RACSAM (2019) 113:2813–2826
https://doi.org/10.1007/s13398-019-00657-w

ORIGINAL PAPER

Expanding the applicability of an a posteriori parameter choice strategy for Tikhonov regularization of nonlinear ill-posed problems

Ioannis K. Argyros[1] · Yeol Je Cho[2,3] · Santhosh George[4] · Yi-bin Xiao[2]

Received: 15 October 2018 / Accepted: 26 February 2019 / Published online: 15 March 2019

Abstract

We expand the applicability of an a posteriori parameter choice strategy for Tikhonov regularization of the nonlinear ill-posed problem presented in Jin and Hou (Numer Math 83:139–159, 1999) by weakening the conditions needed in Jin and Hou [13]. Using a center-type Lipschitz condition instead of a Lipschitz-type condition used in Jin and Hou [13], Scherzer et al. (SIAM J Numer Anal 30:1796–1838, 1993), we obtain a tighter convergence analysis. Numerical examples are presented to show that our results apply but earlier ones do not apply to solve equations.

Keywords Nonlinear ill-posed problems · Tikhonov regularization · Discrepancy principle

Mathematics Subject Classification 65J20 · 65J15 · 47J06

✉ Yeol Je Cho
yjcho@gnu.ac.kr

Ioannis K. Argyros
ioannisa@cameron.edu

Santhosh George
sgeorge@nitk.ac.in

Yi-bin Xiao
xiaoyb9999@hotmail.com

1 Department of Mathematicsal Sciences, Cameron University, Lawton, OK 73505, USA

2 School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu 611731, Sichuan, People's Republic of China

3 Department of Mathematics Education, Gyeongsang National University, Jinju 52828, Korea

4 Department of Mathematical and Computational Sciences, National Institute of Technology Karnataka, Mangalore 757 025, India

1 Introduction

Jin and Hou [13] considered an a posteriori parameter choice strategy for Tikhonov regularization [6–10,14,18–21] of the nonlinear ill-posed problem

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

Here $F : D(F) \subseteq X \to Y$ is a weakly continuous and Fréchet differentiable nonlinear operator between the Hilbert spaces X and Y. It is assumed that $y \in R(F)$, the range of F and that the available data is y^δ with

$$\|y^\delta - y\| \leq \delta.$$

Recall that, in Tikhonov regularization for solving the problem (1.1), the solution x_α^δ of the minimization problem

$$\min_{x \in D(F)} \{\|F(x) - y^\delta\|^2 + \alpha\|x - x_0\|^2\} \tag{1.2}$$

is used to approximate the x_0-minimal norm solution (shortly, x_0-MNS) of the problem (1.1), where $\alpha > 0$ is the regularization parameter and $x_0 \in D(F)$ is an a priori guess of the x_0-MNS $\hat{x}$ of the problem (1.1), i.e.,

$$F(\hat{x}) = y, \quad \|\hat{x} - x_0\| = \min_{x \in D(F)} \{\|x - x_0\| : F(x) = y\}.$$

It is known [9,13] that the regularization parameter α affects not only the convergence of x_α^δ but also the rates of convergence and hence the choice of the regularization parameter is crucial. Many discrepancy principles are considered in the literature for choosing the parameter α (see [8,11–13,17–19]). In [13], the following discripancy principle (developed by Gfrere in [8] for linear ill-posed problems) has been considered for choosing the regularization parameter α.

Rule 1 Let $c \geq 1$ be a given constant and $x_0 \in D(F)$.

1. If $\|F(x_0) - y^\delta\| \leq c\delta^2$, then choose $\alpha = \infty$, i.e., take x_0 as approximation;
2. If $\|F(x_0) - y^\delta\| > c\delta^2$, then choose $\alpha := \alpha(\delta)$ be the root of the equation

$$f(\alpha) := \alpha\langle F(x_\alpha^\delta) - y^\delta, (\alpha I + F'(x_\alpha^\delta)F'(x_\alpha^\delta)^*(F(x_\alpha^\delta) - y^\delta)\rangle = c\delta^2 \tag{1.3}$$

where $F'(x)$ denotes the Fréchet derivative of F at point $x \in D(F)$ and $F'(x)^*$ denote the adjoint of $F'(x)$.

This rule was considered in [19] under a series of restrictive conditions on F (see the assumptions (10)–(14) and (93)–(98) in [19]). In [13], Jin and Hou considered **Rule** 1 under the following weaker and more easier to check conditions.

Assumption 1.1 Let $\hat{x}$ be an x_0-MNS of the problem (1.1) such that there exists a number $p > 3\|x_0 - \hat{x}\|$ such that $B(\hat{x}, p) \subset D(F)$ and there exists a constant K_0 such that, for all $x, z \in B(\hat{x}, p)$ and $v \in X$, there exists $k(x, z, v) \in X$ such that

$$[F'(x) - F'(z)]v = F'(z)k(x, z, v)$$

and

$$\|k(x, z, v)\| \leq K_0\|x - z\|\|v\|.$$

Further, it is assumed that there exist $\nu > 0$ and an element $\omega \in N(F'(\hat{x}))^\perp \subset X$ such that

$$x_0 - \hat{x} = (F'(\hat{x})^* F'(\hat{x}))^{\nu/2}\omega. \tag{1.4}$$

In this paper, we further weaken the above assumption and the assumption (12) in [19] with the following assumptions.

