

A GENERAL CONVERGENCE ANALYSIS OF SOME NEWTON-TYPE METHODS FOR NONLINEAR INVERSE PROBLEMS*

QINIAN JIN[†]

Abstract. We consider the methods $x_{n+1}^\delta = x_n^\delta - g_{\alpha_n}(F'(x_n^\delta)^*F'(x_n^\delta))F'(x_n^\delta)^*(F(x_n^\delta) - y^\delta)$ for solving nonlinear ill-posed inverse problems $F(x) = y$ using the only available noise data y^δ satisfying $\|y^\delta - y\| \leq \delta$ with a given small noise level $\delta > 0$. We terminate the iteration by the discrepancy principle $\|F(x_{n_\delta}^\delta) - y^\delta\| \leq \tau\delta < \|F(x_n^\delta) - y^\delta\|$, $0 \leq n < n_\delta$, with a given number $\tau > 1$. Under certain conditions on $\{\alpha_n\}$ and F , we prove for a large class of spectral filter functions $\{g_\alpha\}$ the convergence of $x_{n_\delta}^\delta$ to a true solution as $\delta \rightarrow 0$. Moreover, we derive the order optimal rates of convergence when certain Hölder source conditions hold. Numerical examples are given to test the theoretical results.

Key words. nonlinear inverse problems, Newton-type methods, discrepancy principle, convergence, order optimal convergence rates

AMS subject classifications. 65J15, 65J20, 47H17

DOI. 10.1137/100804231

1. Introduction. In this paper we consider the nonlinear equations

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

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

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

with a given small noise level $\delta > 0$. Due to the ill-posedness, it is challenging to produce from y^δ a stable approximate solution to x^\dagger , and the regularization techniques must be taken into account.

Many regularization methods have been considered for solving (1.1) in the last two decades. Due to the straightforward implementation, iterative methods are attractive for solving nonlinear inverse problems. In this paper we will consider a class of Newton-type methods. To motivate the methods, we let x_n^δ be a current iterate. We may approximate $F(x)$ by its linearization $F(x_n^\delta) + F'(x_n^\delta)(x - x_n^\delta)$ around x_n^δ . Thus, instead of (1.1) we have the approximate equation

$$(1.3) \quad F'(x_n^\delta)(x - x_n^\delta) = y^\delta - F(x_n^\delta).$$

If $F'(x_n^\delta)$ has bounded inverse, the usual Newton method defines the next iterate by solving (1.3) for x . For nonlinear ill-posed inverse problems, however, $F'(x_n^\delta)$ in

*Received by the editors August 2, 2010; accepted for publication (in revised form) January 3, 2011; published electronically March 22, 2011.

<http://www.siam.org/journals/sinum/49-2/80423.html>

[†]Department of Mathematics, Virginia Tech, Blacksburg, VA 24061 (qnjin@math.vt.edu).

general is not invertible, and (1.3) usually is ill-posed. Therefore, one should use the regularization methods to solve (1.3) approximately. Let $\{g_\alpha\}$ be a family of spectral filter functions. We can apply the linear regularization method defined by $\{g_\alpha\}$ to (1.3) to produce the next iterate. This leads to the Newton-type methods

$$(1.4) \quad x_{n+1}^\delta = x_n^\delta - g_{\alpha_n} (F'(x_n^\delta)^* F'(x_n^\delta)) F'(x_n^\delta)^* (F(x_n^\delta) - y^\delta),$$

where $x_0^\delta := x_0 \in D(F)$ is an initial guess of x^\dagger and $\{\alpha_n\}$ is a sequence of positive numbers. By taking g_α to be various functions, (1.4) then produces the nonlinear Landweber iteration [5], the Levenberg–Marquardt method [3, 7], the exponential Euler iteration [6], and the first-stage Runge–Kutta-type regularization [11].

In this paper we will consider the Newton-type methods (1.4) in a unified way by assuming that $\{\alpha_n\}$ is an a priori given sequence of positive numbers with suitable properties. We will terminate the iteration by the discrepancy principle

$$(1.5) \quad \|F(x_{n_\delta}^\delta) - y^\delta\| \leq \tau\delta < \|F(x_n^\delta) - y^\delta\|, \quad 0 \leq n < n_\delta,$$

with a given number $\tau > 1$ and consider the approximation property of $x_{n_\delta}^\delta$ to x^\dagger as $\delta \rightarrow 0$. For a large class of spectral filter functions $\{g_\alpha\}$ we will establish the convergence of $x_{n_\delta}^\delta$ to x^\dagger as $\delta \rightarrow 0$ and derive the order optimal convergence rates for the methods defined by (1.4) and (1.5). Our work not only reproduces those known results in [5, 7, 6, 11] but also presents new convergence results and new methods.

In the definition of the Newton-type methods (1.4), one may determine the sequence $\{\alpha_n\}$ adaptively during computation. Motivated by the inexact Newton methods in [2] for well-posed problems, the Levenberg–Marquardt scheme was considered in [3] for solving nonlinear inverse problems with $\{\alpha_n\}$ chosen adaptively so that

$$(1.6) \quad \|F(x_n^\delta) - y^\delta + F'(x_n^\delta)(x_{n+1}^\delta - x_n^\delta)\| \leq \mu_n \|F(x_n^\delta) - y^\delta\|$$

at each step, with the forcing terms $\mu_n \in (0, 1)$ being uniformly bounded below 1, and the discrepancy principle was used to terminate the iteration. The order optimal convergence rates were derived recently in [4]. The general methods (1.4) with $\{\alpha_n\}$ chosen adaptively to satisfy (1.6) were considered later in [12, 10], but only suboptimal convergence rates were derived in [13], and the convergence analysis is far from complete. The method of the present paper is essentially different in that the sequence $\{\alpha_n\}$ is given in an a priori way, which has the advantage of saving computational work.

There is another class of Newton-type regularization methods, i.e., the general iteratively regularized Gauss–Newton methods (see [1]) in which the iterates are defined successively by

$$(1.7) \quad x_{n+1}^\delta = x_0 - g_{\alpha_n} (F'(x_n^\delta)^* F'(x_n^\delta)) F'(x_n^\delta)^* (F(x_n^\delta) - y^\delta - F'(x_n^\delta)(x_n^\delta - x_0)).$$

Such methods, coupled with the discrepancy principle (1.5), have been studied extensively in [8]. The methods (1.4) are significantly different from (1.7) in that the methods (1.4) produce the next iterate near the current x_n^δ , while the methods (1.7) define the iterates near the initial guess x_0 .

This paper is organized as follows. In section 2 we first formulate the conditions on $\{\alpha_n\}$, $\{g_\alpha\}$, and F and state the main results on the convergence and rates of convergence for the methods defined by (1.4) and (1.5). We then give several examples of iterative methods that fit into the framework (1.4). In section 3 we prove some

crucial inequalities which are used in the convergence analysis. In section 4 we derive the order optimal convergence rate result when $x_0 - x^\dagger$ satisfies certain source conditions. In section 5 we show the convergence property without assuming any source conditions on $x_0 - x^\dagger$. Finally in section 6 we present numerical examples to test the theoretical results.

2. Main results. In order to carry out the convergence analysis on the methods defined by (1.4) and (1.5), we need to impose suitable conditions on $\{\alpha_n\}$, $\{g_\alpha\}$, and F . For the sequence $\{\alpha_n\}$ of positive numbers, we set

$$(2.1) \quad s_{-1} = 0, \quad s_n := \sum_{j=0}^n \frac{1}{\alpha_j}, \quad n = 0, 1, \dots$$

We will assume that there are constants $c_0 > 1$ and $c_1 > 0$ such that

$$(2.2) \quad \lim_{n \rightarrow \infty} s_n = \infty, \quad s_{n+1} \leq c_0 s_n, \quad \text{and} \quad 0 < \alpha_n \leq c_1, \quad n = 0, 1, \dots$$

For the spectral filter functions $\{g_\alpha\}$, we will assume the following two conditions, where \mathbb{C} denotes the complex plane.

Assumption 2.1. For each $\alpha > 0$, the function

$$\varphi_\alpha(\lambda) := g_\alpha(\lambda) - \frac{1}{\alpha + \lambda}$$

extends 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

$$(2.3) \quad |z| \geq \frac{1}{2}\alpha \quad \text{and} \quad \frac{|z| + \lambda}{|z - \lambda|} \leq b_0 \quad \forall z \in \Gamma_\alpha, \quad \alpha > 0; \quad \text{and} \quad \lambda \in [0, 1];$$

moreover, for all $0 < \alpha \leq c_1$ there holds

$$(2.4) \quad \int_{\Gamma_\alpha} |\varphi_\alpha(z)| |dz| \leq b_1,$$

where b_0 and b_1 are two positive constants independent of $\alpha > 0$.

Assumption 2.2. There is a constant $b_2 > 0$ such that for any sequence $\{\alpha_n\}$ of positive numbers with $\{s_n\}$ defined by (2.1), there hold

$$(2.5) \quad 0 \leq \lambda^\nu \prod_{k=j}^n r_{\alpha_k}(\lambda) \leq (s_n - s_{j-1})^{-\nu},$$

$$(2.6) \quad 0 \leq \lambda^\nu g_{\alpha_j}(\lambda) \prod_{k=j+1}^n r_{\alpha_k}(\lambda) \leq b_2 \frac{1}{\alpha_j} (s_n - s_{j-1})^{-\nu}$$

for $0 \leq \nu \leq 1$, $0 \leq \lambda \leq 1$, and $j = 0, \dots, n$, where $r_\alpha(\lambda) := 1 - \lambda g_\alpha(\lambda)$ is the residual function.

By using the spectral integrals for self-adjoint operators, it follows easily from (2.3) in Assumption 2.1 that for any bounded linear operator A with $\|A\| \leq 1$ there holds

$$(2.7) \quad \|(zI - A^*A)^{-1}(A^*A)^\nu\| \leq \frac{b_0}{|z|^{1-\nu}}$$

for $z \in \Gamma_\alpha$ and $0 \leq \nu \leq 1$. Moreover, since Assumption 2.1 implies that $\varphi_\alpha(z)$ is analytic in D_α for each $\alpha > 0$, there holds the Riesz–Dunford formula (see [1])

$$(2.8) \quad \varphi_\alpha(A^*A) = \frac{1}{2\pi i} \int_{\Gamma_\alpha} \varphi_\alpha(z)(zI - A^*A)^{-1} dz$$

for any linear operator A satisfying $\|A\| \leq 1$.

As a simple consequence of (2.5) in Assumption 2.2, we have for $0 \leq \nu \leq 1$ and $\alpha > 0$ that

$$(2.9) \quad 0 \leq \lambda^\nu (\alpha + \lambda)^{-1} \prod_{k=j+1}^n r_{\alpha_k}(\lambda) \leq 2\alpha^{\nu-1} (1 + \alpha(s_n - s_j))^{-\nu}$$

for all $0 \leq \lambda \leq 1$ and $j = 0, \dots, n$; see [9, Lemma 1].

For the nonlinear operator F , we need the following condition, which has been verified in [5] for several nonlinear inverse problems.

Assumption 2.3. (a) There exists $K_0 \geq 0$ such that

$$(2.10) \quad F'(x) = R(x, \bar{x})F'(\bar{x}) \quad \text{and} \quad \|I - R(x, \bar{x})\| \leq K_0 \|x - \bar{x}\|$$

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

(b) F is properly scaled so that $\|F'(x)\| \leq \min\{1, \sqrt{\alpha_0}\}$ for all $x \in B_\rho(x^\dagger)$.

The condition (a) in Assumption 2.3 clearly implies that $\|F'(x)\|$ is uniformly bounded over $B_\rho(x^\dagger)$. Thus, by multiplying (1.1) by a sufficiently small number, we may assume that F is properly scaled so that condition (b) in Assumption 2.3 is satisfied. A direct consequence of Assumption 2.3 is the inequality

$$(2.11) \quad \|F(x) - F(x^\dagger) - F'(x^\dagger)(x - x^\dagger)\| \leq \frac{1}{2} K_0 \|x - x^\dagger\| \|F'(x^\dagger)(x - x^\dagger)\|$$

for all $x \in B_\rho(x^\dagger)$, which will be frequently used in the convergence analysis.

