

FURTHER CONVERGENCE RESULTS ON THE GENERAL ITERATIVELY REGULARIZED GAUSS-NEWTON METHODS UNDER THE DISCREPANCY PRINCIPLE

QINIAN JIN

ABSTRACT. We consider the general iteratively regularized Gauss-Newton methods

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

for solving nonlinear inverse problems $F(x) = y$ using the only available noise y^δ of y satisfying $\|y^\delta - y\| \leq \delta$ with a given small noise level $\delta > 0$. In order to produce reasonable approximation to the sought solution, we terminate the iteration by the discrepancy principle. Under much weaker conditions we derive some further convergence results which improve the existing ones and thus expand the applied range.

1. INTRODUCTION

In this paper we are interested in the ill-posed equations

$$(1.1) \quad F(x) = y$$

arising from nonlinear inverse problems, where $F : D(F) \subset X \mapsto Y$ is a Fréchet differentiable nonlinear operator between two Hilbert spaces X and Y whose norms and inner products are denoted as $\|\cdot\|$ and (\cdot, \cdot) ; we will use $F'(x)$ and $F'(x)^*$ to denote the Fréchet derivative of F at $x \in D(F)$ and its adjoint respectively. We assume that (1.1) has a solution x^\dagger in the domain $D(F)$ of F , i.e., $F(x^\dagger) = y$. We call (1.1) ill-posed in the sense that its solution does not depend continuously on the right-hand side. Since the data is usually obtained by measurement, instead of y , the only available data is an approximation y^δ satisfying

$$(1.2) \quad \|y^\delta - y\| \leq \delta$$

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 methods should be taken into account.

We consider the general iteratively regularized Gauss-Newton methods ([2, 10])

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

where $x_0^\delta := x_0$ is an initial guess of x^\dagger , $\{\alpha_k\}$ is a given sequence of numbers such that

$$(1.4) \quad \alpha_k > 0, \quad 1 \leq \frac{\alpha_k}{\alpha_{k+1}} \leq q \quad \text{and} \quad \lim_{k \rightarrow \infty} \alpha_k = 0$$

Received by the editor June 30, 2010 and, in revised form, August 22, 2011.

2010 *Mathematics Subject Classification*. Primary 65J15, 65J20; Secondary 65H17.

Key words and phrases. Nonlinear inverse problems, the general iteratively regularized Gauss-Newton methods, the discrepancy principle, convergence, order optimality.

©2012 American Mathematical Society
 Reverts to public domain 28 years from publication

for some constant $q > 1$, and $g_\alpha : [0, \infty) \rightarrow (-\infty, \infty)$ is a family of piecewise continuous spectral filter functions satisfying suitable structure conditions.

In order to produce a reasonable approximation to x^\dagger , the iteration (1.3) must be terminated properly. Due to the practical applications, a posteriori rules, which use only quantities that arise during computation, should be used to choose the stopping index of iteration.

In [8] we considered the discrepancy principle

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

with a given number $\tau > 1$ and we obtained several results concerning the convergence property of $x_{k_\delta}^\delta$ to x^\dagger as $\delta \rightarrow 0$. In particular, we showed in [8] that if F' satisfies the Lipschitz condition, i.e., if there is a number L such that

$$(1.6) \quad \|F'(x) - F'(z)\| \leq L\|x - z\|$$

for all $x, z \in B_\rho(x^\dagger) \subset D(F)$, then the method defined by (1.3) and (1.5) is order optimal for each $1/2 \leq \nu \leq \bar{\nu} - 1/2$, that is, when $x_0 - x^\dagger$ satisfies the source condition

$$(1.7) \quad x_0 - x^\dagger = (F'(x^\dagger)^* F'(x^\dagger))^\nu \omega$$

for some $1/2 \leq \nu \leq \bar{\nu} - 1/2$ and $\omega \in \mathcal{N}(F'(x^\dagger))^\perp \subset X$, there holds

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

for some constant C_ν depending only on ν , where $\bar{\nu} \geq 1$ denotes the qualification of the linear regularization method defined by $\{g_\alpha\}$. In order to carry out the convergence analysis under the source condition (1.7) with $0 \leq \nu < 1/2$, stronger conditions must be imposed on F . In [8] it has been shown that $x_{k_\delta}^\delta$ converges to x^\dagger as $\delta \rightarrow 0$ and the method defined by (1.3) and (1.5) is order optimal for each $0 < \nu \leq 1/2$ if F satisfies the two structure conditions, i.e., there exist constants K_0 and K_1 such that

$$(1.8) \quad \|(F'(x) - F'(z))w\| \leq K_0\|x - z\| \|F'(z)w\| + K_1\|F'(z)(x - z)\| \|w\|$$

and there are bounded linear operators $R(x, z) : X \rightarrow X$ and a constant K_2 such that

$$(1.9) \quad F'(x) = F'(z)R(x, z), \quad \|I - R(x, z)\| \leq K_2\|x - z\|$$

for $x, z \in B_\rho(x^\dagger)$ and $w \in X$. We should point out that, the above results were established under the additional condition

$$(1.10) \quad r_{\alpha_k}(\lambda) \leq c r_{\alpha_{k+1}}(\lambda), \quad \forall k \geq 0 \text{ and } \lambda \in [0, 1]$$

for some constant $c > 1$, where $r_\alpha(\lambda) := 1 - \lambda g_\alpha(\lambda)$. Although it is a direct consequence of (1.4) for some choices of $\{g_\alpha\}$, condition (1.10) in general imposes further restriction on $\{\alpha_k\}$, which could exclude the case that $\{\alpha_k\}$ is a geometric decreasing sequence.

Several important questions arise naturally: Can we drop the additional condition (1.10) on $\{\alpha_k\}$ in the convergence analysis? Is it possible to establish the convergence results under only one of the structure conditions (1.8) and (1.9) on F ? In this paper we give further convergence analysis on the methods defined by (1.3) and the discrepancy principle (1.5). Under (1.8), we show that $x_{k_\delta}^\delta$ converges to x^\dagger as $\delta \rightarrow 0$ and derive the order optimal convergence rate for each $0 < \nu \leq 1/2$.

Our argument relies on neither (1.9) nor (1.10). Thus we expand the applied range of the method defined by (1.3) and (1.5).

This paper is organized as follows. In Section 2 we state several structure conditions on $\{g_\alpha\}$ together with some consequences which enable us to provide a unified treatment on (1.3) and (1.5). In Section 3 we present the convergence analysis. Finally, in Section 4 we provide numerical experiments to test our theoretical results.

2. ASSUMPTIONS

In the formulation of the method (1.3), we can choose $\{g_\alpha\}$ in various ways to produce various iterative methods. In order to give a unified convergence analysis, we need to impose certain conditions on $\{g_\alpha\}$. The function

$$r_\alpha(\lambda) := 1 - \lambda g_\alpha(\lambda)$$

is called the residual function associated with g_α and plays a significant role in the convergence analysis.

Assumption 2.1. *There exist positive constants c_0 and c_1 such that*

$$0 \leq r_\alpha(\lambda) \leq 1, \quad \lambda r_\alpha(\lambda) \leq c_0 \alpha \quad \text{and} \quad 0 \leq g_\alpha(\lambda) \leq c_1 \alpha^{-1}$$

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

Due to the nonlinearity of F , in the convergence analysis we need to deal with terms like $r_\alpha(F'(x)^*F'(x)) - r_\alpha(F'(z)^*F'(z))$ for $x, z \in D(F)$. The following condition on $\{g_\alpha\}$ will be employed to give a unified treatment, where \mathbb{C} denotes the complex plane.

Assumption 2.2. *For each $\alpha > 0$, $g_\alpha(\lambda)$ can be extended to a complex analytic function defined on a simply connected domain $D_\alpha \subset \mathbb{C}$ such that $[0, 1] \subset D_\alpha$, and there is a contour $\Gamma_\alpha \subset D_\alpha$ enclosing $[0, 1]$ such that*

$$(2.1) \quad \frac{|\lambda| + \sigma}{|\lambda - \sigma|} \leq b, \quad \forall \lambda \in \Gamma_\alpha, \alpha > 0 \text{ and } \sigma \in [0, 1],$$

where b is a constant independent of $\alpha > 0$. Moreover, for each $s > 0$ there is a constant b_s such that

$$(2.2) \quad \int_{\Gamma_\alpha} \frac{|r_\alpha(\lambda)|}{|\lambda|^s} |d\lambda| \leq b_s \alpha^{1-s}$$

and

$$(2.3) \quad \int_{\Gamma_\alpha} \frac{|g_\alpha(\lambda)|}{|\lambda|^s} |d\lambda| \leq b_s \alpha^{-s}$$

for all $\alpha > 0$.

By using the spectral integrals for self-adjoint operators, it follows easily from (2.1) in Assumption 2.2 that for any bounded linear operator A with $\|A\| \leq 1$ there holds

$$(2.4) \quad \|(\lambda I - A^*A)^{-1}(A^*A)^\nu\| \leq \frac{b}{|\lambda|^{1-\nu}}$$

for $\lambda \in \Gamma_\alpha$ and $0 \leq \nu \leq 1$. Moreover, since Assumption 2.2 implies that $r_\alpha(\lambda)$ is analytic in D_α for each $\alpha > 0$, there holds the Riesz-Dunford formula (see [3])

$$r_\alpha(A^*A) = \frac{1}{2\pi i} \int_{\Gamma_\alpha} r_\alpha(\lambda)(\lambda I - A^*A)^{-1} d\lambda$$

for any linear operator A satisfying $\|A\| \leq 1$. Similarly, we have the Riesz-Dunford formula for $g_\alpha(A^*A)$.