Assumption 1.2 Suppose that there exist constants $L_0 > 0$, $L_0^* > 0$ such that, for all $x \in B(x_0, r) \subseteq D$ and $w \in X$, there exists $\varphi(x, x_0, w), \varphi_1(x, x_0, w) \in X$ such that

1. $$[F'(x) - F'(x_0)]w = F'(x_0)\varphi(x, x_0, w), \quad \|\varphi(x, x_0, w)\| \le L_0\|x - x_0\|\|w\|$$

   and

2. $$[F'(x)^* - F'(x_0)^*]F'(x_0)w = F'(x_0)^*F'(x_0)\varphi_1(x, x_0, w),$$
   $$\|\varphi_1(x, x_0, w)\| \le L_0^*\|x - x_0\|\|w\|.$$

Assumption 1.3 There exists a continuous and strictly monotonically increasing function $\varphi : (0, a] \to (0, \infty)$ with $a \ge \|F'(x_0)\|^2$ satisfying the following:

1. $\lim_{\lambda\to 0}\varphi(\lambda) = 0$;
2. $\sup_{\lambda\ge 0}\frac{\alpha\varphi(\lambda)}{\lambda+\alpha} \le \varphi(\alpha)$ for all $\lambda \in (0, a]$;
3. there exists $v \in X$ such that

$$x_0 - \hat{x} = \varphi(A_0^*A_0)v. \tag{1.5}$$

Remark 1.4 The hypotheses of Assumption 1.1 may not hold or may be very expensive or impossible to verify in general. In particular, as it is the case for well-posed nonlinear equations the computation of the Lipschitz constant K_0 even if this constant exists is very difficult. Moreover, there are classes of operators for which Assumption 1.1 is not satisfied but the method (1.3) with **Rule** 1 converges.

In the present paper, we expand the applicability of the method (1.3) with **Rule** 1 under less computational cost.

The advantages of the new approach are:

1. Assumption 1.2 is weaker than Assumption 1.1. Notice that there are classes of operators that satisfy Assumption 1.2 but do not satisfy Assumption 1.1;
2. The computational cost of constant L_0 is less than that of constant K_0, even when $K_0 = L_0$;
3. The sufficient convergence criteria are weaker;
4. The computable error bounds on the distances involved (including L_0) are less costly and more precise than the old ones (including K_0);
5. The information on the location of the solution is more precise;

   and

6. The convergence domain of the method (1.3) with **Rule** 1 is larger.

These advantages are also very important in computational mathematics since they provide under less computational cost a wider choice of initial guesses for the method (1.3) with **Rule** 1 and the computation of fewer iterates to achieve a desired error tolerance. Numerical examples for 1–6 are presented in Sect. 3.

Remark 1.5 Note that the source condition (1.4) involves the Fréchet derivative at the exact solution $\hat{x}$ which is unknown in practice. But the source condition (1.5) depends on the Fréchet derivative of F at x_0. It can be seen that the functions

$$\varphi(\lambda) = \lambda^{\nu}, \quad \lambda > 0,$$

for all $0 < \nu \leq 1$ and

$$\varphi(\lambda) = \begin{cases} \left(\ln \frac{1}{\lambda}\right)^{-p}, & 0 < \lambda \leq e^{-(p+1)}, \\ 0, & \text{otherwise} \end{cases}$$

for all $p \geq 0$ satisfy Assumption 1.3 (see [15]).

Now, we give an example which satisfies Assumption 1.2.

Example 1.6 ([19], Example 2.7) Let $F : H^1(0,1) \to L^2(0,1)$ be defined by

$$(Fx)(t) = \int_0^1 k(t,\tau) g(\tau, x(\tau)) d\tau, \tag{1.6}$$

where k is continuous and g sufficiently smooth so that F is Fréchet differentiable with respect to x and

$$F'(x)h(t) = \int_0^1 k(t,\tau) g_x(\tau, x(\tau)) h(\tau) d\tau. \tag{1.7}$$

Let $N : H^1(0,1) \to H^1(0,1)$ be defined by $(Nx)(t) = g_x(t, x(t))$. Assume that N is locally Lipschitz continuous in a neighborhood $U(x_0)$ of x_0 in H^1, i.e., there exists $L = L(U)$ such that

$$\|g_x(\cdot, x(\cdot)) - g_x(\cdot, x_0(\cdot))\|_{H^1} \leq L\|x - x_0\|_{H^1} \tag{1.8}$$

for all $x \in H^1$. Further, we assume that there exists $\kappa > 0$ such that $(Nx_0)(t) = g_x(t, x_0(t)) \geq \kappa$ for all $t \in [0,1]$. Then, for all $x \in B(x_0, \frac{\kappa}{2L})$, we have in turn

$$(Nx)(t) \geq \frac{\kappa}{2} \tag{1.9}$$

for all $t \in [0,1]$. Note that

$$\begin{aligned} [F'(x)h - F'(x_0)h](t) &= \int_0^1 k(t,\tau) g_x(\tau, x_0(\tau)) \left[\frac{g_x(\tau, x(\tau))}{g_x(\tau, x_0(\tau))} - 1 \right] h(\tau) d\tau \\ &= F'(x_0)\varphi(x, x_0, h), \end{aligned} \tag{1.10}$$

where

$$\varphi(x, x_0, h) = \left[\frac{g_x(\tau, x(\tau))}{g_x(\tau, x_0(\tau))} - 1 \right] h(\tau).$$

By (1.8), $g_x(\tau, x(\tau))$ is bounded for all $x \in B(x_0, \frac{\kappa}{2L})$ and hence, by the Banach algebra property [19] of H^1, there exists a constant K such that

$$\begin{aligned} \|\varphi(x, x_0, h)\| &\leq K\|g_x(\tau, x(\tau)) - g_x(\tau, x_0(\tau))\|_{H^1} \left\| \frac{h(\cdot)}{g_x(\cdot, x_0(\cdot))} \right\|_{H^1} \\ &\leq KL\|x - x_0\|_{H^1} \sqrt{\left\| \frac{g_x h_t - h\frac{d}{dt} g_x(\cdot, x(\cdot))}{g_x^2(\cdot, x(\cdot))} \right\|_{L^2}^2 + \left\| \frac{h}{g_x} \right\|_{L^2}^2} \\ &\leq KL\|x - x_0\|_{H^1} \max\left\{ \frac{2}{\kappa}, \frac{4}{\kappa^2} \right\} \\ &\quad \times \sqrt{\|g_x h_t - h\frac{d}{dt} g_x(\cdot, x(\cdot))\|_{L^2}^2 + \|h\|_{L^2}^2}. \end{aligned} \tag{1.11}$$

The above estimate and the the fact that $g_x(\tau, x(\tau))$ is bounded for all $x \in B(x_0, \frac{\kappa}{2L})$ implies that

$$\|\varphi(x, x_0, h)\| \le K_1\|x - x_0\|_{H^1}\|h|_{H^1},$$

where K_1 is independent of $x, h \in H^1$. So F satisfies the condition 1 of Assumption 1.2.

