

Convergence and adaptive discretization of the IRGNM Tikhonov and the IRGNM Ivanov method under a tangential cone condition in Banach space

Barbara Kaltenbacher¹ ·
Mario Luiz Previatti de Souza¹

Received: 24 July 2017 / Revised: 27 February 2018 / Published online: 29 May 2018
© The Author(s) 2018

Abstract In this paper we consider the iteratively regularized Gauss–Newton method (IRGNM) in its classical Tikhonov version as well as two further—Ivanov type and Morozov type—versions. In these two alternative versions, regularization is achieved by imposing bounds on the solution or by minimizing some regularization functional under a constraint on the data misfit, respectively. We do so in a general Banach space setting and under a tangential cone condition, while convergence (without source conditions, thus without rates) has so far only been proven under stronger restrictions on the nonlinearity of the operator and/or on the spaces. Moreover, we provide a convergence result for the discretized problem with an appropriate control on the error and show how to provide the required error bounds by goal oriented weighted dual residual estimators. The results are illustrated for an inverse source problem for a nonlinear elliptic boundary value problem, for the cases of a measure valued and of an L^∞ source. For the latter, we also provide numerical results with the Ivanov type IRGNM.

Mathematics Subject Classification 65F22 · 65N20

Supported by the Austrian Science Fund FWF under the Grant I2271 “Regularization and Discretization of Inverse Problems for PDEs in Banach Spaces”.

✉ Barbara Kaltenbacher
barbara.kaltenbacher@aau.at

Mario Luiz Previatti de Souza
mario.previatti@aau.at

¹ Institute of Mathematics, Alpen-Adria-Universität Klagenfurt, Klagenfurt, Austria

1 Introduction

In this paper we consider a nonlinear ill-posed operator equation

$$F(x) = y, \quad (1)$$

where the possibly nonlinear operator $F : \mathcal{D}(F) \subseteq X \rightarrow Y$ with domain $\mathcal{D}(F)$ maps between real Banach spaces X and Y . We are interested in the ill-posed situation, i.e., F fails to be continuously invertible, and the data are contaminated with noise, thus regularization has to be applied (see, e.g., [8, 27], and references therein).

Throughout this paper we will assume that an exact solution $x^\dagger \in \mathcal{D}(F)$ of (1) exists, i.e., $F(x^\dagger) = y$, and that the noise level δ in the (deterministic) estimate

$$\|y - y^\delta\| \leq \delta \quad (2)$$

is known.

Partially we will also refer to the formulation of the inverse problem as a system of model and observation equation

$$A(x, u) = 0 \quad (3)$$

$$C(u) = y. \quad (4)$$

Here $A : X \times V \rightarrow W^*$ and $C : V \rightarrow Y$ are the model and observation operator, so that with the parameter-to-state map $S : X \rightarrow V$ satisfying $A(x, S(x)) = 0$ and $F = C \circ S$, (1) is equivalent to the all-at-once formulation (3), (4).

Newton type methods for the solution of nonlinear ill-posed problems (1) have been extensively studied in Hilbert spaces (see, e.g., [2, 20] and the references therein) and more recently also in a Banach space setting. In particular, the iteratively regularized Gauss–Newton method [1] can be generalized to a Banach space setting by calculating iterates x_{k+1}^δ in a Tikhonov type variational form as

$$x_{k+1}^\delta \in \operatorname{argmin}_{x \in \mathcal{C}} \|F'(x_k^\delta)(x - x_k^\delta) + F(x_k^\delta) - y^\delta\|^p + \alpha_k \mathcal{R}(x), \quad (5)$$

see, e.g., [11, 16, 17, 21, 28] where $p \in [1, \infty)$, $(\alpha_k)_{k \in \mathbb{N}}$ is a sequence of regularization parameters, and \mathcal{R} is some nonnegative regularization functional. Alternatively, one might introduce regularization by imposing some bound ρ_k on the norm of x , or, again, generally, on a regularization functional of x

$$x_{k+1}^\delta \in \operatorname{argmin}_{x \in \mathcal{C}} \|F'(x_k^\delta)(x - x_k^\delta) + F(x_k^\delta) - y^\delta\| \text{ such that } \mathcal{R}(x) \leq \rho_k, \quad (6)$$

which corresponds to Ivanov regularization or the method of quasi solutions, see, e.g., [7, 13–15, 22, 24, 26]. A third way of incorporating regularization in a Newton type iteration is Morozov regularization, also called the method of the residuals, see, e.g., [9, 22, 23]

$$x_{k+1}^\delta \in \operatorname{argmin}_{x \in \mathcal{C}} \mathcal{R}(x) \text{ such that } \|F'(x_k^\delta)(x - x_k^\delta) + F(x_k^\delta) - y^\delta\| \leq \sigma \|F(x_k^\delta) - y^\delta\|, \tag{7}$$

for some $\sigma \in (0, 1)$, where the choice of the bound in the inequality constraint is very much inspired by the inexact Newton type regularization parameter choice in [10].

We restrict ourselves to the norm in Y as a measure of the data misfit, but the analysis could as well be extended to more general functionals \mathcal{S} satisfying certain conditions, as e.g., in [11, 28]. Here \mathcal{C} is a set (possibly chosen with convenient properties for carrying out the minimization) containing x^\dagger and being contained in $\mathcal{D}(F)$, such that F satisfies additional conditions on \mathcal{C} see (8), (10) below. If F is defined on all of X , then the minimization problem (5) can be posed in an unconstrained way $\mathcal{C} = X$.

As a restriction on the nonlinearity of the forward operator F we impose the tangential cone condition

$$\|F(\tilde{x}) - F(x) - F'(x)(\tilde{x} - x)\| \leq c_{tc} \|F(\tilde{x}) - F(x)\| \text{ for all } \tilde{x}, x \in \mathcal{B}_R \tag{8}$$

(also called Scherzer condition, cf. [25]) for some constant $c_{tc} < 1/3$. Here, for any $r > 0$,

$$\mathcal{B}_r = \{x \in \mathcal{C} : \mathcal{R}(x) \leq r\} \tag{9}$$

is a sublevel set of the regularization functional and R will be specified in the convergence result Theorem 1.

Note that the convergence conditions imposed in [11, 16, 17, 21, 28] in the situation without source condition, namely local invariance of the range of $F'(x)^*$, are slightly stronger, since this adjoint range invariance is sufficient for (8). However, most probably the gap is not very large, as in those application examples where (8) has been verified, the proof of (8) is actually often done via adjoint range invariance. In (5), (6), (7), the bounded linear operator $F'(x)$ is not necessarily a Gâteaux or Fréchet derivative of F , but just some local linearization (in the sense of (8)), satisfying additionally the weak closedness condition

$$\begin{aligned} \forall x \in \mathcal{C}, (x_n)_{n \in \mathbb{N}} \subseteq \mathcal{C} : \left(x_n \xrightarrow{\mathcal{T}_X} \hat{x}, \text{ and } F'(x)x_n \xrightarrow{\mathcal{T}_Y} y \right) \\ \Rightarrow \left(\hat{x} \in \mathcal{C} \text{ and } F'(x)\hat{x} = y \right). \end{aligned} \tag{10}$$

In here, \mathcal{T}_X and \mathcal{T}_Y are topologies on X and Y (e.g., just the weak or weak* topologies) such that bounded sets in Y are \mathcal{T}_Y -compact and the norm in Y is \mathcal{T}_Y -lower semicontinuous.

The remainder of this paper is organized as follows. In Sect. 2 we state and prove convergence results in the continuous and discretized setting. Section 3 shows how to actually obtain the required discretization error estimates by a goal oriented weighted dual residual approach and Sect. 4 illustrates the theoretical findings by an inverse source problem for a nonlinear PDE. In Sect. 5 we provide some numerical results for this model problem and Sect. 6 concludes with some remarks.

2 Convergence

In this section we will study convergence of the IRGNM iterates first of all in a continuous setting, then in the situation of having discretized for computational purposes.

The regularization parameters α_k, ρ_k, σ are chosen a priori

$$\alpha_k = \alpha_0 \theta^k \text{ for some } \theta \in \left(\left(\frac{2c_{tc}}{1-c_{tc}} \right)^p, 1 \right) \tag{11}$$

(note that $(\frac{2c_{tc}}{1-c_{tc}})^p < 1$ for $c_{tc} < 1/3$),

$$\rho_k \equiv \rho \geq \mathcal{R}(x^\dagger), \tag{12}$$

and

$$\sigma \geq \frac{1 + c_{tc}}{\tau} + c_{tc}, \quad \sigma < 1 - 2c_{tc}, \tag{13}$$

with τ as in (14), and the iteration is stopped according to the discrepancy principle

$$k_* = k_*(\delta, y^\delta) = \min\{k \in \mathbb{N}_0 : \|F(x_k^\delta) - y^\delta\| \leq \tau\delta\} \tag{14}$$

with some fixed $\tau > 1$ chosen sufficiently large but independent of δ .

Theorem 1 *Let $\mathcal{R} : X \rightarrow [0, \infty]$ be proper, convex and \mathcal{T}_X lower semicontinuous with $\mathcal{R}(x^\dagger) < \infty$ and let, for all $r \in [\mathcal{R}(x^\dagger), \infty)$ in case of (5), or for all $r \in [\mathcal{R}(x^\dagger), \rho]$ in case of (6), or for $r = \mathcal{R}(x^\dagger)$ in case of (7), the sublevel set (9) be compact with respect to the topology \mathcal{T}_X on X .*

Moreover, let F satisfy (8), (10).

Finally, let the family of data $(y^\delta)_{\delta>0}$ satisfy (2).

