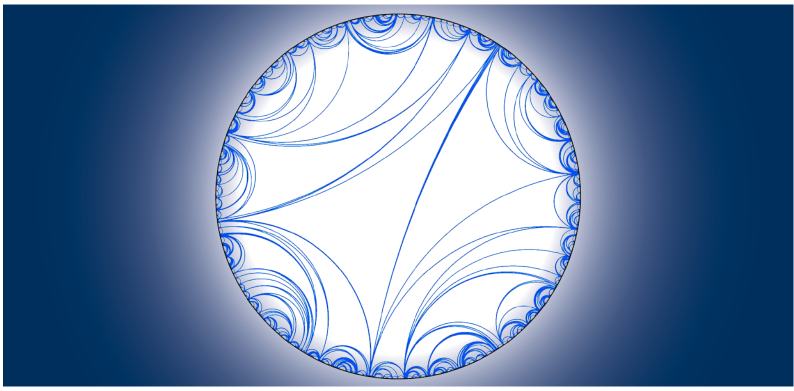
Universität Potsdam



Pornsarp Pornsawad | Christine Böckmann

Modified iterative Runge-Kutta-type methods for nonlinear ill-posed problems

Preprints des Instituts für Mathematik der Universität Potsdam
3 (2014) 7

Preprints des Instituts für Mathematik der Universität Potsdam

Preprints des Instituts für Mathematik der Universität Potsdam
3 (2014) 7

Pornsarp Pornsawad | Christine Böckmann

Modified iterative Runge-Kutta-type methods for nonlinear ill-posed problems

Universitätsverlag Potsdam

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.dnb.de abrufbar.

Universitätsverlag Potsdam 2014
http://verlag.ub.uni-potsdam.de/

Am Neuen Palais 10, 14469 Potsdam
Tel.: +49 (0)331 977 2533 / Fax: 2292
E-Mail: verlag@uni-potsdam.de

Die Schriftenreihe **Preprints des Instituts für Mathematik der Universität Potsdam** wird herausgegeben vom Institut für Mathematik der Universität Potsdam.

ISSN (online) 2193-6943

Kontakt:
Institut für Mathematik
Am Neuen Palais 10
14469 Potsdam
Tel.: +49 (0)331 977 1028
WWW: http://www.math.uni-potsdam.de

Titelabbildungen:
1. Karla Fritze | Institutsgebäude auf dem Campus Neues Palais
2. Nicolas Curien, Wendelin Werner | Random hyperbolic triangulation
Published at: http://arxiv.org/abs/1105.5089

Online veröffentlicht auf dem Publikationsserver der Universität Potsdam
URL http://pub.ub.uni-potsdam.de/volltexte/2014/7083/
URN urn:nbn:de:kobv:517-opus-70834
http://nbn-resolving.de/urn:nbn:de:kobv:517-opus-70834

Modified Iterative Runge-Kutta-Type Methods for Nonlinear Ill-Posed Problems

Pornsarp Pornsawad[a,*], Christine Böckmann[b]

[a] *Department of Mathematics, Faculty of Science, Silpakorn University, Mueng, Nakorn Pathom, 73000, Thailand*
[b] *Institute of Mathematics, University of Potsdam, Am Neuen Palais 10, 14469 Potsdam, Germany*

Abstract

This work is devoted to the convergence analysis of a modified Runge-Kutta-type iterative regularization method for solving nonlinear ill-posed problems under a priori and a posteriori stopping rules. The convergence rate results of the proposed method can be obtained under Hölder-type sourcewise condition if the Fréchet derivative is properly scaled and locally Lipschitz continuous. Numerical results are achieved by using the Levenberg-Marquardt and Radau methods.

Keywords: Nonlinear ill-posed problems, Runge-Kutta methods, Regularization methods, Hölder-type source condition, Stopping rules
AMS Classification: *65J15, 65J22, 47J25, 47J06*

---

1. Introduction

Let X and Y be infinite-dimensional real Hilbert spaces with inner products $\langle \cdot, \cdot \rangle$ and norms $\| \cdot \|$. We consider a Fréchet-differential nonlinear ill-posed operator equation

$$F(w) = g \tag{1}$$

where $F : \mathcal{D}(F) \subset X \to Y$ with a non-closed range $\mathcal{R}(F)$ and a locally uniformly bounded Fréchet derivative $F'(\cdot)$ of F in $\mathcal{D}(F)$. We assume that

*Corresponding author
Email addresses: pornsarp@su.ac.th (Pornsarp Pornsawad), bockmann@rz.uni-potsdam.de (Christine Böckmann)

(1) has a solution w_* for exact data (which need not be unique). Therefore, the element $w_0 \in X$, which is an initial guess, is assumed to be known. We have approximate data g^ε with

$$\|g^\varepsilon - g\|_Y \leq \varepsilon. \tag{2}$$

A family of iterative Runge-Kutta (RK)-type regularization methods for nonlinear ill-posed problems of the form

$$w_{k+1}^\varepsilon = w_k^\varepsilon + \tau_k b^T(\boldsymbol{\delta} + \tau_k A F'(w_k^\varepsilon)^* F'(w_k^\varepsilon))^{-1} \mathbf{1} F'(w_k^\varepsilon)^*(g^\varepsilon - F(w_k^\varepsilon)) \tag{3}$$

has been developed in [5] where the notation $\mathbf{1}$ means the (s×1) vector of identity operators and $\boldsymbol{\delta}$ is the (s×s) diagonal matrix of bounded linear operators with identity operator on the entire diagonal and zero operator outside of the main diagonal with respect to the appropriate spaces. The idea is founded on the asymptotic regularization [24] where an initial value problem

$$\frac{d}{dt}w(t) = F'(w_k^\varepsilon)^*(g^\varepsilon - F(w_k^\varepsilon) + F'(w_k^\varepsilon)(w_k^\varepsilon - w(t))),\ t > 0,\ w(0) = w_k^\varepsilon,$$

has to be solved and, therefore, RK methods are used. The parameter τ_k in (3) is the steplength or also named the relaxation parameter. The (s ×s) matrix A and the (s×1) vectors b are the given parameters which are corresponding to the specific Runge-Kutta method. Whereas in [5] the convergence of the whole family of RK-type regularization methods for nonlinear problems is investigated in [19] the convergence rate analysis for the first stage family is developed. For linear problems those methods are presented in [4, 15, 22]. Both studies for the nonlinear case deal only with an a posteriori parameter choice namely the discrepancy principle.

In [23] the following modification of the Landweber iteration is proposed

$$w_{k+1}^\varepsilon = w_k^\varepsilon - F'(w_k^\varepsilon)^*(F(w_k^\varepsilon) - g^\varepsilon) - \alpha_k(w_k^\varepsilon - \zeta).$$

The additional term $\alpha_k(w_k^\varepsilon - \zeta)$ compared to the classical Landweber iteration is motivated by the iteratively regularized Gauss-Newton method, see e.g. [1, 2, 3, 9],

$$w_{k+1}^\varepsilon = w_k^\varepsilon + (F'(w_k^\varepsilon)^* F'(w_k^\varepsilon) + \alpha_k I)^{-1}(F'(w_k^\varepsilon)^*(g^\varepsilon - F(w_k^\varepsilon)) + \alpha_k(w_0 - w_k^\varepsilon)) \tag{4}$$

where w_0 is an initial guess for the true solution. In [9] this method is investigated under merely the Lipschitz condition of the Fréchet derivative of F using an a posteriori stopping rule. In order to avoid saturation for the convergence rate under Hölder source condition an alternative rule is used

$$\alpha_{k_\varepsilon}\langle F(w^\varepsilon_{k_\varepsilon}) - g^\varepsilon, (F'(w^\varepsilon_{k_\varepsilon})^* F'(w^\varepsilon_{k_\varepsilon}) + \alpha_{k_\varepsilon} I)^{-1}(F(w^\varepsilon_{k_\varepsilon}) - g^\varepsilon)\rangle \leq \lambda^2\varepsilon^2 \ , \lambda > 1,$$

where the termination index k_ε is the first integer fulfilling the above inequality and $\lambda > 1$ is a large number. In [8] the method (4) is examined under logarithmic source conditions. Moreover, in [7] the frozen method of (4), i.e., without updating the Fréchet derivative is regarded using the balancing principle [18] under certain general source conditions.

Different Newton-type methods were studied a lot during the last two decades, see e.g. [10, 13, 21], describing the strategy of outer and inner method. In [10] the method

$$w^\varepsilon_{k+1} = w^\varepsilon_k + g_{\alpha_k}(F'(w^\varepsilon_k)^* F'(w^\varepsilon_k))(F'(w^\varepsilon_k)^*(F(w^\varepsilon_k) - g^\varepsilon)$$

using discrepancy principle for several spectral filter functions $\{g_\alpha\}$ is regarded. Order optimal convergence rate under certain Hölder source conditions is achieved.

General iteratively regularized Gauss-Newton methods

$$w^\varepsilon_{k+1} = w^\varepsilon_k + g_{\alpha_k}(F'(w^\varepsilon_k)^* F'(w^\varepsilon_k))(F'(w^\varepsilon_k)^*(F(w^\varepsilon_k) - g^\varepsilon - F'(w^\varepsilon_k)(w^\varepsilon_k - w_0))$$

are examined in [2, 11, 12, 14, 16]. In [12, 14] this method is investigated under local Lipschitz condition of the Fréchet derivative for Hölder and logarithmic source conditions whereas [16] extended this method to work with a general source condition with an index function φ [17]. Error estimates are given under a priori and a posteriori parameter choices. Under weaker conditions convergence results could be obtained in [11].

In this paper we investigate a modification of (3) as follows

$$w^\varepsilon_{k+1} = w^\varepsilon_k + \tau_k b^T \Pi^{-1}\mathbf{1} F'(w^\varepsilon_k)^*(g^\varepsilon - F(w^\varepsilon_k)) - \tau_k^{-1}(w^\varepsilon_k - \zeta). \tag{5}$$

where Π^{-1} stands for $(\boldsymbol{\delta}+\tau_k A F'(w^\varepsilon_k)^* F'(w^\varepsilon_k))^{-1}$. For the convergence analysis of the modified Runge-Kutta-type regularization (5) the parameter $\alpha_k = 1/\tau_k$ satisfy

$$0 \leq \alpha_k \leq 1. \tag{6}$$

The condition (6) yields useful result [23], i.e., for $l, k \in \mathbb{N}$ and $l < k$

$$1 - \prod_{s=l}^{k}(1-\alpha_s) = \sum_{j=l}^{k}\alpha_j \prod_{s=j+1}^{k}(1-\alpha_s) \leq 1. \tag{7}$$

In addition if

$$\sum_{k=0}^{\infty} \alpha_k < \infty \tag{8}$$

the product $\prod_{k=0}^{\infty}(1-\alpha_k)$ converges and $\prod_{k=l}^{\infty}(1-\alpha_k)$ tends to 1 as l tends to infinity. The condition (8) provides the finite stopping index for $\varepsilon > 0$ and is necessary for the convergence analysis in the noise free case.

This paper is organized as follows. In the next section we investigate the convergence analysis of the proposed modification of the RK-type method (5) under a priori and a posteriori parameter choices. In Section 3 the convergence rate is given for both parameter choices and in Section 4 we show a numerical example and compare different methods.