Now we are ready to state the first main result concerning the rate of convergence of $x_{n_\delta}^\delta$ to x^\dagger as $\delta \rightarrow 0$ when $e_0 := x_0 - x^\dagger$ satisfies the source condition

$$(2.12) \quad x_0 - x^\dagger = (F'(x^\dagger)^* F'(x^\dagger))^\nu \omega$$

for some $0 < \nu \leq 1/2$ and $\omega \in \mathcal{N}(F'(x^\dagger))^\perp \subset X$, where n_δ is the integer determined by the discrepancy principle (1.5) with $\tau > 1$.

THEOREM 2.1. *Let F satisfy Assumption 2.3, let $\{g_\alpha\}$ satisfy Assumptions 2.1 and 2.2, and let $\{\alpha_n\}$ be a sequence of positive numbers satisfying (2.2). If $x_0 - x^\dagger$ satisfies the source condition (2.12) for some $0 < \nu \leq 1/2$ and $\omega \in \mathcal{N}(F'(x^\dagger))^\perp \subset X$, and if $K_0 \|\omega\|$ is suitably small, then the methods given by (1.4) and (1.5) with $\tau > 1$ are well defined and*

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

for the integer n_δ determined by (1.5), where $C_\nu > 0$ is a constant independent of δ and $\|\omega\|$.

Theorem 2.1 shows that the methods (1.4), together with the discrepancy principle (1.5), define order optimal regularization methods for each $0 < \nu \leq 1/2$. This result in particular reproduces the corresponding ones in [5, 6, 7, 11] for various iterative methods even with an improvement by relaxing $\tau > 2$ to $\tau > 1$.

Nevertheless, Theorem 2.1 does not provide the convergence of $x_{n_\delta}^\delta$ to x^\dagger as $\delta \rightarrow 0$ if there is no source condition imposed on $x_0 - x^\dagger$. In the next main result we will show the convergence of $x_{n_\delta}^\delta$ to x^\dagger as $\delta \rightarrow 0$ if $\{\alpha_n\}$ is a sequence of positive numbers satisfying

$$(2.13) \quad d_0 r^n \leq \alpha_n \leq d_1 r^n, \quad n = 0, 1, \dots,$$

for some constants $0 < d_0 \leq d_1 < \infty$ and $0 < r < 1$, which is one of the most important cases in applications.

THEOREM 2.2. *Let F satisfy Assumption 2.3, let $\{g_\alpha\}$ satisfy Assumptions 2.1 and 2.2, and let $\{\alpha_n\}$ be a sequence of positive numbers satisfying (2.13). If $x_0 - x^\dagger \in \mathcal{N}(F'(x^\dagger))^\perp$, and $K_0 \|x_0 - x^\dagger\|$ is suitably small, then the methods given by (1.4) and (1.5) with $\tau > 1$ are well defined, and*

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

for the integer n_δ determined by (1.5).

Theorem 2.2 extends the corresponding result in [7] for the Levenberg–Marquardt method to the general class of methods given by (1.4). We emphasize that the convergence result in Theorem 2.2 requires $\{\alpha_n\}$ to satisfy (2.13). It would be interesting if such a convergence result could be proved for a general sequence $\{\alpha_n\}$ satisfying (2.2) only. This, however, remains open; a new technique needs to be explored. We emphasize also that Theorem 2.2 requires F to satisfy Assumption 2.3. For some specific methods, however, it is possible to prove the convergence under weaker conditions on the nonlinear operator F . For instance, the convergence of the nonlinear Landweber iteration was proved in [5] under the condition

$$\|F(\bar{x}) - F(x) - F'(x)(\bar{x} - x)\| \leq \eta \|F(\bar{x}) - F(x)\|, \quad \bar{x}, x \in B_\rho(x^\dagger),$$

with a small number $\eta < 1/2$. It is not yet clear if the convergence of the general methods (1.4) can be derived under this weaker condition.

2.1. Examples. We conclude this section with several examples of the methods (1.4) in which the spectral filter functions $\{g_\alpha\}$ have been shown (see [9]) to satisfy Assumptions 2.1 and 2.2.

(a) We first consider for $\alpha > 0$ the function g_α given by

$$g_\alpha(\lambda) = \frac{1}{\alpha} \sum_{l=1}^N \alpha^l (\alpha + \lambda)^{-l} = \frac{(\alpha + \lambda)^N - \alpha^N}{\lambda(\alpha + \lambda)^N},$$

where $N \geq 1$ is a fixed integer. This function arises from the iterated Tikhonov regularization of order N for linear ill-posed problems. The corresponding method (1.4) becomes

$$\begin{aligned} u_{n,0} &= x_n^\delta, \\ u_{n,l+1} &= u_{n,l} - (\alpha_n I + F'(x_n^\delta))^* F'(x_n^\delta)^{-1} F'(x_n^\delta)^* (F(x_n^\delta) - y^\delta - F'(x_n^\delta)(x_n^\delta - u_{n,l})), \\ &\quad l = 0, \dots, N - 1, \\ x_{n+1}^\delta &= u_{n,N}. \end{aligned}$$

When $N = 1$, this is the Levenberg–Marquardt method (see [3, 7]).

(b) We consider the methods (1.4) with g_α given by

$$g_\alpha(\lambda) = \frac{1}{\lambda} \left(1 - e^{-\lambda/\alpha}\right),$$

which arises from the asymptotic regularization for linear ill-posed problems. In this method, the iterative sequence $\{x_n^\delta\}$ is equivalently defined as $x_{n+1}^\delta := x(1/\alpha_n)$, where $x(t)$ is the unique solution of the initial value problem

$$\begin{aligned} \frac{d}{dt}x(t) &= F'(x_n^\delta)^* (y^\delta - F(x_n^\delta) + F'(x_n^\delta)(x_n^\delta - x(t))), \quad t > 0, \\ x(0) &= x_n^\delta. \end{aligned}$$

This is the so-called exponential Euler iteration considered in [6]. The initial value problems can be solved by efficient ODE solvers such as the Runge–Kutta-type methods.

(c) For $0 < \alpha \leq 1$ we consider the function

$$(2.14) \quad g_\alpha(\lambda) = \sum_{l=0}^{[1/\alpha]-1} (1-\lambda)^l = \frac{1 - (1-\lambda)^{[1/\alpha]}}{\lambda},$$

which arises from the linear Landweber iteration, where $[1/\alpha]$ denotes the largest integer not greater than $1/\alpha$. The method (1.4) then becomes

$$\begin{aligned} u_{n,0} &= x_n^\delta, \\ u_{n,l+1} &= u_{n,l} - F'(x_n^\delta)^* (F(x_n^\delta) - y^\delta - F'(x_n^\delta)(x_n^\delta - u_{n,l})), \quad 0 \leq l \leq [1/\alpha_n] - 1, \\ x_{n+1}^\delta &= u_{n,[1/\alpha_n]}. \end{aligned}$$

When $\alpha_n = 1$ for all n , this method reduces to the nonlinear Landweber iteration in [5]. It is known that the nonlinear Landweber iteration exhibits slow convergence and requires us to compute a huge number of outer iterations which are rather expensive. By choosing $\{\alpha_n\}$ to satisfy (2.13) in the above method, the computational work could be significantly reduced by computing more cheap steps in each inner iteration.

(d) For $0 < \alpha \leq 1$ we consider the function

$$g_\alpha(\lambda) = \sum_{l=1}^{[1/\alpha]} (1+\lambda)^{-l} = \frac{1 - (1+\lambda)^{-[1/\alpha]}}{\lambda}$$

arising from the Lardy method for linear inverse problems. Then the method (1.4) becomes

$$\begin{aligned} u_{n,0} &= x_n^\delta, \\ u_{n,l+1} &= u_{n,l} - (I + F'(x_n^\delta)^* F'(x_n^\delta))^{-1} F'(x_n^\delta)^* (F(x_n^\delta) - y^\delta - F'(x_n^\delta)(x_n^\delta - u_{n,l})), \\ &\quad l = 0, \dots, [1/\alpha_n] - 1, \\ x_{n+1}^\delta &= u_{n,[1/\alpha_n]}. \end{aligned}$$

When $\alpha_n = 1$ for all n , this is the so-called first-stage Runge–Kutta-type regularization in [11] which again is a slow convergent method. The method can be accelerated with the choice of $\{\alpha_n\}$ to satisfy (2.13).

3. Some crucial inequalities. The following consequence of the above assumptions on F and $\{g_\alpha\}$ plays a crucial role in the convergence analysis.

LEMMA 3.1. *Let $\{g_\alpha\}$ satisfy Assumptions 2.1 and 2.2, let F satisfy Assumption 2.3, and let $\{\alpha_n\}$ be a sequence of positive numbers. Let $T = F'(x^\dagger)$, and for any $x \in B_\rho(x^\dagger)$ let $T_x = F'(x)$. Let $0 \leq a \leq 1/2$. Then for $0 \leq b \leq 1/2 + a$ there holds*

$$(T^*T)^a \prod_{k=j+1}^n r_{\alpha_k}(T^*T) [g_{\alpha_j}(T^*T)T^* - g_{\alpha_j}(T_x^*T_x)T_x^*] = (T^*T)^b S_j(x)$$

for some bounded linear operator $S_j(x) : Y \rightarrow X$ satisfying¹

$$\|S_j(x)\| \lesssim \frac{1}{\alpha_j} (s_n - s_{j-1})^{-1/2-a+b} K_0 \|x - x^\dagger\|,$$

where $j = 0, 1, \dots, n$.

Proof. Let $\eta_\alpha(\lambda) = (\alpha + \lambda)^{-1}$ and $\varphi_\alpha(\lambda) = g_\alpha(\lambda) - (\alpha + \lambda)^{-1}$. We can write

$$(T^*T)^a \prod_{k=j+1}^n r_{\alpha_k}(T^*T) [g_{\alpha_j}(T^*T)T^* - g_{\alpha_j}(T_x^*T_x)T_x^*] = J_1 + J_2 + J_3,$$

where

$$\begin{aligned} J_1 &:= (T^*T)^a \prod_{k=j+1}^n r_{\alpha_k}(T^*T) g_{\alpha_j}(T^*T) [T^* - T_x^*], \\ J_2 &:= (T^*T)^a \prod_{k=j+1}^n r_{\alpha_k}(T^*T) [\eta_{\alpha_j}(T^*T) - \eta_{\alpha_j}(T_x^*T_x)] T_x^*, \\ J_3 &:= (T^*T)^a \prod_{k=j+1}^n r_{\alpha_k}(T^*T) [\varphi_{\alpha_j}(T^*T) - \varphi_{\alpha_j}(T_x^*T_x)] T_x^*. \end{aligned}$$

It suffices to show that for each J_l there holds $J_l = (T^*T)^\nu S_l$ for some bounded linear operator $S_l : Y \rightarrow X$ satisfying the desired estimate. We will use the polar decomposition for linear operators, which implies that $T^* = (T^*T)^{1/2}U$ for some partial isometry $U : Y \rightarrow X$.

By using Assumption 2.3 we have $T^* - T_x^* = T^*(I - R_x)^*$, where $R_x := R(x, x^\dagger)$. This, together with the polar decomposition on T^* , gives

$$(3.1) \quad T^* - T_x^* = (T^*T)^{1/2}U(I - R_x)^*.$$

Consequently we can write $J_1 = (T^*T)^b S_1$ with

$$S_1 = (T^*T)^{1/2+a-b} g_{\alpha_j}(T^*T) \prod_{k=j+1}^n r_{\alpha_k}(T^*T) U(I - R_x)^*.$$

Since $0 \leq 1/2 + a - b \leq 1$, it follows from Assumption 2.2 that

$$\begin{aligned} \|S_1\| &\leq \sup_{0 \leq \lambda \leq 1} \left(\lambda^{1/2+a-b} g_{\alpha_j}(\lambda) \prod_{k=j+1}^n r_{\alpha_k}(\lambda) \right) \|I - R_x\| \\ &\lesssim \frac{1}{\alpha_j} (s_n - s_{j-1})^{-1/2-a+b} K_0 \|x - x^\dagger\|. \end{aligned}$$

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

This shows the desired conclusion on J_1 .