Example 2.3. The following spectral filter functions $\{g_\alpha\}$ satisfy Assumptions 2.1 and 2.2 (see [3, 9]):

(a) $g_\alpha(\lambda) = \frac{(\alpha+\lambda)^m - \alpha^m}{\lambda(\alpha+\lambda)^m}$, arising from the linear iterated Tikhonov regularization of order m , where $m \geq 1$ is a fixed integer.

(b) $g_\alpha(\lambda) = (1 - e^{-\lambda/\alpha})/\lambda$ arising from the linear asymptotic regularization.

(c) $g_\alpha(\lambda) = \sum_{i=0}^{\lceil 1/\alpha \rceil - 1} (1 - \lambda)^i$ arising from the linear Landweber iteration, where $\lceil 1/\alpha \rceil$ denotes the largest integer not greater than $1/\alpha$.

(d) $g_\alpha(\lambda) = \sum_{i=1}^{\lceil 1/\alpha \rceil} (1 + \lambda)^{-i}$ arising from the linear implicit iteration method.

Since our goal is to give some convergence results on the method defined by (1.3) and (1.5) under only condition (1.8) on F , it is necessary to derive some useful consequences. The condition (1.8) clearly implies that $F'(x)$ is uniformly bounded in $B_\rho(x^\dagger)$. Thus, for simplicity of exposition, we assume that the operator F is properly scaled so that

$$(2.5) \quad \|F'(x)\| \leq \min\{1, \sqrt{c_0\alpha_0q}\}, \quad \forall x \in B_\rho(x^\dagger).$$

Lemma 2.4. *Let $\{g_\alpha\}$ satisfy Assumption 2.2 and let F satisfy (1.8). Then there hold¹*

$$(2.6) \quad \|r_\alpha(\mathcal{A}_x) - r_\alpha(\mathcal{A}_z)\| \lesssim K_0\|x - z\| + \frac{1}{\sqrt{\alpha}}K_1(\|F'_x(x - z)\| + \|F'_z(x - z)\|),$$

$$(2.7) \quad \|r_\alpha(\mathcal{B}_x) - r_\alpha(\mathcal{B}_z)\| \lesssim K_0\|x - z\| + \frac{1}{\sqrt{\alpha}}K_1(\|F'_x(x - z)\| + \|F'_z(x - z)\|),$$

$$(2.8) \quad \|F'_x[r_\alpha(\mathcal{A}_x) - r_\alpha(\mathcal{A}_z)]\| \lesssim K_0\|x - z\|\alpha^{1/2} \\ + K_1(\|F'_x(x - z)\| + \|F'_z(x - z)\|),$$

$$(2.9) \quad \|F'_x[g_\alpha(\mathcal{A}_x) - g_\alpha(\mathcal{A}_z)]\| \lesssim \frac{1}{\sqrt{\alpha}}K_0\|x - z\| \\ + \frac{1}{\alpha}K_1(\|F'_x(x - z)\| + \|F'_z(x - z)\|)$$

for all $\alpha > 0$ and $x, z \in B_\rho(x^\dagger)$, where $F'_x := F'(x)$, $\mathcal{A}_x := F_x'^* F'_x$ and $\mathcal{B}_x := F_x' F_x'^*$.

Proof. We will only give the proof of (2.8), since the proofs of the other three inequalities are similar. Since $r_\alpha(\lambda)$ is analytic in D_α for each $\alpha > 0$, it follows from the Riesz-Dunford formula that

$$(2.10) \quad F'_x[r_\alpha(\mathcal{A}_x) - r_\alpha(\mathcal{A}_z)] = \frac{1}{2\pi i} \int_{\Gamma_\alpha} r_\alpha(\lambda)F'_x[(\lambda I - \mathcal{A}_x)^{-1} - (\lambda I - \mathcal{A}_z)^{-1}] d\lambda.$$

¹Throughout this paper we will use C to denote a generic constant independent of δ and k . We will use the convention $\Phi \lesssim \Psi$ to mean $\Phi \leq C\Psi$ for some generic constant C .

Observe that

$$F'_x[(\lambda I - \mathcal{A}_x)^{-1} - (\lambda I - \mathcal{A}_z)^{-1}] = F'_x(\lambda I - \mathcal{A}_x)^{-1}(F'^*_x - F'^*_z)F'_z(\lambda I - \mathcal{A}_z)^{-1} + (\lambda I - \mathcal{B}_x)^{-1}B_x(F'_x - F'_z)(\lambda I - \mathcal{A}_z)^{-1}.$$

Then we have from (2.4) that

$$\|F'_x[(\lambda I - \mathcal{A}_x)^{-1} - (\lambda I - \mathcal{A}_z)^{-1}]\| \lesssim |\lambda|^{-1/2}\|(F'_x - F'_z)(\lambda I - \mathcal{A}_x)^{-1}F'^*_x\| + \|(F'_x - F'_z)(\lambda I - \mathcal{A}_z)^{-1}\|.$$

By (1.8) and (2.4) we then obtain

$$\begin{aligned} \|F'_x[(\lambda I - \mathcal{A}_x)^{-1} - (\lambda I - \mathcal{A}_z)^{-1}]\| &\lesssim |\lambda|^{-1/2}K_0\|x - z\|\|\mathcal{B}_x(\lambda I - \mathcal{B}_x)^{-1}\| \\ &\quad + |\lambda|^{-1/2}K_1\|F'_x(x - z)\|\|(\lambda I - \mathcal{A}_x)^{-1}F'_x\| \\ &\quad + K_0\|x - z\|\|F'_z(\lambda I - \mathcal{A}_z)^{-1}\| \\ &\quad + K_1\|F'_z(x - z)\|\|(\lambda I - \mathcal{A}_z)^{-1}\| \\ &\lesssim K_0\|x - z\|\|\lambda\|^{-1/2} \\ &\quad + K_1(\|F'_x(x - z)\| + \|F'_z(x - z)\|)\|\lambda\|^{-1}. \end{aligned}$$

Combining this with (2.10) yields

$$\begin{aligned} \|F'_x[r_\alpha(\mathcal{A}_x) - r_\alpha(\mathcal{A}_z)]\| &\lesssim K_0\|x - z\|\int_{\Gamma_\alpha} \frac{|r_\alpha(\lambda)|}{|\lambda|^{1/2}}|d\lambda| \\ &\quad + K_1(\|F'_x(x - z)\| + \|F'_z(x - z)\|)\int_{\Gamma_\alpha} \frac{|r_\alpha(\lambda)|}{|\lambda|}|d\lambda|. \end{aligned}$$

With the help of (2.2) in Assumption 2.2, we obtain the desired estimate. □

3. CONVERGENCE ANALYSIS

Several convergence results have been proved in [8] on the Newton type methods (1.3) and the discrepancy principle (1.5). In this section, we derive further convergence results under weaker conditions. In contrast to the counterpart in [8], the new convergence results require neither (1.9) nor (1.10).

Our analysis involves the iterative sequence $\{x_k\}$ defined by (1.3) corresponding to the noise-free case, i.e.,

$$(3.1) \quad x_{k+1} = x_0 - g_{\alpha_k}(F'(x_k)^*F'(x_k))F'(x_k)^*(F(x_k) - y - F'(x_k)(x_k - x_0)).$$

For simplicity of presentation, we use the notations

$$e_k := x_k - x^\dagger, \quad e_k^\delta := x_k^\delta - x^\dagger,$$

and

$$\begin{aligned} \mathcal{A} &:= F'(x^\dagger)^*F'(x^\dagger), & \mathcal{A}_k &:= F'(x_k)^*F'(x_k), & \mathcal{A}_k^\delta &:= F'(x_k^\delta)^*F'(x_k^\delta), \\ \mathcal{B} &:= F'(x^\dagger)F'(x^\dagger)^*, & \mathcal{B}_k &:= F'(x_k)F'(x_k)^*, & \mathcal{B}_k^\delta &:= F'(x_k^\delta)F'(x_k^\delta)^*. \end{aligned}$$

3.1. Justification of the method. In this subsection we show that the method given by (1.3) and (1.5) with $\tau > 1$ is well defined, i.e., $x_k^\delta \in B_\rho(x^\dagger)$ for $0 \leq k \leq k_\delta$ and k_δ is finite. To this end, we introduce the integer \tilde{k}_δ defined by

$$(3.2) \quad \alpha_{\tilde{k}_\delta} \leq \left(\frac{\delta}{\gamma_0 \|e_0\|} \right)^2 < \alpha_k, \quad 0 \leq k < \tilde{k}_\delta,$$

where $\gamma_0 > \sqrt{c_0 q}/(\tau - 1)$ is a fixed number. Due to (1.4), such \tilde{k}_δ exists and is finite.

Lemma 3.1. *Let F satisfy (1.8) and (2.5), let $\{\alpha_k\}$ satisfy (1.4), and let $\{g_\alpha\}$ satisfy Assumption 2.1. Let $\tau > 1$ be a given number. If $(2 + \sqrt{c_1} \gamma_0) \|e_0\| < \rho$ and if $(K_0 + K_1) \|e_0\|$ is suitably small, then*

$$(3.3) \quad \|e_k^\delta\| \leq (2 + \sqrt{c_1} \gamma_0) \|e_0\|, \quad \|F'(x^\dagger) e_k^\delta\| \leq 2\sqrt{q}(\sqrt{c_0} + \gamma_0) \|e_0\| \alpha_k^{1/2}$$

and

$$(3.4) \quad \|F'(x^\dagger) e_k^\delta - y^\delta + y\| \leq \delta + (\sqrt{c_0 q} + C(K_0 + K_1) \|e_0\|) \|e_0\| \alpha_k^{1/2}$$

for $0 \leq k \leq \tilde{k}_\delta$.