2. Convergence Analysis

In this section a convergence analysis of the modified Runge-Kutta-type regularization (5) is provided. Therefore the local property of the nonlinear operator in a ball $B_\rho(w_0)$,

$$\|F(\tilde{w}) - F(w) - F'(w)(\tilde{w} - w)\|_Y \le \eta \|F(\tilde{w}) - F(w)\|_Y \tag{9}$$

for $\tilde{w}, w \in B_\rho(w_0) \subset \mathcal{D}(F)$ with $\eta < 1/2$ is required. It is also useful to deduce from (9) that

$$\frac{1}{\eta+1}\|F'(w)(w-\tilde{w})\|_Y \le \|F(w) - F(\tilde{w})\|_Y \le \frac{1}{1-\eta}\|F'(w)(w-\tilde{w})\|_Y \tag{10}$$

hold for all $w, \tilde{w} \in B_\rho(w_0)$. Due to the boundedness of the operator $b^T H_k^{-1}\mathbf{1}$ and $b^T(A-I)S_k^* H_k^{-1}\mathbf{1}$ with the (s×s) identity matrix I, $S_k = F'(w_k)$, $S_k^* = F'(w_k)^*$ and $H_k^{-1} = (\boldsymbol{\delta} + \tau_k A S_k S_k^*)^{-1}$ there exist constant $c_1, c_2 > 0$ with

$$\frac{c_1}{\sqrt{2}}\|y\|_Y \le \|b^T H_k^{-1}\mathbf{1}y\|_Y \le c_1\|y\|_Y \tag{11}$$

and

$$\|b^T(A-I)S_k^* H_k^{-1}\mathbf{1}y\|_X \le c_2\|y\|_Y. \tag{12}$$

In addition the Fréchet derivative $F'(\cdot)$ of F is bounded in a ball $B_\rho(w_0)$, i.e.,

$$\|F'(w)\| \le L. \tag{13}$$

The approximations (11) and (12) are used to prove that the iterative regularization method (3) converges to a solution. In the presented work the iteration will be terminated by a posteriori and a priori stopping criterion, i.e.,

a posteriori stopping criterion

$$\|g^\varepsilon - F(w^\varepsilon_{k_*})\|_Y \le \lambda\varepsilon < \|g^\varepsilon - F(w^\varepsilon_k)\|_Y, \quad 0 \le k < k_* \tag{14}$$

with

$$\lambda \ge \frac{2c_1(\eta+1)(\tau_k + c_\zeta)}{(\tau_k - 1)(c_1^2 - 2c_1\eta - 2c_1c_2\tau_k L) - 4c_1(\eta+1)} > 0 \tag{15}$$

for some positive number c_ζ.

a priori stopping criterion

$$C\tau_{N_0}^{-1} < \varepsilon \le C\tau_k^{-1}, \quad 0 \le k < N_0, \quad \text{where } C \le \rho/6. \tag{16}$$

Proposition 1. Let w_* be a solution of (1) in $B_{\rho/8}(w_0) \cap B_{\rho/8}(\zeta)$ with $w_0 = w_0^\varepsilon$. Assume that the assumptions (9)-(13) and for some positive number c_ζ the assumption

$$\|F(\zeta) - g^\varepsilon\|_Y \le c_\zeta \varepsilon \tag{17}$$

hold. If $\tau_k \ge 2$ and the termination index k_* is defined by (14) with

$$(\tau_k - 1)(c_1^2 - 2c_1\eta - 2c_1c_2\tau_k L) - 4c_1(\eta+1) - 2c_1(\eta+1)(\tau_k + c_\zeta)/\lambda > 0 \tag{18}$$

holds then $w^\varepsilon_{k+1} \in B_\rho(w_0)$ and

$$\|w_* - w^\varepsilon_{k+1}\|_X \le (1 - \tau_k^{-1})\|w_* - w^\varepsilon_k\|_X + \tau_k^{-1}\rho/8. \tag{19}$$

Moreover the termination index k_* in (14) is finite if $\sum_{k=0}^\infty \tau_k^{-1} < \infty$.

Proof. Let $0 \le k < k_*$. For short we use $f(t) := b^T(I - At)^{-1}\mathbf{1}$, see Appendix for more detail. Using (5) we can show that

$$\begin{aligned}
&\|w^\varepsilon_{k+1} - w_*\|_X^2 \\
&\quad = (1 - \tau_k^{-1})^2\|w^\varepsilon_k - w_*\|_X^2 + \|\tau_k f(-\tau_k S_k^* S_k) F'(w^\varepsilon_k)^*(F(w^\varepsilon_k) - g^\varepsilon)\|_X^2 \\
&\qquad + (\tau_k^{-1})^2\|w_* - \zeta\|_X^2 - 2(1-\tau_k^{-1})\tau_k^{-1}\langle w^\varepsilon_k - w_*, w_* - \zeta\rangle_X \\
&\qquad - 2(1-\tau_k^{-1})\langle w^\varepsilon_k - w_*, \tau_k f(-\tau_k S_k^* S_k) F'(w^\varepsilon_k)^*(F(w^\varepsilon_k) - g^\varepsilon)\rangle_X \\
&\qquad + 2\tau_k^{-1}\langle w_* - \zeta, \tau_k f(-\tau_k S_k^* S_k) F'(w^\varepsilon_k)^*(F(w^\varepsilon_k) - g^\varepsilon)\rangle_X.
\end{aligned} \tag{20}$$

Using (2), (9), (11)-(13) and the consistency property of the RK method, $b^T H_k H_k^{-1}\mathbf{1}y = b^T\mathbf{1}y = y\sum_{i=1}^{s} b_i = y$ for $y \in Y$ the fifth term in (20) becomes

$$\begin{aligned}
&-2(\tau_k-1)\langle w_k^\varepsilon - w_*, f(-\tau_k S_k^* S_k)F'(w_k^\varepsilon)^*(F(w_k^\varepsilon)-g^\varepsilon)\rangle_X \\
&\quad = 2(\tau_k-1)\langle F(w_k^\varepsilon)-g^\varepsilon - F'(w_k^\varepsilon)(w_k^\varepsilon - w_*), f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon)-g^\varepsilon)\rangle_X \\
&\qquad -2(\tau_k-1)\langle F(w_k^\varepsilon)-g^\varepsilon, f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon)-g^\varepsilon)\rangle_X \\
&\quad \le 2(\tau_k-1)\|F(w_k^\varepsilon)-g^\varepsilon - F'(w_k^\varepsilon)(w_k^\varepsilon - w_*)\|_Y \|f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon)-g^\varepsilon)\|_Y \\
&\qquad -2(\tau_k-1)\langle b^T(\boldsymbol{\delta} + \tau_k A S_k S_k^*)H_k^{-1}\mathbf{1}(F(w_k^\varepsilon)-g^\varepsilon), f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon)-g^\varepsilon)\rangle_X \\
&\quad \le 2c_1(\tau_k-1)(\varepsilon + \eta\varepsilon + \eta\|g^\varepsilon - F(w_k^\varepsilon)\|_Y)\|g^\varepsilon - F(w_k^\varepsilon)\|_Y \\
&\qquad -2(\tau_k-1)\|f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon)-g^\varepsilon)\|_Y^2 \\
&\qquad -2(\tau_k-1)\tau_k\langle S_k^* b^T(A-I)H_k^{-1}\mathbf{1}(F(w_k^\varepsilon)-g^\varepsilon), S_k^* f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon)-g^\varepsilon)\rangle_X \\
&\qquad -2(\tau_k-1)\tau_k\langle S_k^* f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon)-g^\varepsilon), S_k^* f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon)-g^\varepsilon)\rangle_X \\
&\quad \le 2c_1(\tau_k-1)(\varepsilon(1+\eta) + \eta\|g^\varepsilon - F(w_k^\varepsilon)\|_Y)\|g^\varepsilon - F(w_k^\varepsilon)\|_Y \\
&\qquad -(\tau_k-1)c_1^2\|F(w_k^\varepsilon)-g^\varepsilon\|_Y^2 + 2(\tau_k-1)\tau_k c_1 c_2 L\|F(w_k^\varepsilon)-g^\varepsilon\|_Y^2 \\
&\qquad -2(\tau_k-1)\tau_k\|S_k^* f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon)-g^\varepsilon)\|_X^2. \qquad (21)
\end{aligned}$$

Using (2), (9), (11) and (17) we can show that the last term in (20) becomes

$$\begin{aligned}
&2\langle w_* - \zeta, f(-\tau_k S_k^* S_k)F'(w_k^\varepsilon)^*(F(w_k^\varepsilon)-g^\varepsilon)\rangle_X \\
&\quad = 2\langle F'(w_k^\varepsilon)(w_* - w_k^\varepsilon), f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon)-g^\varepsilon)\rangle_Y \\
&\qquad +2\langle F'(w_k^\varepsilon)(w_k^\varepsilon - \zeta), f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon)-g^\varepsilon)\rangle_Y \\
&\quad = -2\langle F(w_*) - F(w_k^\varepsilon) - F'(w_k^\varepsilon)(w_* - w_k^\varepsilon), f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon)-g^\varepsilon)\rangle_Y \\
&\qquad +2\langle F(w_*) - F(w_k^\varepsilon), f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon)-g^\varepsilon)\rangle_Y \\
&\qquad +2\langle F(\zeta) - F(w_k^\varepsilon) - F'(w_k^\varepsilon)(\zeta - w_k^\varepsilon), f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon)-g^\varepsilon)\rangle_Y \\
&\qquad -2\langle F(\zeta) - F(w_k^\varepsilon), f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon)-g^\varepsilon)\rangle_Y \\
&\quad \le 2c_1(\eta+1)\|g^\varepsilon - F(w_k^\varepsilon)\|_Y\left[(c_\zeta+1)\varepsilon + 2\|g^\varepsilon - F(w_k^\varepsilon)\|_Y\right]. \qquad (22)
\end{aligned}$$

Applying (21) and (22) into (20) we get

$$\begin{aligned}
&\|w_{k+1}^\varepsilon - w_*\|_X^2 \\
&\quad \le (1-\tau_k^{-1})^2\|w_k^\varepsilon - w_*\|_X^2 + \|\tau_k f(-\tau_k S_k^* S_k)S_k^*(F(w_k^\varepsilon)-g^\varepsilon)\|_X^2 \\
&\qquad +(\tau_k^{-1})^2\|w_* - \zeta\|_X^2 + 2(1-\tau_k^{-1})\tau_k^{-1}\|w_k^\varepsilon - w_*\|_X\|w_* - \zeta\|_X \\
&\qquad +2c_1(\tau_k-1)(\varepsilon(1+\eta) + \eta\|g^\varepsilon - F(w_k^\varepsilon)\|_Y)\|g^\varepsilon - F(w_k^\varepsilon)\|_Y \\
&\qquad -(\tau_k-1)c_1^2\|F(w_k^\varepsilon)-g^\varepsilon\|_Y^2 + 2(\tau_k-1)\tau_k c_1 c_2 L\|F(w_k^\varepsilon)-g^\varepsilon\|_Y^2 \\
&\qquad -2(\tau_k-1)\tau_k\|S_k^* f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon)-g^\varepsilon)\|_X^2 \\
&\qquad +2c_1(\eta+1)\|g^\varepsilon - F(w_k^\varepsilon)\|_Y[(c_\zeta+1)\varepsilon + 2\|g^\varepsilon - F(w_k^\varepsilon)\|_Y]. \qquad (23)
\end{aligned}$$

It follows from the fact that $\tau_k \geq 2$ for all k, (23) becomes

$$\begin{aligned}
&\|w^\varepsilon_{k+1} - w_*\|^2_X \\
&\quad \leq \ ((1-\tau_k^{-1})\|w^\varepsilon_k - w_*\|_X + \tau_k^{-1}\|w_* - \zeta\|_X)^2 \\
&\qquad + \big[\big((\tau_k - 1)(2c_1\eta - c_1^2 + 2\tau_k c_1 c_2 L) + 4c_1(1+\eta)\big)\, \|g^\varepsilon - F(w^\varepsilon_k)\|_Y \\
&\qquad +2c_1(1+\eta)(\tau_k + c_\zeta)\varepsilon\big]\, \|g^\varepsilon - F(w^\varepsilon_k)\|_Y. \qquad (24)
\end{aligned}$$

Thus for the discrepancy principle (14) it follows from (24) that

$$\begin{aligned}
\|w^\varepsilon_{k+1} - w_*\|^2_X \ &\leq \ ((1-\tau_k^{-1})\|w^\varepsilon_k - w_*\|_X + \tau_k^{-1}\|w_* - \zeta\|_X)^2 \\
&\quad + \frac{2}{\lambda} c_1(\eta+1)(\tau_k + c_\zeta)[\lambda\varepsilon - \|F(w^\varepsilon_k) - g^\varepsilon\|_Y]\|F(w^\varepsilon_k) - g^\varepsilon\|_Y \\
&\leq \ ((1-\tau_k^{-1})\|w^\varepsilon_k - w_*\|_X + \tau_k^{-1}\|w_* - \zeta\|_X)^2. \qquad (25)
\end{aligned}$$

The equation (25) together with the assumption that w_* is a solution in $B_{\rho/8}(w_0) \cap B_{\rho/8}(\zeta)$ we get

$$\|w^\varepsilon_{k+1} - w_*\|_X \ \leq \ (1-\tau_k^{-1})\|w^\varepsilon_k - w_*\|_X + \tau_k^{-1}(\rho/8). \qquad (26)$$

From (26) it follows by induction for $0 \leq k < k_*$ that

$$\|w^\varepsilon_{k+1} - w_*\|_X \leq \|w_0 - w_*\|_X \prod_{j=0}^{k}(1-\tau_j^{-1}) + \rho/8 \sum_{j=0}^{k} \tau_j^{-1} \prod_{i=j+1}^{k}(1-\tau_i^{-1}). \quad (27)$$