(i) *Then for fixed δ, y^δ , the iterates according to (5)–(7) are well-defined and satisfy*

$$x_k^\delta \in \mathcal{B}_R \text{ with } R \begin{cases} \text{defined by (23), (19), (20) in case of (5)} \\ = \rho \text{ in case of (6)} \\ = \mathcal{R}(x^\dagger) \text{ in case of (7)} \end{cases} \tag{15}$$

for all $k \leq k_(\delta, y^\delta)$, which denotes the stopping index according to the discrepancy principle (14) with τ sufficiently large, and this stopping index $k_*(\delta, y^\delta)$ is finite.*

(ii) *Moreover, for both methods we have \mathcal{T}_X -subsequential convergence as $\delta \rightarrow 0$ i.e., $(x_{k_*(\delta, y^\delta)}^\delta)_{\delta>0}$ has a \mathcal{T}_X -convergent subsequence and the limit of every \mathcal{T}_X -convergent subsequence solves (1). If the solution x^\dagger of (1) is unique in \mathcal{B}_R , then $x_{k_*(\delta, y^\delta)}^\delta \xrightarrow{\mathcal{T}_X} x^\dagger$ as $\delta \rightarrow 0$.*

(iii) *Additionally, k_* satisfies the asymptotics $k_* = \mathcal{O}(\log(1/\delta))$.*

Proof Existence of minimizers x_{k+1}^δ of (5)–(7) for fixed k, x_k^δ and y^δ follows by the direct method of calculus of variations: In all three cases, the cost functional

$$J_k(x) = \|F'(x_k^\delta)(x - x_k^\delta) + F(x_k^\delta) - y^\delta\|^p + \alpha_k \mathcal{R}(x) \text{ in case of (5),}$$

$$J_k(x) = \frac{1}{2} \|F'(x_k^\delta)(x - x_k^\delta) + F(x_k^\delta) - y^\delta\|^2 \text{ in case of (6),}$$

$$J_k(x) = \mathcal{R} \text{ in case of (7),}$$

is bounded from below and the admissible set

$$X^{\text{ad}} = \mathcal{C} \text{ in case of (5),}$$

$$X^{\text{ad}} = \mathcal{B}_\rho \text{ in case of (6),}$$

$$X^{\text{ad}} = \{x \in \mathcal{C} : \|F'(x_k^\delta)(x - x_k^\delta) + F(x_k^\delta) - y^\delta\| \leq \sigma \|F(x_k^\delta) - y^\delta\|\} \text{ in case of (7)}$$

is nonempty (for (6) this follows from $\rho_k \geq \mathcal{R}(x^\dagger)$ and for (7) from (8), (14) and (13), see (16) below). Hence, there exists a minimizing sequence $(x^l)_{l \in \mathbb{N}} \subseteq X^{\text{ad}} \cap \mathcal{B}_r$ for

$$r = \frac{1}{\alpha_k} J_k(x^\dagger) \text{ in case of (5), } r = \rho \text{ in case of (6), } r = \mathcal{R}(x^\dagger) \text{ in case of (7),}$$

with bounded linearized residuals $\|F'(x_k^\delta)(x^l - x_k^\delta) + F(x_k^\delta) - y^\delta\| \leq s$ for

$$s = J_k(x^\dagger)^{1/p} \text{ in case of (5), (6), } s = \sigma \|F(x_k^\delta) - y^\delta\| \text{ in case of (7),}$$

and $\lim_{l \rightarrow \infty} J_k(x^l) = \inf_{x \in X^{\text{ad}}} J_k(x)$. By \mathcal{T}_X -compactness of \mathcal{B}_r , the sequence $(x^l)_{l \in \mathbb{N}}$ has a \mathcal{T}_X -convergent subsequence $(x^{l_m})_{m \in \mathbb{N}}$ with limit $\bar{x} \in \mathcal{B}_r$. Moreover, \mathcal{T}_Y -compactness of norm bounded sets in Y together with \mathcal{T}_X - \mathcal{T}_Y -closedness of $F'(x_k^\delta)$ and lower \mathcal{T}_Y semicontinuity of the norm in Y , implies that in all three cases $J_k(\bar{x}) \leq \liminf_{m \rightarrow \infty} J_k(x^{l_m}) = \inf_{x \in X^{\text{ad}}} J_k(x)$ and $\bar{x} \in X^{\text{ad}}$, hence \bar{x} is a minimizer.

Note that (ii) follows from (i) by standard arguments and our assumption on \mathcal{T} -compactness of \mathcal{B}_R . Thus it remains to prove (i) and (iii) for the three versions (5), (6), (7) of the IRGNM.

For this purpose we are going to show that for every $\delta > 0$, there exists $k_* = k_*(\delta, y^\delta)$ such that $k_* \sim \log(1/\delta)$, and the stopping criterion according to the discrepancy principle $\|F(x_{k_*}^\delta(\delta, y^\delta)) - y^\delta\| \leq \tau \delta$ is satisfied. For (5), we also need to show that $\mathcal{R}(x_k^\delta) \leq R$ for $k \leq k_*(\delta, y^\delta)$, whereas in (6) this automatically holds by (12). The same holds true for (7): If $x_k^\delta \in \mathcal{B}_R$, then by (8), (14) and (13) we have

$$\begin{aligned} & \|F'(x_k^\delta)(x^\dagger - x_k^\delta) + F(x_k^\delta) - y^\delta\| \leq c_{tc} \|F(x_k^\delta) - y^\delta\| + (1 + c_{tc})\delta \\ & \leq \left(c_{tc} + \frac{1 + c_{tc}}{\tau} \right) \|F(x_k^\delta) - y^\delta\|, \end{aligned} \tag{16}$$

so x^\dagger is admissible, hence $\mathcal{R}(x_{k+1}^\delta) \leq \mathcal{R}(x^\dagger)$, i.e., $x_{k+1}^\delta \in \mathcal{B}_R$.

We start with the Tikhonov version (5) and carry out an induction proof of the following statement: For all $k \in \{0, \dots, k_*(\delta, y^\delta)\}$

$$\mathcal{R}(x_k^\delta) \leq R \text{ and } \forall j \in \{0, \dots, k-1\} : d_{j+1} + \alpha_j \mathcal{R}_{j+1} \leq qd_j + \alpha_j \mathcal{R}^\dagger + C\delta^p, \tag{17}$$

where

$$d_k := 2^{1-p} (1 - c_{tc})^p \|F(x_k^\delta) - y^\delta\|^p, \tag{18}$$

$$\begin{aligned}
 q &:= 2^{p-1}((1 + \gamma)^{p-1} + 1) \left(\frac{c_{tc}}{1 - c_{tc}} \right)^p \in (0, 1), & (19) \\
 \mathcal{R}_k &:= \mathcal{R}(x_k^\delta), \quad \mathcal{R}^\dagger := \mathcal{R}(x^\dagger), \\
 C &:= \left(\frac{1 + \gamma}{\gamma} \right)^{p-1} (1 + c_{tc})^p, & (20)
 \end{aligned}$$

for some fixed small $\gamma \in (0, 1)$. We will require $\frac{q}{\theta} < 1$, which by definition of q (19) is achievable for $\gamma > 0$ sufficiently small, due to $\theta > \left(\frac{2c_{ct}}{1-c_{ct}}\right)^p$, cf. (11). By Lemma 2 (see the ‘‘Appendix’’) the right hand side estimate in (17) implies

$$d_k + \alpha_{k-1} \mathcal{R}_k < q^k d_0 + \left(\frac{1}{1 - \frac{q}{\theta}} \right) \alpha_{k-1} \mathcal{R}^\dagger + \left(\frac{1}{1 - q} \right) C \delta^p. \tag{21}$$

Using the minimality of x_{k+1}^δ and (2), (8) together with $x^\dagger, x_k^\delta \in \mathcal{B}_R$, we have

$$\begin{aligned}
 &\|F'(x_k^\delta)(x_{k+1}^\delta - x_k^\delta) + F(x_k^\delta) - y^\delta\|^p + \alpha_k \mathcal{R}(x_{k+1}^\delta) \\
 &\leq \|F'(x_k^\delta)(x^\dagger - x_k^\delta) + F(x_k^\delta) - y^\delta\|^p + \alpha_k \mathcal{R}(x^\dagger) \\
 &\leq \left(c_{tc} \|F(x_k^\delta) - y^\delta\| + (1 + c_{tc}) \delta \right)^p + \alpha_k \mathcal{R}(x^\dagger), & (22)
 \end{aligned}$$

From (21) and (14) we infer

$$\left(2^{1-p}(1 - c_{tc})^p - \frac{C}{(1 - q)\tau^p} \right) \|F(x_k^\delta) - y^\delta\|^p \leq q^k d_0 + \theta^k \alpha_0 \frac{\mathcal{R}^\dagger}{\theta - q}$$

Using this and again (14) in (22) yields

$$\begin{aligned}
 \mathcal{R}(x_{k+1}^\delta) &\leq \left(c_{ct} + \frac{1 + c_{tc}}{\tau} \right)^p \left(2^{1-p}(1 - c_{tc})^p - \frac{C}{(1 - q)\tau^p} \right)^{-1} \\
 &\quad \times \left(\left(\frac{q}{\theta} \right)^k \frac{d_0}{\alpha_0} + \frac{\mathcal{R}^\dagger}{\theta - q} \right) + \mathcal{R}^\dagger \\
 &\leq \left(\frac{1}{3} + \frac{4}{3\tau} \right)^p \left(\frac{2}{3^p} - \frac{C}{(1 - q)\tau^p} \right)^{-1} \\
 &\quad \times \left(2^{1-p} \frac{\|F(x_0^\delta) - y^\delta\|^p}{\alpha_0} + \frac{\mathcal{R}^\dagger}{\theta - q} \right) + \mathcal{R}^\dagger =: R. & (23)
 \end{aligned}$$

On the other hand, since we have established $x_{k+1}^\delta \in \mathcal{B}_R$, we can apply (8) to the left hand side of (22) to obtain

$$\begin{aligned}
 &\|F'(x_k^\delta)(x_{k+1}^\delta - x_k^\delta) + F(x_k^\delta) - y^\delta\|^p + \alpha_k \mathcal{R}(x_{k+1}^\delta) \\
 &\geq \left| (1 - c_{tc}) \|F(x_{k+1}^\delta) - y^\delta\| - c_{tc} \|F(x_k^\delta) - y^\delta\| \right|^p + \alpha_k \mathcal{R}(x_{k+1}^\delta).
 \end{aligned}$$

see also [11, Lemma 5.2] and [21, proof of Theorem 3].

To handle the power p we make use of the following inequalities that can be proven by solving extremal value problems, see the ‘‘Appendix’’

$$\begin{aligned}
 (a + b)^p &\leq (1 + \gamma)^{p-1} a^p + \left(\frac{1 + \gamma}{\gamma}\right)^{p-1} b^p \text{ and} \\
 (a - b)^p &\geq (1 - \epsilon)^{p-1} a^p - \left(\frac{1 - \epsilon}{\epsilon}\right)^{p-1} b^p,
 \end{aligned}
 \tag{24}$$

for all $a, b > 0, p \geq 1$ and $\gamma, \epsilon \in (0, 1)$, where for the right hand inequality to hold, additionally $a \geq b$ is needed.