Next we consider J_2 . Note that

$$\begin{aligned} \eta_{\alpha_j}(T^*T) - \eta_{\alpha_j}(T_x^*T_x) &= (\alpha_j I + T^*T)^{-1} T^*(T_x - T)(\alpha_j I + T_x^*T_x)^{-1} \\ &\quad + (\alpha_j I + T^*T)^{-1} (T_x^* - T^*)T_x(\alpha_j I + T_x^*T_x)^{-1}. \end{aligned}$$

Plugging this formula into the expression of J_2 , and using the polar decomposition on T^* and the identity (3.1), we have $J_2 = (T^*T)^b S_2$, where

$$\begin{aligned} S_2 &= \prod_{k=j+1}^n r_{\alpha_k} (T^*T)(\alpha_j I + T^*T)^{-1} (T^*T)^{1/2+a-b} U(T_x - T)(\alpha_j I + T_x^*T_x)^{-1} T_x^* \\ &\quad + \prod_{k=j+1}^n r_{\alpha_k} (T^*T)(\alpha_j I + T^*T)^{-1} (T^*T)^{1/2+a-b} U(R_x - I)^* T_x T_x^* (\alpha_j I + T_x^*T_x)^{-1}. \end{aligned}$$

With the help of Assumption 2.3 we have

$$\|(T_x - T)(\alpha_j I + T_x^*T_x)^{-1} T_x^*\| \leq K_0 \|x - x^\dagger\|.$$

Therefore, it follows from (2.9) that

$$\begin{aligned} \|S_2\| &\leq \sup_{0 \leq \lambda \leq 1} \left(\lambda^{1/2+a-b} (\alpha_j + \lambda)^{-1} \prod_{k=j+1}^n r_{\alpha_k}(\lambda) \right) \|(T_x - T)(\alpha_j I + T_x^*T_x)^{-1} T_x^*\| \\ &\quad + \sup_{0 \leq \lambda \leq 1} \left(\lambda^{1/2+a-b} (\alpha_j + \lambda)^{-1} \prod_{k=j+1}^n r_{\alpha_k}(\lambda) \right) \|(R_x - I)^* T_x T_x^* (\alpha_j I + T_x^*T_x)^{-1}\| \\ &\lesssim \alpha_j^{a-b-1/2} (1 + \alpha_j(s_n - s_j))^{-1/2-a+b} K_0 \|x - x^\dagger\| \\ &= \frac{1}{\alpha_j} (s_n - s_{j-1})^{-1/2-a+b} K_0 \|x - x^\dagger\|. \end{aligned}$$

It remains to consider J_3 . Since Assumption 2.1 implies that $\varphi_{\alpha_j}(z)$ is analytic in D_{α_j} , we have from the Riesz–Dunford formula (2.8) that

$$(3.2) \quad J_3 = \frac{1}{2\pi i} \int_{\Gamma_{\alpha_j}} \varphi_{\alpha_j}(z) L_j(z) dz,$$

where

$$L_j(z) := (T^*T)^a \prod_{k=j+1}^n r_{\alpha_k} (T^*T) [(zI - T^*T)^{-1} - (zI - T_x^*T_x)^{-1}] T_x^*.$$

Using the decomposition

$$\begin{aligned} (zI - T^*T)^{-1} - (zI - T_x^*T_x)^{-1} &= (zI - T^*T)^{-1} T^*(T - T_x)(zI - T_x^*T_x)^{-1} \\ &\quad + (zI - T^*T)^{-1} (T^* - T_x^*)T_x(zI - T_x^*T_x)^{-1}, \end{aligned}$$

together with the polar decomposition on T^* and the identity (3.1), we obtain $L_j(z) = (T^*T)^b \tilde{L}_j(z)$, where

$$\begin{aligned} \tilde{L}_j(z) &= \prod_{k=j+1}^n r_{\alpha_k} (T^*T)(zI - T^*T)^{-1} (T^*T)^{1/2+a-b} U(T - T_x)(zI - T_x^*T_x)^{-1} T_x^* \\ &+ \prod_{k=j+1}^n r_{\alpha_k} (T^*T)(zI - T^*T)^{-1} (T^*T)^{1/2+a-b} U(I - R_x)^* T_x T_x^* (zI - T_x T_x^*)^{-1}. \end{aligned}$$

Combining with (3.2) gives $J_3 = (T^*T)^b S_3$, where

$$S_3 = \frac{1}{2\pi i} \int_{\Gamma_{\alpha_j}} \varphi_{\alpha_j}(z) \tilde{L}_j(z) dz.$$

We need to estimate $\|S_3\|$. We first estimate $\tilde{L}_j(z)$ for $z \in \Gamma_{\alpha_j}$. With the help of Assumption 2.3 and (2.7), we have

$$\|(T - T_x)(zI - T_x^*T_x)^{-1} T_x^*\| \lesssim K_0 \|x - x^\dagger\|.$$

Since $|z| \geq \alpha_j/2$ and $|z - \lambda|^{-1} \leq b_0(|z| + \lambda)^{-1}$ for $z \in \Gamma_{\alpha_j}$, we have from (2.9) that

$$\begin{aligned} \|\tilde{L}_j(z)\| &\lesssim \sup_{0 \leq \lambda \leq 1} \left(\lambda^{1/2+a-b} |z - \lambda|^{-1} \prod_{k=j+1}^n r_{\alpha_k}(\lambda) \right) K_0 \|x - x^\dagger\| \\ &\lesssim \sup_{0 \leq \lambda \leq 1} \left(\lambda^{1/2+a-b} (|z| + \lambda)^{-1} \prod_{k=j+1}^n r_{\alpha_k}(\lambda) \right) K_0 \|x - x^\dagger\| \\ &\lesssim |z|^{a-b-1/2} (1 + (s_n - s_j)|z|)^{-1/2-a+b} K_0 \|x - x^\dagger\| \\ &\lesssim \alpha_j^{a-b-1/2} (1 + (s_n - s_j)\alpha_j)^{-1/2-a+b} K_0 \|x - x^\dagger\| \\ &= \frac{1}{\alpha_j} (s_n - s_{j-1})^{-1/2-a+b} K_0 \|x - x^\dagger\|. \end{aligned}$$

Therefore, it follows from Assumption 2.1 that

$$\begin{aligned} \|S_3\| &\lesssim \frac{1}{\alpha_j} (s_n - s_{j-1})^{-1/2-a+b} K_0 \|x - x^\dagger\| \int_{\Gamma_{\alpha_j}} |\varphi_{\alpha_j}(z)| |dz| \\ &\lesssim \frac{1}{\alpha_j} (s_n - s_{j-1})^{-1/2-a+b} K_0 \|x - x^\dagger\|. \end{aligned}$$

The proof is therefore complete. \square

In the proof of Theorem 2.2 we will also need the following inequality, which can be obtained by essentially the same argument used in the proof of Lemma 3.1.

LEMMA 3.2. *Let $\{g_\alpha\}$ satisfy Assumptions 2.1 and 2.2, let F satisfy Assumption 2.3, and let $\{\alpha_n\}$ be a sequence of positive numbers. Let $T = F'(x^\dagger)$, and for any $x \in B_\rho(x^\dagger)$ let $T_x = F'(x)$. Then for $0 \leq \mu \leq 1/2$ there holds*

$$\begin{aligned} &\left\| (T^*T)^\mu \prod_{k=j+1}^n r_{\alpha_k} (T^*T) [g_{\alpha_j}(T_x^*T_x)T_x^* - g_{\alpha_j}(T_{\bar{x}}^*T_{\bar{x}})T_{\bar{x}}^*] \right\| \\ &\lesssim \frac{1}{\alpha_j} (s_n - s_{j-1})^{-\mu-1/2} (1 + K_0 \|x - x^\dagger\|) K_0 \|x - \bar{x}\| \end{aligned}$$

for all $x, \bar{x} \in B_\rho(x^\dagger)$, where $j = 0, 1, \dots, n$.

4. Rates of convergence: Proof of Theorem 2.1. We begin with the following inequality which will be used frequently in the convergence analysis.

LEMMA 4.1. *Let $\{\alpha_n\}$ be a sequence of positive numbers satisfying $\alpha_n \leq c_1$, and let s_n be defined by (2.1). Let $p \geq 0$ and $q \geq 0$ be two numbers. Then we have*

$$\sum_{j=0}^n \frac{1}{\alpha_j} (s_n - s_{j-1})^{-p} s_j^{-q} \leq C_0 s_n^{1-p-q} \begin{cases} 1, & \max\{p, q\} < 1, \\ \log(1 + s_n), & \max\{p, q\} = 1, \\ s_n^{\max\{p, q\}-1}, & \max\{p, q\} > 1, \end{cases}$$

where C_0 is a constant depending only on c_1 , p , and q .

The proof of Lemma 4.1 is essentially contained in [4, Lemma 4.3 and its proof]; a simplified argument can be found in [9]. Taking $p = 1/2$ and $q = 0$ in the above inequality gives

$$(4.1) \quad \sum_{j=0}^n \frac{1}{\alpha_j} (s_n - s_{j-1})^{-1/2} \leq c_2 s_n^{1/2}$$

for some constant c_2 depending only on c_1 .

In this section we will give the proof of Theorem 2.1. For simplicity of presentation, we will use the notation

$$e_n^\delta := x_n^\delta - x^\dagger, \quad T := F'(x^\dagger), \quad \text{and} \quad T_n := F'(x_n^\delta).$$

It follows from (1.4) that

$$e_{n+1}^\delta = e_n^\delta - g_{\alpha_n} (T_n^* T_n) T_n^* (F(x_n^\delta) - y^\delta).$$

Let

$$u_n := F(x_n^\delta) - y - T(x_n^\delta - x^\dagger).$$

Then we can write

$$\begin{aligned} e_{n+1}^\delta &= e_n^\delta - g_{\alpha_n} (T^* T) T^* (F(x_n^\delta) - y^\delta) \\ &\quad - [g_{\alpha_n} (T_n^* T_n) T_n^* - g_{\alpha_n} (T^* T) T^*] (F(x_n^\delta) - y^\delta) \\ &= r_{\alpha_n} (T^* T) e_n^\delta - g_{\alpha_n} (T^* T) T^* (y - y^\delta + u_n) \\ (4.2) \quad &\quad - [g_{\alpha_n} (T_n^* T_n) T_n^* - g_{\alpha_n} (T^* T) T^*] (F(x_n^\delta) - y^\delta). \end{aligned}$$

By induction on (4.2) we can obtain

$$\begin{aligned} e_{n+1}^\delta &= \prod_{j=0}^n r_{\alpha_j} (T^* T) e_0 - \sum_{j=0}^n \prod_{k=j+1}^n r_{\alpha_k} (T^* T) g_{\alpha_j} (T^* T) T^* (y - y^\delta + u_j) \\ (4.3) \quad &\quad - \sum_{j=0}^n \prod_{k=j+1}^n r_{\alpha_k} (T^* T) [g_{\alpha_j} (T_j^* T_j) T_j^* - g_{\alpha_j} (T^* T) T^*] (F(x_j^\delta) - y^\delta). \end{aligned}$$

By multiplying (4.3) by T and noting that

$$(4.4) \quad I - \sum_{j=0}^n \prod_{k=j+1}^n r_{\alpha_k} (T T^*) g_{\alpha_j} (T T^*) T T^* = \prod_{j=0}^n r_{\alpha_j} (T T^*),$$

we can obtain

(4.5)

$$\begin{aligned} &Te_{n+1}^\delta - y^\delta + y \\ &= T \prod_{j=0}^n r_{\alpha_j}(T^*T)e_0 + \prod_{j=0}^n r_{\alpha_j}(TT^*)(y - y^\delta) - \sum_{j=0}^n \prod_{k=j+1}^n r_{\alpha_k}(TT^*)g_{\alpha_j}(TT^*)TT^*u_j \\ &\quad - \sum_{j=0}^n T \prod_{k=j+1}^n r_{\alpha_k}(T^*T) [g_{\alpha_j}(T_j^*T_j)T_j^* - g_{\alpha_j}(T^*T)T^*] (F(x_j^\delta) - y^\delta). \end{aligned}$$

We will employ (4.3) and (4.5) to prove Theorem 2.1 concerning the order optimal convergence rate of $x_{n_\delta}^\delta$ to x^\dagger when $e_0 := x_0 - x^\dagger$ satisfies the source condition (2.12) for some $0 < \nu \leq 1/2$ and $\omega \in \mathcal{N}(F'(x^\dagger))^\perp \subset X$. We will first derive the crucial estimates on $\|e_n^\delta\|$ and $\|Te_n^\delta\|$. To this end, we introduce the integer \tilde{n}_δ satisfying

$$(4.6) \quad s_{\tilde{n}_\delta}^{-\nu-1/2} \leq \frac{(\tau - 1)\delta}{2c_0\|\omega\|} < s_n^{-\nu-1/2}, \quad 0 \leq n < \tilde{n}_\delta,$$

where $c_0 > 1$ is the constant appearing in (2.2). Such an \tilde{n}_δ is well defined since $s_n \rightarrow \infty$ as $n \rightarrow \infty$.