Proof. We first prove (3.3) by induction. In view of the scaling condition (2.5), it is trivial for $k = 0$. Now we assume that (3.3) is true for some $0 \leq k < \tilde{k}_\delta$. We set

$$u_k^\delta := F(x_k^\delta) - y - F'(x_k^\delta) e_k^\delta.$$

Then, it follows from (1.3) that

$$(3.5) \quad e_{k+1}^\delta = r_{\alpha_k} (\mathcal{A}_k^\delta) e_0 - g_{\alpha_k} (\mathcal{A}_k^\delta) F'(x_k^\delta)^* (y - y^\delta + u_k^\delta).$$

In view of (1.8) and the induction hypotheses we have $\|u_k^\delta\| \lesssim (K_0 + K_1) \|e_0\|^2 \alpha_k^{1/2}$. Consequently, by using Assumption 2.1 and $\delta \leq \gamma_0 \|e_0\| \alpha_k^{1/2}$, we obtain

$$\|e_{k+1}^\delta\| \leq \|e_0\| + \sqrt{c_1} \left(\delta \alpha_k^{-1/2} + C(K_0 + K_1) \|e_0\|^2 \right) \leq (2 + \sqrt{c_1} \gamma_0) \|e_0\|$$

if $(K_0 + K_1) \|e_0\|$ is suitably small.

Next, we multiply (3.5) by $F'(x^\dagger)$ to obtain

$$\begin{aligned} F'(x^\dagger) e_{k+1}^\delta - y^\delta + y &= F'(x_k^\delta) r_{\alpha_k} (\mathcal{A}_k^\delta) e_0 + [F'(x^\dagger) - F'(x_k^\delta)] r_{\alpha_k} (\mathcal{A}_k^\delta) e_0 \\ &\quad - [F'(x^\dagger) - F'(x_k^\delta)] g_{\alpha_k} (\mathcal{A}_k^\delta) F'(x_k^\delta)^* (y - y^\delta + u_k^\delta) \\ &\quad - g_{\alpha_k} (\mathcal{B}_k^\delta) \mathcal{B}_k^\delta u_k^\delta - r_{\alpha_k} (\mathcal{B}_k^\delta) (y^\delta - y). \end{aligned}$$

By using (1.8), the induction hypotheses and $\delta \leq \gamma_0 \|e_0\| \alpha_k^{1/2}$, we can follow the argument in [8] to obtain

$$\|F'(x^\dagger) e_{k+1}^\delta - y^\delta + y\| \leq \delta + (\sqrt{c_0} + C(K_0 + K_1) \|e_0\|) \|e_0\| \alpha_k^{1/2}.$$

By using (1.4) we then obtain

$$(3.6) \quad \|F'(x^\dagger) e_{k+1}^\delta - y^\delta + y\| \leq \delta + (\sqrt{c_0 q} + C(K_0 + K_1) \|e_0\|) \|e_0\| \alpha_{k+1}^{1/2}.$$

Thus, using again the fact that $\delta \leq \gamma_0 \|e_0\| \alpha_k^{1/2} \leq \gamma_0 \sqrt{q} \|e_0\| \alpha_{k+1}^{1/2}$, we have for suitably small $(K_0 + K_1) \|e_0\|$ such that

$$\|F'(x^\dagger) e_{k+1}^\delta\| \leq 2\delta + 2\sqrt{c_0 q} \|e_0\| \alpha_{k+1}^{1/2} \leq 2\sqrt{q}(\sqrt{c_0} + \gamma_0) \|e_0\| \alpha_{k+1}^{1/2}.$$

By induction we therefore obtain (3.3) for $0 \leq k \leq \tilde{k}_\delta$. The inequality (3.4) follows immediately from (3.6) since it is trivial for $k = 0$. □

For the noise-free iterates $\{x_k\}$ defined by (3.1), we have

$$(3.7) \quad e_{k+1} = r_{\alpha_k}(\mathcal{A}_k)e_0 - g_{\alpha_k}(\mathcal{A}_k)F'(x_k)^*(F(x_k) - y - F'(x_k)e_k).$$

By using the same argument as above we can derive the following estimates.

Lemma 3.2. *Under the conditions in Lemma 3.1, if $(K_0 + K_1)\|e_0\|$ is suitably small, then*

$$x_k \in B_\rho(x^\dagger), \quad \|e_k\| \lesssim \|e_0\| \quad \text{and} \quad \|F'(x^\dagger)e_k\| \lesssim \alpha_k^{1/2}\|e_0\|$$

for all $k \geq 0$.

By using Lemma 3.1, (1.8), the definition of \tilde{k}_δ , and $\gamma_0 > \sqrt{c_0q}/(\tau - 1)$, we can follow the argument in [8] to derive that

$$\|F(x_{\tilde{k}_\delta}^\delta) - y^\delta\| \leq (1 + \sqrt{c_0q}/\gamma_0 + C(K_0 + K_1)\|e_0\|)\delta \leq \tau\delta$$

if $(K_0 + K_1)\|e_0\|$ is suitably small. Thus the integer k_δ defined by the discrepancy principle (1.5) with $\tau > 1$ must satisfy $k_\delta \leq \tilde{k}_\delta$. This together with Lemma 3.1 shows that the method given by (1.3) and (1.5) is well defined.

3.2. Stability estimates. In this subsection we derive the estimate on $\|x_k^\delta - x_k\|$ together with other useful estimates which will be used in the convergence analysis.

Lemma 3.3. *Let all the conditions in Lemma 3.1 and Assumption 2.2 hold. If $(K_0 + K_1)\|e_0\|$ is suitably small, then*

$$(3.8) \quad \|x_k^\delta - x_k\| \lesssim \frac{\delta}{\sqrt{\alpha_k}}$$

and

$$(3.9) \quad \|F(x_{\tilde{k}_\delta}^\delta) - F(x_k) - y^\delta + y\| \leq (1 + C(K_0 + K_1)\|e_0\|)\delta$$

for all $0 \leq k \leq \tilde{k}_\delta$.

Proof. By induction we first show for $0 \leq k \leq \tilde{k}_\delta$ that

$$(3.10) \quad \|x_k^\delta - x_k\| \leq 2\sqrt{c_1} \frac{\delta}{\sqrt{\alpha_k}} \quad \text{and} \quad \|F'(x^\dagger)(x_k^\delta - x_k)\| \leq 3\delta.$$

It is clear that (3.10) is trivial for $k = 0$. Now we assume that (3.10) is true for some $0 \leq k < \tilde{k}_\delta$. From (1.8) and Lemma 3.2 one can see that

$$\begin{aligned} & \| (F'(x_k) - F'(x^\dagger))(x_k^\delta - x_k) \| \\ & \lesssim K_0\|e_0\| \|F'(x^\dagger)(x_k^\delta - x_k)\| + K_1\|e_0\| \|x_k^\delta - x_k\| \alpha_k^{1/2}. \end{aligned}$$

Thus, with the help of the induction hypotheses, we obtain

$$(3.11) \quad \|F'(x_k)(x_k^\delta - x_k)\| \lesssim K_1\|e_0\|\delta + \|F'(x^\dagger)(x_k^\delta - x_k)\| \lesssim \delta.$$

Similarly, by using Lemma 3.1 we have

$$(3.12) \quad \|F'(x_k^\delta)(x_k^\delta - x_k)\| \lesssim K_1\|e_0\|\delta + \|F'(x^\dagger)(x_k^\delta - x_k)\| \lesssim \delta.$$

Now we introduce the notations

$$u_k := F(x_k) - y - F'(x_k)e_k \quad \text{and} \quad u_k^\delta := F(x_k^\delta) - y - F'(x_k^\delta)e_k^\delta.$$

It then follows from (3.5) and (3.7) that

$$(3.13) \quad x_{k+1}^\delta - x_{k+1} = h_1 + h_2 + h_3 + h_4,$$

where

$$\begin{aligned} h_1 &:= (r_{\alpha_k}(\mathcal{A}_k^\delta) - r_{\alpha_k}(\mathcal{A}_k)) e_0, \\ h_2 &:= g_{\alpha_k}(\mathcal{A}_k^\delta) F'(x_k^\delta)^* (y^\delta - y), \\ h_3 &:= (g_{\alpha_k}(\mathcal{A}_k) F'(x_k)^* - g_{\alpha_k}(\mathcal{A}_k^\delta) F'(x_k^\delta)^*) u_k, \\ h_4 &:= g_{\alpha_k}(\mathcal{A}_k^\delta) F'(x_k^\delta)^* (u_k - u_k^\delta). \end{aligned}$$

By (2.6) and (2.8) in Lemma 2.4, (3.11), (3.12) and the induction hypotheses, we obtain

$$(3.14) \quad \|h_1\| \lesssim (K_0 + K_1) \|e_0\| \frac{\delta}{\sqrt{\alpha_k}}, \quad \|F'(x_k^\delta)h_1\| \lesssim (K_0 + K_1) \|e_0\| \delta.$$

By Assumption 2.1 we have

$$(3.15) \quad \|h_2\| \leq \sqrt{c_1} \frac{\delta}{\sqrt{\alpha_k}}, \quad \|F'(x_k^\delta)h_2 - y^\delta + y\| \leq \delta.$$