Hence, in case $(1 - c_{tc}) \|F(x_{k+1}^\delta) - y^\delta\| \geq c_{tc} \|F(x_k^\delta) - y^\delta\|$ the following general estimate holds

$$\begin{aligned}
 &(1 - \epsilon)^{p-1} (1 - c_{tc})^p \|F(x_{k+1}^\delta) - y^\delta\|^p + \alpha_k \mathcal{R}(x_{k+1}^\delta) \\
 &\leq \left((1 + \gamma)^{p-1} + \left(\frac{1 - \epsilon}{\epsilon}\right)^{p-1} \right) c_{tc}^p \|F(x_k^\delta) - y^\delta\|^p \\
 &\quad + \alpha_k \mathcal{R}(x^\dagger) + \left(\frac{1 + \gamma}{\gamma}\right)^{p-1} (1 + c_{tc})^p \delta^p,
 \end{aligned}
 \tag{25}$$

for $\gamma, \epsilon \in (0, 1)$.

So in order for this recursion to yield geometric decay of $\|F(x_k^\delta) - y^\delta\|$, we need to ensure

$$(1 - \epsilon)^{p-1} (1 - c_{tc})^p > \left((1 + \gamma)^{p-1} + \left(\frac{1 - \epsilon}{\epsilon}\right)^{p-1} \right) c_{tc}^p
 \tag{26}$$

for a proper choice of $\epsilon, \gamma \in (0, 1)$. To obtain the largest possible (and therefore least restrictive) bound on c_{tc} , we rewrite the requirement above as

$$\begin{aligned}
 \left(\frac{c_{tc}}{1 - c_{tc}}\right)^p &< \sup_{\epsilon, \gamma \in (0, 1)} (1 - \epsilon)^{p-1} \left((1 + \gamma)^{p-1} + \left(\frac{1 - \epsilon}{\epsilon}\right)^{p-1} \right)^{-1} \\
 &= \sup_{\epsilon \in (0, 1)} \underbrace{(1 - \epsilon)^{p-1} \left(1 + \left(\frac{1 - \epsilon}{\epsilon}\right)^{p-1} \right)^{-1}}_{=\phi(\epsilon)} = \phi\left(\frac{1}{2}\right) = 2^{-p},
 \end{aligned}$$

as can be found out by evaluating the derivative of ϕ

$$\phi'(\epsilon) = -(p - 1)(1 - \epsilon)^{p-2} \left(1 + \left(\frac{1 - \epsilon}{\epsilon}\right)^{p-1} \right)^{-2} \left(1 - \left(\frac{1 - \epsilon}{\epsilon}\right)^p \right).$$

Thus we will furtheron set $\epsilon = \frac{1}{2}$ and assume that $\gamma > 0$ is sufficiently small so that (26) holds with $\epsilon = \frac{1}{2}$, i.e., (19). Then, using (11), estimate (25) can be written as

$$d_{k+1} + \alpha_k \mathcal{R}_{k+1} \leq qd_k + \alpha_0 \theta^k \mathcal{R}^\dagger + C\delta^p, \tag{27}$$

which we first of all regard as a recursive estimate for d_k .

To derive a similar estimate also in the complementary case $(1 - c_{tc})\|F(x_{k+1}^\delta) - y^\delta\| < c_{tc}\|F(x_k^\delta) - y^\delta\|$, we use that fact that, for d_k as in (18), this inequality just means

$$d_{k+1} < \left(\frac{c_{tc}}{1 - c_{tc}}\right)^p d_k$$

and, using (22) and the left hand part of (24),

$$\alpha_k \mathcal{R}_{k+1} \leq (1 + \gamma)^{p-1} c_{tc}^p \|F(x_k^\delta) - y^\delta\|^p + \alpha_k \mathcal{R}^\dagger + \left(\frac{1 + \gamma}{\gamma}\right)^{p-1} (1 + c_{tc})^p \delta^p,$$

hence after addition we again get (27) (even with a slightly smaller value of $q := (1 + 2^{p-1}(1 + \gamma)^{p-1})\left(\frac{c_{tc}}{1 - c_{tc}}\right)^p$).

Thus in both cases, using Lemma 2 we can conclude that

$$d_{k+1} + \alpha_k \mathcal{R}_{k+1} < q^{k+1} d_0 + \left(\frac{1}{1 - \frac{q}{\theta}}\right) \alpha_k \mathcal{R}^\dagger + \left(\frac{1}{1 - q}\right) C\delta^p. \tag{28}$$

This finishes the induction proof of (17) for all $k \in \{0, \dots, k_*(\delta, y^\delta)\}$.

We next show that the discrepancy stopping criterion from (14), i.e., $d_{k_*} \leq \tilde{\tau} \delta^p$ for $\tilde{\tau} = 2^{1-p}(1 - c_{tc})^p \tau^p$, will be satisfied after finitely many, namely $O(\log(1/\delta))$, steps. For this purpose, note that $\tilde{\tau} > \frac{C}{1-q}$, provided τ is chosen sufficiently large, which we assume to be done. Thus, indeed, using (11), (21), we have

$$d_k \leq d_k + \alpha_{k-1} \mathcal{R}_k < \theta^k \left(d_0 + \frac{\alpha_0}{\theta - q} \mathcal{R}^\dagger\right) + \frac{C}{1 - q} \delta^p, \tag{29}$$

where the right hand side falls below $\tilde{\tau} \delta^p$ as soon as

$$\begin{aligned} k &\geq (\log 1/\theta)^{-1} \left(p \log(1/\delta) + \log \left(d_0 + \frac{\alpha_0}{\theta - q} \mathcal{R}^\dagger \right) - \log \left(\tilde{\tau} - \frac{C}{1 - q} \right) \right) \\ &=: \bar{k}(\delta). \end{aligned}$$

Thus we get the upper estimate $k_*(\delta, y^\delta) \leq \bar{k}(\delta) = O(\log(1/\delta))$.

For the Ivanov version (6), it only remains to show finiteness of the stopping index, as boundedness of the \mathcal{R} values by $R = \rho$ holds by definition. Applying the minimality

argument with x^\dagger being admissible [cf. (12)] to (6) leads to the special case $p = 1$, $\alpha_k = 0$ in (25)

$$(1 - c_{tc})\|F(x_{k+1}^\delta) - y^\delta\| \leq 2c_{tc}\|F(x_k^\delta) - y^\delta\| + (1 + c_{tc})\delta.$$

Our notation becomes

$$\begin{aligned} d_k &:= (1 - c_{tc})\|F(x_k^\delta) - y^\delta\|, \\ q &:= \frac{2c_{tc}}{1 - c_{tc}} \in (0, 1), \\ C &:= (1 + c_{tc}), \end{aligned}$$

which gives

$$d_{k+1} \leq qd_k + C\delta,$$

and by induction, one can conclude

$$d_k < q^k d_0 + \left(\frac{1}{1 - q}\right) C\delta,$$

where the right hand side is smaller than $\tilde{\tau}\delta$ (with $\tilde{\tau} = (1 - c_{tc})\tau$) for all

$$k \geq (\log 1/q)^{-1} \left(p \log(1/\delta) + \log d_0 - \log \left(\tilde{\tau} - \frac{C}{1 - q} \right) \right) =: \bar{k}(\delta),$$

so that we can again conclude $k_*(\delta, y^\delta) \leq \bar{k}(\delta) = O(\log(1/\delta))$.

Finally we consider (7), where boundedness of the \mathcal{R} values by $R = \mathcal{R}(x^\dagger)$ holds by minimality and the fact that x^\dagger is admissible, cf. (16). Geometric decay of the residuals follows by the estimate

$$\begin{aligned} \sigma \|F(x_k^\delta) - y^\delta\| &\geq \|F'(x_k^\delta)(x_{k+1}^\delta - x_k^\delta) + F(x_k^\delta) - y^\delta\| \\ &\geq (1 - c_{tc})\|F(x_{k+1}^\delta) - y^\delta\| - c_{tc}\|F(x_k^\delta) - y^\delta\| \end{aligned} \tag{30}$$

and (13), i.e.,

$$\|F(x_{k+1}^\delta) - y^\delta\| \leq q\|F(x_k^\delta) - y^\delta\|$$

with

$$q = \frac{\sigma + c_{tc}}{1 - c_{tc}},$$

so that similarly to above we end up with a logarithmic estimate for k_* . □

Remark 1 Convergence of $\mathcal{R}(x_{k_*}^\delta(\delta, y^\delta))$ to $\mathcal{R}(x^\dagger)$ as $\delta \rightarrow 0$ holds along the \mathcal{T}_X convergent subsequence according to Theorem 1 (ii), first of all for the Morozov and the Ivanov version of the IRGNM, with the choice $\rho = \mathcal{R}(x^\dagger)$ for the latter, since in both cases $\mathcal{R}(x_{k_*}^\delta(\delta, y^\delta)) \leq \mathcal{R}(x^\dagger)$ holds for all δ and \mathcal{R} is \mathcal{T}_X lower semicontinuous. The same holds true also for the Tikhonov version with the alternative choice of α_k such that

$$\underline{\sigma} \leq \frac{\|F'(x_k^\delta)(x_{k+1}^\delta(\alpha_k) - x_k^\delta) + F(x_k^\delta) - y^\delta\|}{\|F(x_k^\delta) - y^\delta\|} \leq \bar{\sigma}$$

for some constants $\underline{\sigma}, \bar{\sigma}$ satisfying $c_{\tau c} + \frac{1+c_{\tau c}}{\tau} < \underline{\sigma} < \bar{\sigma} < 1$ in place of (11), as can be seen directly from (22). If \mathcal{R} is defined by the norm on a space with the Kadets-Klee property, and \mathcal{T}_X is the weak topology of this space, then this implies norm convergence of $x_{k_*}^\delta(\delta, y^\delta)$ to x^\dagger along the same subsequence.

Remark 2 The fact that x_k^δ stays in \mathcal{B}_R [cf. (15)] is crucial for the applicability of the tangential cone condition (8) in these iterates. If the functional \mathcal{R} quantifies some distance to an a priori guess x_0 , (e.g., $\mathcal{R} = \|x - x_0\|^q$ for some norm $\|\cdot\|$ and some $q > 0$), then $x \in \mathcal{B}_R$ with small R means closeness of x to x_0 in a certain sense. Thus, the smaller R is, the better (8) might get achievable with some $c_{\tau c} < \frac{1}{3}$. On the other hand, making R according to (15) small means closeness of x^\dagger to x_0 . Thus we deal with local convergence, as typical for Newton type methods.