To verify 2 of Assumption 1.2 as in [19], we introduce the Neumann operator $\aleph : H^1 \to (H^1)^*$ by

$$\langle Dx, D\phi\rangle_{L^2} + \langle x, \phi\rangle_{L^2} = \langle\aleph x, \phi\rangle_{(H^1)^* H^1}$$

for all $\phi \in H^1$. Formally, we have $\aleph x = (-\triangle + I)x$. Then the dual $F'(x)^*$ of $F'(x)$ is given by

$$F'(x)^* h = \aleph^{-1}\Big[g_x(\cdot, x(\cdot))\int_0^1 k(t, \cdot)h(t)dt\Big].$$

Then we have

$$\begin{aligned}
&[F'(x)^* - F'(x_0)^*]F'(x_0)h \\
&= \aleph^{-1}\Big[(g_x(\cdot, x(\cdot)) - g_x(\cdot, x_0(\cdot)))\int_0^1 k(t, \cdot)\Big(\int_0^1 k(\cdot, \tau)g_x(\tau, x_0(\tau))h(\tau)d\tau\Big)dt\Big] \\
&= \aleph^{-1}\Big[g_x(\cdot, x_0(\cdot))\int_0^1 k(t, \cdot)\Big(\int_0^1 k(\cdot, \tau)g_x(\tau, x_0(\tau))\Big[\frac{g_x(\cdot, x(\cdot))}{g_x(\cdot, x_0(\cdot))} - 1\Big]h(\tau)d\tau\Big)dt\Big] \\
&= \aleph^{-1}\Big[g_x(\cdot, x_0(\cdot))\int_0^1 k(t, \cdot)\Big(\int_0^1 k(\cdot, \tau)g_x(\tau, x_0(\tau))\varphi_1(x, x_0, h)d\tau\Big)dt\Big],
\end{aligned}$$

where

$$\varphi_1(x, x_0, h) = \Big[\frac{g_x(., x(.))}{g_x(., x_0(.))} - 1\Big]h.$$

As in (1.11), one can prove that

$$\|\varphi_1(x, x_0, h)\| \le K_2\|x - x_0\|_{H^1}\|h|_{H^1}$$

for some constant K_2. Let $\rho := \|x_0 - \hat{x}\| < 1 - \frac{\delta_0}{\sqrt{\alpha_0}}$ for some $\alpha_0 > 0$ and let $\frac{\delta_0}{\sqrt{\alpha_0}} + \rho := r$. Then, for any $\alpha \ge \alpha_0$, we have $x_\alpha^\delta \in B(x_0, r)$.

The organization of this paper is as follows: Convergence analysis and parameter choice strategy are discussed in Sect. 2 and Numerical examples are given in Sect. 3.

2 Error analysis

Considering all notations of the Sect. 1, let

$$r < \frac{\sqrt{(L_0 + L_0^*)^2 + 4L_0L_0^*} - (L_0 + L_0^*)}{2L_0L_0^*}.$$

Theorem 2.1 *Let x_α^δ be as in (1.2) and Assumption 1.3 holds. Then*

$$\|x_\alpha^\delta - \hat{x}\| \le \frac{1 + L_0 r}{1 - q}\Big(\frac{\delta}{\sqrt{\alpha}} + \varphi(\alpha)\Big),$$

where $q = (L_0 + L_0^* + L_0 L_0^* r)r$.

Proof Let $M = \int_0^1 F'(\hat{x} + t(x_\alpha^\delta - \hat{x}))dt$. Then we have

$$F(x_\alpha^\delta) - F(\hat{x}) = M(x_\alpha^\delta - \hat{x})$$

and hence, by (1.2), we have

$$(A_0^* M + \alpha I)(x_\alpha^\delta - \hat{x}) = A_0^*(y^\delta - y) + \alpha(x_0 - \hat{x}) + (F'(x_\alpha^\delta)^* - F'(x_0)^*)(y^\delta - F(x_\alpha^\delta)).$$

Thus we have

$$\begin{aligned} x_\alpha^\delta - \hat{x} &= (A_0^* A_0 + \alpha I)^{-1}[A_0^*(y^\delta - y) + \alpha(x_0 - \hat{x}) + A_0^*(A_0 - M)(x_\alpha^\delta - \hat{x})] \\ &\quad + (A_0^* A_0 + \alpha I)^{-1}(F'(x_\alpha^\delta)^* - F'(x_0)^*)(y^\delta - F(x_\alpha^\delta)) \\ &= s_1 + s_2 + s_3 + s_4, \end{aligned} \tag{2.1}$$

where

$$\begin{aligned} s_1 &:= (A_0^* A_0 + \alpha I)^{-1} A_0^*(y^\delta - y), \quad s_2 := (A_0^* A_0 + \alpha I)^{-1}\alpha(x_0 - \hat{x}), \\ s_3 &:= (A_0^* A_0 + \alpha I)^{-1} A_0^*(A_0 - M)(x_\alpha^\delta - \hat{x}) \end{aligned}$$

and

$$s_4 = (A_0^* A_0 + \alpha I)^{-1}(F'(x_\alpha^\delta)^* - F'(x_0)^*)(y^\delta - F(x_\alpha^\delta)).$$

