x_0 - x^\dagger\| \le \eta_1$, then the method given by* (3.1) *and* (3.2) *is well defined and*

$$\|x_{\alpha(\delta)}^\delta - x^\dagger\| \le C \inf \left\{ \|r_\alpha(A_0^* A_0)(x_0 - x^\dagger)\| + \frac{\delta}{\sqrt{\alpha}} : \alpha > 0 \right\}.$$

Assume, in addition, that $\{g_\alpha\}$ satisfies Assumption 1.4*. Then*

(i) *if $x_0 - x^\dagger \in \mathcal{N}(A_0)^\perp$, then $\lim_{\delta \to 0} x_{\alpha(\delta)}^\delta = x^\dagger$;*

(ii) *if* $x_0 - x^\dagger = (A_0^* A_0)^\nu \omega$ *for some* $\omega \in X$ *and* $0 < \nu \le \bar{\nu}$, *then*

$$\|x_{\alpha(\delta)}^\delta - x^\dagger\| \le C_\nu \|\omega\|^{1/(1+2\nu)} \delta^{2\nu/(1+2\nu)},$$

where C_ν *is a constant depending only on* c_1, τ *and* ν;

(iii) *if* $x_0 - x^\dagger = (-\ln(A_0^* A_0))^{-\mu} \omega$ *for some* $\omega \in X$ *and* $\mu > 0$, *then*

$$\|x_{\alpha(\delta)}^\delta - x^\dagger\| \le C_\mu \|\omega\| \left(1 + \left|\ln \frac{\delta}{\|\omega\|}\right|\right)^{-\mu},$$

where C_μ *is a constant depending only on* c_1, τ *and* μ.

Example 3.1. It is clear that the functions $\{g_\alpha\}$ from Example 2.1 (a) and (c) satisfy Assumption 3.1. Here we give one more example where g_α arises from the *regularized singular value decomposition* and

$$g_\alpha(\lambda) = \begin{cases} 1/\lambda, & \text{if } \lambda \ge \alpha, \\ 1/\alpha, & \text{if } \lambda < \alpha \end{cases} \quad \text{and} \quad r_\alpha(\lambda) = \begin{cases} 0, & \text{if } \lambda \ge \alpha, \\ 1 - \lambda/\alpha, & \text{if } \lambda < \alpha. \end{cases}$$

Assumption 1.1, Assumption 1.4 and Assumption 3.1 hold with $c_1 = 1$ and $\bar{\nu} = \infty$.

Remark 3.1. In [16] Tautenhahn considered a general regularization scheme for (1.1) in which the regularized solution x_α^δ is defined as a fixed point of the nonlinear equation

$$x = x_0 - g_\alpha (F'(x)^* F'(x)) F'(x)^* \left(F(x) - y^\delta - F'(x)(x - x_0)\right), \tag{3.8}$$

and the regularization parameter $\alpha > 0$ is determined by a Morozov's type discrepancy principle. The convergence of the method is established under certain Hölder source conditions. However, the method is not shown to be well defined. This gap can be filled by using the technique of this subsection with a little involved argument.

3.2. **The Lavrentiev type methods.** When a nonlinear inverse problem can be formulated as the nonlinear equation (1.1) with $F : D(F) \subset X \to X$, then we can simplify the method given by (1.8) and (1.9) further if one can find an initial guess x_0 near $x^\dagger$ such that $A_0 := F'(x_0)$ is self-adjoint and positive semi-definite. In this situation, we define the iterates $\{x_k^\delta\}$ by

$$x_k^\delta = x_0 - g_{\alpha_k}(A_0) \left(F(x_k^\delta) - y^\delta - A_0(x_k^\delta - x_0)\right), \tag{3.9}$$

and we terminate the iteration by choosing k_δ as the first integer such that

$$\|\alpha_{k_\delta} (\alpha_{k_\delta} I + A_0)^{-1} \left(F(x_{k_\delta}^\delta) - y^\delta\right)\| \le \tau\delta, \tag{3.10}$$

where $\tau > 1$ is a given number.

The argument in Section 2 applies to this simplified method, but with (2.19) and (2.23) replaced by

$$\|x_k - x^\dagger\| \lesssim \|x_l - x^\dagger\| + \frac{1}{\alpha_l} \|\alpha_k(\alpha_k I + A_0)^{-1}(F(x_k) - y)\| \tag{3.11}$$

for $0 \le k \le l$ and

$$\|x_k^\delta - x_k\| \le 2c_0\delta\alpha_k^{-1}$$

for $0 \le k \le \hat{k}_\delta$, where $\{x_k\}$ are the iterates defined by (3.9) with y^δ replaced by y, and $\hat{k}_\delta$ is defined as the integer satisfying

$$\alpha_{\hat{k}_\delta} \le \frac{(\tau - 1)\delta}{2\|x_0 - x^\dagger\|} < \alpha_k, \quad 0 \le k < \hat{k}_\delta.$$

By the same argument as in the proofs of Theorem 1.1 and Corollary 1.1 we can obtain the following result.