Now we consider the appearance of discretization errors in the numerical solution of (5), (6) arising from restriction of the minimization to finite dimensional subspaces X_h^k and leading to discretized iterates $x_{k,h}^\delta$ and an approximate version F_h^k of the forward operator i.e., we consider the discretized version of Tikhonov-IRGNM (5)

$$x_{k+1,h}^\delta \in \operatorname{argmin}_{x \in \mathcal{C} \cap X_h^k} \|F_h^{k'}(x_{k,h}^\delta)(x - x_{k,h}^\delta) + F_h^k(x_{k,h}^\delta) - y^\delta\|^p + \alpha_k \mathcal{R}(x), \tag{31}$$

of Ivanov-IRGNM (6)

$$x_{k+1,h}^\delta \in \operatorname{argmin}_{x \in \mathcal{C} \cap X_h^k} \|F_h^{k'}(x_{k,h}^\delta)(x - x_{k,h}^\delta) + F_h^k(x_{k,h}^\delta) - y^\delta\| \text{ such that } \mathcal{R}(x) \leq \rho, \tag{32}$$

and of Morozov-IRGNM (7)

$$\begin{aligned} x_{k+1,h}^\delta &\in \operatorname{argmin}_{x \in \mathcal{C} \cap X_h^k} \mathcal{R}(x) \text{ such that } \|F_h^{k'}(x_{k,h}^\delta)(x - x_{k,h}^\delta) + F_h^k(x_{k,h}^\delta) - y^\delta\| \\ &\leq \sigma \|F_h^k(x_{k,h}^\delta) - y^\delta\|, \end{aligned} \tag{33}$$

respectively. Moreover, also in the discrepancy principle, the residual is replaced by its actually computable discretized version

$$k_* = k_*(\delta, y^\delta) = \min\{k \in \mathbb{N}_0 : \|F_h^k(x_{k,h}^\delta) - y^\delta\| \leq \tau \delta\}. \tag{34}$$

We define the auxiliary continuous iterates

$$x_{k+1}^\delta \in \operatorname{argmin}_{x \in \mathcal{C}} \|F'(x_{k,h}^\delta)(x - x_{k,h}^\delta) + F(x_{k,h}^\delta) - y^\delta\|^p + \alpha_k \mathcal{R}(x), \tag{35}$$

$$x_{k+1}^\delta \in \operatorname{argmin}_{x \in \mathcal{C}} \|F'(x_{k,h}^\delta)(x - x_{k,h}^\delta) + F(x_{k,h}^\delta) - y^\delta\| \text{ such that } \mathcal{R}(x) \leq \rho, \tag{36}$$

and

$$x_{k+1}^\delta \in \operatorname{argmin}_{x \in \mathcal{C}} \mathcal{R}(x) \text{ such that } \|F'(x_{k,h}^\delta)(x - x_{k,h}^\delta) + F(x_{k,h}^\delta) - y^\delta\| \leq \sigma \|F(x_{k,h}^\delta) - y^\delta\|, \tag{37}$$

respectively in order to be able to use minimality, i.e., compare with the continuous exact solution x^\dagger . For an illustration we refer to [18, Figure 1].

First of all, we assess how large the discretization errors can be allowed to still enable convergence. Later on, in Sect. 3, we will describe how to really obtain such estimates a posteriori and to achieve the prescribed accuracy by adaptive discretization.

Corollary 1 *Let the assumptions of Theorem 1 be satisfied and assume that the discretization error estimates*

$$\|F(x_{k+1,h}^\delta) - y^\delta\| - \|F(x_{k+1}^\delta) - y^\delta\| \leq \eta_{k+1} \tag{38}$$

$$\left| \|F_h^k(x_{k,h}^\delta) - y^\delta\| - \|F(x_{k,h}^\delta) - y^\delta\| \right| \leq \xi_k \tag{39}$$

$$\mathcal{R}(x_{k,h}^\delta) - \mathcal{R}(x_k^\delta) \leq \zeta_k \tag{40}$$

(note that no absolute value is needed in (38), (40); moreover, (40) is only needed for (5) and (7)) hold with

$$\eta_k \leq c_\eta \|F_h^k(x_{k,h}^\delta) - y^\delta\| \quad \xi_k \leq c_\xi \|F_h^k(x_{k,h}^\delta) - y^\delta\|, \quad \zeta_k \leq \bar{\zeta}. \tag{41}$$

for all $k \leq k_*(\delta, y^\delta)$ and constants $c_\eta, c_\xi > 0$ sufficiently small, $\bar{\zeta} > 0$.

Then the assertions of Theorem 1 remain valid for $x_{k_*(\delta, y^\delta), h}^\delta$ in place of $x_{k_*(\delta, y^\delta)}^\delta$ with (34) in place of (14) and (42) in place of (23).

Proof For the Tikhonov version (31), in order to inductively estimate $\mathcal{R}(x_{k+1,h}^\delta)$, given $x_k^\delta \in \mathcal{B}_R$, note that from (43) with $k + 1$ replaced by k , we get like in (23) that

$$\begin{aligned} \mathcal{R}(x_{k+1,h}^\delta) &\leq \left(\frac{1}{3} + \frac{4}{3\tau}\right)^p \left(\frac{2}{3^p} - \frac{C}{(1-q)\tau^p}\right)^{-1} \\ &\quad \times \left(2^{1-p} \frac{\|F(x_0^\delta) - y^\delta\|^p}{\alpha_0} + \frac{\mathcal{R}^\dagger + \bar{\zeta}}{\theta - q}\right) + \mathcal{R}^\dagger =: R \end{aligned} \tag{42}$$

where

$$\begin{aligned}
 d_{k,h} &:= 2^{1-p}(1 - c_{tc})^p \|F(x_{k,h}^\delta) - y^\delta\|^p, \\
 \tilde{q} &:= 2^{p-1}((1 + \gamma)^{p-1} + (1 + \tilde{\gamma})^{p-1}) \left(\frac{c_{tc}}{1 - c_{tc}}\right)^p, \\
 q &= \tilde{q} + D \frac{c_\eta}{1 - c_\xi} \in (0, 1), \\
 \mathcal{R}_{k,h} &:= \mathcal{R}(x_{k,h}^\delta), \quad \mathcal{R}^\dagger := \mathcal{R}(x^\dagger), \\
 C &:= \left(\frac{1 + \gamma}{\gamma}\right)^{p-1} (1 + c_{tc})^p, \quad D := \left(\frac{1 + \tilde{\gamma}}{\tilde{\gamma}}\right)^{p-1} (1 - c_{tc})^p,
 \end{aligned}$$

for $\gamma, \tilde{\gamma}, c_\eta \in (0, 1)$, which are chosen small enough so that $q < \theta$. As before, from the minimality of x_{k+1}^δ and (2), (8) as well as $x^\dagger \in \mathcal{D}(F)$, we have

$$\begin{aligned}
 &\left((1 - c_{tc})\|F(x_{k+1}^\delta) - y^\delta\| - c_{tc}\|F(x_{k,h}^\delta) - y^\delta\|\right)^p + \alpha_k \mathcal{R}(x_{k+1}^\delta) \\
 &\leq \left(c_{tc}\|F(x_{k,h}^\delta) - y^\delta\| + (1 + c_{tc})\delta\right)^p + \alpha_k \mathcal{R}(x^\dagger),
 \end{aligned}$$

then using (38), (40),

$$\begin{aligned}
 &\left|(1 - c_{tc})(\|F(x_{k+1,h}^\delta) - y^\delta\| - \eta_{k+1}) - c_{tc}\|F(x_{k,h}^\delta) - y^\delta\|\right|^p + \alpha_k \mathcal{R}(x_{k+1,h}^\delta) \\
 &\leq \left(c_{tc}\|F(x_{k,h}^\delta) - y^\delta\| + (1 + c_{tc})\delta\right)^p + \alpha_k \mathcal{R}(x^\dagger) + \alpha_k \zeta_{k+1}.
 \end{aligned}$$

Hence, with the same technique as in the proof of Theorem 1, using (24) with $\epsilon = \frac{1}{2}$, we have

$$\begin{aligned}
 d_{k+1,h} + \alpha_k \mathcal{R}_{k+1,h} &\leq \tilde{q}d_{k,h} + \alpha_0 \theta^k (\mathcal{R}^\dagger + \zeta_{k+1}) + C\delta^p + D\eta_{k+1}^p \\
 &\leq qd_{k,h} + \alpha_0 \theta^k (\mathcal{R}^\dagger + \zeta_{k+1}) + C\delta^p,
 \end{aligned}$$

using (41). From this, by induction we conclude

$$d_{k+1,h} + \alpha_k \mathcal{R}_{k+1,h} \leq q^{k+1}d_0 + \left(\frac{1}{1 - \frac{q}{\theta}}\right) \alpha_k (\mathcal{R}^\dagger + \bar{\zeta}) + \left(\frac{1}{1 - q}\right) C\delta^p \quad (43)$$

Hence, by (39), (41), we have the following estimate

$$\begin{aligned}
 &\|F_h^k(x_{k,h}^\delta) - y^\delta\| \\
 &\leq \left(\delta + \left(\frac{2^{p-1}}{(1 - c_{tc})^p} \left(\theta^k \left(d_0 + \frac{\alpha_0}{\theta - q} (\mathcal{R}^\dagger + \bar{\zeta})\right) + \frac{C}{1 - q} \delta^p\right)\right)^{1/p}\right) \frac{1}{1 - c_\xi},
 \end{aligned}$$

where the right hand side falls below $\tau\delta$ as soon as

$$k \geq (\log 1/\theta)^{-1} \left(p \log(1/\delta) + \log \left(d_0 + \frac{\alpha_0}{\theta - q} (\mathcal{R}^\dagger + \bar{\xi}) \right) - \log \left(\tilde{\tau} - \frac{C}{1 - q} \right) \right) =: \bar{k}(\delta),$$

for $\tilde{\tau} = 2^{1-p}(1 - c_{tc})^p(\tau(1 - c_\xi))^p$. Note that $\tilde{\tau} > \frac{C}{1-q}$, provided τ is chosen sufficiently large, which we assume to be done. That is, we have shown that the discrepancy stopping criterion from (34) will be satisfied after finitely many, namely $O(\log(1/\delta))$, steps.