PROPOSITION 4.2. *Let F satisfy Assumption 2.3, let $\{g_\alpha\}$ satisfy Assumptions 2.1 and 2.2, and let $\{\alpha_n\}$ be a sequence of positive numbers satisfying (2.2). If $x_0 - x^\dagger$ satisfies (2.12) for some $0 < \nu \leq 1/2$ and $\omega \in \mathcal{N}(F'(x^\dagger))^\perp \subset X$, and if $K_0\|\omega\|$ is suitably small, then $x_n^\delta \in B_\rho(x^\dagger)$ for all $0 \leq n \leq \tilde{n}_\delta$, and there exists a generic constant $C_* > 0$ such that*

$$(4.7) \quad \|e_n^\delta\| \leq C_*s_n^{-\nu}\|\omega\| \quad \text{and} \quad \|Te_n^\delta\| \leq C_*s_n^{-\nu-1/2}\|\omega\|$$

and

$$(4.8) \quad \|Te_n^\delta - y^\delta + y\| \leq (c_0 + C_*K_0\|\omega\|)s_n^{-\nu-1/2}\|\omega\| + \delta$$

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

Proof. We will show (4.7) by induction. By using (2.12) and $\|T\| \leq \sqrt{\alpha_0}$ it is easy to see that (4.7) for $n = 0$ holds if $C_* \geq 1$. Next we assume that (4.7) holds for all $0 \leq n \leq l$ for some $l < \tilde{n}_\delta$ and show that (4.7) holds for $n = l + 1$.

With the help of (2.12) we can derive from (4.3) that

$$\begin{aligned} \|e_{l+1}^\delta\| &\leq \left\| \prod_{j=0}^l r_{\alpha_j}(T^*T)(T^*T)^\nu \omega \right\| + \left\| \sum_{j=0}^l \prod_{k=j+1}^l r_{\alpha_k}(T^*T)g_{\alpha_j}(T^*T)T^*(y - y^\delta + u_j) \right\| \\ &\quad + \left\| \sum_{j=0}^l \prod_{k=j+1}^l r_{\alpha_k}(T^*T) [g_{\alpha_j}(T_j^*T_j)T_j^* - g_{\alpha_j}(T^*T)T^*] (F(x_j^\delta) - y^\delta) \right\|. \end{aligned}$$

Thus we may use Assumption 2.2 and Lemma 3.1 with $a = b = 0$ to conclude

$$(4.9) \quad \begin{aligned} \|e_{l+1}^\delta\| &\leq s_l^{-\nu}\|\omega\| + b_2 \sum_{j=0}^l \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1/2} (\delta + \|u_j\|) \\ &\quad + C \sum_{j=0}^l \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1/2} K_0 \|e_j^\delta\| \|F(x_j^\delta) - y^\delta\|. \end{aligned}$$

Moreover, by using (2.12), Assumption 2.2, and Lemma 3.1 with $a = 1/2$ and $b = 0$, we have from (4.5) that

$$(4.10) \quad \begin{aligned} \|Te_{l+1}^\delta - y^\delta + y\| &\leq s_l^{-\nu-1/2}\|\omega\| + \delta + b_2 \sum_{j=0}^l \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1} \|u_j\| \\ &+ C \sum_{j=0}^l \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1} K_0 \|e_j^\delta\| \|F(x_j^\delta) - y^\delta\|. \end{aligned}$$

With the help of Assumption 2.3 and the induction hypotheses, it follows for all $0 \leq j \leq l$ that

$$(4.11) \quad \|u_j\| \leq K_0 \|e_j^\delta\| \|Te_j^\delta\| \lesssim K_0 \|\omega\|^2 s_j^{-2\nu-1/2}.$$

By using the fact that

$$(4.12) \quad \delta \leq \frac{2c_0}{\tau-1} \|\omega\| s_j^{-\nu-1/2}, \quad 0 \leq j \leq l,$$

and the induction hypotheses, we have

$$(4.13) \quad \|F(x_j^\delta) - y^\delta\| \leq \delta + \|Te_j^\delta\| + \|u_j\| \lesssim \|\omega\| s_j^{-\nu-1/2}.$$

In view of (4.1), (4.11), (4.13), and the induction hypothesis on $\|e_j\|$, we have from (4.9) and (4.10) that

$$\|e_{l+1}^\delta\| \leq \|\omega\| s_l^{-\nu} + b_2 c_2 s_l^{1/2} \delta + CK_0 \|\omega\|^2 \sum_{j=0}^l \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1/2} s_j^{-2\nu-1/2}$$

and

$$\|Te_{l+1}^\delta - y^\delta + y\| \leq \|\omega\| s_l^{-\nu-1/2} + \delta + CK_0 \|\omega\|^2 \sum_{j=0}^l \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1} s_j^{-2\nu-1/2}.$$

With the help of Lemma 4.1, $0 < \nu \leq 1/2$, (4.12), and (2.2) with $c_0 > 1$, we have

$$\begin{aligned} \|e_{l+1}^\delta\| &\leq (1 + 2b_2 c_0 c_2 / (\tau - 1) + CK_0 \|\omega\|) \|\omega\| s_l^{-\nu} \\ &\leq \left(c_0^{1/2} + 2b_2 c_0^{3/2} c_2 / (\tau - 1) + CK_0 \|\omega\| \right) \|\omega\| s_{l+1}^{-\nu} \end{aligned}$$

and

$$(4.14) \quad \begin{aligned} \|Te_{l+1}^\delta - y^\delta + y\| &\leq \delta + (1 + CK_0 \|\omega\|) \|\omega\| s_l^{-\nu-1/2} \\ &\leq \delta + (c_0 + CK_0 \|\omega\|) \|\omega\| s_{l+1}^{-\nu-1/2}. \end{aligned}$$

Consequently $\|e_{l+1}^\delta\| \leq C_* \|\omega\| s_{l+1}^{-\nu}$ if $C_* \geq 2c_0^{1/2} + 2b_2 c_0^{3/2} c_2 / (\tau - 1)$ and $K_0 \|\omega\|$ is suitably small. Moreover, from (4.14), (4.12), and (2.2) we also have

$$\begin{aligned} \|Te_{l+1}^\delta\| &\leq 2\delta + (c_0 + CK_0 \|\omega\|) \|\omega\| s_{l+1}^{-\nu-1/2} \\ &\leq (4c_0^2 / (\tau - 1) + c_0 + CK_0 \|\omega\|) \|\omega\| s_{l+1}^{-\nu-1/2} \\ &\leq C_* \|\omega\| s_{l+1}^{-\nu-1/2} \end{aligned}$$

if $C_* \geq 2c_0 + 4c_0^2/(\tau - 1)$ and $K_0\|\omega\|$ is suitably small. We therefore complete the proof of (4.7). In the meantime, (4.14) gives the proof of (4.8). \square

From the proof of Proposition 4.2, we have in fact obtained, under the conditions in Proposition 4.2, that

$$(4.15) \quad \|F(x_n^\delta) - y - Te_n^\delta\| \lesssim K_0\|\omega\|^2 s_n^{-2\nu-1/2}$$

and

$$(4.16) \quad \|F(x_n^\delta) - y^\delta\| \lesssim \|\omega\| s_n^{-\nu-1/2}$$

for $0 \leq n \leq \tilde{n}_\delta$; see (4.11) and (4.13).

LEMMA 4.3. *Let all the conditions in Proposition 4.2 hold. Let $\tau > 1$ be a given number. If $x_0 - x^\dagger$ satisfies (2.12) for some $0 < \nu \leq 1/2$ and $\omega \in \mathcal{N}(F'(x^\dagger))^\perp \subset X$, and if $K_0\|\omega\|$ is suitably small, then the discrepancy principle (1.5) defines a finite integer n_δ satisfying $n_\delta \leq \tilde{n}_\delta$.*

Proof. From Proposition 4.2, (4.15), and $\nu > 0$, it follows for $0 \leq n \leq \tilde{n}_\delta$ that

$$\begin{aligned} \|F(x_n^\delta) - y^\delta\| &\leq \|F(x_n^\delta) - y - Te_n^\delta\| + \|Te_n^\delta - y^\delta + y\| \\ &\leq CK_0\|\omega\|^2 s_n^{-2\nu-1/2} + (c_0 + CK_0\|\omega\|) s_n^{-\nu-1/2}\|\omega\| + \delta \\ &\leq (c_0 + CK_0\|\omega\|) s_n^{-\nu-1/2}\|\omega\| + \delta. \end{aligned}$$

By setting $n = \tilde{n}_\delta$ in the above inequality and using the definition of \tilde{n}_δ we obtain

$$\|F(x_{\tilde{n}_\delta}^\delta) - y^\delta\| \leq \left(1 + \frac{\tau - 1}{2} + CK_0\|\omega\|\right) \delta \leq \tau\delta$$

if $K_0\|\omega\|$ is suitably small. According to the definition of n_δ we have $n_\delta \leq \tilde{n}_\delta$. \square

4.1. Completion of the proof of Theorem 2.1. From Proposition 4.2 and Lemma 4.3 it follows that the methods given by (1.4) and (1.5) with $\tau > 1$ are well defined.