Note that

$$\|s_1\| \le \frac{\delta}{\sqrt{\alpha}}, \tag{2.2}$$

by Assumption 1.3,

$$\|s_2\| \le \varphi(\alpha), \tag{2.3}$$

by Assumption 1.2,

$$\|s_3\| \le L_0 r \|x_\alpha^\delta - \hat{x}\| \tag{2.4}$$

and

$$\begin{aligned} \|s_4\| &\le \|(A_0^* A_0 + \alpha I)^{-1}(F'(x_\alpha^\delta)^* - F'(x_0)^*)(y^\delta - y + F(\hat{x}) - F(x_\alpha^\delta))\| \\ &\le \|(A_0^* A_0 + \alpha I)^{-1}(F'(x_\alpha^\delta)^* - F'(x_0)^*)(y^\delta - y)\| \\ &\quad + \left\|(A_0^* A_0 + \alpha I)^{-1}(F'(x_\alpha^\delta)^* - F'(x_0)^*)(F(\hat{x}) - F(x_\alpha^\delta))\right\| \\ &\le \|(A_0^* A_0 + \alpha I)^{-1}\varphi_1(x_\alpha^\delta, x_0, y^\delta - y)\| \\ &\quad + \left\|(A_0^* A_0 + \alpha I)^{-1}(F'(x_\alpha^\delta)^* - F'(x_0)^*)\int_0^1 F'(\hat{x} + t(\hat{x} - x_\alpha^\delta))dt(\hat{x} - x_\alpha^\delta)\right\| \\ &\le \|(A_0^* A_0 + \alpha I)^{-1} A_0^* \varphi_1(x_\alpha^\delta, x_0, y^\delta - y)\| \\ &\quad + \left\|(A_0^* A_0 + \alpha I)^{-1}(F'(x_\alpha^\delta)^* - F'(x_0)^*)\int_0^1 F'(\hat{x} + t(\hat{x} - x_\alpha^\delta))dt(\hat{x} - x_\alpha^\delta)\right\| \\ &\le L_0^* \|x_\alpha^\delta - x_0\| \frac{\delta}{\sqrt{\alpha}} \end{aligned}$$

$$
\begin{aligned}
&+\left\|(A_0^*A_0+\alpha I)^{-1}(F'(x_\alpha^\delta)^*-F'(x_0)^*)\right.\\
&\left.\times\int_0^1[F'(\hat{x}+t(\hat{x}-x_\alpha^\delta))-F'(x_0)+F'(x_0)]dt(\hat{x}-x_\alpha^\delta)\right\|\\
\le{}& L_0^*r\frac{\delta}{\sqrt{\alpha}}+\|(A_0^*A_0+\alpha I)^{-1}(F'(x_\alpha^\delta)^*-F'(x_0)^*)\\
&\times F'(x_0)\varphi(\hat{x}+t(\hat{x}-x_\alpha^\delta),x_0,\hat{x}-x_\alpha^\delta)\|\\
&+\|(A_0^*A_0+\alpha I)^{-1}(F'(x_\alpha^\delta)^*-F'(x_0)^*)F'(x_0)(\hat{x}-x_\alpha^\delta)\|\\
\le{}& L_0^*r\frac{\delta}{\sqrt{\alpha}}+\|(A_0^*A_0+\alpha I)^{-1}F'(x_0)^*F'(x_0)\\
&\times\varphi_1(x_\alpha^\delta,x_0,\varphi(\hat{x}+t(\hat{x}-x_\alpha^\delta),x_0,\hat{x}-x_\alpha^\delta))\|\\
&+\|(A_0^*A_0+\alpha I)^{-1}F'(x_0)^*F'(x_0)\varphi_1(x_\alpha^\delta,x_0,\hat{x}-x_\alpha^\delta)\|\\
\le{}& L_0^*r\frac{\delta}{\sqrt{\alpha}}+L_0^*L_0\|x_\alpha^\delta-x_0\|\|\hat{x}+t(\hat{x}-x_\alpha^\delta)-x_0\|\|\hat{x}-x_\alpha^\delta\|\\
&+L_0^*\|x_\alpha^\delta-x_0\|\|\hat{x}-x_\alpha^\delta\|\\
\le{}& L_0^*r\frac{\delta}{\sqrt{\alpha}}+L_0^*L_0r^2\|\hat{x}-x_\alpha^\delta\|+L_0^*r\|\hat{x}-x_\alpha^\delta\|\\
\le{}& L_0^*r\frac{\delta}{\sqrt{\alpha}}+(L_0^*L_0r^2+L_0^*r)\|\hat{x}-x_\alpha^\delta\|.
\end{aligned}
\tag{2.5}
$$

The result now follows from (2.1), (2.2), (2.3) and (2.4). This completes the proof. □

In order to show that the method (1.3) has a root, we first show that the function $\alpha\to f(\alpha)$ is continuous. Observe that, if x_α^δ is differentiable with respect to α, then

$$
\|x_\alpha^\delta-x_\beta^\delta\|\le\sup_{\gamma\in[\min\{\alpha,\beta\},\max\{\alpha,\beta\}]}\left\|\frac{dx_\alpha^\delta}{d\alpha}(\gamma)\right\||\beta-\alpha|.
$$

This implies $\lim_{\alpha\to\beta}x_\alpha^\delta=x_\beta^\delta$. Thus our main aim is to show that x_α^δ is differentiable with respect to α. Note that the minimizer of (1.2) satisfies the equation

$$
F'(x_\alpha^\delta)^*(F(x_\alpha^\delta)-y^\delta)+\alpha(x_\alpha^\delta-x_0)=0. \tag{2.6}
$$

The formal differentiation of (2.6) with respect to α yields

$$
F'(x_\alpha^\delta)^*F'(x_\alpha^\delta)\frac{dx_\alpha^\delta}{d\alpha}+(F'(x_\alpha^\delta)^*)'(F(x_\alpha^\delta)-y^\delta)\frac{dx_\alpha^\delta}{d\alpha}+\alpha\frac{dx_\alpha^\delta}{d\alpha}=-(x_\alpha^\delta-x_0) \tag{2.7}
$$

or, equivalently,

$$
\frac{dx_\alpha^\delta}{d\alpha}=-(\alpha I+F'(x_\alpha^\delta)^*F'(x_\alpha^\delta)+(F'(x_\alpha^\delta)^*)'(F(x_\alpha^\delta)-y^\delta))^{-1}(x_\alpha^\delta-x_0) \tag{2.8}
$$

if the operator

$$
(\alpha I+F'(x_\alpha^\delta)^*F'(x_\alpha^\delta)+(F'(x_\alpha^\delta)^*)'(F(x_\alpha^\delta)-y^\delta)) \tag{2.9}
$$

is invertible. Note that the operator (2.9) is invertible (see [19]) if the bilinear form

$$
\alpha(x,x)+\langle F'(x_\alpha^\delta)x,F'(x_\alpha^\delta)x\rangle+\langle F(x_\alpha^\delta)-y^\delta,F''(x_\alpha^\delta)(x,x)\rangle \tag{2.10}
$$

is elliptic.