Theorem 3.2. *Assume that $\tau > 1$ and that $\{\alpha_k\}$, $\{g_\alpha\}$ and F satisfy* (1.4), *Assumption* 1.1, *Assumption* 1.2 *and Assumption* 1.3. *There exist positive constants η_2 and C depending only on c_1, c_2, r and τ such that if $K_0\|x_0 - x^\dagger\| \le \eta_2$ and $A_0 := F'(x_0)$ is self-adjoint and positive semi-definite, then the method given by* (3.9) *and* (3.10) *is well defined and*

$$\|x_{k_\delta}^\delta - x^\dagger\| \le C \inf\left\{\|r_{\alpha_k}(A_0)(x_0 - x^\dagger)\| + \frac{\delta}{\alpha_k} : k = 0, 1, \cdots\right\}.$$

Assume, in addition, that $\{g_\alpha\}$ satisfies Assumption 1.4. *Then*

(i) *if $x_0 - x^\dagger \in \mathcal{N}(A_0)^\perp$, then $\lim_{\delta\to 0} x_{k_\delta}^\delta = x^\dagger$;*

(ii) *if $x_0 - x^\dagger = A_0^\nu \omega$ for some $\omega \in X$ and $0 < \nu \le \bar{\nu}$, then*

$$\|x_{k_\delta}^\delta - x^\dagger\| \le C_\nu \|\omega\|^{1/(1+\nu)} \delta^{\nu/(1+\nu)},$$

where C_ν is a constant depending only on c_1, c_2, r, τ and ν;

(iii) *if $x_0 - x^\dagger = (-\ln A_0)^{-\mu}\,\omega$ for some $\omega \in X$ and $\mu > 0$, then*

$$\|x_{k_\delta}^\delta - x^\dagger\| \le C_\mu \|\omega\| \left(1 + \left|\ln \frac{\delta}{\|\omega\|}\right|\right)^{-\mu},$$

where C_μ is a constant depending only on c_1, c_2, r, τ and μ;

Remark 3.2. The method defined by (3.9) and (3.10) requires only A_0 be self-adjoint and positive semi-definite; it does not require F be monotone everywhere.

Remark 3.3. It is possible to drop the requirement that A_0 be self-adjoint for some special choice of $\{g_\alpha\}$. For instance, if A_0 is only positive semi-definite, then $(\alpha I + A_0)^{-1}$ is well defined for each $\alpha > 0$. Thus, for the function $g_\alpha(\lambda) = (\alpha + \lambda)^{-1}$ the method (3.9) becomes

$$x_{k+1}^\delta = x_0 - (\alpha_k I + A_0)^{-1}\left(F(x_k^\delta) - y^\delta - A_0(x_k^\delta - x_0)\right). \tag{3.12}$$

Note that $\|(\alpha I + A_0)^{-1}\| \le \alpha^{-1}$ and $\|(\alpha I + A_0)^{-1} A_0\| \le 1$ for $\alpha > 0$, the arguments in Section 2 can apply except that three places need to be modified since they involve the theory of spectral integrals which is not available when A_0 is not self-adjoint. The first one is the inequality (2.15) which was obtained from Assumption 1.2. In the current situation, it becomes

$$\|\alpha_k(\alpha_k I + A_0)^{-1} e_0\| \le (1 + r)\|\alpha_{k+1}(\alpha_{k+1} + A_0)^{-1} e_0\|.$$

This can be verified directly by noting that

$$\begin{aligned}\|\alpha_{k+1}^{-1}\alpha_k(\alpha_k I + A_0)^{-1}(\alpha_{k+1} I + A_0)\| &\le 1 + \alpha_{k+1}^{-1}\alpha_k\|(\alpha_k I + A_0)^{-1} A_0\| \\ &\le 1 + \alpha_{k+1}^{-1}\alpha_k \le 1 + r.\end{aligned}$$

The other two are related to the inequality (3.11) for $0 \le k < l$. By checking the proof of Lemma 2.4 one can see that the derivation of (2.21) involves the spectral integral, however, one can use the expression of $r_\alpha(\lambda) = \alpha(\alpha + \lambda)^{-1}$ to get it directly. The other place is to establish

$$\|(\alpha_{l-1} I + A_0)^{-1}(F(x_k) - y)\| \le \frac{1}{\alpha_{l-1}}\|\alpha_k(\alpha_k I + A_0)^{-1}(F(x_k) - y)\|.$$

This can be obtained by showing that

$$\|\alpha_{l-1}\alpha_k^{-1}(\alpha_{l-1}I + A_0)^{-1}(\alpha_k I + A_0)\| \le 1. \tag{3.13}$$

By noting that

$$(\alpha_{l-1}I + A_0)^{-1}(\alpha_k I + A_0) = \left(\frac{\alpha_k}{\alpha_{l-1}} - 1\right)\alpha_{l-1}(\alpha_{l-1}I + A_0)^{-1} + I$$

and the fact $\alpha_k \ge \alpha_{l-1}$, we can obtain (3.13) easily. Therefore, all the results stated in Theorem 3.2 with $\bar{\nu} = 1$, except part (iii), are valid for the method defined by (3.12) and (3.10).

