# A SIMPLIFIED GENERALIZED GAUSS-NEWTON METHOD FOR NONLINEAR ILL-POSED PROBLEMS

PALLAVI MAHALE AND M. THAMBAN NAIR

ABSTRACT. Iterative regularization methods for nonlinear ill-posed equations of the form F(x) = y, where  $F : D(F) \subset X \to Y$  is an operator between Hilbert spaces X and Y, usually involve calculation of the Fréchet derivatives of F at each iterate and at the unknown solution  $x^{\dagger}$ . In this paper, we suggest a modified form of the generalized Gauss-Newton method which requires the Fréchet derivative of F only at an initial approximation  $x_0$  of the solution  $x^{\dagger}$ . The error analysis for this method is done under a general source condition which also involves the Fréchet derivative only at  $x_0$ . The conditions under which the results of this paper hold are weaker than those considered by Kaltenbacher (1998) for an analogous situation for a special case of the source condition.

#### 1. INTRODUCTION

In this paper, we are interested in finding an approximate solution for a nonlinear ill-posed equation

where  $F: D(F) \subset X \to Y$  is an operator between Hilbert spaces X and Y with inner product and corresponding norm denoted by  $\langle \cdot, \cdot \rangle$  and  $\|\cdot\|$ , respectively, and  $y \in Y$ . We assume that (1.1) has a unique solution  $x^{\dagger}$ . For  $\delta > 0$ , let  $y^{\delta} \in Y$  be an available noisy data with

$$(1.2) ||y - y^{\circ}|| \le \delta.$$

As the given operator equation is ill-posed, its solution need not depend continuously on the data; i.e., small perturbations in the data can cause large deviations in the solutions. In order to overcome this problem, regularization methods are used so as to obtain stable approximate solutions. Iterative regularization methods are one such class of regularization methods. An iterative method with iterations defined by

$$x_{k+1}^{\delta} = \Phi(x_0^{\delta}, x_1^{\delta}, \dots, x_k^{\delta}; y^{\delta}), \quad x_0^{\delta} := x_0,$$

for a known function  $\Phi$  together with a stopping rule which determines a stopping index  $k_{\delta} \in \mathbb{N}$  is called an *iterative regularization method* if  $||x_{k_{\delta}}^{\delta} - x^{\dagger}|| \to 0$  as  $\delta \to 0$ . Here,  $x_0 \in D(F)$  is a known initial approximation of the solution  $x^{\dagger}$ .

Assuming that F possesses Fréchet derivatives F'(x) in a neighbourhood of  $x^{\dagger}$ , Bakushinskii [1] proposed an iterative method, namely, the *iteratively regularized* 

 $\textcircled{\sc c}2008$  American Mathematical Society Reverts to public domain 28 years from publication

Received by the editor July 2, 2007 and, in revised form, January 13, 2008.

<sup>2000</sup> Mathematics Subject Classification. Primary 65J20.

Gauss-Newton method, in which the iterations are defined by

(1.3) 
$$x_{k+1}^{\delta} = x_k^{\delta} - (\alpha_k I + A_{k,\delta}^* A_{k,\delta})^{-1} [A_{k,\delta}^* (F(x_k^{\delta}) - y^{\delta}) + \alpha_k (x_k^{\delta} - x_0)], \quad x_0^{\delta} := x_0,$$

where  $A_{k,\delta} := F'(x_k^{\delta})$  and  $(\alpha_k)$  is a sequence of real numbers satisfying

(1.4) 
$$\alpha_k > 0, \quad 1 \le \frac{\alpha_k}{\alpha_{k+1}} \le \mu_1 \quad \text{and} \quad \lim_{k \to 0} \alpha_k = 0$$

for some constant  $\mu_1 > 1$ . The convergence analysis for (1.3) has been done in [1] under a Hölder-type source condition

$$x_0 - x^{\dagger} = A^* A w, \quad A := F'(x^{\dagger})$$

for some  $w \in X$ . In [4], Blaschke et al. carried out an error analysis for the above method with stopping index  $k_{\delta}$  such that

(1.5) 
$$||F(x_{k_{\delta}}^{\delta}) - y^{\delta}|| \le c\delta < ||F(x_{k}^{\delta}) - y^{\delta}||, \quad k = 0, 1, \dots, k_{\delta} - 1$$

for an appropriate c > 1. It is shown in [4] that the *Hölder-type* source condition

(1.6) 
$$x_0 - x^{\dagger} = (A^* A)^{\nu} w, \quad 0 < \nu \le 1.$$

yields the convergence rate

(1.7) 
$$\|x_{k_{\delta}}^{\delta} - x^{\dagger}\| = \begin{cases} o(\delta^{\frac{2\nu}{2\nu+1}}) & \text{if } 0 < \nu < \frac{1}{2}, \\ O(\sqrt{\delta}) & \text{if } \nu = 1/2. \end{cases}$$

In [5] and [6], Hohage also considered the iteratively regularized Gauss-Newton method (1.3) under the *logarithmic-type* source condition

(1.8) 
$$x_0 - x^{\dagger} = f_{\nu}(A^*A) w,$$

where  $f_{\nu}(\lambda) := \log(1/\lambda)^{-\nu}$ ,  $\nu > 0$ , with stopping index  $k_{\delta}$  as in (1.5), and obtained the error bound as

$$\|x_{k_{\delta}}^{\delta} - x^{\dagger}\| = O(\log(1/\delta)^{-\nu})$$

Recently, Langer and Hohage [8] extended the analysis in [5] and [6] by considering (1.3) with the stopping rule (1.5) under a general source condition of the form

(1.9) 
$$x_0 - x^{\dagger} = f(A^*A)w,$$

yielding the error estimate

$$||x_{k_{\delta}}^{\delta} - x^{\dagger}|| = O(f(u^{-1}(\delta))).$$

Here,  $f : [0, ||A||^2] \to [0, \infty)$  is a monotonically increasing continuous function satisfying f(0) = 0 and  $u(\lambda) = \lambda^{1/2} f(\lambda)$ . Here, we want to state that in all the above-mentioned error estimates, the results hold for the limit  $\delta \to 0$ .

Note that the source condition (1.9) includes the cases (1.6) and (1.8).

In [2], Bakushinskii genearalized the procedure in [1] by considering a generalized form of the regularized Gauss-Newton method in which the iterations are defined by

(1.10) 
$$x_{k+1}^{\delta} = x_0 - g_{\alpha_k} (A_{k,\delta}^* A_{k,\delta}) A_{k,\delta}^* [F(x_k^{\delta}) - y^{\delta} - A_{k,\delta} (x_k^{\delta} - x_0)], \quad x_0^{\delta} := x_0,$$

where  $A_{k,\delta} = F'(x_k^{\delta})$  and each  $g_{\alpha}$  for  $\alpha > 0$  is a piecewise continuous function and  $(\alpha_k)$  is a sequence of real numbers satisfying (1.4). In [3], Kaltenbacher considered

the above generalized procedure under the stopping rule in which the stopping index  $k_0$  is chosen such that

(1.11) 
$$\max\{\|F(x_{k_0-1}^{\delta}) - y^{\delta}\|, \ \hat{\beta}_{k_0}\} \le \tau \delta < \max\{\|F(x_{k-1}^{\delta}) - y^{\delta}\|, \ \hat{\beta}_k\}$$

for all  $k \in \{1, 2, \dots, k_0 - 1\}$  and for some  $\tau > 1$ , where

$$\hat{\beta}_k := \|F(x_{k-1}^{\delta}) - y^{\delta} + A_{k-1,\delta}(x_k^{\delta} - x_{k-1}^{\delta})\|, \quad k = 1, 2, ..., k_0,$$

and the error estimate is obtained under the Hölder-type source condition (1.6).