In order to estimate h_3 and $F'(x_k^\delta)h_3$, we note that

$$h_3 = g_{\alpha_k}(\mathcal{A}_k) (F'(x_k)^* - F'(x_k^\delta)^*) u_k + (g_{\alpha_k}(\mathcal{A}_k) - g_{\alpha_k}(\mathcal{A}_k^\delta)) F'(x_k^\delta)^* u_k$$

and

$$F'(x_k^\delta)h_3 = (F'(x_k^\delta) - F'(x_k)) g_{\alpha_k}(\mathcal{A}_k) F'(x_k)^* u_k + (r_{\alpha_k}(\mathcal{B}_k^\delta) - r_{\alpha_k}(\mathcal{B}_k)) u_k.$$

Therefore, it follows from Assumption 2.1, (1.8) and Lemma 2.4 that

$$\begin{aligned} \|h_3\| &\lesssim \frac{1}{\sqrt{\alpha_k}} K_0 \|x_k^\delta - x_k\| \|u_k\| \\ &\quad + \frac{1}{\alpha_k} K_1 (\|F'(x_k)(x_k^\delta - x_k)\| + \|F'(x_k^\delta)(x_k^\delta - x_k)\|) \|u_k\| \end{aligned}$$

and

$$\begin{aligned} \|F'(x_k^\delta)h_3\| &\lesssim K_0 \|x_k^\delta - x_k\| \|u_k\| \\ &\quad + \frac{1}{\sqrt{\alpha_k}} K_1 (\|F'(x_k)(x_k^\delta - x_k)\| + \|F'(x_k^\delta)(x_k^\delta - x_k)\|) \|u_k\|. \end{aligned}$$

Observe that (1.8) and Lemma 3.2 imply $\|u_k\| \lesssim (K_0 + K_1) \|e_0\|^2 \alpha_k^{1/2}$. Therefore, it follows from the induction hypotheses, (3.11) and (3.12) that

$$(3.16) \quad \|h_3\| \lesssim (K_0 + K_1) \|e_0\| \frac{\delta}{\sqrt{\alpha_k}}, \quad \|F'(x_k^\delta)h_3\| \lesssim (K_0 + K_1) \|e_0\| \delta.$$

For h_4 and $F'(x_k^\delta)h_4$ we have from Assumption 2.1 that

$$\|h_4\| \lesssim \frac{1}{\sqrt{\alpha_k}} \|u_k - u_k^\delta\|, \quad \|F'(x_k^\delta)h_4\| \lesssim \|u_k - u_k^\delta\|.$$

By using (1.8), Lemma 3.1, Lemma 3.2, the induction hypothesis, (3.12) and the fact that $\delta/\sqrt{\alpha_k} \lesssim \|e_0\|$, we have

$$\begin{aligned} \|u_k - u_k^\delta\| &\leq \|F(x_k^\delta) - F(x_k) - F'(x_k)(x_k^\delta - x_k)\| + \|(F'(x_k^\delta) - F'(x_k))e_k^\delta\| \\ &\lesssim (K_0 + K_1) \|e_0\| \delta. \end{aligned}$$

Therefore,

$$(3.17) \quad \|h_4\| \lesssim (K_0 + K_1) \|e_0\| \frac{\delta}{\sqrt{\alpha_k}}, \quad \|F'(x_k^\delta)h_4\| \lesssim (K_0 + K_1) \|e_0\| \delta.$$

Combining the estimates (3.14)–(3.17) yields

$$\|x_{k+1}^\delta - x_{k+1}\| \leq (\sqrt{c_1} + C(K_0 + K_1)\|e_0\|) \frac{\delta}{\sqrt{\alpha_k}}$$

and

$$\|F'(x_k^\delta)(x_{k+1}^\delta - x_{k+1}) - y^\delta + y\| \leq (1 + C(K_0 + K_1)\|e_0\|) \delta.$$

Consequently, it follows from (1.8) and Lemma 3.1 that

$$\|(F'(x_k^\delta) - F'(x^\dagger))(x_{k+1}^\delta - x_{k+1})\| \lesssim (K_0 + K_1)\|e_0\|\delta.$$

Hence, if $(K_0 + K_1)\|e_0\|$ is suitably small, we have

$$\|F'(x^\dagger)(x_{k+1}^\delta - x_{k+1})\| \leq \|F'(x_k^\delta)(x_{k+1}^\delta - x_{k+1})\| + C(K_0 + K_1)\|e_0\|\delta \leq 3\delta.$$

We therefore complete the proof of (3.10).

Now we can follow the argument in [8] and derive from (1.8), Lemma 3.1, and (3.8) that

$$\|F'(x_k^\delta)(x_k^\delta - x_k) - y^\delta + y\| \leq (1 + C(K_0 + K_1)\|e_0\|)\delta, \quad 0 \leq k \leq \tilde{k}_\delta,$$

which in turn implies (3.9). □

As immediate consequences of Lemma 3.3, we have

$$(3.18) \quad \|F(x_{k_\delta}) - y\| \lesssim \delta$$

and

$$(3.19) \quad \delta \lesssim \|F(x_k) - y\|, \quad 0 \leq k < k_\delta,$$

if $(K_0 + K_1)\|e_0\|$ is suitably small, where k_δ is the integer determined by the discrepancy principle (1.5) with $\tau > 1$.

3.3. Convergence. In this subsection we show that $x_{k_\delta}^\delta \rightarrow x^\dagger$ as $\delta \rightarrow 0$. In order to achieve this, we first show that $x_k \rightarrow x^\dagger$ as $k \rightarrow \infty$.

Lemma 3.4. *Let all the conditions in Lemma 3.3 be fulfilled and let $x_0 - x^\dagger \in \mathcal{N}(F'(x^\dagger))^\perp$. If $(K_0 + K_1)\|e_0\|$ is suitably small, then*

$$(3.20) \quad \lim_{k \rightarrow \infty} \|e_k\| = 0 \quad \text{and} \quad \lim_{k \rightarrow \infty} \frac{\|F'(x^\dagger)e_k\|}{\sqrt{\alpha_k}} = 0.$$

Proof. From (3.7) it follows that

$$(3.21) \quad e_{k+1} = r_{\alpha_k}(\mathcal{A})e_0 + [r_{\alpha_k}(\mathcal{A}_k) - r_{\alpha_k}(\mathcal{A})]e_0 - g_{\alpha_k}(\mathcal{A}_k)F'(x_k)^*u_k,$$

where $u_k := F(x_k) - y - F'(x_k)e_k$. By using (1.8) and (2.6) in Lemma 2.4 we obtain

$$\begin{aligned} \|e_{k+1} - r_{\alpha_k}(\mathcal{A})e_0\| &\lesssim \frac{1}{\sqrt{\alpha_k}}(K_0 + K_1)\|e_k\|\|F'(x^\dagger)e_k\| + K_0\|e_0\|\|e_k\| \\ &\quad + \frac{1}{\sqrt{\alpha_k}}K_1\|e_0\|(\|F'(x^\dagger)e_k\| + \|F'(x_k)e_k\|). \end{aligned}$$

Note that (1.8) and Lemma 3.2 imply

$$(3.22) \quad \|F'(x_k)e_k\| \leq (1 + C(K_0 + K_1)\|e_0\|)\|F'(x^\dagger)e_k\|.$$

Consequently,

$$(3.23) \quad \|e_{k+1}\| \leq \|r_{\alpha_k}(\mathcal{A})e_0\| + C(K_0 + K_1)\|e_0\| \left(\|e_k\| + \frac{1}{\sqrt{\alpha_k}}\|F'(x^\dagger)e_k\| \right).$$

Next we multiply (3.21) by $F'(x^\dagger)$ and obtain

$$(3.24) \quad \begin{aligned} F'(x^\dagger)e_{k+1} &= F'(x^\dagger)r_{\alpha_k}(\mathcal{A})e_0 + F'(x^\dagger)[r_{\alpha_k}(\mathcal{A}_k) - r_{\alpha_k}(\mathcal{A})]e_0 \\ &\quad + [F'(x^\dagger) - F'(x_k)]g_{\alpha_k}(\mathcal{A}_k)F'(x_k)^*u_k + g_{\alpha_k}(\mathcal{B}_k)\mathcal{B}_k u_k. \end{aligned}$$

By applying Assumption 2.1, (1.8), and (2.8) in Lemma 2.4 we have

$$\begin{aligned} &\|F'(x^\dagger)e_{k+1} - F'(x^\dagger)r_{\alpha_k}(\mathcal{A})e_0\| \\ &\leq K_0\|e_0\|\|e_k\|\alpha_k^{1/2} + K_1\|e_0\|(\|F'(x^\dagger)e_k\| + \|F'(x_k)e_k\|) \\ &\quad + K_0(K_0 + K_1)\|e_k\|^2\|F'(x^\dagger)e_k\| + \frac{1}{\sqrt{\alpha_k}}K_1(K_0 + K_1)\|e_k\|\|F'(x_k)e_k\|^2 \\ &\quad + (K_0 + K_1)\|e_k\|\|F'(x^\dagger)e_k\|. \end{aligned}$$

Thus, we may use Lemma 3.2 and (3.22) to conclude

$$(3.25) \quad \|F'(x^\dagger)e_{k+1}\| \leq \|F'(x^\dagger)r_{\alpha_k}(\mathcal{A})e_0\| + C(K_0 + K_1)\|e_0\| \left(\|e_k\|\alpha_k^{1/2} + \|F'(x^\dagger)e_k\| \right).$$