Remark 3.4. The above simplification applies to the continuous method discussed in subsection 3.1.

4. A numerical example

In this section we present a numerical example to test the convergence result given in Corollary 1.1 on the method defined by (1.8) and (1.9) by considering the estimation of the coefficient a in the two-point boundary value problem

$$\begin{cases} -u'' + au = f, & t \in (0,1), \\ u(0) = g_0, & u(1) = g_1, \end{cases} \tag{4.1}$$

from the L^2 measurement u^δ of the state variable u, where g_0, g_1 and $f \in L^2[0,1]$ are given. This inverse problem reduces to solving the equation (1.1) with the nonlinear operator $F : D(F) \subset L^2[0,1] \mapsto L^2[0,1]$ defined as the parameter-to-solution mapping $F(a) := u(a)$, where $u(a)$ denotes the unique solution of (4.1). It is well known that F is well defined on

$$D(F) := \left\{a \in L^2[0,1] : \|a - \hat{a}\|_{L^2} \le \gamma \text{ for some } \hat{a} \ge 0 \text{ a.e.}\right\}$$

with some $\gamma > 0$. Moreover, F is Fréchet differentiable, the Fréchet derivative and its adjoint are given by

$$F'(a)h = -A(a)^{-1}(hu(a)),$$
$$F'(a)^* w = -u(a)A(a)^{-1}w,$$

where $A(a) : H^2 \cap H_0^1 \mapsto L^2$ is defined by $A(a)u = -u'' + au$. It is known that if, for the desired solution $a^\dagger$, $|u(a^\dagger)(t)| \ge \kappa > 0$ for all $t \in [0,1]$, then (1.12) is satisfied in a neighborhood of $a^\dagger$.

In the following we report some numerical results on the method given by (1.8) and (1.9) with $g_\alpha(\lambda) = (\alpha + \lambda)^{-1}$, which, in the current situation, defines the iterated solutions $\{a_k^\delta\}$ by

$$a_{k+1}^\delta = a_k^\delta - (\alpha_k I + A_0^* A_0)^{-1}\left(A_0^*(F(a_k^\delta) - y^\delta) + \alpha_k(a_k^\delta - a_0)\right) \tag{4.2}$$

and determines the stopping index k_δ as the first integer satisfying

$$\alpha_{k_\delta}\left(F(a_{k_\delta}^\delta) - y^\delta), (\alpha_k I + A_0 A_0^*)^{-1}(F(a_k^\delta) - y^\delta)\right) \le \tau^2\delta^2. \tag{4.3}$$

During the computation, all differential equations are solved approximately by a finite difference method by dividing the interval $[0,1]$ into $n+1$ subintervals with equal length $h = 1/(n+1)$; we take $n = 200$ in our actual computation.

Example 4.1. In this example we consider the estimation of a in (4.1) with $f = 1 + t^2$ and $g_0 = g_1 = 1$. If $u(a^\dagger) = 1$, then $a^\dagger = 1 + t^2$ is the desired solution. When applying the method (4.2)–(4.3), we use the special noise data $u^\delta = 1 + \sqrt{2}\delta \sin(10\pi t)$, $\alpha_k = 1.0 \times 2^{-k}$ and $\tau = 1.2$. In Table 1 we summarize the numerical results corresponding to two different choices of the initial guess

$$a_0 = 1 + t^2 - 2t(1-t)(1+t-t^2) \tag{4.4}$$

and

$$a_0 = 1.2. \tag{4.5}$$

For the a_0 given by (4.4) one can check $a_0 - a^\dagger \in \mathcal{R}(F'(a^\dagger)^* F'(a^\dagger))$. Table 1 indicates that the convergence rate is $O(\delta^{2/3})$, which confirms the theoretical result very well. On the other hand, for the a_0 given by (4.5), $a_0 - a^\dagger$ has no sourcewise representation $a_0 - a^\dagger \in \mathcal{R}((F'(a^\dagger)^* F'(a^\dagger))^\nu)$ with a good $\nu > 0$. Thus a good convergence rate cannot be expected if the method starts from this a_0; Table 1, however, still indicates the convergence of the method.