In the following we will derive the order optimal convergence rate. From (4.3), the source condition (2.12) with $0 < \nu \leq 1/2$, the polar decomposition on T^* , and Lemma 3.1 with $a = 0$ and $b = \nu$, it follows that

$$(4.17) \quad e_{n+1}^\delta = (T^*T)^\nu w_n,$$

where

$$\begin{aligned} w_n &:= \prod_{j=0}^n r_{\alpha_j} (T^*T)\omega - \sum_{j=0}^n S_j(x_j^\delta) (F(x_j^\delta) - y^\delta) \\ &\quad - \sum_{j=0}^n \prod_{k=j+1}^n r_{\alpha_k} (T^*T) g_{\alpha_j} (T^*T) (T^*T)^{1/2-\nu} U(y - y^\delta + u_j). \end{aligned}$$

With the help of Assumption 2.2 and Lemma 3.1 we have

$$\begin{aligned} \|w_n\| &\lesssim \|\omega\| + \sum_{j=0}^n \frac{1}{\alpha_j} (s_n - s_{j-1})^{-1/2+\nu} K_0 \|e_j^\delta\| \|F(x_j^\delta) - y^\delta\| \\ &\quad + \sum_{j=0}^n \frac{1}{\alpha_j} (s_n - s_{j-1})^{-1/2+\nu} (\delta + \|u_j\|). \end{aligned}$$

In view of (4.15), (4.16), Proposition 4.2, Lemma 4.1, and (4.6), for $0 \leq n < \tilde{n}_\delta$ we have

$$\begin{aligned} \|w_n\| &\lesssim \|\omega\| + \delta \sum_{j=0}^n \frac{1}{\alpha_j} (s_n - s_{j-1})^{-1/2+\nu} \\ &\quad + K_0 \|\omega\|^2 \sum_{j=0}^n \frac{1}{\alpha_j} (s_n - s_{j-1})^{-1/2+\nu} s_j^{-2\nu-1/2} \\ &\lesssim \|\omega\| + s_n^{1/2+\nu} \delta \lesssim \|\omega\|. \end{aligned}$$

Since Lemma 4.3 implies that $n_\delta \leq \tilde{n}_\delta$, we have $\|w_{n_\delta-1}\| \lesssim \|\omega\|$. On the other hand, it follows from (4.17), Assumption 2.3, and the definition of n_δ that

$$\|T(T^*T)^\nu w_{n_\delta-1}\| = \|Te_{n_\delta}\| \lesssim \|F(x_{n_\delta}^\delta) - y\| \lesssim \delta.$$

Therefore, by using (4.17) and the above two estimates, we have from the interpolation inequality that

$$\begin{aligned} \|e_{n_\delta}^\delta\| &\leq \|w_{n_\delta-1}\|^{1/(1+2\nu)} \|T(T^*T)^\nu w_{n_\delta-1}\|^{2\nu/(1+2\nu)} \\ &\leq C_\nu \|\omega\|^{1/(1+2\nu)} \delta^{2\nu/(1+2\nu)}. \end{aligned}$$

This gives the desired estimate.

5. Convergence: Proof of Theorem 2.2. In this section we will show Theorem 2.2 concerning the convergence of $x_{n_\delta}^\delta$ to x^\dagger as $\delta \rightarrow 0$ without assuming any source conditions on $e_0 := x_0 - x^\dagger$. The sequence $\{\alpha_n\}$ is now assumed to satisfy (2.13). It is easy to see that

$$(5.1) \quad \sum_{j=0}^n \frac{1}{\alpha_j} (s_n - s_{j-1})^{-1} s_j^{-\mu} \leq C_1 s_n^{-\mu}$$

for each $0 \leq \mu < 1$, where C_1 is a constant depending only on d_0 , d_1 , r , and μ . We remark that (5.1) may not be true for a general sequence $\{\alpha_n\}$ satisfying (2.2).

We first show that the method given by (1.4) and (1.5) is well defined. To this end, we introduce the integer \hat{n}_δ satisfying

$$(5.2) \quad s_{\hat{n}_\delta}^{-1/2} \leq \frac{(\tau-1)\delta}{2c_0\|e_0\|} < s_n^{-1/2}, \quad 0 \leq n < \hat{n}_\delta.$$

Since $s_n \rightarrow \infty$ as $n \rightarrow \infty$, such an \hat{n}_δ is well defined. Since $\{\alpha_n\}$ satisfies (2.13), it is easy to see that $s_n \sim r^{-n}$, and consequently $\hat{n}_\delta = O(1 + |\log \delta|)$.

LEMMA 5.1. *Let F satisfy Assumption 2.3, let $\{g_\alpha\}$ satisfy Assumptions 2.1 and 2.2, and let $\{\alpha_n\}$ be a sequence of positive numbers satisfying (2.13). Let $\tau > 1$ be a given number. If $K_0\|e_0\|$ is suitably small, then there is a constant C_* such that*

$$(5.3) \quad \|e_n^\delta\| \leq C_* \|e_0\| \quad \text{and} \quad \|Te_n^\delta\| \leq C_* \|e_0\| s_n^{-1/2}$$

for $0 \leq n \leq \hat{n}_\delta$, and the discrepancy principle (1.5) defines a finite integer n_δ satisfying $n_\delta \leq \hat{n}_\delta = O(1 + |\log \delta|)$.

Proof. We prove (5.3) by induction. By using $\|T\| \leq \sqrt{\alpha_0}$, it is easy to see that (5.3) is true for $n = 0$ if $C_* \geq 1$. Next we assume that (5.3) holds for all $0 \leq n \leq l$

for some $l < \hat{n}_\delta$ and show that it is also true for $n = l + 1$. By a similar argument in the proof of Proposition 4.2 we obtain

$$(5.4) \quad \|e_{l+1}^\delta\| \leq \|e_0\| + b_2 c_2 s_l^{1/2} \delta + CK_0 \|e_0\|^2 \sum_{j=0}^l \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1/2} s_j^{-1/2}$$

and

$$(5.5) \quad \|Te_{l+1}^\delta - y^\delta + y\| \leq \|e_0\| s_l^{-1/2} + \delta + CK_0 \|e_0\|^2 \sum_{j=0}^l \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1} s_j^{-1/2}.$$

By using (5.2) and Lemma 4.1 we obtain from (5.4) that

$$\|e_{l+1}^\delta\| \leq \left(1 + \frac{2b_2 c_0 c_2}{\tau - 1} + CK_0 \|e_0\|\right) \|e_0\| \leq C_* \|e_0\|$$

if $C_* \geq 2 + 2b_2 c_0 c_2 / (\tau - 1)$ and $K_0 \|e_0\|$ is suitably small. On the other hand, by using (5.1) with $\mu = 1/2$ and (2.2) we obtain from (5.5) that

$$(5.6) \quad \begin{aligned} \|Te_{l+1}^\delta - y^\delta + y\| &\leq \delta + (1 + CK_0 \|e_0\|) \|e_0\| s_l^{-1/2} \\ &\leq \delta + (c_0 + CK_0 \|e_0\|) \|e_0\| s_{l+1}^{-1/2}. \end{aligned}$$

Consequently, we have from (5.2) that

$$\|Te_{l+1}^\delta\| \leq \left(\frac{4c_0^2}{\tau - 1} + c_0 + CK_0 \|e_0\|\right) \|e_0\| s_{l+1}^{-1/2} \leq C_* \|e_0\| s_{l+1}^{-1/2}$$

if $C_* \geq 2c_0 + 4c_0^2 / (\tau - 1)$ and $K_0 \|e_0\|$ is suitably small. We thus complete the proof of (5.3).

Note that the above argument in fact shows also that

$$\|Te_n^\delta - y^\delta + y\| \leq \delta + (c_0 + CK_0 \|e_0\|) \|e_0\| s_n^{-1/2}, \quad 0 \leq n \leq \hat{n}_\delta.$$

Thus, by an argument similar to that in the proof of Lemma 4.3 we can derive $\|F(x_{\hat{n}_\delta}^\delta) - y^\delta\| \leq \tau\delta$ if $K_0 \|e_0\|$ is suitably small. According to the definition of n_δ we obtain $n_\delta \leq \hat{n}_\delta$. \square

In the remaining part of this section we will show $x_{n_\delta}^\delta \rightarrow x^\dagger$ as $\delta \rightarrow 0$. We will achieve this by first considering the noise-free iterative sequence $\{x_n\}$ defined by (1.4) with y^δ replaced by y , i.e.,

$$(5.7) \quad x_{n+1} = x_n - g_{\alpha_n} (F'(x_n)^* F'(x_n)) F'(x_n)^* (F(x_n) - y)$$

and showing that $x_n \rightarrow x^\dagger$ as $n \rightarrow \infty$. We then derive the stability estimate on $\|x_n^\delta - x_n\|$ for $0 \leq n \leq n_\delta$ together with other related estimates. With the help of the definition of n_δ , we will be able to show the convergence of $x_{n_\delta}^\delta$ to x^\dagger as $\delta \rightarrow 0$.

5.1. Convergence of the noise-free iteration. In this subsection we will show the convergence of x_n to x^\dagger as $n \rightarrow \infty$. We first show that if $x_0 - x^\dagger \in \mathcal{R}(F'(x^\dagger)^*)$, then $x_n \rightarrow x^\dagger$ as $n \rightarrow \infty$. We then perturb the initial guess x_0 to be \hat{x}_0 such that $\hat{x}_0 - x^\dagger \in \mathcal{R}(F'(x^\dagger)^*)$ and define $\{\hat{x}_n\}$ by

$$(5.8) \quad \hat{x}_{n+1} = \hat{x}_n - g_{\alpha_n} (F'(\hat{x}_n)^* F'(\hat{x}_n)) F'(\hat{x}_n)^* (F(\hat{x}_n) - y).$$

Since $x_0 - x^\dagger \in \mathcal{N}(F'(x^\dagger))^\perp = \overline{\mathcal{R}(F'(x^\dagger)^*)}$, such an \hat{x}_0 can be chosen as close to x_0 as we want. We then show that $\{x_n\}$ is stable relative to the change of x_0 . This allows us to derive the convergence of $\{x_n\}$.

We start with several lemmas. We first show that x_n is well defined for all n and satisfies certain estimates.

LEMMA 5.2. *Let F satisfy Assumption 2.3, let $\{g_\alpha\}$ satisfy Assumptions 2.1 and 2.2, and let $\{\alpha_n\}$ be a sequence of positive numbers satisfying (2.13). If $K_0\|e_0\|$ is suitably small, then*

$$(5.9) \quad \|e_n\| \leq 2\|e_0\| \quad \text{and} \quad \|Te_n\| \leq 2c_0\|e_0\|s_n^{-1/2}$$

for $n = 0, 1, \dots$, where $e_n := x_n - x^\dagger$.

Proof. This result can be obtained by the same argument as in the proof of Lemma 5.1. \square

LEMMA 5.3. *Let F satisfy Assumption 2.3, let $\{g_\alpha\}$ satisfy Assumptions 2.1 and 2.2, and let $\{\alpha_n\}$ be a sequence of positive numbers satisfying (2.13). If $e_0 = (T^*T)^{1/4}\omega$ for some $\omega \in X$ and $K_0\|e_0\|$ is suitably small, then*

$$(5.10) \quad \|e_n\| \leq 2c_0\|\omega\|s_n^{-1/4} \quad \text{and} \quad \|Te_n\| \leq 2c_0\|\omega\|s_n^{-3/4}$$

for $n = 0, 1, \dots$.

Proof. We prove (5.10) by induction. By using $\|T\| \leq \sqrt{\alpha_0}$ and $e_0 = (T^*T)^{1/4}\omega$, it is easy to see that (5.10) is true for $n = 0$. Next we assume that (5.10) holds for all $0 \leq n \leq l$ and show that it also holds for $n = l + 1$. By an argument similar to that in the proof of Proposition 4.2 we obtain

$$(5.11) \quad \begin{aligned} \|e_{l+1}\| &\leq s_l^{-1/4}\|\omega\| + b_2 \sum_{j=0}^l \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1/2} \|F(x_j) - y - Te_j\| \\ &\quad + C \sum_{j=0}^l \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1/2} K_0 \|e_j\| \|F(x_j) - y\| \end{aligned}$$

and

$$(5.12) \quad \begin{aligned} \|Te_{l+1}\| &\leq s_l^{-3/4}\|\omega\| + b_2 \sum_{j=0}^l \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1} \|F(x_j) - y - Te_j\| \\ &\quad + C \sum_{j=0}^l \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1} K_0 \|e_j\| \|F(x_j) - y\|. \end{aligned}$$

With the help of Assumption 2.3, Lemma 5.2, and the induction hypotheses, we have for $0 \leq j \leq l$ that

$$\begin{aligned} \|F(x_j) - y - Te_j\| &\leq K_0 \|e_j\| \|Te_j\| \leq CK_0 \|e_0\| \|\omega\| s_j^{-3/4}, \\ \|F(x_j) - y\| &\leq \|Te_j\| + \|F(x_j) - y - Te_j\| \lesssim \|\omega\| s_j^{-3/4}. \end{aligned}$$

Therefore, by using Lemma 4.1, we obtain from (5.11) that

$$\begin{aligned} \|e_{l+1}\| &\leq s_l^{-1/4}\|\omega\| + CK_0\|e_0\|\|\omega\| \sum_{j=0}^l \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1/2} s_j^{-3/4} \\ &\leq (1 + CK_0\|e_0\|) \|\omega\| s_l^{-1/4}, \end{aligned}$$

while by using (5.1) with $\mu = 3/4$ we obtain

$$\begin{aligned} \|Te_{l+1}\| &\leq s_l^{-3/4}\|\omega\| + CK_0\|e_0\|\|\omega\| \sum_{j=0}^l \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1} s_j^{-3/4} \\ &\leq (1 + CK_0\|e_0\|) \|\omega\| s_l^{-3/4}. \end{aligned}$$

Thus, by using $s_{l+1} \leq c_0 s_l$, we obtain for suitably small $K_0\|e_0\|$ that $\|e_{l+1}\| \leq 2c_0\|\omega\|s_{l+1}^{-1/4}$ and $\|Te_{l+1}\| \leq 2c_0\|\omega\|s_{l+1}^{-3/4}$. The proof is therefore complete. \square

We remark that the crucial point in Lemma 5.1 is that it requires only the smallness of $K_0\|e_0\|$, which is different from Proposition 4.2, where the smallness of $K_0\|\omega\|$ is needed. This will allow us to carry out the following approximation argument with perturbation on the initial guess.