It follows from the approximation (7) that

$$\|w^\varepsilon_{k+1} - w_*\|_X \ \leq \ \|w_0 - w_*\|_X(1) + (\rho/8)(1) \leq \rho/4$$

holds and thus $\|w^\varepsilon_{k+1} - w_0\|_X \leq \|w^\varepsilon_{k+1} - w_*\|_X + \|w_* - w_0\|_X \leq \rho/4 + \rho/8 < \rho$ which shows $w^\varepsilon_{k+1} \in B_\rho(w_0)$.
Therefore, if the iteration (5) is terminated by (14), we get $\|w^\varepsilon_k - w_*\|_X \leq \rho/4$. Thus

$$2(1-\tau_k^{-1})\tau_k^{-1}\|w^\varepsilon_k - w_*\|_X\|w_* - \zeta\|_X \leq 2\tau_k^{-1}(\rho/4)(\rho/8) = \tau_k^{-1}\rho^2/16 \qquad (28)$$

and

$$(\tau_k^{-1})^2\|w_* - \zeta\|^2_X \leq (\tau_k^{-1})^2(\rho/8)^2 \leq \tau_k^{-1}\rho^2/64 \qquad (29)$$

are obtained. Applying (14) to (24) yields

$$\begin{aligned}
&\|w^\varepsilon_{k+1} - w_*\|^2_X \\
&\quad \le ((1-\tau_k^{-1})\|w^\varepsilon_k - w_*\|_X + \tau_k^{-1}\|w_* - \zeta\|_X)^2 \\
&\qquad +\|g^\varepsilon - F(w^\varepsilon_k)\|^2_Y \left[4c_1(1+\eta) - (\tau_k - 1)(c_1^2 - 2c_1\eta - 2\tau_k c_1 c_2 L)\right. \\
&\qquad \left. +2c_1(1+\eta)(\tau_k + c_\zeta)/\lambda\right] \\
&\quad \le ((1-\tau_k^{-1})\|w^\varepsilon_k - w_*\|_X + \tau_k^{-1}\|w_* - \zeta\|_X)^2 - D\|F(w^\varepsilon_k) - g^\varepsilon\|^2_Y \qquad (30)
\end{aligned}$$

for some positive number D with $0 < D \le (\tau_k - 1)(c_1^2 - 2c_1\eta - 2\tau_k c_1 c_2 L) - 4c_1(1+\eta) - 2c_1(1+\eta)(\tau_k + c_\zeta)/\lambda$. Next the estimations (28)-(29) are applied in (30) to show that

$$\begin{aligned}
&\|w^\varepsilon_{k+1} - w_*\|^2_X + D\|F(w^\varepsilon_k) - g^\varepsilon\|^2_Y \\
&\quad \le \ (1-\tau_k^{-1})^2\|w^\varepsilon_k - w_*\|^2_X + (\tau_k^{-1})^2\|w_* - \zeta\|^2_X \\
&\qquad +2(1-\tau_k^{-1})\tau_k^{-1}\|w^\varepsilon_k - w_*\|_X\|w_* - \zeta\|_X \\
&\quad < \ \|w^\varepsilon_k - w_*\|^2_X + \tau_k^{-1}\rho^2. \qquad (31)
\end{aligned}$$

With the fact that $\sum_{k=0}^{k_*-1}(\|w^\varepsilon_k - w_*\|^2_X - \|w^\varepsilon_{k+1} - w_*\|^2_X) \le \|w_0 - w_*\|^2_X$, the equation (31) provides

$$D\sum_{k=0}^{k_*-1}\|F(w^\varepsilon_k) - g^\varepsilon\|^2_Y \le \|w_0 - w_*\|^2_X + \rho^2\sum_{k=0}^{k_*-1}\tau_k^{-1}.$$

Consequently,

$$D\sum_{k=0}^{k_*-1}\|F(w^\varepsilon_k) - g^\varepsilon\|^2_Y \le \rho^2(1/64 + \sum_{k=0}^{k_*-1}\tau_k^{-1}). \qquad (32)$$

On the other hand, the discrepancy principle (14) yields

$$(\lambda\varepsilon)^2 < \|g^\varepsilon - F(w^\varepsilon_k)\|^2_Y, \quad 0 \le k < k_*. \qquad (33)$$

Adding (33) for $k = 0$ to $k = k_* - 1$, we get

$$(\lambda\varepsilon)^2 k_* < \sum_{k=0}^{k_*-1}\|g^\varepsilon - F(w^\varepsilon_k)\|^2_Y. \qquad (34)$$

Combining (32) and (34) the assertion is obtained. □

Proposition 2. Let the assumptions of Proposition 1 and

$$\max\{c_1 L(\eta+1)(\tau_k + c_\zeta), c_1(\eta+1)(\tau_k + c_\zeta)\} < 9/16 \qquad (35)$$

hold. If the termination index N_0 is chosen according to (16) and

$$(\tau_k - 1)(c_1^2 - 2c_1\eta - 2c_1c_2\tau_k L) - 4c_1(\eta+1) > E > 0 \tag{36}$$

holds for some positive E then $w_{k+1}^\varepsilon \in B_\rho(w_0)$ and

$$\|w_* - w_{k+1}^\varepsilon\|_X \leq (1-\tau_k^{-1})\|w_* - w_k^\varepsilon\|_X + \tau_k^{-1}p \tag{37}$$

for some positive number p with $p < \rho/2$. Moreover

$$E\sum_{k=0}^{N_0} \|F(w_k^\varepsilon) - g^\varepsilon\|_Y^2 \leq \rho^2(1/64 + \sum_{k=0}^{N_0}\tau_k^{-1}). \tag{38}$$

Proof. Let $0 \leq k \leq N_0$. Due to (2) and (13) it follows that

$$\|F(w_k^\varepsilon) - g^\varepsilon\|_Y \leq \varepsilon + \frac{1}{1-\eta}\|F'(w_k^\varepsilon)(w_k^\varepsilon - w_*)\|_X \leq \varepsilon + 2L\|w_k^\varepsilon - w_*\|_X. \tag{39}$$

Applying the condition (16), i.e. $\varepsilon \leq C\tau_k^{-1}$, and the estimation (39) to (24), we get

$$\begin{aligned}
&\|w_{k+1}^\varepsilon - w_*\|_X^2 \\
&\quad\leq (1-\tau_k^{-1})^2\|w_k^\varepsilon - w_*\|_X^2 \\
&\qquad + \big((\tau_k - 1)(2c_1\eta - c_1^2 + 2\tau_k c_1c_2L) + 4c_1(\eta+1)\big)\,\|g^\varepsilon - F(w_k^\varepsilon)\|_Y^2 \\
&\qquad + 2(1-\tau_k^{-1})\tau_k^{-1}\|w_k^\varepsilon - w_*\|_X(\|w_* - \zeta\|_X + \frac{2c_1CL(\eta+1)(\tau_k + c_\zeta)}{1-\tau_k^{-1}}) \\
&\qquad + (\tau_k^{-1})^2(\|w_* - \zeta\|_X^2 + 2c_1(\eta+1)(\tau_k + c_\zeta)C^2).
\end{aligned} \tag{40}$$

Setting

$$p := \max\{\|w_* - \zeta\|_X + \frac{2c_1CL(\eta+1)(\tau_k + c_\zeta)}{1-\tau_k^{-1}}, \sqrt{\|w_* - \zeta\|_X^2 + 2c_1(\eta+1)(\tau_k + c_\zeta)C^2}\}$$

and using (36), (40) becomes

$$\begin{aligned}
&\|w_{k+1}^\varepsilon - w_*\|_X^2 \\
&\quad\leq ((1-\tau_k^{-1})\|w_k^\varepsilon - w_*\|_X + \tau_k^{-1}p)^2 \\
&\qquad - \big((\tau_k - 1)(c_1^2 - 2c_1\eta - 2\tau_k c_1c_2L) - 4c_1(\eta+1)\big)\,\|g^\varepsilon - F(w_k^\varepsilon)\|_Y^2 \\
&\quad\leq ((1-\tau_k^{-1})\|w_k^\varepsilon - w_*\|_X + \tau_k^{-1}p)^2.
\end{aligned}$$

Thus, $$\|w_{k+1}^\varepsilon - w_*\|_X \leq (1-\tau_k^{-1})\|w_k^\varepsilon - w_*\|_X + \tau_k^{-1}p. \tag{41}$$

Note that $p \leq \rho/2$ due to the assumption (35). From (41) it follows by induction for $0 \leq k < N_0$ that

$$\begin{aligned} \|w_{k+1}^\varepsilon - w_*\|_X &\leq \|w_0 - w_*\|_X \prod_{j=0}^{k}(1-\tau_k^{-1}) + p\sum_{j=0}^{k}\tau_j^{-1}\prod_{i=j+1}^{k}(1-\tau_i^{-1}) \\ &\leq \|w_0 - w_*\|_X + p \\ &\leq 5\rho/8. \end{aligned} \tag{42}$$

Therefore,

$$\|w_{k+1}^\varepsilon - w_0\|_X \leq \|w_{k+1}^\varepsilon - w_*\|_X + \|w_* - w_0\|_X \leq 5\rho/8 + \rho/8 < \rho \tag{43}$$

which shows that $w_{k+1}^\varepsilon \in B_\rho(w_0)$. Using (16), (35) and (42) we can show that

$$\begin{aligned} 2\tau_k^{-1}(1-\tau_k^{-1})\|w_k^\varepsilon - w_*\|_X \left(\|w_* - \zeta\|_X + \tfrac{2c_1CL(\eta+1)(\tau_k+c_\zeta)}{1-\tau_k^{-1}}\right)& \\ \leq 5\tau_k^{-1}\rho^2/8 & \end{aligned} \tag{44}$$

and

$$(\tau_k^{-1})^2\left(\|w_* - \zeta\|_X^2 + 2c_1(\eta+1)(\tau_k + c_\zeta)C^2\right) \leq 3(\tau_k^{-1})^2\rho^2/64. \tag{45}$$

Applying (44)-(45) to (40) we get

$$\begin{aligned} &\|w_{k+1}^\varepsilon - w_*\|_X^2 \\ &\quad\leq (1-\tau_k^{-1})^2\|w_k^\varepsilon - w_*\|_X^2 + 5\tau_k^{-1}\rho^2/8 + 3(\tau_k^{-1})^2\rho^2/64 \\ &\qquad + \left((\tau_k - 1)(2c_1\eta - c_1^2 + 2c_1c_2\tau_kL) + 4c_1(\eta+1)\right)\|F(w_k^\varepsilon) - g^\varepsilon\|_Y^2 \\ &\quad< (1-\tau_k^{-1})^2\|w_k^\varepsilon - w_*\|_X^2 + \tau_k^{-1}\rho^2 - E\|F(w_k^\varepsilon) - g^\varepsilon\|_Y^2 \end{aligned} \tag{46}$$

for some positive number E with $0 < E \leq (\tau_k - 1)(c_1^2 - 2c_1\eta - 2c_1c_2\tau_kL) - 4c_1(\eta+1)$. Similar to (32), (46) implies the last assertion, i.e.,

$$E\|F(w_k^\varepsilon) - g^\varepsilon\|_Y^2 \leq \|w_k^\varepsilon - w_*\|_X^2 - \|w_{k+1}^\varepsilon - w_*\|_X^2 + \tau_k^{-1}\rho^2$$

implies

$$E\sum_{k=0}^{N_0}\|F(w_k^\varepsilon) - g^\varepsilon\|_Y^2 \leq \|w_0 - w_*\|_X^2 + \rho^2\sum_{k=0}^{N_0}\tau_k^{-1} \leq \rho^2/64 + \rho^2\sum_{k=0}^{N_0}\tau_k^{-1}.$$

□

We note that the tangential cone condition implies

$$\|F(\zeta) - g^\varepsilon\|_Y \le \|F(\zeta) - F(w_*)\| + \varepsilon \le \frac{1}{1-\eta}\|F'(w_*)(w_* - \zeta)\|_Y + \varepsilon.$$

Using (13) we get

$$\|F(\zeta) - g^\varepsilon\|_Y \le \frac{L}{1-\eta}\|w_* - \zeta\|_X + \varepsilon.$$

If w_* is a solution of (1) in $B_{\rho/8}(w_0) \cap B_{\rho/8}(\zeta)$ and the closeness condition of ζ to the solution w_*, i.e. $\|w_* - \zeta\|_X \le \tilde{c}_\zeta \varepsilon$ for some positive number $\tilde{c}_\zeta$, is satisfied, then the assumption (17) can be achieved.