1.1. The new method and the new stopping rule. We observe that in the iterative procedure (1.3) as well as its generalization (1.10) it is necessary to calculate the Fréchet derivative at each iterate. In this paper we define a new iteration procedure

(1.12) 
$$x_{k+1}^{\delta} = x_0 - g_{\alpha_k} (A_0^* A_0) A_0^* [F(x_k^{\delta}) - y^{\delta} - A_0(x_k^{\delta} - x_0)], \quad x_0^{\delta} := x_0,$$

where  $A_0 := F'(x_0)$ ,  $(\alpha_k)$  is the sequence satisfying (1.4) and each  $g_\alpha$  for  $\alpha > 0$ is a positive real-valued piecewise continuous function defined on [0, M] with  $M \ge ||A_0||^2$ . As iterations in (1.12) involve the Fréchet derivative of F only at one point  $x_0$ , the calculations in (1.12) are simpler than in (1.10). Due to the simplicity of (1.12), we name this iteration as a *simplified generalized Gauss-Newton method*. We choose the stopping index  $k_\delta$  for this iteration as the positive integer which satisfies (1.13)

$$\max\{\|F(x_{k_{\delta}-1}^{\delta}) - y^{\delta}\|, \ \tilde{\beta}_{k_{\delta}}\} \le \tau \delta < \max\{\|F(x_{k-1}^{\delta}) - y^{\delta}\|, \ \tilde{\beta}_{k}\}, \qquad 1 \le k < k_{\delta}.$$

Here  $\tau > 1$  is a sufficiently large constant not depending on  $\delta$ , and

$$\tilde{\beta}_k := \|F(x_{k-1}^{\delta}) - y^{\delta} + A_0(x_k^{\delta} - x_{k-1}^{\delta})\|.$$

We also observe that the source condition (1.9), as well as its special cases (1.6), (1.8), involves the Fréchet derivative at the exact solution  $x^{\dagger}$  which is practically an unknown quantity. So, in analogy to (1.9), we shall consider a source condition which depends on the Fréchet derivative of F only at  $x_0$ .

#### 2. Basic assumptions

In this section we consider some of the basic assumptions under which the results of the subsequent sections hold. First we consider the source condition.

Assumption 2.1. (i) There exists a continuous, strictly monotonically increasing function  $\varphi : (0, M] \to (0, \infty)$  with  $M \ge \|F'(x_0)\|^2$  satisfying  $\lim_{\lambda \to 0} \varphi(\lambda) = 0$  and  $\rho > 0$  such that

(2.14) 
$$x_0 - x^{\dagger} = [\varphi(A_0^*A_0)]^{1/2} w, \quad A_0 := F'(x_0)$$

for some  $w \in X$  with  $||w|| \leq \rho$ .

(ii) The function  $\psi : (0, \varphi(M)] \to (0, M\varphi(M)]$  defined by

(2.15) 
$$\psi(\lambda) = \lambda \varphi^{-1}(\lambda), \quad \lambda \in (0, \varphi(M)],$$

is strictly monotonically increasing and convex.

We observe that the source condition (2.14) represents a class of source conditions and it is also suitable for both mildly and severely ill-posed problems, in particular, *Hölder-type* source conditions, i.e., with  $\varphi(\lambda) = \lambda^{\nu}$ , and logarithmic source conditions, i.e., with  $\varphi(\lambda) = [\log(1/\lambda)]^{-\nu}$  (see [5]). We also note that the source condition (2.14) involves the known quantity  $x_0$  whereas the other source conditions (1.6) and (1.8) as well as their generalization (1.9) require the knowledge of the unknown quantity  $x^{\dagger}$ .

Next we make an assumption on the operator F.

Assumption 2.2. (a) The operator F possesses Fréchet derivatives F'(x) in  $B_r(x_0) := \{ x \in D(F) : ||x - x_0|| < r \} \text{ where } r \ge 2 ||x^{\dagger} - x_0||.$ 

(b) There exists a constant  $C_0 > 0$ , and for each  $u, v \in B_r(x_0)$  there exists a linear operator  $R_u^v: Y \to X$  such that

(2.16) 
$$F'(v) = R_u^v F'(u), \quad ||I - R_u^v|| \le C_0.$$

Assumptions similar to (2.16) are considered by several authors for convergence analysis of the nonlinear ill-posed equations (cf. [3], [4], [5]). It is shown in these references that several parameter identification problems useful in applications satisfy (2.16). But, for many ill-posed problems, it is an open question whether such conditions are satisfied.

For each  $\alpha > 0$ , let  $g_{\alpha} : (0, M] \to (0, \infty)$  be a piecewise continuous function, involved in the method given by (1.12). We shall also assume that  $\varphi$  and  $g_{\alpha}$ ,  $\alpha \geq 0$ , have some additional properties as given in the following two assumptions.

Assumption 2.3. There exists a positive integer  $\mu_2 > 1$  such that

(2.17) 
$$1 \le \frac{\varphi(\alpha_n)}{\varphi(\alpha_{n+1})} \le \mu_2, \quad \forall n \in \mathbb{N}$$

We note that the *Hölder-type* source condition, i.e., with  $\varphi(\lambda) = \lambda^{\nu}$ , and the logarithmic source condition, i.e., with  $\varphi(\lambda) = [\log(1/\lambda)]^{-\nu}$ , satisfy (2.17) for  $\mu_2 =$  $\mu_1^{\nu}$  and  $\mu_2 = 1$ , respectively.

Assumption 2.4. There exist positive real numbers  $\omega > 0$ ,  $\omega_1 > 0$ ,  $c_0 > 0$  such that

- (i)  $\sup_{\substack{0<\lambda\leq M}} |[1-\lambda g_{\alpha}(\lambda)]\sqrt{\varphi(\lambda)}| \leq \omega\sqrt{\varphi(\alpha)},$ (ii)  $\sup_{\substack{0<\lambda\leq M}} |[1-\lambda g_{\alpha}(\lambda)]\sqrt{\lambda\varphi(\lambda)}| \leq \omega_{1}\sqrt{\alpha\varphi(\alpha)},$
- (iii)  $\sup_{\substack{0<\lambda\leq M\\0<\lambda\leq M}}^{-}\sqrt{\lambda}g_{\alpha}(\lambda)\leq\frac{c_{0}}{\sqrt{\alpha}},$ (iv)  $\sup_{\substack{0<\lambda\leq M}}|1-\lambda g_{\alpha}(\lambda)|\leq 1.$

As examples, let us consider some of the well-known regularization methods such as

- (a) Ordinary Tikhonov regularization:  $g_{\alpha}(\lambda) = 1/(\lambda + \alpha)$ ,
- (b) Iterated Tikhonov regularization of order m:  $g_{\alpha}(\lambda) = [1 \alpha^m / (\lambda + \alpha)^m] / \lambda$ ,
- (c) Regularized singular-value decomposition:  $g_{\alpha}(\lambda) = \begin{cases} 1/\lambda, & \lambda \ge \alpha, \\ 1/\alpha, & \lambda \le \alpha, \end{cases}$ (d) A summation in the second seco
- (d) Asymptotical regularization:  $g_{\alpha}(\lambda) = (1 e^{-\lambda/\alpha})/\lambda$

Let us consider the source functions associated with the Hölder-type source conditions, and the *logarithmic-type* source conditions, namely,  $\varphi(\lambda) = \lambda^{\nu}$  and  $\varphi(\lambda) = \lambda^{\nu}$  $[\log(1/\lambda)]^{-\nu}$ , respectively.