We now derive a perturbation result on $\|x_n - \hat{x}_n\|$ and $\|T(x_n - \hat{x}_n)\|$ relative to the change of the initial guess. For simplicity of the presentation we set

$$\hat{e}_n := \hat{x}_n - x^\dagger, \quad T_n = F'(x_n), \quad \hat{T}_n = F'(\hat{x}_n).$$

It follows from (5.7) and (5.8) that

$$\begin{aligned} x_{n+1} - \hat{x}_{n+1} &= x_n - \hat{x}_n - g_{\alpha_n}(T_n^* T_n) T_n^* (F(x_n) - y) + g_{\alpha_n}(\hat{T}_n^* \hat{T}_n) \hat{T}_n^* (F(\hat{x}_n) - y) \\ &= r_{\alpha_n}(T^* T)(x_n - \hat{x}_n) - g_{\alpha_n}(T^* T) T^* (F(x_n) - F(\hat{x}_n) - T(x_n - \hat{x}_n)) \\ &\quad - [g_{\alpha_n}(T_n^* T_n) T_n^* - g_{\alpha_n}(T^* T) T^*] (F(x_n) - F(\hat{x}_n)) \\ &\quad - [g_{\alpha_n}(T_n^* T_n) T_n^* - g_{\alpha_n}(\hat{T}_n^* \hat{T}_n) \hat{T}_n^*] (F(\hat{x}_n) - y). \end{aligned}$$

By induction on this identity we obtain

$$\begin{aligned} (5.13) \quad x_{n+1} - \hat{x}_{n+1} &= \prod_{k=0}^n r_{\alpha_k}(T^* T)(x_0 - \hat{x}_0) \\ &\quad - \sum_{j=0}^n \prod_{k=j+1}^n r_{\alpha_k}(T^* T) g_{\alpha_j}(T^* T) T^* (F(x_j) - F(\hat{x}_j) - T(x_j - \hat{x}_j)) \\ &\quad - \sum_{j=0}^n \prod_{k=j+1}^n r_{\alpha_k}(T^* T) [g_{\alpha_j}(T_j^* T_j) T_j^* - g_{\alpha_j}(T^* T) T^*] (F(x_j) - F(\hat{x}_j)) \\ &\quad - \sum_{j=0}^n \prod_{k=j+1}^n r_{\alpha_k}(T^* T) [g_{\alpha_j}(T_j^* T_j) T_j^* - g_{\alpha_j}(\hat{T}_j^* \hat{T}_j) \hat{T}_j^*] (F(\hat{x}_j) - y). \end{aligned}$$

LEMMA 5.4. *Let F satisfy Assumption 2.3, let $\{g_\alpha\}$ satisfy Assumptions 2.1 and 2.2, and let $\{\alpha_n\}$ be a sequence of positive numbers satisfying (2.13). If $K_0\|e_0\|$ and $K_0\|\hat{e}_0\|$ are suitably small, then*

$$(5.14) \quad \|x_n - \hat{x}_n\| \leq 2\|x_0 - \hat{x}_0\| \quad \text{and} \quad \|T(x_n - \hat{x}_n)\| \leq 2c_0 s_n^{-1/2} \|x_0 - \hat{x}_0\|$$

for $n = 0, 1, \dots$.

Proof. We will show (5.14) by induction. Since $\|T\| \leq \sqrt{\alpha_0}$, (5.14) holds for $n = 0$. In the following we will assume that (5.14) holds for $0 \leq n \leq l$ and show that it is also true for $n = l + 1$.

In view of Assumption 2.2, Lemma 3.1 with $a = b = 0$, and Lemma 3.2 with $\mu = 0$, it follows from (5.13) that

$$\begin{aligned}
 \|x_{l+1} - \hat{x}_{l+1}\| &\leq \|x_0 - \hat{x}_0\| \\
 &+ C \sum_{j=0}^l \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1/2} \|F(x_j) - F(\hat{x}_j) - T(x_j - \hat{x}_j)\| \\
 &+ C \sum_{j=0}^l \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1/2} K_0 \|e_j\| \|F(x_j) - F(\hat{x}_j)\| \\
 (5.15) \quad &+ C \sum_{j=0}^l \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1/2} K_0 \|x_j - \hat{x}_j\| \|F(\hat{x}_j) - y\|.
 \end{aligned}$$

Next we multiply (5.13) by T . By using Assumption 2.2, Lemma 3.1 with $a = 1/2$ and $b = 0$, and Lemma 3.2 with $\mu = 1/2$, we obtain

$$\begin{aligned}
 \|T(x_{l+1} - \hat{x}_{l+1})\| &\leq s_l^{-1/2} \|x_0 - \hat{x}_0\| \\
 &+ C \sum_{j=0}^l \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1} \|F(x_j) - F(\hat{x}_j) - T(x_j - \hat{x}_j)\| \\
 &+ C \sum_{j=0}^l \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1} K_0 \|e_j\| \|F(x_j) - F(\hat{x}_j)\| \\
 (5.16) \quad &+ C \sum_{j=0}^l \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1} K_0 \|x_j - \hat{x}_j\| \|F(\hat{x}_j) - y\|.
 \end{aligned}$$

From Lemma 5.2 and Assumption 2.3 it follows that

$$(5.17) \quad \|F(\hat{x}_j) - y\| \lesssim \|T\hat{e}_j\| + K_0 \|\hat{e}_j\| \|T\hat{e}_j\| \lesssim s_j^{-1/2} \|\hat{e}_0\|.$$

Moreover, by using Assumption 2.3 we have

$$\begin{aligned}
 F(x_j) - F(\hat{x}_j) - T(x_j - \hat{x}_j) &= \int_0^1 [F'(\hat{x}_j + t(x_j - \hat{x}_j)) - T](x_j - \hat{x}_j) dt \\
 &= \int_0^1 [R(\hat{x}_j + t(x_j - \hat{x}_j), x^\dagger) - I] T(x_j - \hat{x}_j) dt.
 \end{aligned}$$

Consequently

$$\begin{aligned}
 \|F(x_j) - F(\hat{x}_j) - T(x_j - \hat{x}_j)\| &\leq \int_0^1 \|R(\hat{x}_j + t(x_j - \hat{x}_j), x^\dagger) - I\| \|T(x_j - \hat{x}_j)\| dt \\
 &\leq \frac{1}{2} K_0 (\|e_j\| + \|\hat{e}_j\|) \|T(x_j - \hat{x}_j)\|.
 \end{aligned}$$

With the help of Lemma 5.2 we obtain

$$(5.18) \quad \|F(x_j) - F(\hat{x}_j) - T(x_j - \hat{x}_j)\| \leq K_0 (\|e_0\| + \|\hat{e}_0\|) \|T(x_j - \hat{x}_j)\|.$$

This in particular implies

$$(5.19) \quad \|F(x_j) - F(\hat{x}_j)\| \leq 2\|T(x_j - \hat{x}_j)\|.$$

By virtue of (5.17), (5.18), (5.19), and the induction hypotheses, we have from (5.15) and (5.16) that

$$(5.20) \quad \begin{aligned} \|x_{l+1} - \hat{x}_{l+1}\| &\leq \|x_0 - \hat{x}_0\| \\ &+ CK_0(\|e_0\| + \|\hat{e}_0\|)\|x_0 - \hat{x}_0\| \sum_{j=0}^l \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1/2} s_j^{-1/2} \end{aligned}$$

and

$$(5.21) \quad \begin{aligned} \|T(x_{l+1} - \hat{x}_{l+1})\| &\leq s_l^{-1/2} \|x_0 - \hat{x}_0\| \\ &+ CK_0(\|e_0\| + \|\hat{e}_0\|)\|x_0 - \hat{x}_0\| \sum_{j=0}^l \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1} s_j^{-1/2}. \end{aligned}$$

With the help of Lemma 4.1 and (5.1) we can derive

$$\begin{aligned} \|x_{l+1} - \hat{x}_{l+1}\| &\leq (1 + CK_0(\|e_0\| + \|\hat{e}_0\|)) \|x_0 - \hat{x}_0\| \\ &\leq 2\|x_0 - \hat{x}_0\| \end{aligned}$$

and

$$\begin{aligned} \|T(x_{l+1} - \hat{x}_{l+1})\| &\leq (1 + CK_0(\|e_0\| + \|\hat{e}_0\|)) s_l^{-1/2} \|x_0 - \hat{x}_0\| \\ &\leq 2c_0 s_{l+1}^{-1/2} \|x_0 - \hat{x}_0\| \end{aligned}$$

if $K_0\|e_0\|$ and $K_0\|\hat{e}_0\|$ are suitably small. The proof is thus complete. \square

THEOREM 5.5. *Let F satisfy Assumption 2.3, let $\{g_\alpha\}$ satisfy Assumptions 2.1 and 2.2, and let $\{\alpha_n\}$ be a sequence of positive numbers satisfying (2.13). If $e_0 \in \mathcal{N}(T)^\perp$ and $K_0\|e_0\|$ is suitably small, then*

$$(5.22) \quad \lim_{n \rightarrow \infty} \|x_n - x^\dagger\| = 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} s_n^{1/2} \|T(x_n - x^\dagger)\| = 0$$

for the sequence $\{x_n\}$ defined by (5.7).

Proof. Let $0 < \varepsilon < \|e_0\|$ be an arbitrarily small number. Since $e_0 \in \mathcal{N}(T)^\perp = \overline{\mathcal{R}(T^*)}$, there is an $\hat{x}_0 \in X$ such that $\hat{e}_0 := x_0 - \hat{x}_0 \in \mathcal{R}(T^*)$ and $\|x_0 - \hat{x}_0\| < \varepsilon$. Note that $K_0\|\hat{e}_0\| \leq 2K_0\|e_0\|$. Thus, if $K_0\|e_0\|$ is suitably small, then for the sequence $\{\hat{x}_n\}$ defined by (5.8), it follows from Lemma 5.4 that

$$\|x_n - \hat{x}_n\| \leq 2\|x_0 - \hat{x}_0\| < 2\varepsilon$$

and

$$s_n^{1/2} \|T(x_n - \hat{x}_n)\| \leq 2c_0\|x_0 - \hat{x}_0\| < 2c_0\varepsilon$$

for all $n \geq 0$. On the other hand, since $\hat{e}_0 \in \mathcal{R}(T^*) = \mathcal{R}((T^*T)^{1/2}) \subset \mathcal{R}((T^*T)^{1/4})$, from Lemma 5.3 we have $\|\hat{e}_n\| \rightarrow 0$ and $s_n^{1/2} \|T\hat{e}_n\| \rightarrow 0$ as $n \rightarrow \infty$. Thus, there is an n_0 such that $\|\hat{e}_n\| < \varepsilon$ and $s_n^{1/2} \|T\hat{e}_n\| < c_0\varepsilon$ for all $n \geq n_0$. Consequently

$$\|e_n\| \leq \|x_n - \hat{x}_n\| + \|\hat{e}_n\| < 3\varepsilon$$

and

$$s_n^{1/2} \|Te_n\| \leq s_n^{1/2} \|T(x_n - \hat{x}_n)\| + s_n^{1/2} \|T\hat{e}_n\| < 3c_0\varepsilon$$

for all $n \geq n_0$. Since $\varepsilon > 0$ is arbitrarily small, we therefore obtain (5.22). \square

5.2. Stability estimates. In this subsection we will derive the stability estimates on $\|x_n^\delta - x_n\|$ for $0 \leq n \leq \hat{n}_\delta$, where \hat{n}_δ is defined by (5.2). We will use the notation

$$\mathcal{A} := F'(x^\dagger)^* F'(x^\dagger), \quad \mathcal{A}_n := F'(x_n)^* F'(x_n), \quad \mathcal{A}_n^\delta := F'(x_n^\delta)^* F'(x_n^\delta).$$

The main result is as follows.