Proposition 3. If $\sum_{k=0}^{\infty} \tau_k^{-1} < \infty$ and (36) hold for all $k \in \mathbb{N}_0$, then in the noise free case,

$$\sum_{k=0}^{\infty} \|F(w_k) - g\|_Y^2 < \infty. \tag{47}$$

Proof. It follows from (24) that

$$\begin{aligned}
&\|w_{k+1} - w_*\|_X^2 \\
&\quad \le ((1-\tau_k^{-1})\|w_k - w_*\|_X + \tau_k^{-1}\|w_* - \zeta\|_X)^2 \\
&\qquad + \big((\tau_k - 1)(2c_1\eta - c_1^2 + 2\tau_k c_1 c_2 L) + 4c_1(1+\eta)\big)\,\|g - F(w_k)\|_Y^2 \\
&\quad \le ((1-\tau_k^{-1})\|w_k - w_*\|_X + \tau_k^{-1}\|w_* - \zeta\|_X)^2 \\
&\qquad - E\|g - F(w_k)\|_Y^2
\end{aligned} \tag{48}$$

where $E > 0$ is defined by (36). Thus

$$\|w_{k+1} - w_*\|_X \le (1-\tau_k^{-1})\|w_k - w_*\|_X + \tau_k^{-1}\|w_* - \zeta\|_X.$$

Note that w_* is a solution of (1) in $B_{\rho/8}(w_0) \cap B_{\rho/8}(\zeta)$ which yields

$$\|w_{k+1} - w_*\|_X \le (1-\tau_k^{-1})\|w_k - w_*\|_X + \tau_k^{-1}\rho/8.$$

It follows by induction and (7) that

$$\begin{aligned}
\|w_{k+1} - w_*\|_X &\le \|w_0 - w_*\|_X \prod_{j=0}^{k}(1-\tau_j^{-1}) + \rho/8 \sum_{j=0}^{k} \tau_j^{-1} \prod_{i=j+1}^{k} (1-\tau_i^{-1}) \\
&\le \rho/4.
\end{aligned} \tag{49}$$

Using (49) we get

$$(\tau_k^{-1})^2\|w_* - \zeta\|_X^2 + 2(1-\tau_k^{-1})\tau_k^{-1}\|w_k - w_*\|_X\|w_* - \zeta\|_X < \tau_k^{-1}\rho^2. \tag{50}$$

Applying (50) to (48) we can show that

$$\begin{aligned} E\|F(w_k) - g\|_Y^2 & \\ \le\ & (1-\tau_k^{-1})^2\|w_k - w_*\|_X^2 - \|w_{k+1} - w_*\|_X^2 \\ & +(\tau_k^{-1})^2\|w_* - \zeta\|_X^2 + 2(1-\tau_k^{-1})\tau_k^{-1}\|w_k - w_*\|_X\|w_* - \zeta\|_X \\ <\ & (1-\tau_k^{-1})^2\|w_k - w_*\|_X^2 - \|w_{k+1} - w_*\|_X^2 + \tau_k^{-1}\rho^2 \\ <\ & \|w_k - w_*\|_X^2 - \|w_{k+1} - w_*\|_X^2 + \tau_k^{-1}\rho^2. \end{aligned} \tag{51}$$

Similar to (32), (51) implies

$$E\sum_{k=0}^{\infty}\|F(w_k) - g\|_Y^2 < \|w_0 - w_*\|_X^2 + \rho^2\sum_{k=0}^{\infty}\tau_k^{-1} < \rho^2(1/64 + \sum_{k=0}^{\infty}\tau_k^{-1}).$$

□

Proposition 4. Let $\varepsilon = 0$ and (6) hold. If the conditions (9)-(13) are satisfied and if $F(w) = g$ is solvable in $B_\rho(w_0)$ as well as the relaxation parameters τ_k are bounded, then w_k converges to a solution of $F(w) = g$ in $B_\rho(w_0)$. Moreover, if $w_0 = \zeta$ and $w^\dagger$ is the unique w_0-minimum-norm solution and if $\mathcal{N}(F'(w^\dagger)) \subset \mathcal{N}(F'(w))$ for all $w \in B_\rho(w_0)$, then w_k converges to $w^\dagger$.

Proof. Let $\tilde{w}_*$ be a solution of $F(w) = g$ in $B_{\rho/8}(w_0) \cap B_{\rho/8}(\zeta)$ and denote $e_k := w_k - \tilde{w}_*$. Then, for each $n \ge m$,

$$\begin{aligned} e_{n+1} &= e_m\prod_{j=m}^{n}(1-\tau_j^{-1}) - \sum_{k=m}^{n}\tau_k^{-1}\prod_{j=k+1}^{n}(1-\tau_j^{-1})(\tilde{w}_* - \zeta) \\ &\quad - \sum_{k=m}^{n}\tau_k f(-\tau_k S_k^* S_k)S_k^*(F(w_k) - g)\prod_{j=k+1}^{n}(1-\tau_j^{-1}). \end{aligned} \tag{52}$$

If $\varepsilon = 0$, (48) implies that

$$\|w_{k+1} - \tilde{w}_*\|_X \le (1-\tau_k^{-1})\|w_k - \tilde{w}_*\|_X + \tau_k^{-1}\|\tilde{w}_* - \zeta\|_X. \tag{53}$$

Thus

$$\begin{aligned} \|e_{k+1}\|_X &\le (1-\tau_k^{-1})\|e_k\|_X + \tau_k^{-1}\|\tilde{w}_* - \zeta\|_X \\ &\le (1-\tau_k^{-1})\|e_k\|_X + \tau_k^{-1}\rho/8. \end{aligned} \tag{54}$$

By induction we can show that

$$\|e_n\|_X \le \|e_m\|_X \prod_{j=m}^{n-1}(1-\tau_j^{-1}) + \rho/8 \sum_{j=m}^{n-1} \tau_j^{-1} \prod_{r=j+1}^{n-1}(1-\tau_r^{-1}). \tag{55}$$

Proposition 2 implies that the sequence $\|e_k\|_X$ is bounded, and thus it has a convergent subsequence $\|e_{n(k)}\|_X$ to some $\omega \ge 0$. Let $\|e_{s(l)}\|_X$ be a subsequence of $\|e_k\|_X$. For given sufficiently large $l \in \mathbb{N}$, we denote $\tilde{k}(l)$ for the maximal index k such that $s(l) \ge n(k)$. Now (55) can be written by

$$\|e_{s(l)}\|_X \le \|e_{n(\tilde{k}(l))}\|_X \prod_{j=n(\tilde{k}(l))}^{s(l)-1}(1-\tau_j^{-1}) + \rho/8 \sum_{j=n(\tilde{k}(l))}^{s(l)-1} \tau_j^{-1} \prod_{r=j+1}^{s(l)-1}(1-\tau_r^{-1}). \tag{56}$$

If $\sum_{k=0}^{\infty}\tau_k^{-1} < \infty$ the product $\prod_{k=l}^{\infty}(1-\tau_k^{-1}) \to 1$ as $l \to \infty$ which implies that $\prod_{j=n(\tilde{k}(l))}^{s(l)-1}(1-\tau_k^{-1}) \to 1$ as $l \to \infty$. Thus (56) implies that

$$\|e_{s(l)}\|_X - \|e_{n(\tilde{k})}\|_X \le \rho/8 \sum_{j=n(\tilde{k}(l))}^{s(l)-1} \tau_j^{-1} \prod_{r=j+1}^{s(l)-1}(1-\tau_r^{-1}) = \rho/8 \left[1 - \prod_{j=n(\tilde{k}(l))}^{s(l)-1}(1-\tau_j^{-1})\right].$$

Consequently,

$$\limsup_{l\to\infty}\left(\|e_{s(l)}\|_X - \|e_{n(\tilde{k})}\|_X\right) \le 0. \tag{57}$$

Similary,

$$\limsup_{l\to\infty}\left(\|e_{n(\tilde{k})}\|_X - \|e_{s(l)}\|_X\right) \le 0. \tag{58}$$

Using (57)-(58) together with the fact that $\|e_{n(\tilde{k})}\|_X \to \omega$ as $l \to \infty$, we get $\|e_{s(l)}\|_X \to \omega$ as $l \to \infty$. Thus $\|e_k\|_X \to \omega$ as $k \to \infty$. For $j \ge k$ we choose l with $j \ge l \ge k$ such that

$$\|g - F(w_l)\|_Y \le \|g - F(w_r)\|_Y, \quad k \le r \le j. \tag{59}$$

Using (52) we get

$$\begin{aligned} e_j \;=\; & e_l \prod_{r=l}^{j-1}(1-\tau_j^{-1}) - \sum_{r=l}^{j-1}\tau_r^{-1}\prod_{s=r+1}^{j-1}(1-\tau_s^{-1})(\tilde{w}_* - \zeta) \\ & - \sum_{r=l}^{j-1}\tau_r f(-\tau_r S_r^* S_r) S_r^*(F(w_r) - g) \prod_{s=r+1}^{j-1}(1-\tau_s^{-1}) \end{aligned}$$

which provides

$$
\begin{aligned}
& \langle e_l - e_j, e_l\rangle_X \\
= \; & \langle e_l - e_l \prod_{r=l}^{j-1}(1-\tau_j^{-1}), e_l\rangle_X + \langle \sum_{r=l}^{j-1} \tau_r^{-1} \prod_{s=r+1}^{j-1} (1-\tau_s^{-1})(\tilde{w}_* - \zeta), e_l\rangle_X \\
& + \langle \sum_{r=l}^{j-1} \tau_r f(-\tau_r S_r^* S_r) S_r^*(F(w_r) - g) \prod_{s=r+1}^{j-1} (1-\tau_s^{-1}), e_l\rangle_X \\
= \; & \left(1 - \prod_{r=l}^{j-1}(1-\tau_r^{-1})\right) \|e_l\|_X^2 + \sum_{r=l}^{j-1} \tau_r^{-1} \prod_{s=r+1}^{j-1} (1-\tau_s^{-1}) \langle \tilde{w}_* - \zeta, w_l - \tilde{w}_*\rangle_X \\
& + \sum_{r=l}^{j-1} \prod_{s=r+1}^{j-1} (1-\tau_s^{-1}) \langle \tau_r f(-\tau_r S_r S_r^*)(F(w_r) - g), F'(w_r) e_l\rangle_X . \qquad (60)
\end{aligned}
$$

Using (9), (11) and (59) we get

$$
\begin{aligned}
& \left| \sum_{r=l}^{j-1} \prod_{s=r+1}^{j-1} (1-\tau_s^{-1}) \langle \tau_r f(-\tau_r S_r S_r^*)(F(w_r) - g), F'(w_r)(w_l - \tilde{w}_*)\rangle_X \right| \\
\leq \; & \sum_{r=l}^{j-1} \tau_r c_1 \|F(w_r) - g\|_Y (\|F'(w_r)(w_l - w_r)\|_Y + \|F'(w_r)(w_r - \tilde{w}_*)\|_Y) \\
\leq \; & \sum_{r=l}^{j-1} \tau_r c_1 \|F(w_r) - g\|_Y ((1+\eta)\|F(w_l) - F(w_r)\|_Y \\
& + (1+\eta)\|F(w_r) - F(\tilde{w}_*)\|_Y) \\
\leq \; & 3\tau c_1 (1+\eta) \sum_{r=l}^{j-1} \|F(w_r) - g\|^2 \qquad (61)
\end{aligned}
$$

for some τ with $\tau_r \leq \tau$ for all $r \in \mathbb{N}_0$. Using (61) and Proposition 3 the third term on the right hand side of (60) tends to zero. The fact that $\|e_l\|_X$ is bounded and

$$
1 \geq \prod_{r=l}^{j-1}(1-\tau_r^{-1}) \geq \prod_{r=l}^{\infty}(1-\tau_r^{-1}) \to 1 \qquad (62)
$$

imply $\left(1-\prod_{r=l}^{j-1}(1-\tau_r^{-1})\right)\|e_l\|_X^2$ tends to zero. Moreover the second term on the right hand side of (60) which is written as

$$\begin{aligned}&\left|\sum_{r=l}^{j-1}\tau_r^{-1}\prod_{s=r+1}^{j-1}(1-\tau_s^{-1})\langle\tilde{w}_*-\zeta, w_l-\tilde{w}_*\rangle_X\right|\\ \le\ &\left(1-\prod_{i=l}^{j-1}(1-\tau_i^{-1})\right)\|\tilde{w}_*-\zeta\|_X\|e_l\|_X,\end{aligned}$$

tends to zero due to (62), $\tilde{w}_*\in B_{\rho/8}(w_0)\cap B_{\rho/8}(\zeta)$ and $\|e_l\|_X\to\omega$ as $l\to\infty$. With the above imformation we get $|\langle e_l-e_j,e_l\rangle_X|\to 0$ as $l\to\infty$. In the same manner, $|\langle e_l-e_k,e_l\rangle_X|\to 0$ as $l\to\infty$. Thus the equations

$$\|e_j-e_l\|_X^2=2\langle e_l-e_j,e_l\rangle_X+\|e_j\|_X^2-\|e_l\|_X^2$$

and

$$\|e_l-e_k\|_X^2=2\langle e_l-e_k,e_l\rangle_X+\|e_k\|_X^2-\|e_l\|_X^2$$

provide that $\|e_j-e_l\|_X$ and $\|e_l-e_k\|_X$ tend to zero. This means that $\|e_j-e_k\|_X\le\|e_j-e_l\|_X+\|e_l-e_k\|_X$ tends to zero. Thus e_k and w_k are Cauchy sequences. Denote the limit of w_k by w_*. The continuity of F provide that $F(w_k)\to F(w_*)$ as $k\to\infty$. By Proposition 3 the residual $g-F(w_k)$ converges to zero. Thus w_* is a solution of $F(w)=g$.