Now we set

$$\eta_k := \|e_k\| + \frac{\|F'(x^\dagger)e_k\|}{\sqrt{\alpha_k}}, \quad \varepsilon_k := \|r_{\alpha_k}(\mathcal{A})e_0\| + \frac{\|F'(x^\dagger)r_{\alpha_k}(\mathcal{A})e_0\|}{\sqrt{\alpha_k}}.$$

Then it follows from (3.23), (3.25), and (1.4) that

$$(3.26) \quad \eta_{k+1} \leq \sqrt{q}\varepsilon_k + \sigma\eta_k, \quad k = 0, 1, \dots,$$

where $\sigma := C(K_0 + K_1)\|e_0\|$ for a generic constant C . We may assume $\sigma < 1$ by taking $(K_0 + K_1)\|e_0\|$ to be suitably small. By iterating this inequality we get

$$\eta_k \leq \sqrt{q} \sum_{j=0}^{k-1} \varepsilon_j \sigma^{k-1-j} + \sigma^{k-1} \eta_0, \quad k = 1, 2, \dots.$$

Since $e_0 \in \mathcal{N}(F'(x^\dagger))^\perp$, from Assumption 2.1 we have $\varepsilon_j \rightarrow 0$ as $j \rightarrow \infty$. Thus for any $\varepsilon > 0$ there is $k_0 > 0$ such that $\varepsilon_j \leq \varepsilon$ for all $j > k_0$. Consequently, for $k > k_0$,

$$\eta_k \leq \sqrt{q}\varepsilon \sum_{j=k_0+1}^k \sigma^{k-1-j} + C \sum_{j=0}^{k_0} \sigma^{k-1-j} \leq \frac{\varepsilon}{1-\sigma} + C \frac{\sigma^{k-1-k_0}}{1-\sigma}.$$

Recall that $\sigma < 1$, we therefore have $\eta_k \leq C\varepsilon$ if k is sufficiently large. Since $\varepsilon > 0$ is arbitrary, we obtain $\eta_k \rightarrow 0$ as $k \rightarrow \infty$. □

We are now ready to give the main result concerning the convergence of the method (1.3) under the discrepancy principle (1.5).

Theorem 3.5. *Let F satisfy (1.8), let $\{g_\alpha\}$ satisfy Assumptions 2.1 and 2.2, and let $\{\alpha_k\}$ satisfy (1.4). Let $\tau > 1$ be a given number. If $(K_0 + K_1)\|e_0\|$ is suitably small, then the method given by (1.3) and (1.5) is well defined, and if $e_0 \in \mathcal{N}(F'(x^\dagger))^\perp$ and $\mathcal{N}(F'(x^\dagger)) \subset \mathcal{N}(F'(x))$ for all $x \in B_\rho(x^\dagger)$, then*

$$\lim_{\delta \rightarrow 0} x_{k_\delta}^\delta = x^\dagger$$

for the integer k_δ defined by the discrepancy principle (1.5).

Proof. We have shown that the method given by (1.3) and (1.5) is well defined. It remains only to show that $x_{k_\delta}^\delta \rightarrow x^\dagger$ as $\delta \rightarrow 0$.

Assume first that there is a sequence $\delta_n \searrow 0$ such that $k_n := k_{\delta_n} \rightarrow k$ as $n \rightarrow \infty$ for some finite integer k . Without loss of generality, we can assume that $k_n = k$ for all n . Since (3.18) implies

$$\|F(x_{k_n}) - y\| \lesssim \delta_n,$$

by sending $n \rightarrow \infty$ we obtain $F(x_k) = y$. This together with (1.8) gives $F'(x^\dagger)e_k = 0$, i.e., $e_k \in \mathcal{N}(F'(x^\dagger))$. On the other hand, since $\mathcal{N}(F'(x^\dagger)) \subset \mathcal{N}(F'(x))$ for all $x \in B_\rho(x^\dagger)$, we can see from the definition of x_k and the condition $e_0 \in \mathcal{N}(F'(x^\dagger))^\perp$ that $e_k \in \mathcal{N}(F'(x^\dagger))^\perp$. Therefore $x_k = x^\dagger$, which together with Lemma 3.3 implies $x_{k_n}^{\delta_n} \rightarrow x^\dagger$ as $n \rightarrow \infty$.

Assume next that there is a sequence $\delta_n \searrow 0$ such that $k_n := k_{\delta_n} \rightarrow \infty$ as $n \rightarrow \infty$. Then Lemma 3.4 implies that $\|e_{k_n}\| \rightarrow 0$ and $\|F'(x^\dagger)e_{k_n-1}\|/\sqrt{\alpha_{k_n}} \rightarrow 0$ as $n \rightarrow \infty$. But from (3.19) we have

$$\delta_n \lesssim \|F(x_{k_n-1}) - y\| \lesssim \|F'(x^\dagger)e_{k_n-1}\|.$$

Consequently, $\delta_n/\sqrt{\alpha_{k_n}} \rightarrow 0$ as $n \rightarrow \infty$. By Lemma 3.3 we again obtain $x_{k_n}^{\delta_n} \rightarrow x^\dagger$ as $n \rightarrow \infty$. □

3.4. Rates of convergence. Although Theorem 3.5 gives the convergence of $x_{k_\delta}^\delta$ to x^\dagger as $\delta \rightarrow 0$, it does not give the convergence speed. For the ill-posed problem, the convergence in general could be arbitrarily slow if there is no further source condition on $e_0 := x_0 - x^\dagger$. In this section we derive the order optimal convergence rates for the method given by (1.3) and (1.5) when e_0 satisfies the source condition (1.7) for some $0 < \nu \leq 1/2$ and $\omega \in X$.

Lemma 3.6. *Let F satisfy (1.8), let $\{g_\alpha\}$ satisfy Assumptions 2.1 and 2.2, and let $\{\alpha_k\}$ satisfy (1.4). If $(K_0 + K_1)\|e_0\|$ is suitably small and if $e_0 = \mathcal{A}^\nu \omega$ for some $0 < \nu \leq 1/2$ and $\omega \in X$, then*

$$(3.27) \quad \|e_k\| \leq C_\nu \|\omega\| \alpha_k^\nu, \quad \|F'(x^\dagger)e_k\| \leq C_\nu \|\omega\| \alpha_k^{\nu+1/2}$$

for all $k \geq 0$, and

$$(3.28) \quad \delta \leq C_\nu \alpha_{k_\delta}^{\nu+1/2} \|\omega\|$$

for the integer k_δ determined by the discrepancy principle (1.5) with $\tau > 1$, where C_ν is a positive constant depending only on ν .

Proof. We will prove (3.27) by using (3.26). We may assume $\sigma \leq 1/(2\sqrt{q})$ by taking $(K_0 + K_1)\|e_0\|$ to be suitably small. Since $e_0 = \mathcal{A}^\nu \omega$ for some $0 < \nu \leq 1/2$, we have from Assumption 2.1 that

$$\varepsilon_k \leq C'_\nu \alpha_k^\nu \|\omega\|, \quad k \geq 0,$$

where $C'_\nu := c_0^{\nu/2} + c_0^{\nu/2+1/4}$. Consequently, it follows from (3.26) that

$$\eta_{k+1} \leq \sqrt{q} C'_\nu \alpha_k^\nu \|\omega\| + \frac{1}{2\sqrt{q}} \eta_k, \quad k = 0, 1, \dots$$

By using (1.4), we can derive by induction that

$$\eta_k \leq 2q C'_\nu \alpha_k^\nu \|\omega\|, \quad k = 0, 1, \dots$$

Recalling the definition of η_k , we therefore obtain (3.27).

For the inequality (3.28), we note that (3.19) and (3.27) imply

$$\delta \lesssim \|F'(x^\dagger)e_{k_\delta-1}\| \leq C_\nu \alpha_{k_\delta-1}^{\nu+1/2} \|\omega\|.$$

Consequently, by using (1.4) we obtain the desired estimate. □

In the following we assume e_0 satisfies the source condition (1.7) with $0 < \nu \leq 1/2$ and derive the estimate on $\|e_{k_\delta}^\delta\|$. It follows from (3.8) and (3.28) that

$$(3.29) \quad \|e_{k_\delta}^\delta\| \lesssim \|e_{k_\delta}\| + \frac{\delta}{\sqrt{\alpha_{k_\delta}}} \leq \|e_{k_\delta}\| + C_\nu \|\omega\|^{1/(1+2\nu)} \delta^{2\nu/(1+2\nu)}.$$

Thus, in order to obtain the order optimality, it suffices to derive estimate on $\|e_{k_\delta}\|$.

We first consider the method (1.3) with $g_\alpha(\lambda) = (\alpha + \lambda)^{-1}$, which is exactly the iteratively regularized Gauss-Newton method (see [1])

$$(3.30) \quad x_{k+1}^\delta = x_k^\delta - (\alpha_k I + F'(x_k^\delta)^* F'(x_k^\delta))^{-1} (F'(x_k^\delta)^* (F(x_k^\delta) - y^\delta) + \alpha_k (x_k^\delta - x_0)).$$

The order optimal rates of convergence have been derived in [4] when the iteration is terminated by the discrepancy principle (1.5) with sufficiently large τ . The following result improves the one in [4] in the sense that it requires only $\tau > 1$, which is significant for the accuracy in numerical computation.