Now, we prove that the bilinear form (2.10) is elliptic.

Lemma 2.2 *The bilinear form (2.10) is elliptic. Further, if $L_0(L_0+1)r \leq 1$, then*

$$\|(\alpha I + F'(x_\alpha^\delta)^* F'(x_\alpha^\delta) + (F'(x_\alpha^\delta)^*)'(F(x_\alpha^\delta) - y^\delta))^{-1}\| \leq \frac{1}{\alpha}.$$

Proof Observe that

$$\begin{aligned}
F''(x_\alpha^\delta)(x,x) &= \lim_{t\to 0} \frac{F'(x_\alpha^\delta + tx) - F'(x_\alpha^\delta)}{t} \\
&= \lim_{t\to 0} \frac{F'(x_\alpha^\delta + tx) - F'(x_0) + F'(x_0) - F'(x_\alpha^\delta)}{t} \\
&= \lim_{t\to 0} \frac{F'(x_0)[\varphi(x_\alpha^\delta + tx, x_0, x) - \varphi(x_\alpha^\delta, x_0, x)]}{t} \\
&= \lim_{t\to 0} \frac{F'(x_0)\varphi(x_\alpha^\delta + tx, x_\alpha^\delta, x)}{t}
\end{aligned}$$

and so

$$\begin{aligned}
&\left|\left\langle F(x_\alpha^\delta) - y^\delta, F''(x_\alpha^\delta)(x,x)\right\rangle\right| \\
&= \lim_{t\to 0} \left|\left\langle F(x_\alpha^\delta) - y^\delta, \frac{F'(x_0)\varphi(x_\alpha^\delta + tx, x_\alpha^\delta, x)}{t}\right\rangle\right| \\
&= \lim_{t\to 0} \left|\left\langle F(x_\alpha^\delta) - y^\delta, \frac{[F'(x_0) - F'(x_\alpha^\delta) + F'(x_\alpha^\delta)]\varphi(x_\alpha^\delta + tx, x_\alpha^\delta, x)}{t}\right\rangle\right| \\
&= \lim_{t\to 0} \left|\left\langle F'(x_\alpha^\delta)^*[F(x_\alpha^\delta) - y^\delta], \frac{\varphi(x_0, x_\alpha^\delta, \varphi(x_\alpha^\delta + tx, x_\alpha^\delta, x))}{t}\right\rangle\right| \\
&\quad + \lim_{t\to 0} \left|\left\langle F'(x_\alpha^\delta)^*[F(x_\alpha^\delta) - y^\delta], \frac{\varphi(x_\alpha^\delta + tx, x_\alpha^\delta, x)}{t}\right\rangle\right| \\
&= \lim_{t\to 0} \alpha \left|\left\langle x_\alpha^\delta - x_0, \frac{\varphi(x_0, x_\alpha^\delta, \varphi(x_\alpha^\delta + tx, x_\alpha^\delta, x))}{t}\right\rangle\right| \\
&\quad + \lim_{t\to 0} \alpha \left|\left\langle x_\alpha^\delta - x_0, \frac{\varphi(x_\alpha^\delta + tx, x_\alpha^\delta, x)}{t}\right\rangle\right| \\
&\leq \alpha L_0(L_0+1)\|x_\alpha^\delta - x_0\|\|x\|^2. \qquad (2.11)
\end{aligned}$$

This implies the bilinear form is elliptic. Again, by (2.11), we have

$$|\langle F(x_\alpha^\delta) - y^\delta, F''(x_\alpha^\delta)(x,x)\rangle| \leq \alpha\|x\|^2.$$

Therefore, we have (see [19])

$$\|(\alpha I + F'(x_\alpha^\delta)^* F'(x_\alpha^\delta) + (F'(x_\alpha^\delta)^*)'(F(x_\alpha^\delta) - y^\delta))^{-1}\| \leq \frac{1}{\alpha}.$$

This completes the proof. □

Lemma 2.3 *Let y^δ be fixed, $c > 2$ and $a := \frac{c/2-1}{\|x_0-\hat{x}\|^2}\delta^2$. Suppose that the assumptions in Lemma 2.2 holds. Then $f : (a,\infty) \to \mathbb{R}$ is continuous. For any $\alpha > a$ sufficiently small,*

$$f(\alpha) < c\delta^2.$$

Moreover, we have

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

Proof Note that

$$\begin{aligned}\|x_\alpha^\delta - x_\beta^\delta\| &\le \sup_{\gamma\in[\min\{\alpha,\beta\},\max\{\alpha,\beta\}]} \left\|\frac{dx_\alpha^\delta}{d\alpha}(\gamma)\right\| |\beta-\alpha| \\ &\le \frac{1}{\min\{\alpha,\beta\}}\|x_\alpha^\delta - x_0\||\beta-\alpha|.\end{aligned}$$

This implies $\|F(x_\alpha^\delta) - F(x_\beta^\delta)\| \to 0$ as $\alpha \to \beta$ and so the function f is continuous. The rest of the proof follows as in Lemma 3.8 in [19]. This completes the proof. □

Lemma 2.4 *Let $\chi(\lambda) := \lambda\sqrt{\varphi^{-1}(\lambda)}$. If α is chosen according to the method (1.3), then*

$$\alpha \ge \varphi^{-1}\chi^{-1}(C\delta),$$

where $C = \dfrac{\sqrt{c} - \frac{(L_0 r+1)^2}{1-q} - 1}{\frac{(L_0 r+1)^2}{1-q}}$.