Next we will show that w_* is the unique w_0-minimum-norm solution of $F(w)=g$. Let $w^\dagger$ be the unique w_0-minimum-norm solution of $F(w)=g$. Thus $w^\dagger-w_0\in\mathcal{N}(F'(w^\dagger))^\perp$. If $w_0=\zeta$ and $\mathcal{N}(F'(w^\dagger))\subset\mathcal{N}(F'(w))$ for all $w\in B_\rho(w_0)$, (5) provides

$$w_{k+1}-w_k=\tau_kF'(w_k)^*b^TH_k^{-1}\mathbf{1}(g-F(w_k))-\tau_k^{-1}(w_k-\zeta)\in\mathcal{R}(F'(w_k)^*)\subset\mathcal{N}(F'(x_k))^\perp.$$

Thus, $w_k-w_0\in\mathcal{N}(F'(x^\dagger))^\perp$ which implies

$$w^\dagger-w_*=w^\dagger-w_0+w_0-w_*\in\mathcal{N}(F'(x^\dagger))^\perp. \tag{63}$$

The tangential cone condition together with (63) provide the assertion. □

Theorem 1. Let the assumptions of Proposition 1 and 2 hold.

(i) If $k_*=k_*(\varepsilon,g^\varepsilon)$ is chosen according to (14) to stop the Runge-Kutta-type method (5) , then $w_{k_*}^\varepsilon\to w_*$ as $\varepsilon\to 0$.

(ii) If $N_0(\varepsilon)$ is chosen according to (16) to stop the Runge-Kutta-type method (5), then $w_{N_0(\varepsilon)}^\varepsilon\to w_*$ as $\varepsilon\to 0$.

Proof. The proof is analogous to the one in [23].

3. Convergence Rate Analysis

In this section we assume that F is properly scaled with a Lipschitz-continuous Frechéchet derivative in $B_\rho(w_0)$, i.e.

$$\|F'(w) - F'(\tilde{w})\| \le \tilde{L}\|w - \tilde{w}\|_X, \quad w, \tilde{w} \in B_\rho(w_0) \tag{64}$$

with $\tilde{L} \le 1$. Additionally, the Hölder type sourcewise condition

$$w_* - \zeta = F'(w_*)^* v, \quad v \in X. \tag{65}$$

is considered. For fixed $0 < \psi < 1$ the relaxation parameter is chosen as

$$\tau_k = (k + l_0)^\psi, \quad k \in \mathbb{N}_0 \tag{66}$$

where $l_0 \in \mathbb{N}_0$ is sufficiently large, i.e. such that $\tau_k > \max\{2, 1 + 1/c_1\}$ and in addition for $k \ge l_0$

$$\Psi(k) := \frac{1}{(1 + 1/k)^\psi} \frac{1 - (1 + 1/k)^\psi}{1/k} \frac{1}{k^{1-\psi}} + 1 \ge \tilde{L}^2. \tag{67}$$

Theorem 2. Let w_* be a solution of (1) in $B_{\rho/2}(w_0)$. Assume that (64) and the following closeness conditions, i.e.,

$$\tau_0 \|w_0 - w_*\|_X^2 \le \tilde{C} \le \min\{1, \tau_0 \rho^2/4\} \tag{68}$$

with $0 < \tilde{C}(\tau_k - 1) < 1$ and

$$2\|w_* - \zeta\|_X^2 + 2(1 - \tau_k^{-1})^2 \|v\|_X^2 + 17c_1^2 C^2 (\tau_k - 1)^2 \le \tilde{L}^2 \tilde{C}/2 \tag{69}$$

hold. Moreover, assume that ζ satisfies (65) and the relaxation parameter is chosen according to (66). If (5) is terminated by the discrepancy principle (14) where λ satisfies

$$1 + \frac{1}{\lambda^2} + \frac{2c_1}{\lambda}(\tau_k - 1) + 2\tau_k c_1^2 L^2 + (\tau_k - 1)(2\tau_k c_1 c_2 L - \frac{1}{2}c_1^2) < 0 \tag{70}$$

then in the case of noisy data

$$\|w_{k_*(\varepsilon)}^\varepsilon - w_*\|_X = O(k_*^{-\psi/2}). \tag{71}$$

If (5) is terminated by a-priori stopping rule (16) with

$$\frac{17}{16} + 2\tau_k c_1^2 L^2 + (\tau_k - 1)(2\tau_k c_1 c_2 L - \frac{1}{2}c_1^2) < 0 \tag{72}$$

then in the case of noisy data

$$\|w_{N_0}^{\varepsilon} - w_*\|_X = O(\varepsilon^{1/2}). \tag{73}$$

For the noise free case if (72) holds then

$$\|w_k - w_*\|_X = O(k^{-\psi/2}). \tag{74}$$

Proof. For short we use $f(t) := b^T(I - At)^{-1}\mathbf{1}$, see Appendix for more detail. Firstly we find the estimate of $\|w_{k+1}^{\varepsilon} - w_*\|_X$. Using (5) we can show that

$$\begin{aligned} \|w_{k+1}^{\varepsilon} - w_*\|_X^2 &= (1-\tau_k^{-1})^2\|w_k^{\varepsilon} - w_*\|_X^2 + (\tau_k^{-1})^2\|w_* - \zeta\|_X^2 \\ &+\|\tau_k f(-\tau_k S_k^* S_k) S_k^*(F(w_k^{\varepsilon}) - g^{\varepsilon})\|_X^2 \\ &-2(1-\tau_k^{-1})\tau_k^{-1}\langle w_k^{\varepsilon} - w_*, w_* - \zeta\rangle_X \\ &-2(1-\tau_k^{-1})\langle w_k^{\varepsilon} - w_*, \tau_k f(-\tau_k S_k^* S_k) S_k^*(F(w_k^{\varepsilon}) - g^{\varepsilon})\rangle_X \\ &+2\tau_k^{-1}\langle w_* - \zeta, \tau_k f(-\tau_k S_k^* S_k) S_k^*(F(w_k^{\varepsilon}) - g^{\varepsilon})\rangle_X. \end{aligned} \tag{75}$$

Using the fact that $2\langle A, B\rangle \le \|A\|^2 + \|B\|^2$ for the last term on the right-hand side of (75) we get

$$\begin{aligned} &2\tau_k^{-1}\langle w_* - \zeta, \tau_k f(-\tau_k S_k^* S_k) S_k^*(F(w_k^{\varepsilon}) - g^{\varepsilon})\rangle_X \\ &\quad\le (\tau_k^{-1})^2\|w_* - \zeta\|_X^2 + \|\tau_k f(-\tau_k S_k^* S_k) S_k^*(F(w_k^{\varepsilon}) - g^{\varepsilon})\|_X^2. \end{aligned} \tag{76}$$

Inserting (76) into (75) we obtain

$$\begin{aligned} \|w_{k+1}^{\varepsilon} - w_*\|_X^2 \le\ & (1-\tau_k^{-1})^2\|w_k^{\varepsilon} - w_*\|_X^2 + 2(\tau_k^{-1})^2\|w_* - \zeta\|_X^2 \\ &+2\|\tau_k f(-\tau_k S_k^* S_k) S_k^*(F(w_k^{\varepsilon}) - g^{\varepsilon})\|_X^2 \\ &-2(1-\tau_k^{-1})\tau_k^{-1}\langle w_k^{\varepsilon} - w_*, w_* - \zeta\rangle_X \\ &-2(1-\tau_k^{-1})\tau_k\langle w_k^{\varepsilon} - w_*, f(-\tau_k S_k^* S_k) S_k^*(F(w_k^{\varepsilon}) - g^{\varepsilon})\rangle_X. \end{aligned} \tag{77}$$

Due to the Lipschitz condition (64) it follows that

$$\|F(w) - F(\tilde{w}) - F'(\tilde{w})(w - \tilde{w})\|_Y \le \frac{1}{2}\tilde{L}\|w - \tilde{w}\|_X^2, \quad w, \tilde{w} \in B_{2\rho}(w_0). \tag{78}$$

Using (2) and (78) we can show that

$$\|F(w_k^{\varepsilon}) - g^{\varepsilon} - F'(w_*)(w_k^{\varepsilon} - w_*)\|_Y \le \frac{\tilde{L}}{2}\|w_k^{\varepsilon} - w_*\|_X^2 + \varepsilon$$

and consequently

$$\|F'(w_*)(w_k^\varepsilon - w_*)\|_X \le \|F(w_k^\varepsilon) - g^\varepsilon\|_Y + \frac{\tilde{L}}{2}\|w_k^\varepsilon - w_*\|_X^2 + \varepsilon. \tag{79}$$

The sourcewise representation (65) and the estimate (79) yield

$$\begin{aligned}
&-2(1-\tau_k^{-1})\tau_k^{-1}\langle w_k^\varepsilon - w_*, w_* - \zeta\rangle_X \\
&\quad= -2(1-\tau_k^{-1})\tau_k^{-1}\langle w_k^\varepsilon - w_*, F'(w_*)^* v\rangle_X \\
&\quad\le 2(1-\tau_k^{-1})\tau_k^{-1}\|F'(w_*)(w_k^\varepsilon - w_*)\|_Y\|v\|_X \\
&\quad\le 2(1-\tau_k^{-1})\tau_k^{-1}\|F(w_k^\varepsilon) - g^\varepsilon\|_Y\|v\|_X \\
&\qquad +(1-\tau_k^{-1})\tau_k^{-1}\tilde{L}\|w_k^\varepsilon - w_*\|_X^2\|v\|_X + 2(1-\tau_k^{-1})\tau_k^{-1}\varepsilon\|v\|_X.
\end{aligned} \tag{80}$$

The fact that $2AB \le A^2 + B^2$ leads to

$$\begin{aligned}
&2\tilde{L}\|w_k^\varepsilon - w_*\|_X^2\|f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon) - g^\varepsilon)\|_Y \\
&\le \|f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon) - g^\varepsilon)\|_Y^2 + \tilde{L}^2\|w_k^\varepsilon - w_*\|_X^4.
\end{aligned} \tag{81}$$

Using (2), (78) and (81) we can show that

$$\begin{aligned}
&-2(1-\tau_k^{-1})\tau_k\langle w_k^\varepsilon - w_*, f(-\tau_k S_k^* S_k)S_k^*(F(w_k^\varepsilon) - g^\varepsilon)\rangle_X \\
&\quad= 2(\tau_k - 1)\langle F(w_k^\varepsilon) - g^\varepsilon - F'(w_k^\varepsilon)(w_k^\varepsilon - w_*), f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon) - g^\varepsilon)\rangle_Y \\
&\qquad -2(\tau_k - 1)\langle F(w_k^\varepsilon) - g^\varepsilon, f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon) - g^\varepsilon)\rangle_Y \\
&\quad\le 2(\tau_k - 1)\left(\|g - F(w_k^\varepsilon) - F'(w_k^\varepsilon)(w_* - w_k^\varepsilon)\|_Y + \varepsilon\right) \\
&\qquad \times\|f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon) - g^\varepsilon)\|_Y \\
&\qquad -2(\tau_k - 1)\langle F(w_k^\varepsilon) - g^\varepsilon, f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon) - g^\varepsilon)\rangle_Y \\
&\quad= (\tau_k - 1)\tilde{L}\|w_k^\varepsilon - w_*\|_X^2\|f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon) - g^\varepsilon)\|_Y \\
&\qquad +2(\tau_k - 1)\varepsilon\|f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon) - g^\varepsilon)\|_Y \\
&\qquad -2(\tau_k - 1)\langle F(w_k^\varepsilon) - g^\varepsilon, f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon) - g^\varepsilon)\rangle_Y \\
&\quad\le \frac{1}{2}(\tau_k - 1)\|f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon) - g^\varepsilon)\|_Y^2 + \frac{1}{2}(\tau_k - 1)\tilde{L}^2\|w_k^\varepsilon - w_*\|_X^4 \\
&\qquad +2(\tau_k - 1)\varepsilon\|f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon) - g^\varepsilon)\|_Y \\
&\qquad -2(\tau_k - 1)\langle F(w_k^\varepsilon) - g^\varepsilon, f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon) - g^\varepsilon)\rangle_Y.
\end{aligned} \tag{82}$$