Theorem 3.7. *Let F satisfy (1.8), let $\{\alpha_k\}$ satisfy (1.4), and let $\tau > 1$ be a given number. If $(K_0 + K_1)\|e_0\|$ is suitably small, and if $e_0 = \mathcal{A}^\nu \omega$ for some $0 < \nu \leq 1/2$ and $\omega \in \mathcal{N}(F'(x^\dagger))^\perp \subset X$, then for the sequence $\{x_k^\delta\}$ defined by (3.30) and the integer k_δ determined by (1.5) there holds*

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

where C_ν is a constant depending only on ν .

Proof. We first use (3.26) to give an estimate on $\|e_k\|$. Note that $r_\alpha(\lambda) = \alpha(\alpha + \lambda)^{-1}$. It follows from (1.4) that $r_{\alpha_{k+1}}(\lambda) \leq r_{\alpha_k}(\lambda) \leq q r_{\alpha_{k+1}}(\lambda)$ for all $\lambda \geq 0$. Thus,

$$\|r_{\alpha_{k+1}}(\mathcal{A})e_0\| \leq \|r_{\alpha_k}(\mathcal{A})e_0\| \leq q \|r_{\alpha_{k+1}}(\mathcal{A})e_0\|$$

and

$$(3.31) \quad \|F'(x^\dagger)r_{\alpha_{k+1}}(\mathcal{A})e_0\| \leq \|F'(x^\dagger)r_{\alpha_k}(\mathcal{A})e_0\| \leq q \|F'(x^\dagger)r_{\alpha_{k+1}}(\mathcal{A})e_0\|.$$

Consequently, $\varepsilon_k \leq q\varepsilon_{k+1}$ for all $k \geq 0$. Hence, if $(K_0 + K_1)\|e_0\|$ is so small that $\sigma \leq 1/(2q)$, we can obtain from (3.26) that $\eta_k \leq 2q\varepsilon_k$ for all $k \geq 0$, i.e.,

$$(3.32) \quad \|e_k\| \lesssim \|r_{\alpha_k}(\mathcal{A})e_0\| + \frac{\|F'(x^\dagger)r_{\alpha_k}(\mathcal{A})e_0\|}{\sqrt{\alpha_k}}, \quad k = 0, 1, \dots$$

Next we use the special form of $r_\alpha(\lambda)$ to give an estimate on $\|F'(x^\dagger)r_{\alpha_k}(\mathcal{A})e_0\|$. From (3.24), (1.8), and Lemma 3.2 it follows that

$$(3.33) \quad \begin{aligned} \|F'(x^\dagger)e_{k+1} - F'(x^\dagger)r_{\alpha_k}(\mathcal{A})e_0\| &\leq \|F'(x^\dagger)(r_{\alpha_k}(\mathcal{A}_k) - r_{\alpha_k}(\mathcal{A}))e_0\| \\ &+ C(K_0 + K_1)\|e_0\| \|F'(x^\dagger)e_k\|. \end{aligned}$$

We use $r_\alpha(\lambda) = \alpha(\alpha + \lambda)^{-1}$ and write

$$F'(x^\dagger)(r_{\alpha_k}(\mathcal{A}_k) - r_{\alpha_k}(\mathcal{A}))e_0 = u_1 + u_2,$$

where

$$\begin{aligned} u_1 &:= F'(x^\dagger)(\alpha_k I + \mathcal{A}_k)^{-1} F'(x_k)^* (F'(x^\dagger) - F'(x_k)) r_{\alpha_k}(\mathcal{A}) e_0, \\ u_2 &:= F'(x^\dagger)(\alpha_k I + \mathcal{A}_k)^{-1} (F'(x^\dagger)^* - F'(x_k)^*) F'(x^\dagger) r_{\alpha_k}(\mathcal{A}) e_0. \end{aligned}$$

By using (1.8) and Lemma 3.2 we have

$$\|u_1\| \leq \beta_k (K_0 \|e_0\| \|F'(x^\dagger) r_{\alpha_k}(\mathcal{A}) e_0\| + K_1 \|e_0\| \|F'(x^\dagger) e_k\|),$$

where

$$\beta_k := \|F'(x^\dagger)(\alpha_k I + \mathcal{A}_k)^{-1} F'(x_k)^*\|.$$

In view of (1.8) and Lemma 3.2 we obtain

$$\begin{aligned} \beta_k &\leq \|(\alpha_k I + \mathcal{B}_k)^{-1} \mathcal{B}_k\| + \|(F'(x^\dagger) - F'(x_k)) (\alpha_k I + \mathcal{A}_k)^{-1} F'(x_k)^*\| \\ (3.34) \quad &\leq 1 + K_0 \|e_k\| + \frac{1}{2\sqrt{\alpha_k}} K_1 \|F'(x_k) e_k\| \lesssim 1. \end{aligned}$$

Therefore,

$$\|u_1\| \lesssim K_0 \|e_0\| \|F'(x^\dagger) r_{\alpha_k}(\mathcal{A}) e_0\| + K_1 \|e_0\| \|F'(x^\dagger) e_k\|.$$

In order to estimate u_2 , we note that (1.8) and (3.34) imply

$$\begin{aligned} &\|(F'(x^\dagger) - F'(x_k))(\alpha_k I + \mathcal{A}_k)^{-1} F'(x^\dagger)^*\| \\ &\leq K_0 \|e_k\| \beta_k + K_1 \|F'(x_k) e_k\| \|(\alpha_k I + \mathcal{A}_k)^{-1} F'(x^\dagger)^*\| \\ &\lesssim K_0 \|e_k\| + K_1 \|F'(x^\dagger) e_k\| \|F'(x^\dagger)(\alpha_k I + \mathcal{A}_k)^{-1}\|. \end{aligned}$$

Similar to the derivation of (3.34) we can show that

$$\|F'(x^\dagger)(\alpha_k I + \mathcal{A}_k)^{-1}\| \lesssim 1/\sqrt{\alpha_k}.$$

Therefore, with the help of Lemma 3.2, we obtain

$$\|(F'(x^\dagger) - F'(x_k))(\alpha_k I + \mathcal{A}_k)^{-1} F'(x^\dagger)^*\| \lesssim (K_0 + K_1) \|e_0\|.$$

Hence

$$\|u_2\| \lesssim (K_0 + K_1) \|e_0\| \|F'(x^\dagger) r_{\alpha_k}(\mathcal{A}) e_0\|.$$

Combining the estimates on u_1 and u_2 we obtain

$$\begin{aligned} &\|F'(x^\dagger) (r_{\alpha_k}(\mathcal{A}_k) - r_{\alpha_k}(\mathcal{A})) e_0\| \\ &\lesssim (K_0 + K_1) \|e_0\| \|F'(x^\dagger) r_{\alpha_k}(\mathcal{A}) e_0\| + K_1 \|e_0\| \|F'(x^\dagger) e_k\|. \end{aligned}$$

Plugging this estimate into (3.33) yields

$$\begin{aligned} &\|F'(x^\dagger) e_{k+1} - F'(x^\dagger) r_{\alpha_k}(\mathcal{A}) e_0\| \lesssim (K_0 + K_1) \|e_0\| \|F'(x^\dagger) r_{\alpha_k}(\mathcal{A}) e_0\| \\ (3.35) \quad &+ (K_0 + K_1) \|e_0\| \|F'(x^\dagger) e_k\|. \end{aligned}$$

This in particular implies

$$(3.36) \quad \|F'(x^\dagger) e_{k+1}\| \lesssim \|F'(x^\dagger) r_{\alpha_k}(\mathcal{A}) e_0\| + (K_0 + K_1) \|e_0\| \|F'(x^\dagger) e_k\|.$$

By using (3.31), we can conclude from (3.36) that if $(K_0 + K_1) \|e_0\|$ is suitably small, then

$$\|F'(x^\dagger) e_k\| \lesssim \|F'(x^\dagger) r_{\alpha_k}(\mathcal{A}) e_0\|, \quad k = 0, 1, \dots.$$

Combining this estimate with (3.35) we obtain for $k \geq 1$ that

$$\begin{aligned} \|F'(x^\dagger)r_{\alpha_{k-1}}(\mathcal{A})e_0\| &\lesssim (K_0 + K_1)\|e_0\|\|F'(x^\dagger)r_{\alpha_{k-1}}(\mathcal{A})e_0\| + \|F'(x^\dagger)e_k\| \\ &\quad + (K_0 + K_1)\|e_0\|\|F'(x^\dagger)e_{k-1}\| \\ &\lesssim (K_0 + K_1)\|e_0\|\|F'(x^\dagger)r_{\alpha_{k-1}}(\mathcal{A})e_0\| + \|F'(x^\dagger)e_k\|. \end{aligned}$$

Thus, if $(K_0 + K_1)\|e_0\|$ is suitably small, this together with (3.31) implies

$$(3.37) \quad \|F'(x^\dagger)r_{\alpha_k}(\mathcal{A})e_0\| \lesssim \|F'(x^\dagger)e_k\|, \quad k = 0, 1, \dots,$$

since it is trivial for $k = 0$.