First let us consider the case of  $\varphi(\lambda) = \lambda^{\nu}, \nu > 0$ . It can be seen that (i) and (ii) in Assumption 2.4 hold for (a), (b), (c) with  $\omega = 1 = \omega_1$ , and for (d) with

 $\omega = (\nu/2)^{\nu/2} e^{-\nu/2}$  and  $\omega_1 = ((\nu+1)/2)^{(\nu+1)/2} e^{-(\nu+1)/2}$ , and with  $\nu$  satisfying  $0 < \nu \leq 1$  for (a),  $0 < \nu \leq 2m - 1$  for (b), and  $\nu > 0$  for (c) and (d). It can also be verified that Assumption 2.4 (iii) holds for (a), (c), (d) with  $c_0 = 1$ , and for (b) with  $c_0 = m$ . Now, in view of Lemma 3.13 in [6], which is the same as Lemma 4 in [7], we can also assert that the conditions in Assumption 2.4 are satisfied for the source function  $\varphi(\lambda) = [\log(1/\lambda)]^{-\nu}$  for all  $\nu > 0$ . We may also observe that regularization methods (a), (b), (c), (d) satisfy Assumption 2.4 (iv).

#### 3. Error analysis

3.1. **Background results.** Now we discuss some of the results which are essential for the error analysis of the simplified generalized Gauss-Newton method.

Throughout this section we use the following notation:

$$\begin{array}{lll}
e_{k}^{\delta} & := & x_{k}^{\delta} - x^{\dagger}, \\
g_{k}(\lambda) & := & g_{\alpha_{k}}(\lambda), \\
r_{k}(\lambda) & := & 1 - \lambda g_{k}(\lambda), \\
\beta_{k} & := & \|A_{0}r_{k-1}(A_{0}^{*}A_{0})(x_{0} - x^{\dagger})\|.
\end{array}$$

First we observe from Taylor's formula that for  $u, v \in B_r(x_0)$ ,

(3.18) 
$$F(v) - F(u) - F'(x_0)(v-u) = \int_0^1 [F'(u+t(v-u)) - F'(x_0)](v-u)dt.$$

Hence, by the Assumption 2.2, it follows that

$$F(v) - F(u) - F'(x_0)(v - u) = \int_0^1 (R_{x_0}^{u + t(v - u)} - I)F'(x_0)(v - u)dt$$

and

(3.19) 
$$||F(v) - F(u) - F'(x_0)(v - u)|| \le C_0 ||F'(x_0)(v - u)||.$$

**Lemma 3.1.** Let (1.2) hold and let the iterates  $x_k^{\delta}$  be defined by (1.12) with a real sequence  $(\alpha_k)$  satisfying (1.4). Moreover, let Assumptions 2.2, 2.3, 2.4 hold and let  $k_{\delta}$  be chosen according to the stopping rule (1.13). Then

$$(3.20) \qquad \qquad |\beta_k - \tilde{\beta}_k| \le \delta + C_0 \|A_0 e_{k-1}^{\delta}\|$$

and

(3.21) 
$$\delta \le c_1 \beta_k + c_2 \|A_0 e_{k-1}^{\delta}\|, \quad k \in \{0, 1, \dots, k_{\delta} - 1\},$$

where  $c_1 = 1/(\tau - 1)$ ,  $c_2 = (1 + C_0)/(\tau - 1)$ ,

$$\beta_k = \|A_0 r_{k-1} (A_0^* A_0) (x_0 - x^{\dagger})\| \quad and \quad \tilde{\beta}_k = \|F(x_{k-1}^{\delta}) - y^{\delta} + A_0 (x_k^{\delta} - x_{k-1}^{\delta})\|.$$

*Proof.* We observe that

$$\begin{split} \tilde{\beta}_k &= \|F(x_{k-1}^{\delta}) - y^{\delta} + A_0(x_k^{\delta} - x_{k-1}^{\delta})\| \\ &= \|r_{k-1}(A_0A_0^*)(F(x_{k-1}^{\delta}) - y^{\delta} - A_0(x_{k-1}^{\delta} - x_0))\|. \end{split}$$

Thus

$$\begin{split} |\tilde{\beta}_{k} - \beta_{k}| &\leq |\|r_{k-1}(A_{0}A_{0}^{*})(F(x_{k-1}^{\delta}) - y^{\delta} - A_{0}(x_{k-1}^{\delta} - x_{0}))\| \\ &- \|A_{0}r_{k-1}(A_{0}^{*}A_{0})(x_{0} - x^{\dagger})\|| \\ &= \|r_{k-1}(A_{0}A_{0}^{*})(F(x_{k-1}^{\delta}) - y^{\delta} + A_{0}(x^{\dagger} - x_{k-1}^{\delta}))\| \\ &\leq \|r_{k-1}(A_{0}A_{0}^{*})(y - y^{\delta})\| + \|r_{k-1}(A_{0}A_{0}^{*})[F(x_{k-1}^{\delta}) - y - A_{0}(x_{k-1}^{\delta} - x^{\dagger})]\|. \end{split}$$

Hence, using (1.2), Assumption 2.4 (iv) and (3.19), we get  $|\tilde{\beta}_k - \beta_k| \leq \delta + C_0 ||A_0 e_{k-1}||$ , proving (3.20).

To prove (3.21) we consider two cases.

Case 1: Suppose  $||F(x_{k-1}^{\delta}) - y^{\delta}|| \geq \tilde{\beta}_k$ . As the iteration is stopped according to the rule (1.13), we have

$$\tau \delta < \|F(x_{k-1}^{\delta}) - y^{\delta}\|$$
  
$$\leq \delta + (1+C_0) \|A_0 e_{k-1}^{\delta}\|.$$

Thus, we get

(3.22) 
$$\delta < \frac{(1+C_0)\|A_0 e_{k-1}^{\delta}\|}{\tau - 1}$$

Case 2: Suppose  $\tilde{\beta}_k \ge \|F(x_{k-1}^{\delta}) - y^{\delta}\|$ . Then, by (3.20), we obtain

(3.23)  $\tau \delta \leq \tilde{\beta}_k \leq \delta + C_0 \|A_0 e_{k-1}^{\delta}\| + \beta_k$ 

so that

$$(\tau - 1)\delta \le C_0 ||A_0 e_{k-1}^{\delta}|| + \beta_k$$

which gives

(3.24) 
$$\delta \le \frac{\beta_k}{\tau - 1} + \frac{C_0 \|A_0 e_{k-1}^{\delta}\|}{\tau - 1}.$$

From (3.22), (3.24), we get  $\delta \leq c_1 \beta_k + c_2 ||A_0 e_{k-1}||$ , as required in (3.21).

The following technical lemma is used in due course.