Note that the consistency property of the RK method and the assumptions (11)-(13) imply

$$-2(\tau_k - 1)\langle F(w_k^\varepsilon) - g^\varepsilon, f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon) - g^\varepsilon)\rangle_Y$$

$$
\begin{aligned}
&\le 2(\tau_k - 1)\big(-\|f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon) - g^\varepsilon)\|_Y^2 \\
&\quad +\tau_k\|S_k^* b^T(A-I)H_k^{-1}\mathbf{1}(F(w_k^\varepsilon) - g^\varepsilon)\|_X \|S_k^* f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon) - g^\varepsilon)\|_X \\
&\quad -\tau_k\|S_k^* f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon) - g^\varepsilon)\|_X^2\big) \\
&\le -(\tau_k - 1)c_1^2\|F(w_k^\varepsilon) - g^\varepsilon\|_Y^2 + 2(\tau_k - 1)\tau_k c_1 c_2 L\|F(w_k^\varepsilon) - g^\varepsilon\|_Y^2 \\
&\quad -2(\tau_k - 1)\tau_k\|S_k^* f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon) - g^\varepsilon)\|_X^2. \qquad (83)
\end{aligned}
$$

Combining (82) and (83) leads to

$$
\begin{aligned}
&-2(1-\tau_k^{-1})\tau_k\langle w_k^\varepsilon - w_*, f(-\tau_k S_k^* S_k)S_k^*(F(w_k^\varepsilon) - g^\varepsilon)\rangle_X \\
&\le \frac{1}{2}c_1^2(\tau_k - 1)\|F(w_k^\varepsilon) - g^\varepsilon\|_Y^2 + \frac{1}{2}(\tau_k - 1)\tilde{L}^2\|w_k^\varepsilon - w_*\|_X^4 \\
&\quad +2c_1(\tau_k - 1)\varepsilon\|F(w_k^\varepsilon) - g^\varepsilon\|_Y \\
&\quad -(\tau_k - 1)c_1^2\|F(w_k^\varepsilon) - g^\varepsilon\|_Y^2 + 2(\tau_k - 1)\tau_k c_1 c_2 L\|F(w_k^\varepsilon) - g^\varepsilon\|_Y^2 \\
&\quad -2(\tau_k - 1)\tau_k\|S_k^* f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon) - g^\varepsilon)\|_X^2 \\
&= (2\tau_k c_1 c_2 L - \frac{1}{2}c_1^2)(\tau_k - 1)\|F(w_k^\varepsilon) - g^\varepsilon\|_Y^2 + \frac{1}{2}(\tau_k - 1)\tilde{L}^2\|w_k^\varepsilon - w_*\|_X^4 \\
&\quad +2c_1(\tau_k - 1)\varepsilon\|F(w_k^\varepsilon) - g^\varepsilon\|_Y \\
&\quad -2(\tau_k - 1)\tau_k\|S_k^* f(-\tau_k S_k S_k^*)(F(w_k^\varepsilon) - g^\varepsilon)\|_X^2. \qquad (84)
\end{aligned}
$$

Applying (80) and (84) into (77) we get

$$
\begin{aligned}
\|w_{k+1}^\varepsilon - w_*\|_X^2 \le\ & (1-\tau_k^{-1})\|w_k^\varepsilon - w_*\|_X^2(1 - \tau_k^{-1}(1 - \tilde{L}\|v\|_X)) \\
& +\frac{1}{2}(\tau_k - 1)\tilde{L}^2\|w_k^\varepsilon - w_*\|_X^4 \\
& +2\tau_k^{-1}(\tau_k^{-1}\|w_* - \zeta\|_X^2 + (1-\tau_k^{-1})\|F(w_k^\varepsilon) - g^\varepsilon\|_Y\|v\|_X) \\
& +2\varepsilon((1-\tau_k^{-1})\tau_k^{-1}\|v\|_X + c_1(\tau_k - 1)\|F(w_k^\varepsilon) - g^\varepsilon\|_Y) \\
& +\|F(w_k^\varepsilon) - g^\varepsilon\|_Y^2(2\tau_k c_1^2 L^2 + (\tau_k - 1)(2\tau_k c_1 c_2 L - \frac{1}{2}c_1^2)). \quad (85)
\end{aligned}
$$

To prove by induction that $\tau_k\|w_k^\varepsilon - w_*\|_X^2 \le \tilde{C}$ for a posteriori parameter choice rule (14) we observe that the assumption (68) implies the case $k = 0$. Let $k < k_*(\varepsilon)$ and assume that $\tau_k\|w_k^\varepsilon - w_*\|_X^2 \le \tilde{C}$. It follows from the discrepancy principle (14) and the property $2AB \le A^2 + B^2$ that

$$
\begin{aligned}
2\varepsilon(1-\tau_k^{-1})\tau_k^{-1}\|v\|_X &\le 2(1-\tau_k^{-1})\tau_k^{-1}\|v\|_X\frac{\|g^\varepsilon - F(w_k^\varepsilon)\|_Y}{\lambda} \\
&\le \frac{1}{\lambda^2}\|g^\varepsilon - F(w_k^\varepsilon)\|_Y^2 + (1-\tau_k^{-1})^2(\tau_k^{-1})^2\|v\|_X^2 \qquad (86)
\end{aligned}
$$

and

$$2\varepsilon c_1(\tau_k-1)\|g^\varepsilon - F(w_k^\varepsilon)\|_Y \le 2c_1(\tau_k-1)\frac{\|g^\varepsilon - F(w_k^\varepsilon)\|_Y^2}{\lambda}. \qquad (87)$$

Similarly,

$$2(1-\tau_k^{-1})\tau_k^{-1}\|v\|_X\|g^\varepsilon - F(w_k^\varepsilon)\|_Y \le \|g^\varepsilon - F(w_k^\varepsilon)\|_Y^2 + (1-\tau_k^{-1})^2(\tau_k^{-1})^2\|v\|_X^2. (88)$$

Inserting (86)-(88) into (85) we obtain

$$\begin{aligned}
&\|w_{k+1}^\varepsilon - w_*\|_X^2 \\
&\quad\le (1-\tau_k^{-1})\|w_k^\varepsilon - w_*\|_X^2(1-\tau_k^{-1}(1-\tilde{L}\|v\|_X)) \\
&\qquad + \frac{1}{2}(\tau_k-1)\tilde{L}^2\|w_k^\varepsilon - w_*\|_X^4 \\
&\qquad + \left(1+\frac{1}{\lambda^2}+\frac{2c_1}{\lambda}(\tau_k-1)+2\tau_k c_1^2L^2+(\tau_k-1)(2\tau_k c_1 c_2 L - \frac{1}{2}c_1^2)\right) \\
&\qquad \times\|F(w_k^\varepsilon)-g^\varepsilon\|_Y^2 \\
&\qquad + (\tau_k^{-1})^2(2\|w_* - \zeta\|_X^2 + 2(1-\tau_k^{-1})^2\|v\|_X^2). \qquad (89)
\end{aligned}$$

Note that the assumption (69) and $\tau_k \ge 2$ provide

$$\|v\|_X^2 \le \frac{\tau_k^2}{4(\tau_k-1)^2}\tilde{L}^2\tilde{C} - \frac{\tau_k^2}{(\tau_k-1)^2}\|w_*-\zeta\|_X^2 - \frac{17}{2}c_1^2C^2\tau_k^2 \le \frac{\tau_k^2}{4(\tau_k-1)^2}\tilde{L}^2\tilde{C}.$$

Thus we can conclude from the assumption (64) and (68) that

$$\tilde{L}^2\|v\|_X^2 \le \tilde{L}^4\tilde{C} \le 1. \qquad (90)$$

Due to the assumptions (70) and (90) we have

$$\begin{aligned}
\|w_{k+1}^\varepsilon - w_*\|_X^2 &\le (1-\tau_k^{-1})\|w_k^\varepsilon - w_*\|_X^2 + \frac{1}{2}(\tau_k-1)\tilde{L}^2\|w_k^\varepsilon - w_*\|_X^4 \\
&\quad + (\tau_k^{-1})^2(2\|w_*-\zeta\|_X^2 + 2(1-\tau_k^{-1})^2\|v\|_X^2). \qquad (91)
\end{aligned}$$

Define $\beta_k := \tau_k\|w_k^\varepsilon - w_*\|_X^2$ and $B := 2\|w_*-\zeta\|_X^2 + 2(1-\tau_k^{-1})^2\|v\|_X^2$. Thus (91) can be written as

$$\beta_{k+1} \le \beta_k\frac{\tau_{k+1}}{\tau_k^2}(\tau_k-1) + \frac{\tau_{k+1}}{2\tau_k^2}(\tau_k-1)\tilde{L}^2\beta_k^2 + \frac{\tau_{k+1}}{\tau_k^2}B. \qquad (92)$$

Define $J(\beta_k) := \frac{\tau_{k+1}}{2\tau_k^2}(\tau_k-1)\tilde{L}^2\beta_k^2 + \beta_k\frac{\tau_{k+1}}{\tau_k^2}(\tau_k-1) + \frac{\tau_{k+1}}{\tau_k^2}B$. Using the closeness assumption (69) we can see that

$$B \le \tilde{L}^2\tilde{C}/2. \qquad (93)$$

The definition of J yields

$$
\begin{aligned}
J(\tilde{C}) - \tilde{C} &= \frac{\tau_{k+1}}{2\tau_k^2}(\tau_k - 1)\tilde{L}^2\tilde{C}^2 + \tilde{C}\frac{\tau_{k+1}}{\tau_k^2}(\tau_k - 1) + \frac{\tau_{k+1}}{\tau_k^2}B - \tilde{C} \\
&= -\tilde{C}\left(1 - \frac{\tau_{k+1}}{\tau_k^2}(\tau_k - 1)\right) + \frac{\tau_{k+1}}{\tau_k^2}\left(\tilde{C}^2\frac{\tilde{L}^2}{2}(\tau_k - 1) + B\right). \qquad (94)
\end{aligned}
$$

Note that from the assumption (68) we have $\tilde{C}^2(\tau_k - 1) < \tilde{C}$. This estimate combines with (93) provide

$$\frac{1}{2}\tilde{C}^2\tilde{L}^2(\tau_k - 1) + B < \frac{1}{2}\tilde{C}\tilde{L}^2 + B < \frac{1}{2}\tilde{C}\tilde{L}^2 + \frac{1}{2}\tilde{C}\tilde{L}^2 = \tilde{C}\tilde{L}^2. \qquad (95)$$

Note that

$$
\begin{aligned}
\Psi(k + l_0) &= \frac{(k + l_0)^\psi}{(k + l_0 + 1)^\psi}\left((k + l_0)^\psi - (k + l_0 + 1)^\psi\right) + 1 \\
&= \tau_{k+1}^{-1}(\tau_k^2 - \tau_{k+1}(\tau_k - 1)). \qquad (96)
\end{aligned}
$$

Applying (67) and (96) into (95) we obtain

$$\frac{1}{2}\tilde{C}^2\tilde{L}^2(\tau_k - 1) + B < \tilde{C}\tilde{L}^2 \leq \Psi(k + l_0)\tilde{C} = \tau_{k+1}^{-1}\left(\tau_k^2 - \tau_{k+1}(\tau_k - 1)\right)\tilde{C}. \qquad (97)$$

Therefore (94) becomes

$$J(\tilde{C}) - \tilde{C} \leq -\tilde{C}\left(1 - \frac{\tau_{k+1}}{\tau_k^2}(\tau_k - 1)\right) + \frac{\tau_{k+1}}{\tau_k^2}\frac{\tau_k^2 - \tau_{k+1}(\tau_k - 1)}{\tau_{k+1}}\tilde{C} = 0. \qquad (98)$$

By the assumption of the induction and the fact that $J(\beta_k)$ is monotonically increasing we can see that

$$\beta_{k+1} \leq J(\beta_k) \leq J(\tilde{C}) \leq \tilde{C}.$$

Thus the induction is complete. Due to the monotonicity of $\{\tau_k^{-1}\}$ and (68) we can see that $w_{k+1}^\varepsilon \in B_{\rho/2}(w_*)$.