Proof Observe that

$$\begin{aligned}\sqrt{c}\delta &= \|\sqrt{\alpha}(A_0A_0^* + \alpha I)^{-1/2}(F(x_\alpha^\delta) - y^\delta)\| \\ &\le \|\sqrt{\alpha}(A_0A_0^* + \alpha I)^{-1/2}(F(x_\alpha^\delta) - F(\hat{x}))\| + \delta \\ &\le \left\|\sqrt{\alpha}(A_0A_0^* + \alpha I)^{-1/2}\int_0^1 [F'(\hat{x} + t(x_\alpha^\delta - \hat{x})) \right. \\ &\qquad \left. - F'(x_0) + F'(x_0)]dt(x_\alpha^\delta - \hat{x})\right\| + \delta \\ &\le \left\|\sqrt{\alpha}(A_0A_0^* + \alpha I)^{-1/2}A_0\left[\int_0^1 \varphi(\hat{x} + t(x_\alpha^\delta - \hat{x}), x_0, x_\alpha^\delta - \hat{x})dt \right.\right. \\ &\qquad \left.\left. + (x_\alpha^\delta - \hat{x})\right]\right\| + \delta \\ &\le \sqrt{\alpha}(L_0 r + 1)\|x_\alpha^\delta - \hat{x}\| + \delta \\ &\le \sqrt{\alpha}\frac{(L_0r+1)^2}{1-q}\left(\frac{\delta}{\sqrt{\alpha}} + \varphi(\alpha)\right) + \delta. \qquad (2.12)\end{aligned}$$

Thus we have

$$\left(\sqrt{c} - \frac{(L_0r+1)^2}{1-q} - 1\right)\delta \le \frac{(L_0r+1)^2}{1-q}\sqrt{\alpha}\varphi(\alpha).$$

This completes the proof. □

Theorem 2.5 *If α is chosen according to the method (1.3) and $\chi^{-1}(c\lambda) = \chi^{-1}(c)\chi^{-1}(\lambda)$, then we have*

$$\|x_\alpha^\delta - \hat{x}\| = O(\chi^{-1}(\delta)).$$

Proof Let

$$F(x_\alpha^\delta)^*(F(x_\alpha^\delta) - y^\delta) + \alpha(x_\alpha^\delta - x_0) = 0.$$

Then we have

$$x_\alpha^\delta - \hat{x} = (A_0^*A_0 + \alpha I)^{-1}\alpha(x_0 - \hat{x}) - (A_0^*A_0 + \alpha I)^{-1}(s_\alpha + r_\alpha),$$

where

$$r_\alpha = (F'(x_\alpha^\delta)^* - F'(x_0)^*)(F(x_\alpha^\delta) - y^\delta)$$

and

$$s_\alpha = F'(x_0)^*[F(x_\alpha^\delta) - y - F'(x_0)(x_\alpha^\delta - x_0)].$$

Note that

$$[\alpha(A_0^*A_0 + \alpha I)^{-1} - \alpha_0(A_0^*A_0 + \alpha_0 I)^{-1}](x_0 - \hat{x}) = I_1 + I_2 + I_3,$$

where

$$\begin{aligned}
I_1 &= \left(1 - \frac{\alpha_0}{\alpha}\right)(A_0^*A_0 + \alpha_0 I)^{-1}A_0^*(F(x_\alpha^\delta) - y^\delta),\\
I_2 &= \left(1 - \frac{\alpha_0}{\alpha}\right)(A_0^*A_0 + \alpha_0 I)^{-1}A_0^*[A_0(x_\alpha^\delta - \hat{x}) - (F(x_\alpha^\delta) - y^\delta)]
\end{aligned}$$

and

$$I_3 = \left(1 - \frac{\alpha_0}{\alpha}\right)(A_0^*A_0 + \alpha_0 I)^{-1}A_0^*A_0[\alpha(A_0^*A_0 + \alpha I)^{-1}(x_0 - \hat{x}) - (x_\alpha^\delta - \hat{x})].$$

Observe that

$$\begin{aligned}
\|I_1\| &\le \frac{1}{\sqrt{\alpha_0}}\|\alpha^{1/2}(A_0^*A_0 + \alpha I)^{-1/2}(F(x_\alpha^\delta) - y^\delta)\|,\\
\|I_2\| &\le \|(A_0^*A_0 + \alpha_0 I)^{-1}A_0^*[A_0(x_\alpha^\delta - \hat{x}) - (F(x_\alpha^\delta) - F(\hat{x}) + y - y^\delta)]\|\\
&\le \|(A_0^*A_0 + \alpha_0 I)^{-1}A_0^*[A_0(x_\alpha^\delta - \hat{x}) - (F(x_\alpha^\delta) - F(\hat{x})\|\\
&\quad + \|(A_0^*A_0 + \alpha_0 I)^{-1}A_0^*(y - y^\delta)\|\\
&\le \left\|(A_0^*A_0 + \alpha_0 I)^{-1}A_0^*\int_0^1 [A_0 - F'(\hat{x} + t(x_\alpha^\delta - \hat{x})]dt(x_\alpha^\delta - \hat{x})\right\|\\
&\quad + \|(A_0^*A_0 + \alpha_0 I)^{-1}A_0^*(y - y^\delta)\|\\
&\le \left\|(A_0^*A_0 + \alpha_0 I)^{-1}A_0^*A_0\int_0^1 \varphi(\hat{x} + t(x_\alpha^\delta - \hat{x}), x_0, x_\alpha^\delta - \hat{x})\right\| + \frac{\delta}{\sqrt{\alpha_0}}\\
&\le L_0 r\|x_\alpha^\delta - \hat{x}\| + \frac{\delta}{\sqrt{\alpha_0}},\\
\|I_3\| &= \|(A_0^*A_0 + \alpha I)^{-1}(s_\alpha + r_\alpha)\|,
\end{aligned} \tag{2.13}$$

where

$$\begin{aligned}
s_\alpha &= A_0^*(F(x_\alpha^\delta) - y^\delta - A_0(x_\alpha^\delta - \hat{x}),\\
r_\alpha &= (F'(x_\alpha^\delta)^* - A_0^*)(F(x_\alpha^\delta) - y^\delta),\\
\|(A_0^*A_0 + \alpha I)^{-1}s_\alpha\| &= \|(A_0^*A_0 + \alpha I)^{-1}A_0^*(F(x_\alpha^\delta) - F(\hat{x}) - A_0(x_\alpha^\delta - \hat{x})\|\\
&\quad + \|(A_0^*A_0 + \alpha I)^{-1}A_0^*(y - y^\delta)\| \le L_0 r\|x_\alpha^\delta - \hat{x}\| + \frac{\delta}{\sqrt{\alpha}}
\end{aligned}$$

and

$$
\begin{aligned}
\|(A_0^*A_0+\alpha I)^{-1}r_\alpha\| &= \|(A_0^*A_0+\alpha I)^{-1}(F'(x_\alpha^\delta)^*-A_0^*)(F(x_\alpha^\delta)-y^\delta)\| \\
&= s_4 \\
&\le L_0^* r\frac{\delta}{\sqrt{\alpha}}+(L_0^*L_0r^2+L_0^*r)\|\hat{x}-x_\alpha^\delta\|
\end{aligned}
$$