**Lemma 3.2.** Suppose  $\tau > 1 + \sqrt{\mu_1 \mu_2}$  and

$$0 < C_0 < \min\{1/2, \ [\tau - 1 - \sqrt{\mu_1 \mu_2}]/\tau \sqrt{\mu_1 \mu_2}\}$$

Then

$$b := \sqrt{\mu_1 \mu_2} (c_2 + C_0) < 1$$

with  $c_2$  as in Lemma 3.1.

*Proof.* We have

$$b = (\mu_1 \mu_2)^{1/2} (c_2 + C_0)$$
  

$$\leq (\mu_1 \mu_2)^{1/2} \left( \frac{(1+C_0)}{(\tau-1)} + C_0 \right)$$
  

$$= (\mu_1 \mu_2)^{1/2} (1 + \tau C_0) / (\tau-1)$$
  

$$< \frac{(\mu_1 \mu_2)^{1/2}}{(\tau-1)} \left( 1 + \frac{(\tau-1 - \sqrt{\mu_1 \mu_2})}{\sqrt{\mu_1 \mu_2}} \right)$$
  

$$= 1.$$

The next proposition prescribes conditions which ensure, among other useful estimates,  $x_k^{\delta} \in B_r(x_0)$  for  $k \in \{0, 1, \dots, k_{\delta}\}$ .

**Proposition 3.3.** Let the assumptions of Lemma 3.1 and Assumption 2.1 hold. Let

 $a := \omega_1 \sqrt{\mu_1 \mu_2} (1+c_1) \rho, \quad b = \sqrt{\mu_1 \mu_2} (c_2 + C_0), \quad C_1 := a/\rho(1-b)$ 

with  $c_1$  and  $c_2$  as in Lemma 3.1. Assume further that  $\tau$  and  $C_0$  satisfy the conditions in Lemma 3.2, and

(3.25) 
$$(1-b)\sqrt{Mr} \le 2a, \quad 2[\omega + c_0(c_1\omega_1 + c_2C_1 + C_0C_1)]\varphi(\alpha_0)^{1/2}\rho < r.$$

Then for  $k \in \{1, \ldots, k_{\delta}\}$  with  $k_{\delta}$  as in (1.13),

(3.26) 
$$||A_0(x_k^{\delta} - x^{\dagger})|| \le C_1 \rho \sqrt{\alpha_k \varphi(\alpha_k)}$$

(3.27)  $||x_k^{\delta} - x^{\dagger}|| < r/2,$ 

(3.28) 
$$\beta^2 \delta^2 / \rho^2 \le \alpha_k \varphi(\alpha_{k-1}),$$

where  $\beta = 1/(\omega_1 c_1 + c_2 C_1)$ .

*Proof.* From (1.12), we have

$$e_{k}^{\delta} = x_{k}^{\delta} - x^{\dagger}$$
  
=  $x_{0} - x^{\dagger} - g_{k-1}(A_{0}^{*}A_{0})A_{0}^{*}\{F(x_{k-1}^{\delta}) - y^{\delta} - A_{0}(x_{k-1} - x_{0})\}$   
(3.29) =  $r_{k-1}(A_{0}^{*}A_{0})(x_{0} - x^{\dagger}) - g_{k-1}(A_{0}^{*}A_{0})A_{0}^{*}\{F(x_{k-1}^{\delta}) - y^{\delta} - A_{0}e_{k-1}^{\delta}\}.$ 

Hence,

$$A_0 e_k^{\delta} = A_0 r_{k-1} (A_0^* A_0) (x_0 - x^{\dagger}) - g_{k-1} (A_0 A_0^*) A_0 A_0^* \{ F(x_{k-1}^{\delta}) - y^{\delta} - A_0 e_{k-1}^{\delta} \}.$$

Thus, we get

 $\begin{aligned} \|A_0 e_k^{\delta}\| &\leq \|A_0 r_{k-1} (A_0^* A_0) (x_0 - x^{\dagger})\| + \|g_{k-1} (A_0 A_0^*) A_0 A_0^*\| \| \{F(x_{k-1}^{\delta}) - y^{\delta} - A_0 e_{k-1}^{\delta} \} \|. \end{aligned}$ Using (1.2), Assumption 2.4 (ii), and (3.19), we get

$$\|A_0 e_k^{\delta}\| \le \omega_1 \sqrt{\alpha_{k-1} \varphi(\alpha_{k-1})} \rho + \delta + C_0 \|A_0 e_{k-1}^{\delta}\|.$$

From (3.21), we have

$$\|A_0 e_k^{\delta}\| \le \omega_1 \sqrt{\alpha_{k-1} \varphi(\alpha_{k-1})} \rho + c_1 \beta_k + c_2 \|A_0 e_{k-1}^{\delta}\| + C_0 \|A_0 e_{k-1}^{\delta}\|.$$

Again using Assumptions 2.2, 2.4 (ii) we get

(3.30) 
$$||A_0 e_k^{\delta}|| \le \omega_1 (1+c_1) \sqrt{\alpha_{k-1} \varphi(\alpha_{k-1})} \rho + c_2 ||A_0 e_{k-1}^{\delta}|| + C_0 ||A_0 e_{k-1}^{\delta}||.$$

Dividing both sides of (3.30) by  $\sqrt{\alpha_k \varphi(\alpha_k)}$ , we get

$$\frac{\|A_0 e_k^{\delta}\|}{\sqrt{\alpha_k \varphi(\alpha_k)}} \le \frac{\omega_1 (1+c_1) \sqrt{\alpha_{k-1} \varphi(\alpha_{k-1})} \rho}{\sqrt{\alpha_k \varphi(\alpha_k)}} + \frac{(c_2+C_0) \|A_0 e_{k-1}^{\delta}\| \sqrt{\alpha_{k-1} \varphi(\alpha_{k-1})}}{\sqrt{\alpha_k \varphi(\alpha_k)} \sqrt{\alpha_{k-1} \varphi(\alpha_{k-1})}}$$

Using (1.4) and (2.17), we get

$$\frac{\|A_0 e_k^{\delta}\|}{\sqrt{\alpha_k \varphi(\alpha_k)}} \le \omega_1 \sqrt{\mu_1 \mu_2} (1+c_1)\rho + \frac{(c_2+C_0)\|A_0 e_{k-1}^{\delta}\|\sqrt{\mu_1 \mu_2}}{\sqrt{\alpha_{k-1} \varphi(\alpha_{k-1})}}.$$

Denoting  $\gamma_k = ||A_0 e_k^{\delta}|| / \sqrt{\alpha_k \varphi(\alpha_k)}$ , we get  $\gamma_k \le \omega_1 \sqrt{\mu_1 \mu_2} (1+c_1) \rho + (c_2 + C_0) \sqrt{\mu_1 \mu_2} \gamma_{k-1} = a + b \gamma_{k-1}, \qquad k = 0, 1, \dots, k_{\delta}.$  Now by Lemma 3.2 we have b < 1 and the condition  $(1-b)r\sqrt{M} \leq 2a$  implies  $||A_0e_0|| \leq a/(1-b)$ . Hence, it follows that

$$\gamma_k \le a/(1-b)$$
  $k \in \{0, 1, ..., k_\delta\},$ 

which gives

(3.31) 
$$\|A_0 e_k^{\delta}\| \le C_1 \sqrt{\alpha_k \varphi(\alpha_k)} \rho, \quad k \in \{0, 1, \dots, k_{\delta}\}.$$