PROPOSITION 5.6. *Let F satisfy Assumption 2.3, let $\{g_\alpha\}$ satisfy Assumptions 2.1 and 2.2, and let $\{\alpha_n\}$ be a sequence of positive numbers satisfying (2.13). If $K_0\|e_0\|$ is suitably small, then*

$$(5.23) \quad \|x_n^\delta - x_n\| \lesssim s_n^{1/2} \delta$$

and

$$(5.24) \quad \|F(x_n^\delta) - F(x_n) - y^\delta + y\| \leq (1 + CK_0\|e_0\|) \delta$$

for $0 \leq n \leq \hat{n}_\delta$.

Proof. We first show (5.23) by establishing

$$(5.25) \quad \|x_n^\delta - x_n\| \leq 2b_2c_2s_n^{1/2}\delta \quad \text{and} \quad \|T(x_n^\delta - x_n)\| \leq 3\delta$$

for $0 \leq n \leq \hat{n}_\delta$, where b_2 and c_2 are the constants appearing in (2.6) and (4.1), respectively. It is clear that (5.25) is true for $n = 0$. Now we assume that (5.25) is true for all $0 \leq n \leq l$ for some $l < \hat{n}_\delta$ and show that it is also true for $n = l + 1$. We set

$$\begin{aligned} v_n &:= F(x_n^\delta) - F(x_n) - T(x_n^\delta - x_n), \\ w_n &:= F(x_n^\delta) - F(x_n) - y^\delta + y. \end{aligned}$$

It then follows from the definition of $\{x_n^\delta\}$ and $\{x_n\}$ that

$$\begin{aligned} x_{n+1}^\delta - x_{n+1} &= x_n^\delta - x_n - g_{\alpha_n}(\mathcal{A}_n^\delta)F'(x_n^\delta)^*(F(x_n^\delta) - y^\delta) \\ &\quad + g_{\alpha_n}(\mathcal{A}_n)F'(x_n)^*(F(x_n) - y) \\ &= r_{\alpha_n}(\mathcal{A})(x_n^\delta - x_n) - g_{\alpha_n}(\mathcal{A})F'(x^\dagger)^*(v_n - y^\delta + y) \\ &\quad - [g_{\alpha_n}(\mathcal{A}_n)F'(x_n)^* - g_{\alpha_n}(\mathcal{A})F'(x^\dagger)^*]w_n \\ &\quad - [g_{\alpha_n}(\mathcal{A}_n^\delta)F'(x_n^\delta)^* - g_{\alpha_n}(\mathcal{A}_n)F'(x_n)^*](F(x_n^\delta) - y^\delta). \end{aligned}$$

By induction on the above equation and noting that $x_0^\delta = x_0$ we obtain

$$\begin{aligned} (5.26) \quad x_{l+1}^\delta - x_{l+1} &= \sum_{j=0}^l \prod_{k=j+1}^l r_{\alpha_k}(\mathcal{A})g_{\alpha_j}(\mathcal{A})F'(x^\dagger)^*(y^\delta - y - v_j) \\ &\quad - \sum_{j=0}^l \prod_{k=j+1}^l r_{\alpha_k}(\mathcal{A})[g_{\alpha_j}(\mathcal{A}_j^\delta)F'(x_j^\delta)^* - g_{\alpha_j}(\mathcal{A}_j)F'(x_j)^*](F(x_j^\delta) - y^\delta) \\ &\quad - \sum_{j=0}^l \prod_{k=j+1}^l r_{\alpha_k}(\mathcal{A})[g_{\alpha_j}(\mathcal{A}_j)F'(x_j)^* - g_{\alpha_j}(\mathcal{A})F'(x^\dagger)^*]w_j. \end{aligned}$$

In view of Assumption 2.1, Lemma 3.1, and Lemma 3.2 it follows that

$$\begin{aligned}
 \|x_{l+1}^\delta - x_{l+1}\| &\leq b_2 \sum_{j=0}^l \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1/2} (\delta + \|v_j\|) \\
 &\quad + C \sum_{j=0}^l \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1/2} K_0 \|x_j^\delta - x_j\| \|F(x_j^\delta) - y^\delta\| \\
 (5.27) \quad &\quad + C \sum_{j=0}^l \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1/2} K_0 \|e_j\| \|w_j\|.
 \end{aligned}$$

By multiplying (5.26) by T and using (4.4), we obtain with $\mathcal{B} := TT^*$ that

$$\begin{aligned}
 &T(x_{l+1}^\delta - x_{l+1}) - y^\delta + y \\
 &= \prod_{j=0}^l r_{\alpha_j}(\mathcal{B})(y - y^\delta) - \sum_{j=0}^l \prod_{k=j+1}^l r_{\alpha_k}(\mathcal{B})g_{\alpha_j}(\mathcal{B})\mathcal{B}v_j \\
 &\quad - \sum_{j=0}^l T \prod_{k=j+1}^l r_{\alpha_k}(\mathcal{A}) [g_{\alpha_j}(\mathcal{A}_j)F'(x_j^\delta)^* - g_{\alpha_j}(\mathcal{A}_j)F'(x_j)^*] (F(x_j^\delta) - y^\delta) \\
 &\quad - \sum_{j=0}^l T \prod_{k=j+1}^l r_{\alpha_k}(\mathcal{A}) [g_{\alpha_j}(\mathcal{A}_j)F'(x_j)^* - g_{\alpha_j}(\mathcal{A})F'(x^\dagger)^*] w_j.
 \end{aligned}$$

It follows from Assumption 2.1, Lemma 3.1, and Lemma 3.2 that

$$\begin{aligned}
 &\|T(x_{l+1}^\delta - x_{l+1}) - y^\delta + y\| \\
 &\leq \delta + C \sum_{j=0}^l \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1} \|v_j\| + C \sum_{j=0}^l \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1} K_0 \|e_j\| \|w_j\| \\
 (5.28) \quad &\quad + C \sum_{j=0}^l \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1} K_0 \|x_j^\delta - x_j\| \|F(x_j^\delta) - y^\delta\|.
 \end{aligned}$$

With the help of Assumption 2.3, Lemma 5.1, Lemma 5.2, and the induction hypotheses we have

$$\|v_j\| \lesssim K_0 (\|e_j\| + \|e_j^\delta\|) \|T(x_j^\delta - x_j)\| \lesssim K_0 \|e_0\| \delta$$

and

$$\|w_j\| \leq \delta + \|T(x_j^\delta - x_j)\| + \|v_j\| \lesssim \delta.$$

Moreover, by using Assumption 2.3, Lemma 5.1, and (5.2) we have for $0 \leq j \leq l$

$$\|F(x_j^\delta) - y^\delta\| \leq \delta + \|Te_j^\delta\| + K_0 \|e_j^\delta\| \|Te_j^\delta\| \lesssim \delta + s_j^{-1/2} \|e_0\| \lesssim s_j^{-1/2} \|e_0\|.$$

Combining the above three inequalities with (5.27) and (5.28) and using the induction hypothesis $\|x_j^\delta - x_j\| \lesssim s_j^{1/2} \delta$ for $0 \leq j \leq l$, it follows that

$$(5.29) \quad \|x_{l+1}^\delta - x_{l+1}\| \leq (b_2 + CK_0 \|e_0\|) \delta \sum_{j=0}^l \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1/2}$$

and

$$(5.30) \quad \|T(x_{l+1}^\delta - x_{l+1}) - y^\delta + y\| \leq \delta + K_0 \|e_0\| \delta \sum_{j=0}^l \frac{1}{\alpha_j} (s_l - s_{j-1})^{-1}.$$

By Lemma 4.1 and the fact $s_l \leq s_{l+1}$ we obtain for small $K_0 \|e_0\|$ that

$$\|x_{l+1}^\delta - x_{l+1}\| \leq (b_2 c_2 + CK_0 \|e_0\|) s_l^{1/2} \delta \leq 2b_2 c_2 s_{l+1}^{1/2} \delta.$$

Moreover, with the help of (5.1) we can derive

$$\|T(x_{l+1}^\delta - x_{l+1}) - y^\delta + y\| \leq (1 + CK_0 \|e_0\|) \delta.$$

Thus $\|T(x_{l+1}^\delta - x_{l+1})\| \leq 3\delta$ if $K_0 \|e_0\|$ is suitably small. We therefore complete the proof of (5.25).

Next we will prove (5.24). From the above proof we in fact obtain

$$\|T(x_n^\delta - x_n) - y^\delta + y\| \leq (1 + CK_0 \|e_0\|) \delta, \quad 0 \leq n \leq \hat{n}_\delta.$$

Therefore, it follows from Assumption 2.3, Lemma 5.1, and Lemma 5.2 that

$$\begin{aligned} & \|F(x_n^\delta) - F(x_n) - y^\delta + y\| \\ & \leq \|F(x_n^\delta) - F(x_n) - T(x_n^\delta - x_n)\| + \|T(x_n^\delta - x_n) - y^\delta + y\| \\ & \leq K_0 (\|e_n^\delta\| + \|e_n\|) \|T(x_n^\delta - x_n)\| + (1 + CK_0 \|e_0\|) \delta \\ & \leq (1 + CK_0 \|e_0\|) \delta. \end{aligned}$$

The proof is thus complete. \square

5.3. Completion of the proof of Theorem 2.2. We have shown in Lemma 5.1 that $n_\delta \leq \hat{n}_\delta$. Thus we may use the definition of n_δ and Proposition 5.6 to obtain

$$(5.31) \quad \|F(x_{n_\delta}) - y\| \leq \|F(x_{n_\delta}^\delta) - y^\delta\| + \|F(x_{n_\delta}^\delta) - F(x_{n_\delta}) - y^\delta + y\| \lesssim \delta,$$

and for $0 \leq n < n_\delta$,

$$\begin{aligned} \tau \delta & \leq \|F(x_n^\delta) - y^\delta\| \leq \|F(x_n^\delta) - F(x_n) - y^\delta + y\| + \|F(x_n) - y\| \\ & \leq (1 + CK_0 \|e_0\|) \delta + \|F(x_n) - y\|. \end{aligned}$$

Since $\tau > 1$, if $K_0 \|e_0\|$ is suitably small, then

$$(5.32) \quad \delta \lesssim \|F(x_n) - y\| \lesssim \|Te_n\|, \quad 0 \leq n < n_\delta.$$

We now prove the convergence of $x_{n_\delta}^\delta$ to x^\dagger as $\delta \rightarrow 0$. Assume first that there is a sequence $\delta_k \searrow 0$ such that $n_k := n_{\delta_k} \rightarrow n$ as $k \rightarrow \infty$ for some finite integer n . Without loss of generality, we can assume that $n_k = n$ for all k . It then follows from (5.31) that $F(x_n) = y$. Thus, from (5.7) we can conclude that $x_j = x_n$ for all $j \geq n$. Since Theorem 5.5 implies $x_j \rightarrow x^\dagger$ as $j \rightarrow \infty$, we must have $x_n = x^\dagger$, which together with Proposition 5.6 implies $x_{n_k}^{\delta_k} \rightarrow x^\dagger$ as $k \rightarrow \infty$.