Now we take $k = k_\delta$ in (3.37) and use (3.18) to obtain

$$\|F'(x^\dagger)r_{\alpha_{k_\delta}}(\mathcal{A})e_0\| \lesssim \delta.$$

With the help of the source condition $e_0 = \mathcal{A}^\nu\omega$ with $0 < \nu \leq 1/2$ and the interpolation inequality we obtain

$$\begin{aligned} \|r_{\alpha_{k_\delta}}(\mathcal{A})e_0\| &\leq \|r_{\alpha_{k_\delta}}(\mathcal{A})\omega\|^{1/(1+2\nu)}\|F'(x^\dagger)r_{\alpha_{k_\delta}}(\mathcal{A})e_0\|^{2\nu/(1+2\nu)} \\ &\leq C_\nu\|\omega\|^{1/(1+2\nu)}\delta^{2\nu/(1+2\nu)}. \end{aligned}$$

Combining the above two estimates with (3.32) and using (3.28) we obtain

$$\|e_{k_\delta}\| \lesssim C_\nu\|\omega\|^{1/(1+2\nu)}\delta^{2\nu/(1+2\nu)} + \frac{\delta}{\sqrt{\alpha_{k_\delta}}} \leq C_\nu\|\omega\|^{1/(1+2\nu)}\delta^{2\nu/(1+2\nu)}.$$

This together with (3.29) then completes the proof. □

Next we return to the general method (1.3) together with the discrepancy principle (1.5). We need the following additional condition.

Assumption 3.8. *There exist positive constants $0 < m \leq M < \infty$ such that*

$$m\|F'(x^\dagger)h\| \leq \|F'(x)h\| \leq M\|F'(x^\dagger)h\|, \quad h \in X$$

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

The following result enables us to use Assumption 3.8 in the derivation of convergence rates.

Proposition 3.9. *Let $T, S : X \rightarrow X$ be two self-adjoint bounded linear operators. If $\|Sh\| \leq C_0\|Th\|$ for all $h \in X$, then $R(S) \subset R(T)$ and $\|T^{-1}S\| \leq C_0$.*

Proof. This is [5, Proposition 2.1]. □

Theorem 3.10. *Let F satisfy (1.8) and Assumption 3.8, let $\{g_\alpha\}$ satisfy Assumptions 2.1 and 2.2, and let $\{\alpha_k\}$ satisfy (1.4). Let $\{x_k^\delta\}$ be defined by (1.3) and let k_δ be the integer determined by (1.5) with $\tau > 1$. If $(K_0 + K_1)\|e_0\|$ is suitably small and if $e_0 = \mathcal{A}^\nu\omega$ for some $0 < \nu \leq 1/2$ and $\omega \in \mathcal{N}(F'(x^\dagger))^\perp \subset X$, then*

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

where C_ν is a constant depending only on ν .

Proof. We will use (3.7). By Assumption 3.8 we have

$$m\|\mathcal{A}^{1/2}h\| \leq \|\mathcal{A}_k^{1/2}h\| \leq M\|\mathcal{A}^{1/2}h\|, \quad h \in X,$$

which implies that $m^2\mathcal{A} \leq \mathcal{A}_k \leq M^2\mathcal{A}$. Since the function $t \rightarrow t^{2\nu}$ is operator monotone for $0 \leq \nu \leq 1/2$, we have $m^{4\nu}\mathcal{A}^{2\nu} \leq \mathcal{A}_k^{2\nu} \leq M^{4\nu}\mathcal{A}^{2\nu}$, which implies

$$m^{2\nu}\|\mathcal{A}^\nu h\| \leq \|\mathcal{A}_k^\nu h\| \leq M^{2\nu}\|\mathcal{A}^\nu h\|, \quad h \in X.$$

In view of of Proposition 3.9 we obtain

$$R(\mathcal{A}^\nu) = R(\mathcal{A}_k^\nu), \quad \|\mathcal{A}_k^{-\nu}\mathcal{A}^\nu\| \lesssim 1 \quad \text{and} \quad \|\mathcal{A}^{-\nu}\mathcal{A}_k^\nu\| \lesssim 1.$$

Since $e_0 = \mathcal{A}^\nu\omega$, we must have $\omega_k \in X$ such that

$$e_0 = \mathcal{A}_k^\nu\omega_k \quad \text{and} \quad \|\omega_k\| \lesssim \|\omega\|.$$

Therefore, it follows from (3.7) that

$$e_{k+1} = r_{\alpha_k}(\mathcal{A}_k)\mathcal{A}_k^\nu\omega_k - g_{\alpha_k}(\mathcal{A}_k)F'(x_k)^*(F(x_k) - y - F'(x_k)e_k).$$

By the polar factorization we can write $F'(x_k)^* = \mathcal{A}_k^{1/2}U$ for some partial isometry $U : Y \rightarrow X$. Therefore,

$$(3.38) \quad e_{k+1} = \mathcal{A}_k^\nu v_k,$$

where

$$v_k = r_{\alpha_k}(\mathcal{A}_k)\omega_k - g_{\alpha_k}(\mathcal{A}_k)\mathcal{A}_k^{1/2-\nu}U(F(x_k) - y - F'(x_k)e_k).$$

By using Assumption 2.1, (1.8), Lemma 3.2 and Lemma 3.6, for $0 \leq k \leq k_\delta$ we have

$$\begin{aligned} \|v_k\| &\lesssim \|\omega_k\| + \alpha_k^{-\nu-1/2}\|F(x_k) - y - F'(x_k)e_k\| \\ &\lesssim \|\omega_k\| + \alpha_k^{-\nu-1/2}(K_0 + K_1)\|e_k\|\|F'(x_k)^\dagger e_k\| \\ &\lesssim \|\omega\| + (K_0 + K_1)\|e_0\|\|\omega\| \lesssim \|\omega\|. \end{aligned}$$

By using (3.38), Assumption 3.8, (1.8), and (3.18) we also have

$$\begin{aligned} \|F'(x_{k_\delta-1})\mathcal{A}_{k_\delta-1}^\nu v_{k_\delta-1}\| &= \|F'(x_{k_\delta-1})e_{k_\delta}\| \lesssim \|F'(x_{k_\delta-1})^\dagger e_{k_\delta}\| \\ &\lesssim \|F(x_{k_\delta}) - y\| \lesssim \delta. \end{aligned}$$

Therefore, by using (3.38) and the two estimates above, we have from the interpolation inequality that

$$\begin{aligned} \|e_{k_\delta}\| &= \|\mathcal{A}_{k_\delta-1}^\nu v_{k_\delta-1}\| \leq \|v_{k_\delta-1}\|^{1/(1+2\nu)}\|F'(x_{k_\delta-1})\mathcal{A}_{k_\delta-1}^\nu v_{k_\delta-1}\|^{2\nu/(1+2\nu)} \\ &\leq C_\nu\|\omega\|^{1/(1+2\nu)}\delta^{2\nu/(1+2\nu)}. \end{aligned}$$

Combining this with (3.29) gives the desired estimate. □

4. NUMERICAL EXPERIMENTS

In this section we will present the numerical results for a parameter identification problem in differential equations.

Consider the identification of the diffusion parameter a in

$$(4.1) \quad \begin{cases} -(au')' = f & \text{in } (0, 1), \\ u(0) = u_0, \quad u(1) = u_1 \end{cases}$$

from the L^2 measurement u^δ of u on $(0, 1)$ satisfying $\|u^\delta - u\|_{L^2} \leq \delta$, where u_0, u_1 , and $f \in L^2(0, 1)$ are given. It is well known that for $a \in L^\infty(0, 1)$ with $a \geq \underline{a} > 0$ on $(0, 1)$, (4.1) has a unique solution $u = u(a) \in H^2(0, 1)$. In order to put the problem

into the framework of Hilbert space, we assume $a^\dagger \in H^1(0, 1)$ is the sought solution with $a^\dagger \geq 2\underline{a} > 0$ on $(0, 1)$. We then define F as

$$F : H^1(0, 1) \rightarrow L^2(0, 1), \quad F(a) := u(a)$$

with

$$D(F) := \{a \in H^1(0, 1) : a \geq \underline{a} > 0 \text{ on } (0, 1)\}.$$

Since $H^1(0, 1)$ embeds into $L^\infty(0, 1)$, F is well defined. It is well known that F is Fréchet differentiable on $D(F)$ with

$$\begin{aligned} F'(a)h &= A(a)^{-1}((h(u(a)))'), \\ F'(a)^*w &= -B^{-1}((u(a))'(A(a)^{-1}w)'), \end{aligned}$$

for all $a \in D(F)$, $h \in H^1(0, 1)$ and $w \in L^2(0, 1)$, where, for each $a \in D(F)$, $A(a) : V := H_0^1 \cap H^2(0, 1) \rightarrow L^2(0, 1)$ is an isomorphism defined by $A(a)\eta = -(a\eta)'$ for $\eta \in H_0^1 \cap H^2$, and $B : D(B) := \{\varphi \in H^2(0, 1) : \varphi'(0) = \varphi'(1) = 0\} \rightarrow L^2(0, 1)$ is defined by $B\varphi := -\varphi'' + \varphi$. It was shown in [6, Example 11.1] that there is a ball $B_\rho(a^\dagger)$ around a^\dagger such that

$$(4.2) \quad \|F(a_2) - F(a_1) - F'(a_1)(a_2 - a_1)\| \leq \frac{1}{2}\|F(a_2) - F(a_1)\|, \quad a_1, a_2 \in B_\rho(a^\dagger).$$

In the following we will verify (1.8). It is easy to see that for $a_1, a_2 \in B_\rho(a^\dagger)$ and $h \in H^1$ we get

$$(F'(a_2) - F'(a_1))h = A(a_2)^{-1} \left[((a_2 - a_1)(F'(a_1)h)')' + (h(u(a_2) - u(a_1)))' \right].$$

Let V' denote the anti-dual of V with respect to the bilinear form $\int_\Omega \varphi \psi dx$. Recall that $A(a)$ can be extended as an isomorphism $A(a) : L^2 \rightarrow V'$ so that $A(a)^{-1} : V' \rightarrow L^2$ is uniformly bounded around a^\dagger . We have from the above equation that

$$(4.3) \quad \begin{aligned} \|(F'(a_2) - F'(a_1))h\|_{L^2} &\lesssim \|((a_2 - a_1)(F'(a_1)h)')'\|_{V'} \\ &\quad + \|(h(u(a_2) - u(a_1)))'\|_{V'}. \end{aligned}$$

We claim that for any $h \in H^1$ and $\varphi \in H_0^1 \cap H^2$ there holds

$$(4.4) \quad \|(h\varphi)'\|_{V'} \lesssim \|h\|_{H^1} \|\varphi\|_{L^2}.$$

To see this, by the divergence theorem we have for $\psi \in V$, that

$$\int_0^1 (h\varphi)'\psi dx = \int_0^1 \varphi(h\psi)' dx \leq \|\varphi\|_{L^2} \|(h\psi)'\|_{L^2}.$$

Since $H^1 \hookrightarrow L^\infty$, we have

$$\|(h\psi)'\|_{L^2} \leq \|h\|_{L^\infty} \|\psi''\|_{L^2} + \|h'\|_{L^2} \|\psi'\|_{L^\infty} \lesssim \|h\|_{H^1} \|\psi\|_V.$$

Therefore,

$$\int_0^1 (h\varphi)'\psi dx \lesssim \|h\|_{H^1} \|\varphi\|_{L^2} \|\psi\|_V, \quad \forall \psi \in V,$$

which implies the claim (4.4).