From (3.29), we have

$$e_{k}^{\delta} = r_{k-1}(A_{0}^{*}A_{0})(x_{0} - x^{\dagger}) - g_{k-1}(A_{0}^{*}A_{0})A_{0}^{*}\{F(x_{k-1}^{\delta}) - y^{\delta} - A_{0}e_{k-1}^{\delta}\}$$
  
$$= r_{k-1}(A_{0}^{*}A_{0})\varphi(A_{0}^{*}A_{0})^{1/2}w - g_{k-1}(A_{0}^{*}A_{0})A_{0}^{*}\{F(x_{k-1}^{\delta}) - y^{\delta} - A_{0}e_{k-1}^{\delta}\}.$$

Using (1.2), Assumptions 2.2, (i) and (iii) in 2.4, we get

(3.32) 
$$\|e_k^{\delta}\| \le \omega \sqrt{\varphi(\alpha_{k-1})}\rho + \frac{c_0\delta}{\sqrt{\alpha_{k-1}}} + \frac{c_0C_0\|A_0e_{k-1}^{\delta}\|}{\sqrt{\alpha_{k-1}}}$$

Now from (3.21), (3.31) we have

$$\delta \leq c_1 \beta_k + c_2 \|A_0 e_{k-1}^{\delta}\|$$
  
$$\leq \omega_1 c_1 \sqrt{\alpha_{k-1} \varphi(\alpha_{k-1})} \rho + c_2 C_1 \sqrt{\alpha_{k-1} \varphi(\alpha_{k-1})} \rho$$
  
$$\leq (\omega_1 c_1 + c_2 C_1) \sqrt{\alpha_{k-1} \varphi(\alpha_{k-1})} \rho,$$

which gives

(3.33) 
$$\frac{\delta}{\sqrt{\alpha_{k-1}}} \le (\omega_1 c_1 + c_2 C_1) \sqrt{\varphi(\alpha_{k-1})} \rho.$$

Using (3.31) and (3.33) in (3.32) we get

$$\begin{aligned} \|e_k^{\delta}\| &\leq \omega\sqrt{\varphi(\alpha_{k-1})}\rho + c_0(\omega_1c_1 + c_2C_1)\sqrt{\varphi(\alpha_{k-1})}\rho + \frac{(c_0C_0)C_1\sqrt{\alpha_{k-1}}\varphi(\alpha_{k-1})\rho}{\sqrt{\alpha_{k-1}}} \\ &= \omega\sqrt{\varphi(\alpha_{k-1})}\rho + c_0(\omega_1c_1 + c_2C_1)\sqrt{\varphi(\alpha_{k-1})}\rho + c_0C_0C_1\sqrt{\varphi(\alpha_{k-1})}\rho. \end{aligned}$$

Thus, using the condition (3.25), we have

 $||x_k^{\delta} - x^{\dagger}|| = ||e_k^{\delta}|| \le [\omega + c_0(\omega_1 c_1 + c_2 C_1 + C_0 C_1)]\sqrt{\varphi(\alpha_0)}\rho < r/2, \quad k \in \{1, \dots, k_{\delta}\}.$ From (3.21), we have

$$\delta \le (c_1 \beta_k + c_2 \| A_0 e_{k-1}^{\delta} \|).$$

Using Assumption 2.4 (ii) and (3.31), we get  $\delta \leq c_1 \omega_1 \rho \sqrt{\alpha_{k-1} \varphi(\alpha_{k-1})} + c_2 C_1 \rho \sqrt{\alpha_{k-1} \varphi(\alpha_{k-1})} = (\omega_1 c_1 + c_2 C_1) \rho \sqrt{\alpha_{k-1} \varphi(\alpha_{k-1})},$ 

which gives 
$$\beta^2 \delta^2 / \rho^2 \leq \alpha_{k-1} \varphi(\alpha_{k-1})$$
 as required.

Remark 3.4. (a) We observe from the definition of  $\psi$  in (2.15) and the relation (3.28) that

$$\delta \le \rho \sqrt{\psi(\varphi(\alpha_{k-1}))} / \beta,$$

which gives

(3.34) 
$$\varphi^{-1}\psi^{-1}\left(\frac{\beta^2\delta^2}{\rho^2}\right) \le \alpha_{k-1} \qquad k \in \{1, 2, ..., k_\delta\}.$$

(b) We also note that the assumptions in Proposition 3.3 do not require any a priori knowledge of the exact solution  $x^{\dagger}$  except that  $x^{\dagger} - x_0$  satisfies the source

condition (2.14). Thus, the assumptions of Proposition 3.3 are more realistic than the assumptions considered in the literature (see, e.g., [6], [4], [3]).

(c) We observe that for the particular case of  $\varphi(\lambda) = \lambda^{2\nu}$ , results in Lemma 3.1 and Proposition 3.3 are analogous to the results of Lemma 2.1 in [3]. But, if we compare the assumptions of both under the framework of the present paper, we note that the conditions in Lemma 2.1 in [3] are more stringent than the conditions in Lemma 3.1 and Proposition 3.3. For example, in Lemma 3.1, we get the estimate (3.21) with

$$c_1 = 1/(\tau - 1)$$
 and  $c_2 = (1 + C_0)/(\tau - 1)$ .

In [3], an estimate similar to (3.21) is obtained by replacing both  $c_1$  and  $c_2$  above by  $c_2$ . Note that  $c_1 < c_2$ . We note that in Proposition 3.3 we need  $b = \sqrt{\mu_1 \mu_2}(c_2 + C_0) < 1$ , where  $\mu_2$  can be taken as  $\mu_1^{2\nu}$ . In place of the above inequality, Lemma 2.1 in [3] uses the inequality

(3.35) 
$$\tilde{b} := \sqrt{\mu_1 \mu_2} \left( \frac{c_2}{1 - C_0} + \frac{C_0 (1 + C_0)}{1 - C_0} \right) < 1.$$

Note that

$$(3.36) b < b.$$

Also, in Proposition 3.3 we use the condition

$$\theta := 2(\|x_0 - x^{\dagger}\| + c_0(c_1\omega_1 + c_2C_1 + C_0C_1)\alpha_0^{\nu}\rho) < r,$$

whereas Lemma 2.1 in [3] uses the condition

$$\tilde{\theta} := 2(\|x_0 - x^{\dagger}\| + c_0(C_0(1 + C_0)\tilde{C}_1 + c_2\tilde{C}_1 + (c_2\omega_1)/(1 - C_0))\alpha_0^{\nu}\rho) < r,$$

where  $C_1$  is as in Proposition 3.3 and  $\tilde{C}_1 = \tilde{a}/(1-\tilde{b})\rho$  with

$$\tilde{a} := \frac{\omega_1 \sqrt{\mu_1 \mu_2}}{1 - C_0} (1 + c_2 / (1 - C_0))\rho,$$

and  $\tilde{b}$  is as in (3.35). From (3.36) and the fact that  $C_1 < \tilde{C}_1$ , we have  $\theta < \tilde{\theta}$ . Thus, under the setting of the present paper, the conditions of Lemma 2.1 of [3] are stronger than the conditions of Lemma 3.1 and Proposition 3.3.