Thus we have

$$
\begin{aligned}
[1-((2L_0+L_0^*)r+L_0L_0^*r^2)]\|x_\alpha^\delta-\hat{x}\| &\le \frac{\delta}{\alpha_0}+\varphi(\alpha_0)+\frac{\delta}{\alpha_0} \\
&\quad +\|\alpha^{1/2}(A_0^*A_0+\alpha I)^{-1/2}(F(x_\alpha^\delta)-y^\delta)\| \\
&\le (\sqrt{c}+1)\frac{\delta}{\alpha_0}+\varphi(\alpha_0)+\frac{\delta}{\alpha} \\
&\le (\sqrt{c}+2)\frac{\delta}{\alpha_0}+\varphi(\alpha_0) \\
&\le (\sqrt{c}+2)\left(\frac{\delta}{\alpha_0}+\varphi(\alpha_0)\right). \qquad (2.14)
\end{aligned}
$$

Thus the result follows by choosing $\alpha_0=\varphi^{-1}\chi^{-1}(C\delta)$, where C is as in Lemma 2.2. This completes the proof. □

Remark 2.6 Let us denote by $\bar{q}, \bar{c}$ the crucial constants obtained in the convergence analysis in [13] obtained using Assumption 1.1 instead of Assumption 1.2 (with K_0 replacing L_0 and L_0^*). Then, we have

$$
\frac{\bar{q}}{q}\to 0,\quad \frac{\bar{c}}{c}\to 0 \ \text{ as } \ \frac{L_0}{K_0}\to 0,\quad \frac{L_0^*}{K_0}\to 0,\quad \text{respectively.}
$$

These estimates show by how many times our new estimates can be better than the ones in [13]. A similar favorable comparison can be given using the rest of the constants introduced in our convergence analysis. The rest of the advantages of our approach have already been stated in the Abstract and the Introduction of this paper.

3 Numerical examples

In the next two cases, we present examples for nonlinear equations where Assumption 1.2 is satisfied but not Assumption 1.1.

Example 3.1 Let $X=Y=\mathbb{R}$, $D=[0,\infty)$, $x_0=1$ and define a function F on D by

$$
F(x)=\frac{x^{1+\frac{1}{i}}}{1+\frac{1}{i}}+c_1x+c_2, \qquad (3.1)
$$

where c_1, c_2 are real parameters and $i>2$ is an integer. Then $F'(x)=x^{1/i}+c_1$ is not Lipschitz on D. Hence Assumption 1.1 is not satisfied. However, the central Lipschitz condition in Assumption 1.2 holds for $K_0=1$. Indeed, we have

$$\begin{aligned}\|F'(x) - F'(x_0)\| &= |x^{1/i} - x_0^{1/i}| \\ &= \frac{|x - x_0|}{x_0^{\frac{i-1}{i}} + \cdots + x^{\frac{i-1}{i}}}\end{aligned}$$

and so

$$\|F'(x) - F'(x_0)\| \leq \overline{K}_0 |x - x_0|.$$

Example 3.2 We consider the integral equations

$$u(s) = f(s) + \lambda \int_a^b G(s,t) u(t)^{1+1/n} dt \tag{3.2}$$

for each $n \in \mathbb{N}$. Here f is a given continuous function satisfying $f(s) > 0$ for any $s \in [a, b]$, λ is a real number and the kernel G is continuous and positive in $[a, b] \times [a, b]$.

For example, when $G(s, t)$ is the Green kernel, the corresponding integral equation is equivalent to the boundary value problem

$$\begin{cases} u'' = \lambda u^{1+1/n}, \\ u(a) = f(a), \quad u(b) = f(b). \end{cases}$$

These type of problems have been considered in [1–5]. The equation of the form (3.2) generalize equations of the form

$$u(s) = \int_a^b G(s,t) u(t)^n dt \tag{3.3}$$

studied in [1–5]. Instead of (3.2), we can try to solve the equation $F(u) = 0$, where

$$F : \Omega \subseteq C[a,b] \to C[a,b], \Omega = \{u \in C[a,b] : u(s) \geq 0, s \in [a,b]\},$$

and

$$F(u)(s) = u(s) - f(s) - \lambda \int_a^b G(s,t) u(t)^{1+1/n} dt.$$

The norm we consider is the max-norm. The derivative F' is given by

$$F'(u)v(s) = v(s) - \lambda \left(1 + \frac{1}{n}\right) \int_a^b G(s,t) u(t)^{1/n} v(t) dt$$

for all $v \in \Omega$.

First of all, we notice that F' does not satisfy a Lipschitz-type condition in Ω. Let us consider, for instance, $[a, b] = [0, 1]$, $G(s, t) = 1$ and $y(t) = 0$. Then $F'(y)v(s) = v(s)$ and

$$\|F'(x) - F'(y)\| = |\lambda| \left(1 + \frac{1}{n}\right) \int_a^b x(t)^{1/n} dt.$$

If F' were a Lipschitz function, then we have

$$\|F'(x) - F'(y)\| \leq L_1 \|x - y\|,$$

or, equivalently, the inequality

$$\int_0^1 x(t)^{1/n} dt \leq L_2 \max_{x \in [0,1]} x(s) \tag{3.4}$$

would hold for all $x \in \Omega$ and for a constant L_2. But this is not true. Consider, for example, the functions

$$x_j(t) = \frac{t}{j}$$

for each $j \geq 1$ and $t \in [0, 1]$.