It remains to show the rate of the modified RK-type method where the termination index is chosen according to a-priori parameter choice rule (16). Similarly to the previous case we will show that $\tau_k\|w_k^\varepsilon - w_*\|_X^2 \leq \tilde{C}$. Clearly the assumption (68) implies the case $k = 0$. Let $k < N_0$ and assume that

$\tau_k\|w_k^\varepsilon - w_*\|_X^2 \le \tilde{C}$. It follows from (16) and the property $2AB \le A^2 + B^2$ that

$$\begin{aligned}2\varepsilon c_1(\tau_k - 1)\|F(w_k^\varepsilon) - g^\varepsilon\|_Y &\le 8\tau_k^{-1}Cc_1(\tau_k-1)\frac{\|F(w_k^\varepsilon) - g^\varepsilon\|_Y}{4} \\ &\le 16C^2c_1^2(1-\tau_k^{-1})^2 + \frac{1}{16}\|F(w_k^\varepsilon) - g^\varepsilon\|_Y^2 \quad (99)\end{aligned}$$

and

$$2\varepsilon(1-\tau_k^{-1})\tau_k^{-1}\|v\|_X \le 2C(1-\tau_k^{-1})(\tau_k^{-1})^2\|v\|_X. \quad (100)$$

Applying (88), (99)-(100) into (85) we obtain

$$\begin{aligned}&\|w_{k+1}^\varepsilon - w_*\|_X^2 \\ &\le (1-\tau_k^{-1})\|w_k^\varepsilon - w_*\|_X^2(1-\tau_k^{-1}(1-\tilde{L}\|v\|_X)) + \frac{1}{2}(\tau_k-1)\tilde{L}^2\|w_k^\varepsilon - w_*\|_X^4 \\ &\quad +\|F(w_k^\varepsilon) - g^\varepsilon\|_Y^2\left(\frac{17}{16} + 2\tau_k c_1^2L^2 + (\tau_k-1)(2\tau_k c_1c_2L - \frac{1}{2}c_1^2)\right) \\ &\quad +(\tau_k^{-1})^2\big(2\|w_* - \zeta\|_X^2 \\ &\quad +(1-\tau_k^{-1})^2\|v\|_X^2 + 16C^2c_1^2(\tau_k-1)^2 + 2C(1-\tau_k^{-1})\|v\|_X\big). \quad (101)\end{aligned}$$

Due to the assumptions (72) and (90) it follows that

$$\begin{aligned}&\|w_{k+1}^\varepsilon - w_*\|_X^2 \\ &\le (1-\tau_k^{-1})\|w_k^\varepsilon - w_*\|_X^2(1-\tau_k^{-1}(1-\tilde{L}\|v\|_X)) + \frac{1}{2}(\tau_k-1)\tilde{L}^2\|w_k^\varepsilon - w_*\|_X^4 \\ &\quad +(\tau_k^{-1})^2\big(2\|w_* - \zeta\|_X^2 \\ &\quad +(1-\tau_k^{-1})^2\|v\|_X^2 + 16C^2c_1^2(\tau_k-1)^2 + 2C(1-\tau_k^{-1})\|v\|_X\big). \quad (102)\end{aligned}$$

Define $\beta_k := \tau_k\|w_k^\varepsilon - w_*\|_X^2$ and $\tilde{B} := 2\|w_* - \zeta\|_X^2 + (1-\tau_k^{-1})^2\|v\|_X^2 + 16C^2c_1^2(\tau_k-1)^2 + 2C(1-\tau_k^{-1})\|v\|_X$. Hence (102) can be written as

$$\beta_{k+1} \le \beta_k\frac{\tau_{k+1}}{\tau_k^2}(\tau_k-1) + \frac{\tau_{k+1}}{2\tau_k^2}(\tau_k-1)\tilde{L}^2\beta_k^2 + \frac{\tau_{k+1}}{\tau_k^2}\tilde{B}. \quad (103)$$

Define $\tilde{J}(\beta_k) := \frac{\tau_{k+1}}{2\tau_k^2}(\tau_k-1)\tilde{L}^2\beta_k^2 + \frac{\tau_{k+1}}{\tau_k^2}(\tau_k-1)\beta_k + \frac{\tau_{k+1}}{\tau_k^2}\tilde{B}$. Similar to the previous case we can show that $\beta_{k+1} \le \tilde{J}(\beta_k) \le \tilde{J}(\tilde{C}) \le \tilde{C}$. The monotonicity of $\{\tau_k^{-1}\}$ and (68) deduce that $w_{k+1}^\varepsilon \in B_{\rho/2}(w_*)$. Moreover, the rate result for the noise free case is proven, since $k_*(0) = \infty$. Due to (16) the assertion for the noisy case is obtained. □

Table 1: The values of c_1 and c_2 for some RK methods.

	Landweber	Levenberg-Marquardt	Radau	Implicit trapezoidal
c_1	1	1	1	1
c_2	$\|S_k\|$	-	$0.3\tau_k^{-1/2}$	$0.5(2\tau_k)^{-1/2}$

4. Numerical Example

In this section, we consider the nonlinear operator $F : L^2[0,1] \to L^2[0,1]$ defined by

$$[F(w)](s) = \exp \int_0^1 k(s,t)w(t)\mathrm{d}t \tag{104}$$

with the noisy data $g^\varepsilon(s) = \exp((s^4 - 2s^3 + s)/12) + \varepsilon \cos(100s)$, $s \in [0,1]$ and the kernel function defined by

$$k(s,t) = \begin{cases} s(1-t) & \text{if } s < t; \\ t(1-s) & \text{if } t \leq s. \end{cases} \tag{105}$$

Note that $\sup_{0\leq k,s\leq 1} |k(s,t)| \leq 1$. The operator F is Fréchet differentiable and its Fréchet differential is given by [20]

$$F'(w)h = F(w) \int_0^1 k(s,t)h(t)\mathrm{d}t. \tag{106}$$

Moreover $\|F'(w)\| \leq \exp(1/\sqrt{30})$ if $w_0 = 0$.

This example was considered in [5] where the values of c_1 and c_2 for some RK methods are presented in the Table 1. It has been shown that the number of iteration steps are rapidly reduced if the relaxation parameter increases for the Levenberg-Marquardt (LM) and Radau methods [5] using the iterative RK-type method defined by (3). In order to make the implementations for the modified RK-type method (5) more clear we briefly recall useful notations and formulas in a finite space.

Let $w_m^\varepsilon(t) \in \mathcal{R}(P_m)$ be a solution of the problem $F_m(w_m^\varepsilon) = Q_m g^\varepsilon$ in a finite space with the orthogonal projections of X and Y onto X_m and Y_m denoted by P_m and Q_m, respectively. For each pair of the finite space X_m and Y_m it is convinence to denote that $\tau_k := \tau_{m,k}$, $w_{k+1}^\varepsilon := w_{m,k+1}^\varepsilon$ and $w_k^\varepsilon := w_{m,k}^\varepsilon$. For the Hilbert spaces in Example 1 and 2, i.e. X and Y

are $L^2[0,1]$, we choose orthogonal bases $\{\varphi_1^{(m)} \dots \varphi_m^{(m)}\}$ and $\{\psi_1^{(m)} \dots \psi_m^{(m)}\}$ which is defined by the piecewise continuous function with $\varphi(t) = 1$ for $t \in [t_j, t_{j+1}]$, $\psi(t) = 1$ for $s \in [s_j, s_{j+1}]$ and $\varphi(t) = 0, \psi(t) = 0$ otherwise. Due to X_m =span$\{\varphi_j^{(m)}\}_{j=1,\dots,m}$ and Y_m =span$\{\psi_j^{(m)}\}_{j=1,\dots,m}$ we let $w_k^\varepsilon(t) = \sum_{j=1}^m u_j^{(m)}\varphi_j^{(m)}(t)$, $w_{k+1}^\varepsilon(t) = \sum_{j=1}^m v_j^{(m)}\varphi_j^{(m)}(t)$ and $\zeta(t) = \sum_{j=1}^m z_j^{(m)}\varphi_j^{(m)}(t)$ for some $U^{(m)} = (u_1^{(m)} \dots u_m^{(m)})^T$, $V^{(m)} = (v_1^{(m)} \dots v_m^{(m)})^T$ and $Z^{(m)} = (z_1^{(m)} \dots z_m^{(m)})^T$. For the 1- and 2-stage modified RK-type regularization which is defined in the finite space by

$$\begin{aligned} w_{k+1}^\varepsilon &= w_k^\varepsilon + \tau_k b^T(\boldsymbol{\delta}_m + \tau_k A F_m'(w_k^\varepsilon)^* F_m'(w_k^\varepsilon))^{-1}\mathbf{1}F_m'(w_k^\varepsilon)^*(Q_m g^\varepsilon - F_m(w_k^\varepsilon)) \\ &\quad -\tau_k^{-1}(w_k^\varepsilon - \zeta) \end{aligned} \tag{107}$$

the following propositions provide the coefficient vector V^m.

Proposition 6. For the 1-stage Runge-Kutta method $A = (a_{11})$, $b = (b_1)$ and $\mathbf{1} = (1)$ if $I + (\tau_k/h^{(m)})a_{11}\Phi^{(m)}$ is nonsingular, then the coefficient vector $V^{(m)}$ corresponding to (107) is given by

$$V^{(m)} = (1-\tau_k^{-1})U^{(m)} + \tau_k^{-1}Z^{(m)} + \tau_k b_1(I + (\tau_k/h^{(m)})a_{11}\Phi^{(m)})^{-1}Q^{(m)} \tag{108}$$

where the matrix $\Phi^{(m)} = (\phi_{ij}^{(m)})_{i,j=1,2,\dots,m}$ and the vector $Q^{(m)} = (q_1^{(m)} \dots q_m^m)$ are defined by

$$\phi_{ij}^{(m)} = \langle \varphi_i^{(m)(t)}, F_m'(w_k^\varepsilon)^* F_m'(w_k^\varepsilon)\varphi_j^{(m)}(t)\rangle_X \tag{109}$$

and

$$q_j^{(m)} = \langle \varphi_j^{(m)}(t), F_m'(w_k^\varepsilon)^*(Q_m g^\varepsilon - F_m(w_k^\varepsilon))\rangle_X / h^{(m)} \tag{110}$$

Proof. The idea of the proof is analogous to Proposition 5 in [5]. □

Proposition 7. For the 2-stage Runge-Kutta method $A = (a_{ij})_{i,j=1,2}$, $b = (b_1 \quad b_2)^T$ and $\mathbf{1} = (1 \quad 1)^T$ if $\mathbf{S} = I + (\tau_k/h^{(m)})(\mathrm{Tr A})\Phi^{(\mathrm{m})} + (\tau_\mathrm{k}/\mathrm{h}^{(\mathrm{m})})^2(\det \mathrm{A})(\Phi^{(\mathrm{m})})^2$ and $\mathbf{R} = I + (\tau_k/h^{(m)})a_{11}\Phi^{(m)}$ are nonsingular, then the coefficient vector $V^{(m)}$ corresponding to (107) is given by

$$V^{(m)} = (1 - \tau_k^{-1})U^{(m)} + \tau_k^{-1}Z^{(m)} + \tau_k b_1 R^{(m)} + \tau_k b_2 S^{(m)} \tag{111}$$

where the matrix $\Phi^{(m)}$ and the vector $Q^{(m)}$ are defined by (109) and (110), respectively. Moreover,

$$S^{(m)} = \mathbf{S}^{-1}(I + (\tau_k/h^{(m)})(a_{11} - a_{21})\Phi^{(m)})Q^{(m)} \tag{112}$$

Table 2: The relative L^2-error e_{k_*} for the Levenberg-Marquardt, Radau and Gauss-Newton methods with $m = 65, \lambda = 1.1, \tau_k = 0.5k^2$ and $w_0(t) = \zeta(t) = \varepsilon w_*(t)$. The termination index is chosen according to the discrepancy principle (14).