**Lemma 3.5.** Let (1.2) and Assumption 2.2 be satisfied and let  $C_0 < 1/2$ . Let the iterations (1.12) be stopped according to the stopping rule (1.13) and for  $k \in \{1, \ldots, k_{\delta}\}$ , let

$$f(\delta, k) := \max\{\|F(x_{k-1}^{\delta}) - y^{\delta}\|, \|F(x_{k}^{\delta}) - y^{\delta}\|\}.$$

Then

$$(3.37) f(\delta, k_{\delta}) \le \tau_0 \delta_{\delta}$$

and for  $k \in \{1, ..., k_{\delta} - 1\}$ ,

$$(3.38) f(\delta,k) > \tau_1\delta,$$

where

$$\tau_0 = \frac{2C_0 + \tau}{1 - 2C_0}, \qquad \tau_1 = \frac{\tau(1 - C_0) - 2C_0}{1 + C_0}.$$

*Proof.* Let  $k \in \{1, \ldots, k_{\delta}\}$ . We observe that

$$\begin{aligned} ||F(x_{k}^{\delta}) - y^{\delta}|| &- \tilde{\beta}_{k}| &= |||F(x_{k}^{\delta}) - y^{\delta}|| - ||F(x_{k-1}^{\delta}) + A_{0}(x_{k}^{\delta} - x_{k-1}^{\delta}) - y^{\delta}|| \\ &\leq ||F(x_{k}^{\delta}) - F(x_{k-1}^{\delta}) - A_{0}(x_{k}^{\delta} - x_{k-1}^{\delta})|| \\ &\leq ||F(x_{k}^{\delta}) - F(x^{\dagger}) - A_{0}e_{k}^{\delta}|| \\ &+ ||F(x_{k-1}^{\delta}) - F(x^{\dagger}) - A_{0}e_{k-1}^{\delta}|| \\ \end{aligned}$$

$$(3.39) \leq C_{0}(||A_{0}e_{k-1}^{\delta}|| + ||A_{0}e_{k}^{\delta}||).$$

(0.00)

From (3.19), for any  $u, v \in X$ , we have

$$||A_0(v-u)|| - ||F(v) - F(u)|| \le C_0 ||A_0(v-u)||,$$

which implies

$$(1 - C_0) || ||A_0(v - u)|| \le ||F(v) - F(u)||;$$

equivalently,

(3.40) 
$$||A_0(v-u)|| \le \frac{||F(v) - F(u)||}{1 - C_0}$$

Using (3.40) in (3.39), we have

$$\begin{aligned} |||F(x_k^{\delta}) - y^{\delta}|| &- \tilde{\beta}_k| &\leq \frac{C_0 ||F(x_{k-1}^{\delta}) - F(x^{\dagger})||}{1 - C_0} + \frac{C_0 ||F(x_k^{\delta}) - F(x^{\dagger})||}{1 - C_0} \\ &= C_0'(||F(x_k^{\delta}) - F(x^{\dagger})|| + ||F(x_{k-1}^{\delta}) - F(x^{\dagger})||) \\ &\leq C_0'(||F(x_k^{\delta}) - y^{\delta}|| + ||F(x_{k-1}^{\delta}) - y^{\delta}|| + 2\delta), \end{aligned}$$

where  $C'_0 := C_0/(1 - C_0)$ . From this we have

$$\tilde{\beta}_{k} \leq (1 + C'_{0}) \|F(x_{k}^{\delta}) - y^{\delta}\| + C'_{0} \|F(x_{k-1}^{\delta}) - y^{\delta}\| + 2C'_{0}\delta \leq (1 + 2C'_{0})f(\delta, k) + 2C'_{0}\delta.$$

We also observe that

(3.42)  $||F(x_{k-1}^{\delta}) - y^{\delta}|| \le (1 + C'_0) ||F(x_{k-1}^{\delta}) - y^{\delta}|| + C'_0 ||F(x_k^{\delta}) - y^{\delta}|| + 2C'_0 \delta.$ Combining (3.41) and (3.42), we get (3.43)  $\max\{\tilde{\beta}_k, ||F(x_{k-1}^{\delta}) - y^{\delta}||\} \le (1 + 2C'_0)f(\delta, k) + 2C'_0 \delta.$ 

We also have

$$-C_0'(\|F(x_k^{\delta}) - y^{\delta}\| + \|F(x_{k-1}^{\delta}) - y^{\delta}\| - 2\delta) \le \tilde{\beta}_k - \|F(x_k^{\delta}) - y^{\delta}\|,$$

which gives

$$(1 - C'_0) \|F(x_k^{\delta}) - y^{\delta}\| - 2C'_0 \delta \le \tilde{\beta}_k + C'_0 \|F(x_{k-1}^{\delta}) - y^{\delta}\|$$

and

(3.41)

(3.44) 
$$\frac{(1-C_0')\|F(x_k^{\delta})-y^{\delta}\|}{(1+C_0')} - \frac{2C_0'\delta}{1+C_0'} \le \max\{\tilde{\beta}_k, \|F(x_{k-1}^{\delta})-y^{\delta}\|\}.$$

We also observe that

$$\frac{(1-C_0')\|F(x_{k-1}^{\delta})-y^{\delta}\|}{(1+C_0')} - \frac{2C_0'\delta}{(1+C_0')} \le \|F(x_{k-1}^{\delta})-y^{\delta}\|,$$

which gives

(3.45) 
$$\frac{(1-C_0')\|F(x_{k-1}^{\delta})-y^{\delta}\|}{(1+C_0')} - \frac{2C_0'\delta}{(1+C_0')} \le \max\{\tilde{\beta}_k, \|F(x_{k-1}^{\delta})-y^{\delta}\|\}.$$

From (3.44) and (3.45) we have

(3.46) 
$$\frac{(1-C'_0)}{(1+C'_0)}f(\delta,k) - \frac{2C'_0\delta}{(1+C'_0)} \le \max\{\tilde{\beta}_k, \|F(x_{k-1}^{\delta}) - y^{\delta}\|\}.$$

From (3.43) and (3.46) we get

$$\frac{(1-C'_0)}{(1+C'_0)}f(\delta,k) - \frac{2C'_0\delta}{(1+C'_0)} \leq \max\{\tilde{\beta}_k, \|F(x^{\delta}_{k-1}) - y^{\delta}\|\} \leq (1+2C'_0)f(\delta,k) + 2C'_0\delta.$$

Now, using the stopping rule (1.13),

(3.47) 
$$\frac{(1-C'_0)}{(1+C'_0)}f(\delta,k_{\delta}) - \frac{2C'_0\delta}{(1+C'_0)} \le \max\{\tilde{\beta}_{k_{\delta}}, \|F(x^{\delta}_{k_{\delta}-1}) - y^{\delta}\|\}$$

and

(3.48) 
$$\tau \delta < \max\{\tilde{\beta}_k, \|F(x_{k-1}^{\delta}) - y^{\delta}\|\} \le (1 + 2C_0')f(\delta, k) + 2C_0'\delta.$$

From (3.47) and (3.48), we have

$$\frac{(1-C_0')}{(1+C_0')}f(\delta,k_{\delta}) - \frac{2C_0'\delta}{(1+C_0')} \le \tau\delta < (1+2C_0')f(\delta,k) + 2C_0'\delta.$$

Now, let

$$\tau_0 = \left(\frac{2C'_0}{1+C'_0} + \tau\right) \frac{(1+C'_0)}{(1-C'_0)} = \frac{2C_0 + \tau}{1-2C_0},$$
  
$$\tau_1 = (\tau - 2C'_0)/(1+2C'_0) = \frac{\tau(1-C_0) - 2C_0}{1+C_0}.$$

Note that, by the condition  $0 < C_0 < 1/2$  and  $\tau > 1 + \sqrt{\mu_1 \mu_2}$ , we have  $\tau_0 > 0$  and  $\tau_1 > 0$ . Thus, we have

$$f(\delta, k_{\delta}) \le \tau_0 \delta, \quad \tau_1 \delta < f(\delta, k).$$

This completes the proof.