On the other hand, the continuous discrepancy at the iterate defined by the discretized discrepancy principle (34) by (39), (41) satisfies

$$\|F(x_{k,h}^\delta) - y^\delta\| \leq \tau(1 + c_\xi)\delta.$$

To estimate $\mathcal{R}(x_{k_*(\delta, y^\delta), h}^\delta)$, note that according to our notation, from (43), we get, like in (23), that for all $k \in \{1, \dots, k_*(\delta, y^\delta)\}$

$$\mathcal{R}_k \leq \theta \left(\frac{d_0}{\alpha_0} + \frac{\mathcal{R}^\dagger + \bar{\xi}}{\theta - q} \right) \left(1 + \frac{C}{1 - q} \left(\tilde{\tau} - \frac{C}{1 - q} \right)^{-1} \right) =: R.$$

Now we show finiteness of the stopping index for the discretized Ivanov-IRGNM (32). By minimality of x_{k+1}^δ and (38), for this problem we have

$$(1 - c_{tc})\|F(x_{k+1,h}^\delta) - y^\delta\| \leq 2c_{tc}\|F(x_{k,h}^\delta) - y^\delta\| + (1 + c_{tc})\delta + (1 - c_{tc})\eta_{k+1}.$$

which with

$$\begin{aligned} d_{k,h} &:= (1 - c_{tc})\|F(x_{k,h}^\delta) - y^\delta\|, \\ \tilde{q} &:= \frac{2c_{tc}}{1 - c_{tc}}, \quad q = \tilde{q} + D \frac{c_\eta}{1 - c_\xi} \in (0, 1), \\ C &:= (1 + c_{tc}), \quad D := (1 - c_{tc}), \end{aligned}$$

by induction, (39) and (41) gives

$$\|F_h^k(x_{k,h}^\delta) - y^\delta\| \leq \frac{1}{1 - c_{tc}}d_{k,h} + \xi_k \leq \frac{1}{1 - c_{tc}} \left(q^k d_0 + \frac{C}{1 - q} \delta \right) + \hat{\tau}\delta,$$

where the right hand side is smaller than $\tau\delta$ for all

$$k \geq (\log 1/q)^{-1} \left(\log(1/\delta) + \log d_0 - \log \left(\tilde{\tau} - \frac{C}{1 - q} \right) \right) =: \bar{k}(\delta),$$

with $\tilde{\tau} = (1 - c_{tc})\tau(1 - c_\xi)$, so that we can again conclude $k_*(\delta, y^\delta) \leq \bar{k}(\delta) = O(\log(1/\delta))$.

It remains to show finiteness of the stopping index for the discretized Morozov-IRGNM (33). By minimality of x_{k+1}^δ we have (30) with x_k^δ replaced by $x_{k,h}^\delta$, thus the inequalities (38) and (41) yield

$$\|F(x_{k+1,h}^\delta) - y^\delta\| \leq \frac{\sigma + c_{tc}}{\left(1 - \frac{c_\eta}{1-c_\xi}\right)(1 - c_{tc})} \|F(x_{k,h}^\delta) - y^\delta\|$$

then, by (39) and induction

$$\|F_h^k(x_{k,h}^\delta) - y^\delta\| \leq q^k \|F(x_{0,h}^\delta) - y^\delta\| \tag{44}$$

where

$$q := \frac{\sigma + c_{tc}}{\left(1 - \frac{c_\eta}{1-c_\xi}\right)(1 - c_{tc})} \in (0, 1),$$

and the right hand side of (44) falls below $\tau\delta$ for all

$$k \geq (\log 1/q)^{-1} (\log(1/\delta) + \log d_0 - \log \tilde{\tau}) =: \bar{k}(\delta),$$

where $\tilde{\tau} = \tau(1 - c_\xi)$, and we can again conclude $k_*(\delta, y^\delta) \leq \bar{k}(\delta) = O(\log(1/\delta))$. Boundedness of the \mathcal{R} values for (33) by $\mathcal{R}(x^\dagger) + \zeta$ follows like in the proof of Theorem 1 together with (40), (41). \square

3 Error estimators for adaptive discretization

The error estimators η_k, ξ_k and ζ_k can be quantified, e.g., by means of a goal oriented dual weighted residual (DWR) approach [3], applied to the minimization problems

$$\begin{aligned} (x_{k+1,h}^\delta, v_{k,h}^\delta, u_{k+1}^\delta, u_{k,h}^\delta) &\in \operatorname{argmin}_{(x,v,u,\tilde{u}) \in \mathcal{C} \times V^3} \|C'(\tilde{u})v + C(\tilde{u}) - y^\delta\|^p + \alpha_k \mathcal{R}(x) \\ \text{s.t. } \forall w \in W : &\langle A'_x(x_{k,h}^\delta, \tilde{u})(x - x_{k,h}^\delta) + A'_u(x_{k,h}^\delta, \tilde{u})v, w \rangle_{W^*, W} = 0, \\ &\langle A(x_{k,h}^\delta, \tilde{u}), w \rangle_{W^*, W} = 0, \quad \langle A(x, u), w \rangle_{W^*, W} = 0, \end{aligned} \tag{45}$$

(note that the last constraint is added in order to enable computation of I_2^k below)

$$\begin{aligned} (x_{k+1,h}^\delta, v_{k,h}^\delta, u_{k+1}^\delta, u_{k,h}^\delta) &\in \operatorname{argmin}_{(x,v,u,\tilde{u}) \in \mathcal{C} \times V^3} \frac{1}{2} \|C'(\tilde{u})v + C(\tilde{u}) - y^\delta\|^2 \\ \text{s.t. } \mathcal{R}(x) &\leq \rho_k, \\ \text{and } \forall w \in W : &\langle A'_x(x_{k,h}^\delta, \tilde{u})(x - x_{k,h}^\delta) + A'_u(x_{k,h}^\delta, \tilde{u})v, w \rangle_{W^*, W} = 0, \\ &\langle A(x_{k,h}^\delta, \tilde{u}), w \rangle_{W^*, W} = 0, \quad \langle A(x, u), w \rangle_{W^*, W} = 0, \end{aligned} \tag{46}$$

and

$$\begin{aligned}
 (x_{k+1,h}^\delta, v_{k,h}^\delta, u_{k+1}^\delta, u_{k,h}^\delta) &\in \operatorname{argmin}_{(x,v,u,\tilde{u}) \in \mathcal{C} \times \mathcal{V}^3} \mathcal{R}(x) \\
 \text{s.t. } &\|C'(\tilde{u})v + C(\tilde{u}) - y^\delta\| \leq \sigma \|C(\tilde{u}) - y^\delta\|, \\
 \text{and } \forall w \in W : &\langle A'_x(x_{k,h}^\delta, \tilde{u})(x - x_{k,h}^\delta) + A'_u(x_{k,h}^\delta, \tilde{u})v, w \rangle_{W^*, W} = 0, \\
 &\langle A(x_{k,h}^\delta, \tilde{u}), w \rangle_{W^*, W} = 0, \quad \langle A(x, u), w \rangle_{W^*, W} = 0,
 \end{aligned}
 \tag{47}$$

which are equivalent to (5), (6), and (7), respectively, with

$$\begin{aligned}
 I_1^k(x, v, u, \tilde{u}) &= \|C(\tilde{u}) - y^\delta\|, \quad I_2^k(x, v, u, \tilde{u}) = \|C(u) - y^\delta\|, \\
 I_3^k(x, v, u, \tilde{u}) &= \mathcal{R}(x)
 \end{aligned}$$

as quantities of interest [where I_3^k is only needed for (5) and (7)]. We assume that C , \mathcal{R} and the norms can be evaluated without discretization error, so the discretized versions of I_i^k only arise due to discreteness of the arguments. Indeed, it is easy to see that the left hand sides of (38) and (39) can be bounded (at least approximately) by combinations of I_1^k and I_2^k , using the triangle inequality:

$$\begin{aligned}
 &\|F(x_{k+1,h}^\delta) - y^\delta\| - \|F(x_{k+1}^\delta) - y^\delta\| \\
 &= I_1^{k+1}(x_{k+2}^\delta, v_{k+1}^\delta, u_{k+2}^\delta, \tilde{u}_{k+1}^\delta) - I_1^{k+1}(x_{k+2,h}^\delta, v_{k+1,h}^\delta, u_{k+2,h}^\delta, \tilde{u}_{k+1,h}^\delta) \\
 &\quad - (I_2^k(x_{k+1}^\delta, v_k^\delta, u_{k+1}^\delta, \tilde{u}_k^\delta) - I_2^k(x_{k+1,h}^\delta, v_{k,h}^\delta, u_{k+1,h}^\delta, \tilde{u}_{k,h}^\delta)) + R_\eta^{k+1}; \tag{48} \\
 &\|F_h^k(x_{k,h}^\delta) - y^\delta\| - \|F(x_{k,h}^\delta) - y^\delta\| \\
 &= I_1^k(x_{k+1,h}^\delta, v_{k,h}^\delta, u_{k+1,h}^\delta, \tilde{u}_{k,h}^\delta) - I_1^k(x_{k+1}^\delta, v_k^\delta, u_{k+1}^\delta, \tilde{u}_k^\delta), \tag{49}
 \end{aligned}$$

where we will neglect $R_\eta^{k+1} = \|F_h^{k+1}(x_{k+1,h}^\delta) - y^\delta\| - \|F_h^k(x_{k+1,h}^\delta) - y^\delta\|$.

It is important to note that $I_{1,h}^{k+1}$ is not equal to $I_{2,h}^k$, see [18].

The computation of the a posteriori error estimators η_k, ξ_k, ζ_k is done as in [18]. These error estimators can be used within the following adaptive algorithm for error control and mesh refinement: We start on a coarse mesh, solve the discretized optimization problem and evaluate the error estimator. Thereafter, we refine the current mesh using local information obtained from the error estimator, reducing the error with respect to the quantity of interest. This procedure is iterated until the value of the error estimator is below the given tolerance (41), cf. [3].

In this case, all the variables x, v, u, \tilde{u} are subject to a new discretization. For better readability we will partially omit the iteration index k and the discretization index h . The previous iterate x_k^δ is fixed and not subject to a new discretization.