Observe that both $F'(a_1)h$ and $u(a_2) - u(a_1)$ are in $H_0^1 \cap H^2$, we may apply the claim (4.4) to estimate the right-hand side of (4.3). Thus,

$$\|(F'(a_2) - F'(a_1))h\|_{L^2} \lesssim \|a_2 - a_1\|_{H^1} \|F'(a_1)h\|_{L^2} + \|F(a_2) - F(a_1)\|_{L^2} \|h\|_{H^1}.$$

Since (4.2) implies $\|F(a_2) - F(a_1)\|_{L^2} \lesssim \|F'(a_1)(a_2 - a_1)\|_{L^2}$, (1.8) is thus satisfied.

In the following we report some numerical results on the method given by (1.3) and (1.5) with $g_\alpha(\lambda) = (\alpha + \lambda)^{-1}$ which, in the current context, defines the iterative solutions $\{a_k^\delta\}$ by

$$(4.5) \quad a_{k+1}^\delta = a_k^\delta - (\alpha_k I + F'(a_k^\delta))^* F'(a_k^\delta)^{-1} (F'(a_k^\delta)^* (F(a_k^\delta) - y^\delta) + \alpha_k (a_k^\delta - a_0))$$

and determines the stopping index k_δ by

$$(4.6) \quad \|F(a_{k_\delta}^\delta) - y^\delta\| \leq \tau\delta < \|F(a_k^\delta) - y^\delta\|, \quad 0 \leq k < k_\delta.$$

During the computation, all differential equations are solved approximately by 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.

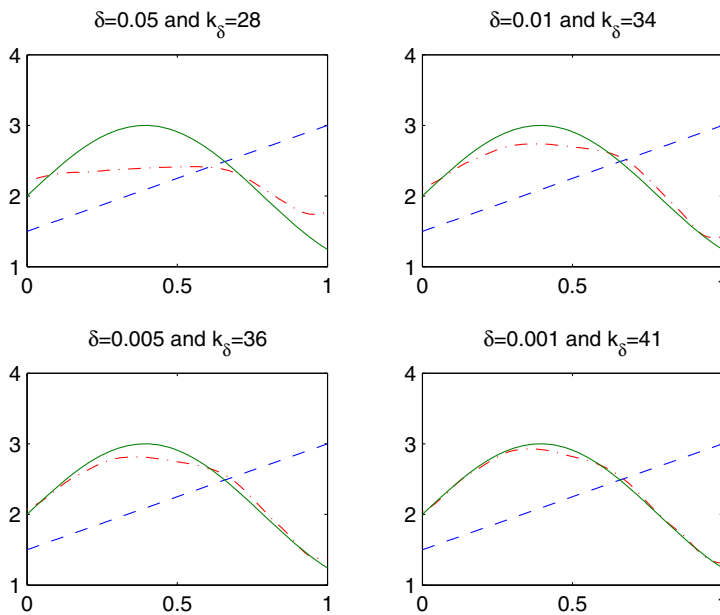


FIGURE 1. Numerical results corresponding to the initial guess given by (4.7)

We consider the estimation of a in (4.1) with $f = 4 - 4 \cos(4t) + 2 \sin(4t) + 8t \cos(4t)$ and $u_0 = u_1 = 0$. If $u = t(1 - t)$, then $a^\dagger = 2 + \sin(4t)$ is the desired solution. When applying the method (4.5)–(4.6), we take $\alpha_k = (1.5)^{-k}$ and $\tau = 1.01$ and use random noise data u^δ satisfying $\|u^\delta - u\|_{L^2[0,1]} = \delta$ with noise level $\delta > 0$. In Figure 1 and Figure 2 we plot the numerical results corresponding to two different choices of the initial guess:

$$(4.7) \quad a_0 = 3 - t$$

and

$$(4.8) \quad a_0 = 2 + \sin(4t) + 3 \left(9 - 4t + 4t^2 + \frac{4}{1 - e} e^t + \frac{4e}{1 - e} e^{-t} \right)$$

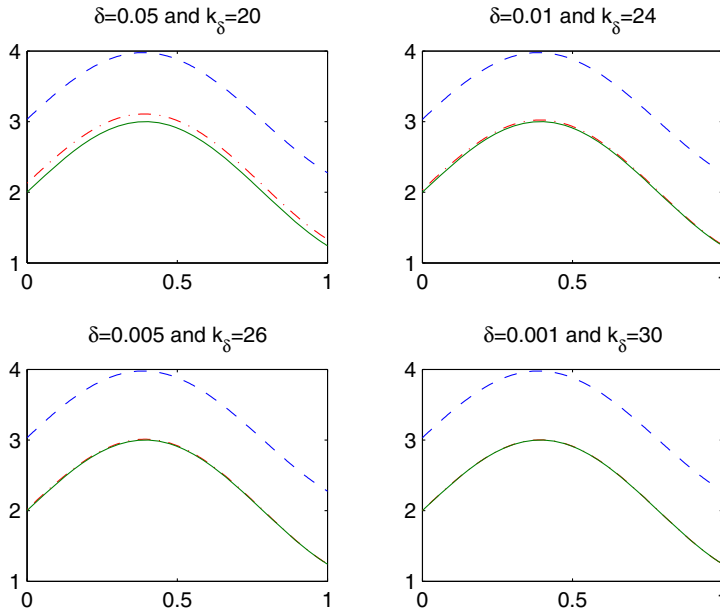


FIGURE 2. Numerical results corresponding to the initial guess given by (4.8)

with various values of the noise level $\delta > 0$, where the solid, dashed, and dashed-dot curves denote the exact solution a^\dagger , the initial guess a_0 , and the computed solution $a_{k_\delta}^\delta$, respectively. For the a_0 given by (4.7), $a_0 - a^\dagger$ has no sourcewise representation $a_0 - a^\dagger \in \mathcal{R}((F'(a^\dagger))^* F'(a^\dagger))^\nu$ with a good $\nu > 0$. Thus no good convergence rate can be expected if the method starts from this a_0 . Theorem 3.5, however, guarantees the convergence of $a_{k_\delta}^\delta$ to a^\dagger as $\delta \rightarrow 0$, which is confirmed by Figure 1. On the other hand, for the a_0 given by (4.8) one can check $a_0 - a^\dagger \in R(F'(a^\dagger)^*)$. Thus, according to Theorem 3.7, one can expect $\|a_{k_\delta}^\delta - a^\dagger\|_{H^1} = O(\delta^{1/2})$. This shows the much better accuracy which is confirmed by Figure 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 for solving nonlinear operator equations without regularity. A new approach*, Russ. Acad. Sci. Dokl. Math., 47 (1993), 451–454. MR1241957 (94m:65099)
3. A. B. Bakushinsky and M. Yu. Kokurin, *Iterative Methods for Approximate Solutions of Inverse Problems*, Mathematics and its applications, Springer, 2004. MR2133802 (2006e:47025)
4. 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)
5. A. Böttcher, B. Hofmann, U. Tautenhahn and M. Yamamoto, *Convergence rates for Tikhonov regularization from different kinds of smoothness conditions*, Applicable Analysis, 85 (2006), 555–578. MR2213075 (2006k:65150)

6. 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)
7. M. Hanke, A. Neubauer and O. Scherzer, *A convergence analysis of the Landweber iteration for nonlinear ill-posed problems*, Numer. Math., 72 (1995), 21–37. MR1359706 (96i:65046)
8. Q. 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 (2010a:65090)
9. Q. Jin and U. Tautenhahn, *Inexact Newton regularization methods in Hilbert scales*, Numer. Math., 117 (2011), 555–579. MR2772419 (2012b:65078)
10. B. Kaltenbacher, *Some Newton-type methods for the regularization of nonlinear ill-posed problems*, Inverse Problems 13 (1997), 729–753. MR1451018 (98h:65025)

DEPARTMENT OF MATHEMATICS, VIRGINIA TECH, BLACKSBURG, VIRGINIA 24061

E-mail address: qnjin@math.vt.edu

Current address: Mathematical Sciences Institute, The Australian National University, Canberra, ACT 0200, Australia

E-mail address: Qinian.Jin@anu.edu.au