3.2. Main theorem. Now we prove the main theorem of this paper, which gives an error estimate for the simplified generalized Gauss-Newton method (1.12) under the stopping rule (1.13).

**Theorem 3.6.** Let the assumptions of Proposition 3.3 hold and let the iteration be stopped according to the stopping rule (1.13). Then

$$\|x_{k_{\delta}}^{\delta} - x^{\dagger}\| \le \eta \rho \sqrt{\psi^{-1}(\kappa \delta^2 / \rho^2)},$$

where

$$\begin{split} \kappa &:= \max\{1, \beta^2\}, \qquad \eta := c_0/\beta + \xi + c_0 C_0 (1+\tau_0)/\beta (1-C_0) \\ \text{with } \xi &:= 1 + (1+C_0)[(1+\tau_0)/(1-C_0)] \text{ and } \beta \text{ as in Proposition 3.3.} \end{split}$$

*Proof.* We observe that

(3.49) 
$$e_{k_{\delta}}^{\delta} = x_{k_{\delta}}^{\delta} - x^{\dagger} = r_{k_{\delta}-1}(A_{0}^{*}A_{0})(x_{0} - x^{\dagger}) - g_{k_{\delta}-1}(A_{0}^{*}A_{0})A_{0}^{*}\{F(x_{k_{\delta}-1}^{\delta}) - y^{\delta} - A_{0}e_{k_{\delta}-1}^{\delta}\}.$$

Thus,

$$\|e_{k_{\delta}}^{\delta}\| \leq \|r_{k_{\delta}-1}(A_{0}^{*}A_{0})(x_{0}-x^{\dagger})\| + \|g_{k_{\delta}-1}(A_{0}^{*}A_{0})A_{0}^{*}(y^{\delta}-y)\| + \|g_{k_{\delta}-1}(A_{0}^{*}A_{0})A_{0}^{*}\|(F(x_{k_{\delta}-1}^{\delta})-F(x^{\dagger})-A_{0}e_{k_{\delta}-1}^{\delta})\| \leq \|r_{k_{\delta}-1}(A_{0}^{*}A_{0})(x_{0}-x^{\dagger})\| + \frac{c_{0}\delta}{\sqrt{\alpha_{k_{\delta}-1}}} + \frac{c_{0}C_{0}\|A_{0}e_{k_{\delta}-1}^{\delta}\|}{\sqrt{\alpha_{k_{\delta}-1}}}.$$

First, we estimate  $||r_{k_{\delta}-1}(A_0^*A_0)(x_0 - x^{\dagger})||$ . For this, we observe, by the convexity of  $\psi$  and Jensen's inequality that

$$\psi\left(\frac{\|r_{k_{\delta}-1}(A_{0}^{*}A_{0})(x_{0}-x^{\dagger})\|^{2}}{\|w\|^{2}}\right) = \psi\left(\frac{\int_{0}^{M}\varphi(\lambda)r_{k_{\delta}-1}^{2}(\lambda)\mathrm{d}\|E_{\lambda}w\|^{2}}{\int_{0}^{M}\mathrm{d}\|E_{\lambda}w\|^{2}}\right) \\ \leq \frac{\int_{0}^{M}\psi(\varphi(\lambda)r_{k_{\delta}-1}^{2}(\lambda))\mathrm{d}\|E_{\lambda}w\|^{2}}{\int_{0}^{M}\mathrm{d}\|E_{\lambda}w\|^{2}} \\ \leq \frac{\int_{0}^{M}\varphi(\lambda)r_{k_{\delta}-1}^{2}(\lambda)\varphi^{-1}(\varphi(\lambda)r_{k_{\delta}-1}^{2}(\lambda))\mathrm{d}\|E_{\lambda}w\|^{2}}{\int_{0}^{M}\mathrm{d}\|E_{\lambda}w\|^{2}}$$

$$(3.51)$$

By the relation  $1 - \lambda g_{\alpha}(\lambda) \leq 1$  and monotonicity of  $\varphi^{-1}$  we have  $\varphi^{-1}(\varphi(\lambda)r_{k-1}^2(\lambda)) \leq \lambda$ . Hence, from (3.51) we get

Now, we estimate  $||A_0r_{k_{\delta}-1}(A_0^*A_0)(x_0-x^{\dagger})||$ . For this, first we observe from (3.49) that

$$A_0 e_{k_{\delta}}^{\delta} = A_0 r_{k_{\delta}-1} (A_0^* A_0) (x_0 - x^{\dagger}) - g_{k_{\delta}-1} (A_0 A_0^*) A_0 A_0^* \{ F(x_{k_{\delta}-1}^{\delta}) - y^{\delta} - A_0 e_{k_{\delta}-1}^{\delta} \}.$$
  
So, we have

$$A_0 r_{k_{\delta}-1} (A_0^* A_0) (x_0 - x^{\dagger}) = A_0 e_{k_{\delta}}^{\delta} + g_{k_{\delta}-1} (A_0 A_0^*) A_0 A_0^* \{ F(x_{k_{\delta}-1}^{\delta}) - y^{\delta} - A_0 e_{k_{\delta}-1}^{\delta} \}$$
  
Thus,

(3.53) 
$$||A_0r_{k_{\delta}-1}(A_0^*A_0)(x_0-x^{\dagger})|| \le ||A_0e_{k_{\delta}}^{\delta}|| + ||\{F(x_{k_{\delta}-1}^{\delta})-y^{\delta}-A_0e_{k_{\delta}-1}^{\delta}\}||.$$
  
By (3.18), for  $j \in \{k_{\delta}-1, k_{\delta}\}$ , we have

$$\|A_0 e_j^{\delta}\| = \|F(x_j^{\delta}) - F(x^{\dagger}) - \int_0^1 (F'(x^{\dagger} + t(x_j^{\delta} - x^{\dagger})) - A_0)(x_j^{\delta} - x^{\dagger})dt\|.$$

Using Assumption 2.2, we have

$$||A_0 e_j^{\delta}|| \le ||y - y^{\delta}|| + ||F(x_j^{\delta}) - y^{\delta}|| + C_0 ||A_0 e_j^{\delta}||$$

so that by (1.2) and Lemma 3.5, we get

(3.54) 
$$||A_0 e_j^{\delta}|| \le \frac{(1+\tau_0)\delta}{(1-C_0)}$$

Using (1.2), (3.19) and (3.54) in (3.53), we get

(3.55) 
$$\|A_0 r_{k_{\delta}-1} (A_0^* A_0) (x_0 - x^{\dagger})\| \le \frac{(1+\tau_0)\delta}{(1-C_0)} + \delta + \frac{(1+\tau_0)C_0\delta}{(1-C_0)} = \xi \delta.$$

where  $\xi = \left( (1 + C_0) \frac{(1 + \tau_0)}{1 - C_0} + 1 \right)$ . Thus, using (3.55) in (3.52), we have

(3.56) 
$$\psi\left(\frac{\|r_{k_{\delta}-1}(A_{0}^{*}A_{0})(x_{0}-x^{\dagger})\|^{2}}{\|w\|^{2}}\right) \leq \frac{\xi^{2}\delta^{2}}{\|w\|^{2}}.$$