Consider now the cost functional for (45)

$$J(x, v, \tilde{u}) = \|C'(\tilde{u})v + C(\tilde{u}) - y^\delta\|^p + \alpha_k \mathcal{R}(x)$$

and define the Langrangian functional

$$L(x, v, u, \tilde{u}, \lambda, \tilde{\mu}, \mu) := J(x, v, \tilde{u}) + \langle A'_x(x_k^\delta, \tilde{u})(x - x_k^\delta) + A'_u(x_k^\delta, \tilde{u})v, \lambda \rangle_{W^*, W} + \langle A(x_k^\delta, \tilde{u}), \tilde{\mu} \rangle_{W^*, W} + \langle A(x, u), \mu \rangle_{W^*, W}, \tag{50}$$

neglecting for simplicity (cf. Remark 2) the constraints defined by \mathcal{C} . The first-order necessary optimality conditions for (45) are given by stationarity for the Lagrangian L . Setting $z = (x, v, u, \tilde{u}, \lambda, \tilde{\mu}, \mu)$, they read

$$L'(z)(dz) = 0, \forall dz \in Z = X \times V \times V \times V \times W \times W \times W$$

and for the discretized problem,

$$L'(z_h)(dz_h) = 0, \forall dz_h \in Z_h = X_h \times V_h \times V_h \times V_h \times W_h \times W_h \times W_h.$$

To derive a posteriori error estimators for the error with respect to the quantities of interest (I_1, I_2, I_3) , we introduce auxiliary functionals M_i :

$$M_i(z, \bar{z}) = I_i(z) + L'(z)\bar{z}, \quad z, \bar{z} \in Z, \quad i = 1, 2, 3,$$

Let $\tilde{z} = (z, \bar{z}) \in \tilde{Z} = Z \times Z$ and $\tilde{z}_h = (z_h, \bar{z}_h) \in \tilde{Z}_h = Z_h \times Z_h$ be continuous and discrete stationary points of M_i satisfying

$$M'(\tilde{z})(d\tilde{z}) = 0, \forall d\tilde{z} \in Z \quad M'(\tilde{z}_h)(d\tilde{z}_h) = 0, \forall d\tilde{z}_h \in Z_h,$$

respectively. Then, z, z_h are continuous and discrete stationary points of L and there holds $I_i(z) = M_i(\tilde{z}), i = 1, 2, 3$. Thus the z part, as computed already during the numerical solution of the minimization problem (45) (or (46)) remains fixed for all $i \in \{1, 2, 3, \}$. Moreover, after computing the discrete stationary point z_h for L (e.g., by applying Newton’s method), it requires only one more Newton step to compute the \bar{z} coordinate of the stationary point for M from

$$L''(z_h)(\bar{z}_{i,h}, d\bar{z}) = -I'_i(z_h)d\bar{z}, \forall d\bar{z}_h \in Z_h.$$

According to [3], there holds

$$I_i(x, v, \tilde{u}) - I_i(x_h, v_h, \tilde{u}_h) = \frac{1}{2}M'(\tilde{z}_h)(\tilde{z} - \hat{z}_h) + R, \quad \forall \hat{z}_h \in Z_h \quad i = 1, 2, 3,$$

with a remainder term R of order $O(\|\tilde{z} - \tilde{z}_h\|^3)$ that is therefore neglected. Thus we use

$$I_i^k(z) - I_i^k(z_h) \approx \frac{1}{2}M'_i(z_h, \bar{z}_{i,h})(\pi_h \tilde{z}_{i,h} - \tilde{z}_{i,h}) = \varepsilon_i^k,$$

where π_h is an operator defined such that $(\pi_h \tilde{z}_{i,h} - \tilde{z}_{i,h})$ approximates the interpolation error as in [18], typically defined by local averaging, to define the estimators η_k, ξ_k, ζ_k according to the rule

$$\eta_{k+1} = \varepsilon_1^{k+1} + \varepsilon_2^k, \quad \xi_k = \varepsilon_1^k, \quad \zeta_k = \varepsilon_3^k; \tag{51}$$

cf. (48), (49). The estimators obtained by this procedure can be used to trigger local mesh refinement until the requirements (41) are met cf. [3].

Explicitly, for $p = 2$ (for simplicity) such a stationary point $z = (x, v, u, \tilde{u}, \lambda, \tilde{\mu})$ can be computed by solving the following system of equations (analogously for the discrete stationary point of L)

$$-(A'_x(x, u)^* \mu + A'_x(x_k^\delta, \tilde{u})^* \lambda) \in \alpha_k \partial \mathcal{R}(x); \tag{52}$$

$$2\langle C'(\tilde{u})(dv), C'(\tilde{u})v + C(\tilde{u}) - y^\delta \rangle + \langle A'_u(x_k^\delta, \tilde{u})(dv), \lambda \rangle = 0, \quad \forall dv \in V; \tag{53}$$

$$\langle A'_u(x, u)(du), \mu \rangle = 0, \quad \forall du \in V; \tag{54}$$

$$\langle A''_{xu}(x_k^\delta, \tilde{u})(x - x_k^\delta, d\tilde{u}) + A''_{uu}(x_k^\delta, \tilde{u})(v, d\tilde{u}), \lambda \rangle + \langle A'_u(x_k^\delta, \tilde{u})(d\tilde{u}), \tilde{\mu} \rangle + 2\langle C''(\tilde{u})(d\tilde{u}, v) + C'(\tilde{u})(d\tilde{u}), C'(\tilde{u})v + C(\tilde{u}) - y^\delta \rangle = 0, \quad \forall d\tilde{u} \in V; \tag{55}$$

$$\langle A'_x(x_k^\delta, \tilde{u})(x - x_k^\delta) + A'_u(x_k^\delta, \tilde{u})v, d\lambda \rangle = 0, \quad \forall d\lambda \in W; \tag{56}$$

$$\langle A(x_k^\delta, \tilde{u}), d\tilde{\mu} \rangle = 0, \quad \forall d\tilde{\mu} \in W; \tag{57}$$

$$\langle A(x, u), d\mu \rangle = 0, \quad \forall d\mu \in W. \tag{58}$$

Note that (58) is decoupled from the other equations and that if $A'_u(x, u)^*$ is injective, Eq. (54) implies $\mu = 0$.

Summarizing, since we have a convex minimization problem, after solving a non-linear system of seven equations to find the minimizer, we need only one more Newton step to compute the error estimators to check whether we need a refinement on the mesh or not.

Regarding the problem (46) related to the Ivanov-IRGNM, we have the Lagrangian functional (50) with the cost functional defined by

$$J(x, v, \tilde{u}) = \frac{1}{2} \|C'(\tilde{u})v + C(\tilde{u}) - y^\delta\|^2 + I_{(-\infty, 0]}(\mathcal{R}(x) - \rho),$$

and the indicator functional $I_{(-\infty, 0]}(\mathcal{R}(x) - \rho)$ takes the role of a regularization functional. The resulting optimality system is the same as above, cf. (52)-(58), just with (52) replaced by

$$-(A'_x(x, u)^* \mu + A'_x(x_k^\delta, \tilde{u})^* \lambda) \in \partial I_{(-\infty, 0]}(\mathcal{R}(x) - \rho). \tag{59}$$

Similarly for (47) for Morozov-IRGNM, with the cost function

$$J(x, v, \tilde{u}) = \mathcal{R}(x) + \underbrace{I_{(-\infty, 0]}(\|C'(\tilde{u})v + C(\tilde{u}) - y^\delta\| - \sigma \|C(\tilde{u}) - y^\delta\|)}_{=: \mathcal{Q}(\tilde{u}, v)},$$

we end up with an optimality system by setting $\alpha_k = 1$ and replacing (53), (55) in (52)–(58) by

$$-A'_u(x_k^\delta, \tilde{u})^* \lambda \in \partial_v Q(\tilde{u}, v) \tag{60}$$

$$-A''_{xu}(x_k^\delta, \tilde{u})(x - x_k^\delta) + A''_{uu}(x_k^\delta, \tilde{u})v^* \lambda - A'_u(x_k^\delta, \tilde{u})^* \tilde{\mu} \in \partial_{\tilde{u}} Q(\tilde{u}, v) \tag{61}$$

respectively.

Note that the bound on I_2 only appears—via (51)—in connection to the assumption $\eta_k \leq c_\eta \|F_h^k(x_{k,h}^\delta) - y^\delta\|$, for $k \leq k_*(\delta, y^\delta)$ in (41). This may be satisfied in practice without refining explicitly with respect to η_k , but simply by refining with respect to the other error estimators ξ_k (and ζ_k in the Tikhonov or Morozov case). The fact that $I_{1,h}^k$ and $I_{2,h}^{k-1}$ only differ in the discretization level, motivates the assumption that for small h , we have $I_{1,h}^k \approx I_{2,h}^{k-1}$ and $\eta_{k-1} \approx \xi_k$. Therefore, the algorithm used in actual computations will be built neglecting I_2 and hence skipping the constraint $\langle A(x, u), w \rangle_{W^*, W} = 0, \forall w \in W$ in (45), (46), (47), which implies a modification of the Lagrangian (50) accordingly. Therefore, the corresponding optimality systems for $p = 2$ in the Tikhonov case is given by

$$-A'_x(x_k^\delta, \tilde{u})^* \lambda \in \alpha_k \partial \mathcal{R}(x); \tag{62}$$

$$2\langle C'(\tilde{u})(dv), C'(\tilde{u})v + C(\tilde{u}) - y^\delta \rangle + \langle A'_u(x_k^\delta, \tilde{u})(dv), \lambda \rangle = 0, \quad \forall dv \in V; \tag{63}$$

$$\langle A''_{xu}(x_k^\delta, \tilde{u})(x - x_k^\delta, d\tilde{u}) + A''_{uu}(x_k^\delta, \tilde{u})(v, d\tilde{u}), \lambda \rangle + \langle A'_u(x_k^\delta, \tilde{u})(d\tilde{u}), \tilde{\mu} \rangle + 2\langle C''(\tilde{u})(d\tilde{u}, v) + C'(\tilde{u})(d\tilde{u}), C'(\tilde{u})v + C(\tilde{u}) - y^\delta \rangle = 0, \quad \forall d\tilde{u} \in V; \tag{64}$$

$$\langle A'_x(x_k^\delta, \tilde{u})(x - x_k^\delta) + A'_u(x_k^\delta, \tilde{u})v, d\lambda \rangle = 0, \quad \forall d\lambda \in W; \tag{65}$$

$$\langle A(x_k^\delta, \tilde{u}), d\tilde{\mu} \rangle = 0, \quad \forall d\tilde{\mu} \in W. \tag{66}$$

Note that Eq. (66) is decoupled from the others. Therefore, the strategy is to solve (66) for \tilde{u} first, then solve the linear system (62), (63), (65) for (x, v, λ) , and finally compute $\tilde{\mu}$ via the linear equation (64). Here, the system (62), (63), (65) can be interpreted as the optimality conditions for the following problem

$$(x_{k+1,h}^\delta, v_{k,h}^\delta) \in \operatorname{argmin}_{(x,v) \in C \times V} \|C'(\tilde{u})v + C(\tilde{u}) - y^\delta\|^2 + \alpha_k \mathcal{R}(x) \\ \text{s.t. } \forall w \in W : \langle A'_x(x_{k,h}^\delta, \tilde{u})(x - x_{k,h}^\delta) + A'_u(x_{k,h}^\delta, \tilde{u})v, w \rangle_{W^*, W} = 0.$$