If these are substituted into (3.4), then we have

$$\frac{1}{j^{1/n}(1+1/n)} \leq \frac{L_2}{j} \iff j^{1-1/n} \leq L_2(1+1/n)$$

for each $j \geq 1$. This inequality is not true when $j \to \infty$. Therefore, the condition (3.4) is not satisfied in this case. Hence Assumption 1.1 is not satisfied. However, the central Lischitz condition in Assumption 1.2 holds. To show this, let $x_0(t) = f(t)$, $\gamma = \min_{s\in[a,b]} f(s)$ and $\alpha > 0$. Then, for any $v \in \Omega$, we have

$$\begin{aligned}\|[F'(x) - F'(x_0)]v\| &= |\lambda|\Big(1+\frac{1}{n}\Big)\max_{s\in[a,b]}\Big|\int_a^b G(s,t)(x(t)^{1/n} - f(t)^{1/n})v(t)dt\Big| \\ &\leq |\lambda|\Big(1+\frac{1}{n}\Big)\max_{s\in[a,b]} G_n(s,t),\end{aligned}$$

where

$$G_n(s,t) = \frac{G(s,t)|x(t) - f(t)|}{x(t)^{(n-1)/n} + x(t)^{(n-2)/n}f(t)^{1/n} + \cdots + f(t)^{(n-1)/n}}\|v\|.$$

Hence we have

$$\begin{aligned}\|[F'(x) - F'(x_0)]v\| &= \frac{|\lambda|(1+1/n)}{\gamma^{(n-1)/n}}\max_{s\in[a,b]}\int_a^b G(s,t)dt\|x - x_0\| \\ &\leq \overline{K}_0\|x - x_0\|,\end{aligned}$$

where

$$\overline{K}_0 = \frac{|\lambda|(1+1/n)}{\gamma^{(n-1)/n}}N, \quad N = \max_{s\in[a,b]}\int_a^b G(s,t)dt.$$

Then Assumption 1.2 holds for sufficiently small λ.

In the last example, we show that $\frac{K_0}{\overline{K}_0}$ can be arbitrarily large in certain nonlinear equation.

Example 3.3 Let $X = D(F) = \mathbb{R}$, $x_0 = 0$ and define a function F on $D(F)$ by

$$F(x) = d_0x + d_1 + d_2\sin e^{d_3x}, \tag{3.5}$$

where d_i for each $i = 0, 1, 2, 3$ are given parameters. Then it can easily be seen that, for d_3 sufficiently large and d_2 sufficiently small, $\frac{K_0}{\overline{K}_0}$ can be arbitrarily large.

Acknowledgements The research work was supported by the National Natural Science Foundation of China (11771067) and the Applied Basic Project of Sichuan Province (2019YJ0204).

References

1. Argyros, I.K.: Convergence and Application of Newton-type Iterations. Springer, Berlin (2008)
2. Argyros, I.K.: Approximating solutions of equations using Newton's method with a modified Newton's method iterate as a starting point. Rev. Anal. Numer. Theor. Approx. **36**, 123–138 (2007)
3. Argyros, I.K.: A Semilocal convergence for directional Newton methods. Math. Comput. (AMS) **80**, 327–343 (2011)
4. Argyros, I.K., Hilout, S.: Weaker conditions for the convergence of Newton's method. J. Complex. **28**, 364–387 (2012)
5. Argyros, I.K., Cho, Y.J., Hilout, S.: Numerical Methods for Equations and its Applications. CRC Press, Taylor and Francis, New York (2012)
6. Engl, H.W., Hanke, M., Neubauer, A.: Tikhonov regularization of nonlinear differential equations. In: Sabatier, P.C. (ed.) Inverse Methods in Action, pp. 92–105. Springer, New York (1990)
7. Engl, H.W., Kunisch, K., Neubauer, A.: Convrgence rates for Tikhonov regularization of nonlinear ill-posed problems. Inverse Probl. **5**, 523–540 (1989)
8. Gfrerer, H.: An a posteriori parameter choice for ordinary and iterated Tikhonov regularization leading to optimal convergence rates. Math. Comput. **49**, 507–522 (1987)
9. Groetsch, C.W.: Theory of Tikhonov Regularization for Fredholm Equation of the First Kind. Pitmann Books, London (1984)
10. Groetsch, C.W., Guacaneme, J.E.: Arcangeli's method for Fredhom equations of the first kind. Proc. Am. Math. Soc. **99**, 256–260 (1987)
11. George, S., Nair, M.T.: An a posteriori parameter choice for simplified regularization of ill-posed problems. Integral Equ. Oper. Theory **16**, 392–399 (1993)
12. George, S., Nair, M.T.: On a generalized Arcangeli's method for Tikhonov regularization with inexact data. Numer. Funct. Anal. Optim. **19**, 773–787 (1998)
13. Jin, Q.N., Hou, Z.Y.: On an a posteriori parameter choice strategy for Tikhonov regularization of nonlinear ill-posed problems. Numer. Math. **83**, 139–159 (1999)
14. Mair, B.A.: Tikhonov regularization for finitely and infinitely smoothing operators. SIAM J. Math. Anal. **25**, 135–147 (1994)
15. Nair, M.T., Ravishankar, P.: Regularized versions of continuous Newton's method and continuous modified Newton's method under general source conditions. Numer. Funct. Anal. Optim. **29**, 1140–1165 (2008)
16. Pereverzev, S., Schock, E.: On the adaptive selection of the parameter in regularization of ill-posed problems. SIAM J. Numer. Anal. **43**, 2060–2076 (2005)
17. Raus, T.: On the discrepancy principle for the solution of ill-posed problems. Acta Comment. Univ. Tartuensis **672**, 16–26 (1984)
18. Scherzer, O.: A parameter choice for Tikhonov regularization for solving nonlinear inverse problems leading to optimal rates. Appl. Math. **38**, 479–487 (1993)
19. Scherzer, O., Engl, H.W., Kunisch, K.: Optimal a posteriori parameter choice for Tikhonov regularization for solving nonlinear ill-posed problems. SIAM J. Numer. Anal. **30**, 1796–1838 (1993)
20. Scherzer, O.: The use of Tikhonov regularization in the identification of electrical conductivities from overdetermined problems. Inverse Probl. **5**, 227–238 (1989)
21. Schock, E.: On the asymptotic order of accuracy of Tikhonov regularization. J. Optim. Theory Appl. **44**, 95–104 (1984)