Now, the relation (3.56) together with the monotonicity of  $\varphi^{-1}$  implies

$$\varphi^{-1} \left( \frac{\|r_{k_{\delta}-1}(A_{0}^{*}A_{0})(x_{0}-x^{\dagger})\|^{2}}{\xi^{2}\rho^{2}} \right) \leq \varphi^{-1} \left( \frac{\|r_{k_{\delta}-1}(A_{0}^{*}A_{0})(x_{0}-x^{\dagger})\|^{2}}{\|w\|^{2}} \right)$$
$$= \frac{\|w\|^{2}}{\|r_{k_{\delta}-1}(A_{0}^{*}A_{0})(x_{0}-x^{\dagger})\|^{2}} \psi \left( \frac{\|r_{k_{\delta}-1}(A_{0}^{*}A_{0})(x_{0}-x^{\dagger})\|^{2}}{\|w\|^{2}} \right)$$
$$\leq \frac{\xi^{2}\delta^{2}}{\|r_{k_{\delta}-1}(A_{0}^{*}A_{0})(x_{0}-x^{\dagger})\|^{2}}.$$

Thus,

$$\psi\left(\frac{\|r_{k_{\delta}-1}(A_{0}^{*}A_{0})(x_{0}-x^{\dagger})\|^{2}}{\xi^{2}\rho^{2}}\right) \leq \frac{\delta^{2}}{\rho^{2}},$$

and hence we get

(3.57) 
$$||r_{k_{\delta}-1}(A_0^*A_0)(x_0-x^{\dagger})|| \le \xi \rho \sqrt{\psi^{-1}(\delta^2/\rho^2)}.$$

Using (3.57) in (3.50), we get

$$\|e_{k_{\delta}}^{\delta}\| \leq \xi \rho \sqrt{\psi^{-1}(\delta^{2}/\rho^{2})} + \frac{c_{0}\delta}{\sqrt{\alpha_{k_{\delta}-1}}} + \frac{c_{0}C_{0}\|A_{0}e_{k_{\delta}}^{\delta}-1\|}{\sqrt{\alpha_{k_{\delta}-1}}}.$$

Using (3.54) and the inequality (3.34), we obtain

(3.58)  
$$\|e_{k_{\delta}}^{\delta}\| \leq \xi \rho \sqrt{\psi^{-1}(\delta^{2}/\rho^{2})} + \frac{c_{0}\delta}{\sqrt{\varphi^{-1}\psi^{-1}(\beta^{2}\delta^{2}/\rho^{2})}} + \frac{c_{0}C_{0}(1+\tau_{0})\delta}{(1-C_{0})\sqrt{\varphi^{-1}\psi^{-1}(\beta^{2}\delta^{2}/\rho^{2})}}.$$

Observe that

$$\varphi^{-1}\psi^{-1}(\beta^2\delta^2/\rho^2) = \frac{\beta^2\delta^2}{\rho^2\psi^{-1}(\beta\delta^2/\rho^2)}$$

so that

(3.59) 
$$\frac{\delta}{\sqrt{\varphi^{-1}\psi^{-1}(\beta^2\delta^2/\rho^2)}} = \frac{\rho\sqrt{\psi^{-1}(\beta^2\delta^2/\rho^2)}}{\beta}.$$

Using (3.59) in (3.58), we get

(3.60) 
$$\|e_{k_{\delta}}^{\delta}\| \leq \xi \rho \sqrt{\psi^{-1}(\delta^{2}/\rho^{2})} + c_{0} \frac{\rho \sqrt{\psi^{-1}(\beta^{2}\delta^{2}/\rho^{2})}}{\beta} + c_{0}C_{0} \frac{(1+\tau_{0})\rho \sqrt{\psi^{-1}(\beta^{2}\delta^{2}/\rho^{2})}}{(1-C_{0})\beta}$$

Thus,

(3.61) 
$$\|e_{ks}^{\delta}\| \le \eta \rho \sqrt{\psi^{-1}(\kappa \delta^2 / \rho^2)},$$

where  $\kappa = \max\{1, \beta^2\}$  and  $\eta = c_0/\beta + \xi + c_0C_0(1 + \tau_0)/\beta(1 - C_0)$ .

### 4. Concluding remarks

 $\square$ 

In this paper, we have considered a simplified generalized Gauss-Newton method (1.12) under a general source condition of the form (2.14). We obtained an order optimal estimate, in the sense that an improved order estimate which is applicable for the case of linear ill-posed problems as well is not possible (cf. [9]). As the iterations (1.12) and the source condition involve the Fréchet derivative only at the initial approximation  $x_0$  of the exact solution  $x^{\dagger}$  of (1.1), the calculations in this method become simpler than the generalized regularized Gauss-Newton method (1.10), where it is required to calculate the Fréchet derivative at each iterate. In [3], the iterations are defined by (1.10), and the stopping rule (1.11) is used to get the approximate solution for  $x^{\dagger}$ . But in the example of a parameter identification problem in [3], for illustration,  $F'(x_k^{\delta})$  is replaced by an operator independent of k. The same example in [3] illustrates the procedure of the present paper as well.

## References

- Bakushinskii, A.B. (1992): The problem of the convergence of the iteratively regularised Gauss-Newton method, Comput. Math. Phys., 32, 1353-1359. MR1185952 (93k:65049)
- Bakushinskii, A.B. (1995): Iterative methods without saturation for solving degenerate nonlinear operator equations, Dokl. Akad. Nauk, 344: 7-8. MR1361018
- Kaltenbacher, Barbara (1998): A posteriori parameter choice strategies for some Newton type methods for the regularization of nonlinear ill-posed problems, Numerische Mathematik, 79, 501-528. MR1631677 (99f:65091)
- [4] Blaschke, B., Neubauer, A., Scherzer, O. (1997): On convergence rates for the iteratively regularized Gauss-Newton method, IMA J. Numer. Anal., 17, 421-436. MR1459331 (98f:65066)
- Hohage, T. (1997): Logarithmic convergence rates of the iteratively regularized Gauss-Newton method for an inverse potential and an inverse scattering problem, Inverse Problems, 13, 1279-1299. MR1474369 (98k:65031)
- [6] Hohage, T. (1999): Iterative methods in inverse obstacle scattering: Regularization theory of linear and nonlinear exponentially ill-posed problems, Ph.D. Thesis, Johannes Kepler University, Linz, Austria.
- Hohage, T. (2000): Regularization of exponentially ill-posed problems, Numer. Funct. Anal. & Optim., 21, 439-464. MR1769885 (2001e:65095)
- [8] Langer, S., Hohage, T. (2007): Convergence analysis of an inexact iteratively regulaized Gauss-Newton method under general source conditions, J. of Inverse & Ill-Posed Problems, 15, 19-35. MR2337589
- Mathé, P., Pereverzev, S. (2003): Geometry of ill-posed problems in variable Hilbert scales, Inverse Problems, 19, 789-803. MR1984890 (2004i:47021)

DEPARTMENT OF MATHEMATICS, IIT MADRAS, CHENNAI 600036, INDIA *E-mail address*: pallavimahale@iitm.ac.in

DEPARTMENT OF MATHEMATICS, IIT MADRAS, CHENNAI 600036, INDIA E-mail address: mtnair@iitm.ac.in