For the Ivanov case, we have to solve (63)–(66) with

$$-A'_x(x_k^\delta, \tilde{u})^* \lambda \in \partial I_{(-\infty, 0]}(\mathcal{R}(x) - \rho) \tag{67}$$

in place of (62), hence again (66) is decoupled from the other equations, (64) is linear with respect to $\tilde{\mu}$, once (x, v, λ) has been computed, and the remaining system for (x, v, λ) can be interpreted as the optimality conditions for the following problem

$$\begin{aligned}
 (x_{k+1,h}^\delta, v_{k,h}^\delta) &\in \operatorname{argmin}_{(x,v) \in \mathcal{C} \times V} \frac{1}{2} \|C'(\tilde{u})v + C(\tilde{u}) - y^\delta\|^2 \\
 \text{s.t. } &\mathcal{R}(x) \leq \rho_k, \\
 \text{and } \forall w \in W : &\langle A'_x(x_{k,h}^\delta, \tilde{u})(x - x_{k,h}^\delta) + A'_u(x_{k,h}^\delta, \tilde{u})v, w \rangle_{W^*, W} = 0.
 \end{aligned}$$

The Morozov case requires solution of (62) (with $\alpha_k = 1$), (65), (66), (60), (61). Thus again, we first solve (66) for \tilde{u} , then the system (62) (with $\alpha_k = 1$), (65), (60), which is the first order optimality condition for

$$\begin{aligned}
 (x_{k+1,h}^\delta, v_{k,h}^\delta) &\in \operatorname{argmin}_{(x,v) \in \mathcal{C} \times V} \mathcal{R}(x) \\
 \text{s.t. } &\|C'(\tilde{u})v + C(\tilde{u}) - y^\delta\| \leq \sigma \|C(\tilde{u}) - y^\delta\|, \\
 \text{and } \forall w \in W : &\langle A'_x(x_{k,h}^\delta, \tilde{u})(x - x_{k,h}^\delta) + A'_u(x_{k,h}^\delta, \tilde{u})v, w \rangle_{W^*, W} = 0,
 \end{aligned}$$

with Lagrange multiplier λ for the equality constraint, and finally the (now possibly nonlinear) inclusion (61) for $\tilde{\mu}$.

Remark 3 Since DWR estimators are based on residuals which are computed in the optimization process, the additional costs for estimation are very low, which makes this approach attractive for our purposes. However, although these error estimators are known to work efficiently in practice (see [3]), they are not reliable, i.e., the conditions $I_i^k(z) - I_i^k(z_h) \leq \epsilon_i^k, i = 1, 2, 3$ can not be guaranteed in a strict sense in the computations, since we neglect the remainder term R and use an approximation for $\tilde{z} - \hat{z}_h$. As our analysis in Theorem 1 is kept rather general, it is not restricted to DWR estimators and would also work with different (e.g., reliable) error estimators.

4 Model examples

We present a model example to illustrate the abstract setting from the previous section. Consider the following inverse source problem for a semilinear elliptic PDE, where the model and observation equations are given by

$$-\Delta u + \kappa u^3 = \chi_{\omega_c} x \text{ in } \Omega \subset \mathbb{R}^d, \tag{68}$$

$$u = 0 \text{ on } \partial\Omega, \tag{69}$$

$$C(u) = u|_{\omega_o}, \|y - y^\delta\|_{L^2(\omega_o)} \leq \delta, \tag{70}$$

where χ_{ω_c} denotes the extension by zero of a function on ω_c to a function on all of Ω . We first of all consider Tikhonov regularization and, aiming for a sparsely supported source, therefore use the space of Radon measures $\mathcal{M}(\omega_c)$ as a preimage space X . Thus we define the operators $A : \mathcal{M}(\omega_c) \times W_0^{1,q'}(\Omega) \rightarrow W^{-1,q}(\Omega), A(x, u) = -\Delta u + \kappa u^3 - \chi_{\omega_c} x, \kappa \in \mathbb{R}$ and the injection $C : W_0^{1,q'}(\Omega) \rightarrow L^2(\omega_o) = Y, q > d$, where Ω is a bounded domain in \mathbb{R}^d with $d = 2$ or 3 , with Lipschitz boundary $\partial\Omega$ and $\omega_c, \omega_o \subset \Omega$ are the control domain and the observation domain, respectively.

A monotonicity argument yields well posedness of the above semilinear boundary value problem, i.e., well-definedness of $u \in W_0^{1,q'}(\Omega)$ as a solution to the elliptic

boundary value problem (68), (69), as long as we can guarantee that $u^3 \in W^{-1,q}(\Omega)$ for any $u \in W_0^{1,q'}(\Omega)$, i.e., the embeddings $W_0^{1,q'}(\Omega) \rightarrow L^{3r}(\Omega)$ and $L^r(\Omega) \rightarrow W^{-1,q}(\Omega)$ are continuous for some $r \in [1, \infty]$, which (by duality) is the case iff $W_0^{1,q'}(\Omega)$ embeds continuously both into $L^{3r}(\Omega)$ and $L^{r'}(\Omega)$. By Sobolev’s Embedding Theorem, this boils down to the inequalities

$$1 - \frac{d}{q'} \geq -\frac{d}{3r} \text{ and } 1 - \frac{d}{q'} \geq -\frac{d}{r'},$$

which by elementary computations turns out to be equivalent to

$$\frac{dq}{q+d} \leq r \leq \frac{dq}{3(dq-q-d)}, \tag{71}$$

where the left hand side is larger than one and the denominator on the right hand side is positive due to the fact that for $d \geq 2$ we have $q > d \geq d' = \frac{d}{d-1}$. Taking the extremal bounds for $q > d$ —note that the lower bound is increasing and the upper bound is decreasing with q —in (72) we get

$$\frac{d}{2} < r < \frac{d}{3(d-2)}. \tag{72}$$

Thus, as a by-product, we get that for any $t \in [1, \bar{t})$ there exists $q > d$ such that $W_0^{1,q'}(\Omega)$ continuously embeds into L^t , with

$$\bar{t} = \infty \text{ in case } d = 2 \text{ and } \bar{t} = 3 \text{ in case } d = 3. \tag{73}$$

For the regularization functional $\mathcal{R}(x) = \|x\|_{\mathcal{M}(\omega_c)}$, the IRGNM-Tikhonov minimization step is given by (ignoring h in the notation)

$$\begin{aligned} (x_{k+1}^\delta, v_k^\delta, u_k^\delta) &\in \operatorname{argmin}_{(x,v,\tilde{u}) \in \mathcal{M}(\omega_c) \times (W_0^{1,q'}(\Omega))^2} \|v + \tilde{u} - y^\delta\|_{L^2(\omega_o)}^2 \\ &\quad + \alpha_k \|x\|_{\mathcal{M}(\omega_c)} \\ \text{s.t. } \forall w \in W_0^{1,q'}(\Omega) : &\int_{\Omega} (\nabla v \nabla w + 3\kappa \tilde{u}^2 v w) d\Omega = \int_{\omega_c} w d(x - x_k^\delta), \\ &\int_{\Omega} (\nabla \tilde{u} \nabla w + \kappa \tilde{u}^3 w) d\Omega = \int_{\omega_c} w dx_k^\delta. \end{aligned}$$

Here and below $\int_{\Omega} d\Omega$ and $\int_{\omega_c} dx$ denote the integrals with respect to the Lebesgue measure and with respect to the measure x , respectively.

Therefore, to compute this Gauss–Newton step, one first needs to solve the nonlinear equation

$$-\Delta \tilde{u} + \kappa \tilde{u}^3 = \chi_{\omega_c} x_k^\delta \tag{74}$$

for $\tilde{u} = u_k^\delta$, then solve the following optimality system with respect to (x, v, λ) (written in a strong formulation)

$$\begin{aligned} \|\lambda\|_{C_b(\omega_c)} \leq \alpha_k \text{ and } \int_{\Omega} (x^* - \lambda) dx \leq 0, \forall x^* \in B_{\alpha_k}^{C_b(\omega_c)} \\ -\Delta\lambda + 3\kappa(u_k^\delta)^2\lambda + 2v + 2u_k^\delta = 2\chi_{\omega_o}y^\delta \\ -\Delta v + 3\kappa(u_k^\delta)^2v = \chi_{\omega_c}(x - x_k^\delta), \end{aligned}$$

which can be interpreted as the optimality system for the minimization problem

$$\begin{aligned} (x_{k+1}^\delta, v_k^\delta) \in \operatorname{argmin}_{(x,v) \in \mathcal{M}(\omega_c) \times W_0^{1,q'}(\Omega)} \|u_k^\delta + v - y^\delta\|_{L^2(\omega_o)}^2 + \alpha_k \|x\|_{\mathcal{M}(\omega_c)} \\ \text{s.t. } -\Delta v + 3\kappa(u_k^\delta)^2v = \chi_{\omega_c}(x - x_k^\delta), \end{aligned} \tag{75}$$

with Lagrange multiplier λ for the equality constraint, and finally, compute $\tilde{\mu}$ by solving

$$-\Delta\tilde{\mu} + 3\kappa(u_k^\delta)^2\tilde{\mu} = -6\kappa u_k^\delta v \lambda - 2(v + u_k^\delta - \chi_{\omega_o}y^\delta). \tag{76}$$

For carrying out the IRGNM iteration, $\tilde{\mu}$ is not required, but we need it for evaluating the error estimators.

For the Ivanov case, we consider the same model and observation equations (68), (69), (70) but now we intend to regularize by imposing L^∞ bounds and thus use the slightly different function space setting, $A : L^\infty(\omega_c) \times H_0^1(\Omega) \rightarrow H^{-1}(\Omega)$, $A(x, u) = -\Delta u + \kappa u^3 - x$, $\kappa \in \mathbb{R}$ and the injection $C : H_0^1(\Omega) \rightarrow L^2(\omega_o)$.

The IRGNM-Ivanov minimization step with the regularization functional $\mathcal{R}(x) = \|x\|_{L^\infty(\omega_c)}$ is given by

$$\begin{aligned} (x_{k+1}^\delta, v_k^\delta, u_k^\delta) \in \operatorname{argmin}_{(x,v,\tilde{u}) \in L^\infty(\omega_c) \times (H_0^1(\Omega))^2} \|v + \tilde{u} - y^\delta\|_{L^2(\omega_o)}^2 \\ \text{s.t. } \|x\|_{L^\infty(\omega_c)} \leq \rho \\ \text{and } \forall w \in H_0^1(\Omega) : \int_{\Omega} (\nabla v \nabla w + 3\kappa \tilde{u}^2 v w) d\Omega = \int_{\omega_c} w(x - x_k^\delta) d\Omega, \\ \int_{\Omega} (\nabla \tilde{u} \nabla w + \kappa \tilde{u}^3 w) d\Omega = \int_{\omega_c} w x_k^\delta d\Omega. \end{aligned}$$

For the Gauss–Newton step, one needs to first solve the nonlinear equation (74) for $\tilde{u} = u_k^\delta$, and then solve the following optimality system with respect to (x, v, λ)

$$\begin{aligned} \|x\|_{L^\infty(\omega_c)} \leq \rho \text{ and } \int_{\omega_c} (x^* - x)\lambda d\Omega \leq 0, \forall x^* \in B_\rho^{L^\infty(\omega_c)} \\ -\Delta\lambda + 3\kappa(u_k^\delta)^2\lambda + 2v + 2u_k^\delta = 2\chi_{\omega_o}y^\delta \\ -\Delta v + 3\kappa(u_k^\delta)^2v = \chi_{\omega_c}(x - x_k^\delta), \end{aligned}$$

which can be interpreted as the optimality system for the minimization problem

$$\begin{aligned}
 (x_{k+1}^\delta, v_k^\delta) &\in \operatorname{argmin}_{(x,v) \in L^\infty(\omega_c) \times H_0^1(\Omega)} \frac{1}{2} \|u_k^\delta + v - y^\delta\|_{L^2(\omega_o)}^2 \\
 \text{s.t. } &\|x\|_{L^\infty(\omega_c)} \leq \rho \\
 &-\Delta v + 3\kappa(u_k^\delta)^2 v = \chi_{\omega_c}(x - x_k^\delta)
 \end{aligned} \tag{77}$$

with Lagrange multiplier λ for the equality constraint. Finally, $\tilde{\mu}$ is computed from (76).