Table 1. Numerical results for example 1, where $error := \|a^\delta_{k_\delta} - a^\dagger\|_{L^2}$

	$a_0 = 1 + t^2 - 2t(1-t)(1+t-t^2)$			$a_0 = 1.2$	
δ	k_δ	$error$	$error/\delta^{2/3}$	k_δ	$error$
$5.0e-2$	6	$3.43e-1$	2.53	1	$3.24e-1$
$1.0e-2$	10	$8.34e-2$	1.80	8	$3.03e-1$
$5.0e-3$	11	$4.74e-2$	1.62	11	$2.51e-1$
$1.0e-3$	12	$2.57e-2$	2.57	15	$1.70e-1$
$5.0e-4$	13	$1.39e-2$	2.21	17	$1.39e-1$
$1.0e-4$	15	$4.40e-3$	2.04	20	$1.05e-1$
$5.0e-5$	16	$2.80e-3$	2.04	21	$9.20e-2$

References

1. A. B. Bakushinskii, *The problems of the convergence of the iteratively regularized Gauss-Newton method*, Comput. Math. Math. Phys., 32 (1992), 1353–1359. MR1185952 (93k:65049)
2. A. B. Bakushinskii, *Iterative methods without saturation for solving degenerate nonlinear operator equations*, Dokl. Akad. Nauk., 344 (1995), 7–8. MR1361018
3. B. Blaschke, A. Neubauer and O. Scherzer, *On convergence rates for the iteratively regularized Gauss-Newton method*, IMA J. Numer. Anal., 17 (1997), 421–436. MR1459331 (98f:65066)
4. D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, 2nd edition, Applied Mathematical Sciences, 93, Springer-Verlag, Berlin, 1998. MR1635980 (99c:35181)
5. H. W. Engl, M. Hanke and A. Neubauer, *Regularization of Inverse Problems*, Mathematics and its Applications, 375, Kluwer Academic Publishers Group, Dordrecht, 1996. MR1408680 (97k:65145)
6. F. Hettlich and W. Rundell, *A second degree method for nonlinear inverse problems*, SIAM J. Numer. Anal., 37 (2000), no. 2, 587–620. MR1740766 (2000j:65064)
7. T. Hohage, *Logarithmic convergence rates of the iteratively regularized Gauss-Newton method for an inverse potential and an inverse scattering problem*, Inverse Problems, 13 (1997), no. 5, 1279–1299. MR1474369 (98k:65031)
8. V. Isakov, *Inverse Problems for Partial Differential Equations*, 2nd edition, Applied Mathematical Sciences, 127, Springer, New York, 2006. MR2193218 (2006h:35279)
9. Q. N. Jin, *On the iteratively regularized Gauss-Newton method for solving nonlinear ill-posed problems*, Math. Comp., 69 (2000), no. 232, 1603–1623. MR1680887 (2001a:65073)
10. Q. N. Jin, *A convergence analysis of the iteratively regularized Gauss-Newton method under the Lipschitz condition*, Inverse Problems, 24 (2008), no. 4, article 045002. MR2425869

11. Q. N. Jin and U. Tautenhahn, *On the discrepancy principle for some Newton type methods for solving nonlinear inverse problems*, Numer. Math., 111 (2009), 509–558. MR2471609
12. B. Kaltenbacher, *A posteriori choice strategies for some Newton type methods for the regularization of nonlinear ill-posed problems*, Numer. Math., 79 (1998), 501-528. MR1631677 (99f:65091)
13. B. Kaltenbacher, A. Neubauer and O. Scherzer, *Iterative Regularization Methods for Nonlinear Ill-Posed Problems*, Berlin, de Gruyter, 2008. MR2459012
14. P. Mahale and M. T. Nair, *A simplified generalized Gauss-Newton method for nonlinear ill-posed problems*, Math. Comp., 78 (2009), no. 265, 171–184. MR2448702 (2009h:65087)
15. P. Stefanov and G. Uhlmann, *Linearizing non-linear inverse problems and an application to inverse backscattering*, J. Funct. Anal. 256 (2009), no. 9, 2842–2866. MR2502425
16. U. Tautenhahn, *On a general regularization scheme for nonlinear ill-posed problems*, Inverse Problems, 13 (1997), no. 5, 1427–1437. MR1474375 (98j:65045)
17. G. M. Vainikko and A. Y. Veretennikov, *Iteration Procedures in Ill-Posed Problems*, Moscow, Nauka, 1986 (In Russian). MR859375 (88c:47019)

Department of Mathematics, The University of Texas at Austin, Austin, Texas 78712
E-mail address: qjin@math.utexas.edu
Current address: Department of Mathematics, Virginia Tech, Blacksburg, Virginia 24061
E-mail address: qnjin@math.vt.edu