Method	ε	k_*	e_{k_*}	s_{k_*}	τ_{k_*}
LM	1.0e-3	14	0.6543	0.9298	98
	1.0e-4	26	0.0750	0.9084	338
	1.0e-5	34	0.0390	0.8947	578
	1.0e-6	101	0.0167	0.9634	5100.5
Radau	1.0e-3	14	0.6536	0.9288	98
	1.0e-4	26	0.0738	0.8875	338
	1.0e-5	34	0.0389	0.8728	578
	1.0e-6	101	0.0167	0.9602	5100.5
GN	1.0e-3	31	0.7013	0.9962	480.5
	1.0e-4	103	0.0393	0.0763	5304.5

and

$$R^{(m)} = \mathbf{R}^{-1}(Q^{(m)} - (\tau_k/h^{(m)})a_{12}\Phi^{(m)}S^{(m)}). \tag{113}$$

Proof. The idea of the proof is analogous to Proposition 5 in [5]. □

We note that the choice of τ_k in Theorem 2 does not satisfy the condition in Proposition 1. We set $\tau_k = 0.5k^2$ in Table 2 which yields $\sum_k^\infty \tau_k^{-1} < \infty$ according to Proposition 1. Therefore the iterates (5) converges even if the sourcewise representation is not satisfied. However the choice $\zeta(t) = \varepsilon w_*(t)$ satisfies the condition (17). Let $Kx := \int_0^1 k(s,t)x(t)\mathrm{d}t$. Thus the adjoint of K is defined by $K^*z := \int_0^1 k(s,t)z(s)\mathrm{d}s$. Due to (106) the adjoint operator of $F'(w)$ is defined by $F'(w)^*y := K^*(F(w)\cdot y)$. For the example we have $w_*(t) = t(1-t)$ and $F(w_*) = \exp(\frac{s^4-2s^3+s}{12})$. If $v(s) = c\exp(-\frac{s^4-2s^3+s}{12})$ for some number $c \in \mathbb{R}$, we can see that

$$F'(w_*)^*v = K^*(F(w_*)\cdot v) = K^*(c) = c\int_0^1 k(s,t)\mathrm{d}s = \frac{c}{2}t(1-t) = 0.5cw_*(t).$$

This means that for a priori guess $\zeta(t) = (1-0.5c)w_*(t)$ we have $w_* - \zeta \in \mathcal{R}(F'(w_*)^*)$. Therefore $\zeta(t) = \varepsilon w_*(t)$ is used in Table 3 for the choice $\tau_k = (k+10)^{0.99}$.

Table 3: The relative L^2-error e_{k_*} for the Levenberg-Marquardt, Radau and Gauss-Newton methods with $m = 65, \lambda = 1.1, \tau_k = (k+10)^{0.99}$ and $w_0(t) = \zeta(t) = \varepsilon w_*(t)$. The termination index is chosen according to the discrepancy principle (14).

Method	ε	k_*	e_{k_*}	s_{k_*}	τ_{k_*}	$k_*^{0.495} e_{k_*}$
LM	1.0e-3	26	0.7021	0.9972	34.7328	4.0648
	1.0e-4	128	0.0796	0.9527	131.3652	0.8905
	1.0e-5	435	0.0269	0.9981	418.6744	0.5346
	1.0e-6	1473	0.0148	0.9995	1378.5729	0.5312
Radau	1.0e-3	26	0.7018	0.9969	34.7328	4.0634
	1.0e-4	128	0.0793	0.9879	131.3652	0.8872
	1.0e-5	432	0.0270	0.9973	415.8800	0.5351
	1.0e-6	1442	0.0149	0.9997	1350.0410	0.5273
GN	1.0e-3	504	0.7038	0.9998	482.8959	14.9962
	1.0e-4	1824	0.0795	0.9803	1701.2390	3.1598

In Tables 2 and 3 the iterates are stopped by the discrepancy principle (14). However as mention in [22] a large step size τ_k can lead to an overshot solution. To avoid the problem with a large step size the quotient $s_k := \|g^\varepsilon - F(w_k^\varepsilon)\|_Y / \lambda\varepsilon$ is used and $w_{k_*}^\varepsilon$ is accepted if $s_{k_*} \approx 1$. The values of s_{k_*} reported in Tables 2 and 3, are very close to 1 as expected. Except the iteratively regularized Gauss-Newton (GN) method with $\varepsilon = 1.0e-4$ in table 2. Here the iteratively regularized Gauss-Newton method [3] is given by

$$w_{k+1}^\varepsilon = w_k^\varepsilon + (\tau_k^{-1} I + F'(w_k^\varepsilon)^* F'(w_k^\varepsilon))^{-1} (F'(w_k^\varepsilon)^* (g^\varepsilon - F(w_k^\varepsilon)) + \tau_k^{-1}(w_0 - w_k^\varepsilon)).$$

While it is not possible to show the rate result in table 2, table 3 indicates that for the Levenberg-Marquardt and Radau methods w_k^ε converges to w_* with a rate $O(k_*^{-0.495})$ as expected if k_* is chosen according to the discrepancy principle where τ_k satisfies (66).

5. Conclusions

In this paper, we proposed the modified Runge-Kutta-type regularization for nonlinear ill-posed problems. The proposed iterative method is terminate by a priori and a posteriori parameter choice rules under the certain condition on a priori guess ζ, the step size τ_k and on the nonlinear operator F.

For the noise case the termination index $k_*(\varepsilon, g^\varepsilon)$, chosen according to the discrepancy principle, is finite if $\sum_{k=1}^{\infty} \tau_k^{-1}$ is finite. However if $\sum_{k=1}^{\infty} \tau_k^{-1}$ diverges the sourcewise representation of $w_* - \zeta$ guarantees that $w_k^\varepsilon \in B_{\rho/2}$. The convergence rate results for the modified RK-type regularization can be obtained under the sourcewise representation of $w_* - \zeta$ if the Fréchet derivative is properly scaled and is locally Lipschitz continuous. The numerical example shows that the Levenberg-Marquardt and Radau methods converge to a solution with a few iteration steps and a large step size without the problem of an overshot solution.

6. Acknowledgments

This work was supported by Department of Mathematics, Faculty of Science, Silpakorn University and by the European Commission with respect to the EARLINET-ASOS project under grant 025991 (RICA) and ITaRS project under grant 289923.

7. Appendix

From the spectral theory and functional calculus if $(\sigma_j; v_j, u_j)$ is a singular system for a compact linear operator K, an integral with respect to a spectral family E_λ can be defined by

$$f(K^*K)x := \int f(\lambda) \mathrm{d}E_\lambda x := \sum_{j=1}^{\infty} f(\sigma_j^2)\langle x, v_j\rangle v_j \tag{114}$$

where f is a (picewise) continuous function and we integrate over the whole domain. Here the parameter λ is used for the spectral family only. Thus for all (picewise) continuous function

$$f(K^*K)K^*x = \sum_{j=1}^{\infty} f(\sigma_j^2)\langle K^*x, v_j\rangle v_j = \sum_{j=1}^{\infty} f(\sigma_j^2)\langle x, u_j\rangle K^*u_j = K^*f(KK^*) \tag{115}$$

where $f(KK^*)$ is defined analogously to $f(K^*K)$. If $f(K^*K)$ is defined via (114), then for $x, y \in X$

$$\langle f(K^*K)x, y\rangle = \int f(\lambda)\mathrm{d}\langle E_\lambda x, y\rangle \tag{116}$$

and

$$\|f(K^*K)x\|^2 = \int f^2(\lambda)\mathrm{d}\|E_\lambda x\|^2. \tag{117}$$

The formular (115)-(117) remain valid if $K : X \to Y$ is a linear bounded operator (for more detail see page 47 in [6]) where the function of a selfadjoint operator $T := K^*K$ in X with spectral family $\{E_\lambda\}_{\lambda\in\mathbb{R}}$ is defined by

$$f(T)x := \int f(\lambda)\mathrm{d}E_\lambda x, \quad x \in \mathcal{D}(f(T)). \tag{118}$$

Let $f : \mathcal{D}(f) \subset X \to Y$ be defined by $f(\lambda) = b^T(I - A\lambda)^{-1}\mathbf{1}$ where I is an (s× s) identity matirx and the matrix A and vector b are given by the Runge-Kutta method. Here the notation $\mathbf{1}$ stands for the (s× 1) vector of 1. Thus, f is a (s× s) matrix of picewise continuous function. The definition (118) provide that (116) holds and thus

$$b^T\left(\boldsymbol{\delta} + \tau_k AS_k^*S_k\right)^{-1}\mathbf{1}S_k^* = f(S_k^*S_k)S_k^* = S_k^*f(S_kS_k^*) = S_k^*b^T\left(\boldsymbol{\delta} + \tau_k AS_kS_k^*\right)^{-1}\mathbf{1}$$

for the linear bounded operator $S_k := F'(w_k)$. By the linearity of the operator $S_kS_k^*$ we have $S_kS_k^*b^TA = S_kS_k^*(\sum_{i=1}^s b_ia_{ij})_{j=1,2,\ldots,s} = (\sum_{i=1}^s b_ia_{ij}S_kS_k^*)_{j=1,2,\ldots,s} = b^TAS_kS_k^*$. Analogously, the assumption (12) can be represented by

$$\|S_k^*b^T(A-I)H_k^{-1}\mathbf{1}y\|_X = \|b^T(A-I)S_k^*H_k^{-1}\mathbf{1}y\|_X \le c_2\|y\|_Y.$$

8. References

[1] A.B. Bakushinsky. The problem of the convergence of the iteratively regularized Gauss-Newton method. *Comput. Math. Math. Phys.*, 32:1353–1359, 1992.

[2] A.B. Bakushinsky and M.Y. Kokurin. *Iterative Methods for Approximate Solution of Inverse Problems.* Dordrecht: Springer, 2004.

[3] B. Blaschke, A. Neubauer, and O. Scherzer. On convergence rates for the iteratively regularized Gauss-Newton method. *IMA J Numer Anal*, 17(3):421–436, 1997.

[4] C. Böckmann. Runge-Kutta type methods for ill-posed problems and the retrieval of aerosol size distributions. *PAMM*, 1(1):486–487, 2002.

[5] C. Böckmann and P. Pornsawad. Iterative Runge-Kutta-type methods for nonlinear ill-posed problems. *Inverse Probl*, 24(2), 2008.

[6] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of inverse problem*. Kluwer Academic Publishers, 1996.

[7] S. George. On convergence of regularized modified Newton's method for nonlinear ill-posed problems. *J Inverse Ill-Posed Probl*, 18(2):133–243, 2010.

[8] T. Hohage. Logarithmic convergence rates of the iteratively regularized Gauss - Newton method for an inverse potential and an inverse scattering problem. *Inverse Probl*, 13(5):1279, 1997.

[9] Q. Jin. A convergence analysis of the iteratively regularized Gauss-Newton method under the Lipschitz condition. *Inverse Probl*, 24(4):045002, 2008.

[10] Q. Jin. A general convergence analysis of some Newton-type methods for nonlinear inverse problems. *SIAM SIAM J Numer Anal*, 49(2):549–573, 2011.

[11] Q. Jin. Further convergence results on the genearal iteratively regularized Gauss-Newton methods under the discrepancy principle. *Math Comput*, 82(283):1647–1665, 2013.

[12] Q. Jin and U. Tautenhahn. On the discrepancy principle for some Newton type methods for solving nonlinear inverse problems. *Numer Math*, 111(4):509–558, 2009.

[13] B. Kaltenbacher. A posteriori parameter choice strategies for some Newton type methods for the regularization of nonlinearill-posed problems. *Numer Math*, 79(4):501–528, 1998.

[14] B. Kaltenbacher, A. Neubauer, and O. Scherzer. *Iterative Regularization Methods for Nonlinear Ill-Posed Problems*. Berlin, Boston: De Gruyter, 2008.

[15] A. Kirsche and C. Böckmann. Padé iteration method for regularization. *Appl Math Comput*, 180(2):648–663, 2006.

[16] P. Mahale and T. Nair. General source conditions for nonlinear ill-posed equations. *Numer Func Anal Opt*, 28(1-2):111–126, 2007.

[17] P. Mathé and B. Hofmann. How general are general source conditions? *Inverse Probl*, 24(1):015009, 2008.

[18] S. Pereverzev and E. Schock. On the adaptive selection of the parameter in the regularization of ill-posed problems. *SIAM J Numer Anal*, 43:2060–2076, 2005.

[19] P. Pornsawad and C. Böckmann. Convergence rate analysis of the first-stage Runge-Kutta-type regularizations. *Inverse Probl*, 26(3), 2010.

[20] R. Ramlau. *Modified Landweber iterations for inverse problems*. PhD thesis, University of Potsdam, 1997.

[21] A. Rieder. On convergence rates of inexact Newton regularizations. *Numer Math*, 88(2):347–365, 2001.

[22] A. Rieder. Runge-Kutta integrators yield optimal regularization schemes. *Inverse Probl*, 21:453–471, 2005.

[23] O. Scherzer. A modified Landweber iteration for solving parameter estimation problems. *Appl Math Optim*, 38(1):45–68, 1998.

[24] U. Tautenhahn. On the asymptotical regularization of nonlinear ill-posed problems. *Inverse Probl*, 10(6):1405, 1994.