Assume next that there is a sequence $\delta_k \searrow 0$ such that $n_k := n_{\delta_k} \rightarrow \infty$ as $k \rightarrow \infty$. Then Theorem 5.5 and (5.32) imply that $\|e_{n_k}\| \rightarrow 0$ and $s_{n_k}^{1/2} \delta_k \rightarrow 0$ as $k \rightarrow \infty$. Consequently, by Proposition 5.6 we again obtain $x_{n_k}^{\delta_k} \rightarrow x^\dagger$ as $k \rightarrow \infty$.

6. Numerical results. In this section we present some numerical results to test the theoretical conclusions given in Theorems 2.1 and 2.2. We consider the estimation of the coefficient c in the two-point boundary value problem

$$(6.1) \quad \begin{cases} -u'' + cu = f & \text{in } (0, 1), \\ u(0) = g_0, & u(1) = g_1 \end{cases}$$

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 (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 $F(c) := u(c)$, where $u(c)$ denotes the unique solution of (6.1). It is well known that F is well defined on

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

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

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

where $A(c) : H^2 \cap H_0^1 \mapsto L^2$ is defined by $A(c)u = -u'' + cu$. It has been shown in [5] that if, for the sought solution c^\dagger , $|u(c^\dagger)| \geq \kappa > 0$ on $[0, 1]$, then Assumption 2.3 is satisfied in a neighborhood of c^\dagger .

In the following we report some numerical results on the methods given by (1.4) and (1.5) with g_α defined by (2.14), which, in the current context, defines the iterative solutions $\{c_n^\delta\}$ by

$$\begin{aligned} u_{n,0} &= c_n^\delta, \\ u_{n,l+1} &= u_{n,l} - F'(c_n^\delta)^*(F(c_n^\delta) - u^\delta - F'(c_n^\delta)(c_n^\delta - u_{n,l})), \quad 0 \leq l \leq [1/\alpha_n] - 1, \\ c_{n+1}^\delta &= u_{n,[1/\alpha_n]} \end{aligned}$$

and determines the stopping index n_δ by

$$(6.2) \quad \|F(c_{n_\delta}^\delta) - u^\delta\| \leq \tau\delta < \|F(c_n^\delta) - u^\delta\|, \quad 0 \leq n < n_\delta.$$

During the computation, all differential equations are solved approximately by a finite difference method by dividing the interval $[0, 1]$ into $m + 1$ subintervals with equal length $h = 1/(m + 1)$; we take $m = 100$ in our actual computation.

Example 1. We estimate c in (6.1) by assuming $f(t) = (1+t)(1+t-0.8\sin(2\pi t))$, $g_0 = 1$ and $g_1 = 2$. If $u(c^\dagger) = 1+t$, then $c^\dagger = 1+t-0.8\sin(2\pi t)$ is the sought solution. When applying the above method, we take $\alpha_n = 2^{-n}$ and use random noise data u^δ satisfying $\|u^\delta - u(c^\dagger)\|_{L^2[0,1]} = \delta$ with noise level $\delta > 0$. As an initial guess we choose $c_0 = 1+t$. One can show that $c_0 - c^\dagger \in \mathcal{R}(F'(c^\dagger)^*)$. Thus, according to Theorem 2.1, the expected rate of convergence should be $O(\delta^{1/2})$.

The numerical result is reported in Table 6.1. In order to see the effect of τ in the discrepancy principle (6.2), we consider the three distinct values $\tau = 1.1, 2$, and 4 . In order to indicate the dependence of the convergence rates on the noise level, different values of δ are selected. The rates in Table 6.1 coincide with Theorem 2.1 very well. Table 6.1 indicates also that the absolute error increases with respect to τ . Thus, in numerical computation, one should use smaller τ if possible.

TABLE 6.1

Numerical results for Example 1 with $\alpha_n = 2^{-n}$ and three distinct values of τ , where n_δ is determined by (6.2), $\text{error} := \|c_{n_\delta}^\delta - c^\dagger\|_{L^2}$, and $\text{ratio} := \text{error}/\delta^{1/2}$.

δ	$\tau = 1.1$			$\tau = 2.0$			$\tau = 4.0$		
	k_δ	Error	Ratio	k_δ	Error	Ratio	k_δ	Error	Ratio
10^{-2}	12	$4.67e-2$	0.47	9	$2.90e-1$	2.90	1	$5.65e-1$	5.65
10^{-3}	14	$1.47e-2$	0.47	12	$3.89e-2$	1.23	12	$3.89e-2$	1.23
10^{-4}	16	$4.30e-3$	0.43	15	$5.30e-3$	0.53	14	$8.70e-3$	0.87
10^{-5}	18	$1.30e-3$	0.40	17	$1.80e-3$	0.56	16	$2.80e-3$	0.87
10^{-6}	21	$4.45e-4$	0.44	19	$6.41e-4$	0.64	18	$1.00e-3$	1.03

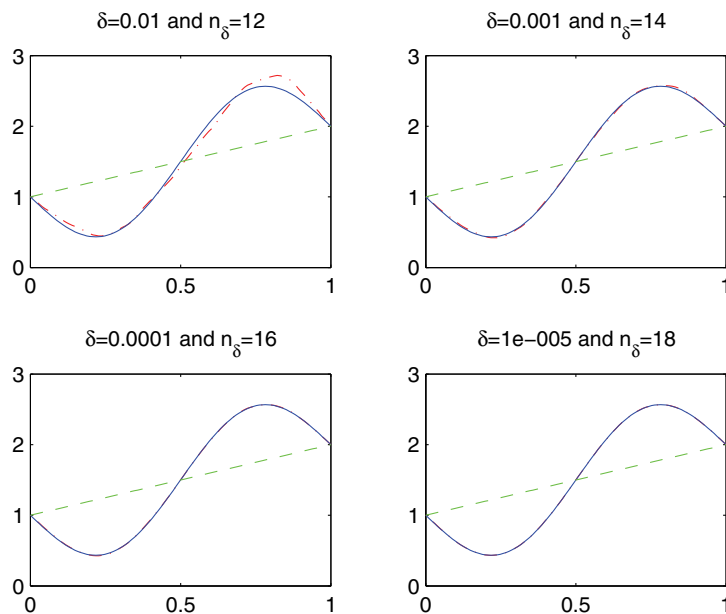


FIG. 6.1. Comparison on the computed and exact solution for Example 1 with $\tau = 1.1$.

In order to visualize the computed solutions, we plot in Figure 6.1 the results obtained for $\tau = 1.1$ and various values of the noise level δ , where the solid, dashed, and dash-dotted curves denote the exact solution c^\dagger , the initial guess c_0 , and the computed solution $c_{n_\delta}^\delta$ respectively. It clearly indicates the fast convergence as $\delta \rightarrow 0$ as reported in Table 6.1.

Example 2. We repeat Example 1 but with $\tau = 1.1$ and the initial guess $c_0 = 2 - t$. Now $c_0 - c^\dagger \notin \mathcal{R}(F'(c^\dagger)^*)$, and in fact $c_0 - c^\dagger$ has no sourcewise representation $c_0 - c^\dagger \in \mathcal{R}((F'(c^\dagger)^*F'(c^\dagger))^\nu)$ with a good $\nu > 0$ since $c_0(0) \neq c^\dagger(0)$ and $c_0(1) \neq c^\dagger(1)$. However, Theorem 2.2 asserts that $\|c_{n_\delta}^\delta - c^\dagger\|_{L^2[0,1]} \rightarrow 0$ as $\delta \rightarrow 0$. Figure 6.2 clearly indicates such a convergence, although the convergence speed could be quite slow which is typical for inverse problems.

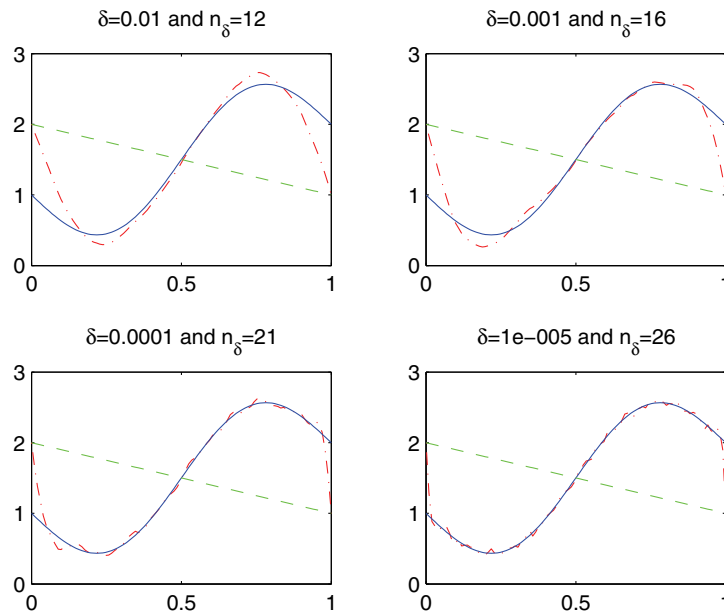


FIG. 6.2. Comparison on the computed and exact solutions for Example 2 with $\tau = 1.1$.

REFERENCES

- [1] A. B. BAKUSHINSKY AND M. YU. KOKURIN, *Iterative Methods for Approximate Solutions of Inverse Problems*, Math. Appl. (N.Y.) 577, Springer, Dordrecht, 2004.
- [2] R. S. DEMBO, S. C. EISENSTAT, AND T. STEIHAUG, *Inexact Newton methods*, SIAM J. Numer. Anal., 19 (1982), pp. 400–408.
- [3] M. HANKE, *A regularizing Levenberg-Marquardt scheme with applications to inverse ground-water filtration problems*, Inverse Problems, 13 (1997), pp. 79–95.
- [4] M. HANKE, *The regularizing Levenberg-Marquardt scheme is of optimal order*, J. Integ. Equations Appl., 22 (2010), pp. 259–283.
- [5] M. HANKE, A. NEUBAUER, AND O. SCHERZER, *A convergence analysis of the Landweber iteration for nonlinear ill-posed problems*, Numer. Math., 72 (1995), pp. 21–37.
- [6] M. HOCHBRUCK, M. HÖNIG, AND A. OSTERMANN, *A convergence analysis of the exponential Euler iteration for nonlinear ill-posed problems*, Inverse Problems, 25 (2009), article no. 075009.
- [7] Q. JIN, *On a regularized Levenberg-Marquardt method for solving nonlinear inverse problems*, Numer. Math., 115 (2010), pp. 229–259.
- [8] Q. JIN AND U. TAUTENHAHN, *On the discrepancy principle for some Newton type methods for solving nonlinear inverse problems*, Numer. Math., 111 (2009), pp. 509–558.
- [9] Q. JIN AND U. TAUTENHAHN, *Inexact Newton regularization methods in Hilbert scales*, Numer. Math., 117 (2011), pp. 555–579.
- [10] A. LECHLEITER AND A. RIEDER, *Towards a general convergence theory for inexact Newton regularizations*, Numer. Math., 114 (2010), pp. 521–548.
- [11] P. PORNSAWAD AND C. BÖCKMANN, *Convergence rate analysis of the first-stage Runge-Kutta type regularizations*, Inverse Problems, 26 (2010), article no. 035005.
- [12] A. RIEDER, *On the regularization of nonlinear ill-posed problems via inexact Newton iterations*, Inverse Problems, 15 (1999), pp. 309–327.
- [13] A. RIEDER, *On convergence rates of inexact Newton regularizations*, Numer. Math., 88 (2001), pp. 347–365.