For the IRGNM-Morozov case, using for simplicity the regularization functional $\mathcal{R}(x) = \frac{1}{2} \|x\|_{L^2(\omega_c)}^2$, and leaving the rest of the setting as in the IRGNM-Ivanov case, the step is defined by

$$\begin{aligned}
 (x_{k+1}^\delta, v_k^\delta, u_k^\delta) &\in \operatorname{argmin}_{(x,v,\tilde{u}) \in L^\infty(\omega_c) \times (H_0^1(\Omega))^2} \frac{1}{2} \|x\|_{L^2(\omega_c)}^2 \\
 \text{s.t. } &\|v + \tilde{u} - y^\delta\|_{L^2(\omega_o)}^2 \leq \sigma \|\tilde{u} - y^\delta\|_{L^2(\omega_o)}^2 \\
 \text{and } \forall w \in H_0^1(\Omega) : &\int_{\Omega} (\nabla v \nabla w + 3\kappa \tilde{u}^2 v w) d\Omega = \int_{\omega_c} w(x - x_k^\delta) d\Omega, \\
 &\int_{\Omega} (\nabla \tilde{u} \nabla w + \kappa \tilde{u}^3 w) d\Omega = \int_{\omega_c} w x_k^\delta d\Omega.
 \end{aligned}$$

So again we first solve (74) for $\tilde{u} = u_k^\delta$, then the minimization problem

$$\begin{aligned}
 (x_{k+1}^\delta, v_k^\delta) &\in \operatorname{argmin}_{(x,v) \in L^\infty(\omega_c) \times H_0^1(\Omega)} \frac{1}{2} \|x\|_{L^2(\omega_c)}^2 \\
 \text{s.t. } &\|v + u_k^\delta - y^\delta\|_{L^2(\omega_o)}^2 \leq \sigma \|u_k^\delta - y^\delta\|_{L^2(\omega_o)}^2 \\
 &-\Delta v + 3\kappa(u_k^\delta)^2 v = \chi_{\omega_c}(x - x_k^\delta),
 \end{aligned}$$

or actually its first order optimality system

$$\begin{aligned}
 \lambda|_{\omega_c} &= x \\
 \phi \geq 0, &\|v + u_k^\delta - y^\delta\|_{L^2(\omega_o)}^2 \leq \sigma \|u_k^\delta - y^\delta\|_{L^2(\omega_o)}^2, \\
 \phi &(\|v + u_k^\delta - y^\delta\|_{L^2(\omega_o)}^2 - \sigma \|u_k^\delta - y^\delta\|_{L^2(\omega_o)}^2) = 0 \\
 -\Delta v &+ 3\kappa(u_k^\delta)^2 v = \chi_{\omega_c}(x - x_k^\delta),
 \end{aligned}$$

for $(x_{k+1}^\delta, v_k^\delta, \phi, \lambda)$, and finally,

$$-\Delta \tilde{\mu} + 3\kappa(u_k^\delta)^2 \tilde{\mu} = -6\kappa u_k^\delta v \lambda - \phi(v + (1 - \sigma)(u_k^\delta - \chi_{\omega_o} y^\delta)).$$

for $\tilde{\mu}$.

For numerically efficient methods to solve the minimization problems (75) and (77) we refer to e.g., [4–6] and the references therein.

We finally check the tangential cone condition in case $\omega_o = \Omega$ and, for simplicity also $\omega_c = \Omega$, in both settings

$$X = \mathcal{M}(\omega_c), \quad V = W_0^{1,q'}(\Omega), \quad W = W_0^{1,q}(\Omega)$$

(where we will have to restrict ourselves to $d = 2$) and

$$X = L^\infty(\omega_c) \text{ or } X = L^2(\omega_c), \quad V = W = H_0^1(\Omega).$$

For this purpose, we use the fact that with the notation $F(\tilde{x}) = \tilde{u}|_{\omega_o}$, $F(x) = u|_{\omega_o}$, $F(\tilde{x}) - F(x) = v|_{\omega_o}$ and $F(\tilde{x}) - F(x) - F'(x)(\tilde{x} - x) = w|_{\omega_o}$, the functions $v, w \in W_0^{1,q'}(\Omega)$ satisfy the homogeneous Dirichlet boundary value problems for the equations

$$\begin{aligned} -\Delta v + \kappa(\tilde{u}^2 + \tilde{u}u + u^2)v &= \tilde{x} - x \\ -\Delta w + \kappa u^2 w &= -\kappa(\tilde{u} + 2u)v^2. \end{aligned}$$

Using an Aubin-Nitsche type duality trick, we can estimate the L^2 norm of w via the adjoint state $p \in W_0^{1,n}(\Omega)$, which solves

$$-\Delta p + \kappa u^2 p = w,$$

with homogeneous Dirichlet boundary conditions, so that by Hölder’s inequality

$$\begin{aligned} \|w\|_{L^2(\Omega)}^2 &= \langle w, (-\Delta + \kappa u^2 \text{id})p \rangle = \langle (-\Delta + \kappa u^2 \text{id})w, p \rangle \\ &= -\kappa \langle (\tilde{u} + 2u)v^2, p \rangle \leq \kappa \|v\|_{L^2(\Omega)} \|\tilde{u} + 2u\|_{L^m(\Omega)} \|v\|_{L^m(\Omega)} \|p\|_{L^{\frac{2m}{m-4}}(\Omega)} \\ &\leq \tilde{C}\kappa \|v\|_{L^2(\Omega)} \|\tilde{u} + 2u\|_{L^m(\Omega)} \|v\|_{L^m(\Omega)} \|w\|_{L^2(\Omega)}, \end{aligned}$$

where we aim at choosing $m \in [4, \infty]$, $n \in [1, \infty]$ such that indeed

$$\|p\|_{W_0^{1,n}(\Omega)} \leq C \|w\|_{W^{-1,n}(\Omega)} \leq \tilde{C} \|w\|_{L^2(\Omega)}$$

and the embeddings $V \rightarrow L^m(\Omega)$, $W^{1,n}(\Omega) \rightarrow L^{\frac{2m}{m-4}}(\Omega)$, $L^2(\Omega) \rightarrow W^{-1,n}(\Omega)$ are continuous. If we succeed in doing so, we can bound $\tilde{C}\kappa \|\tilde{u} + 2u\|_{L^m(\Omega)} \|v\|_{L^m(\Omega)}$ by some constant c_{tc} , which will be small provided $\|\tilde{x} - x\|_X$ and hence $\|v\|_{L^m(\Omega)}$ is small. Thus, the numbers n, m are limited by the requirements

$$V \subseteq L^m(\Omega) \text{ and } W^{1,n}(\Omega) \subseteq L^{\frac{2m}{m-4}}(\Omega) \text{ and } m \geq 4, \tag{78}$$

$L^2(\Omega) \subseteq W^{-1,n}(\Omega)$, i.e., by duality,

$$W_0^{1,n'}(\Omega) \subseteq L^2(\Omega), \tag{79}$$

and the fact that $\kappa u^2 p \in L^o(\Omega)$ should be contained in $W^{-1,n'}(\Omega)$ for $u \in V \subseteq L^t(\Omega)$, and $p \in W^{1,n}(\Omega)$, which via Hölder’s inequality in

$$\left(\int_{\Omega} (u^2 p)^o d\Omega \right)^{1/o} \leq \|u\|_{L^t(\Omega)}^2 \|p\|_{L^{\frac{ot}{t-2o}}(\Omega)}$$

and duality leads to the requirements

$$W_0^{1,n}(\Omega) \subseteq L^{o'}(\Omega) \text{ and } V \subseteq L^t(\Omega) \text{ and } W_0^{1,n}(\Omega) \subseteq L^{\frac{ot}{t-2o}}(\Omega) \text{ and } o \leq \frac{t}{2} \quad (80)$$

In case $V = W_0^{1,q'}(\Omega)$ with $q > d$ and $d = 3$, (78) will not work out, since according to (73), m cannot be chosen larger or equal to four.

In case $V = W_0^{1,q'}(\Omega)$ with $q > d$ and $d = 2$, we can choose, e.g., $t = m = n = 6$, $o = 2$ to satisfy (78), (79), (80) as well as $t, m < \bar{t}$ as in (73).

The same choice is possible in case $V = H_0^1(\Omega)$ with $d \in \{2, 3\}$.

5 Numerical tests

In this section, we provide some numerical illustration of the IRGNM Ivanov method applied to the example from Sect. 4, i.e., each Newton step consists of solving (74) and subsequently (77). For the numerical solution of (74) we apply a damped Newton iteration to the equation $\Phi(\tilde{u}) = 0$ where

$$\begin{aligned} \Phi : H_0^1(\Omega) &\rightarrow H^{-1}(\Omega), \quad \Phi(\tilde{u}) = -\Delta\tilde{u} + \kappa\tilde{u}^3 - x_k^\delta, \\ \tilde{u}^{l+1} &= \tilde{u}^l - \left(-\Delta\tilde{u} + 3\kappa(\tilde{u}^l)^2\right)^{-1} \left(-\Delta\tilde{u} + \kappa(\tilde{u}^l)^3 - x_k^\delta\right), \end{aligned}$$

which is stopped as soon as $\|\Phi(\tilde{u}^l)\|_{H^{-1}(\Omega)}$ has been reduced by a factor of $1.e-4$. The sources x and states u are discretized by piecewise linear finite elements, hence after elimination of the state via the linear equality constraint, (77) becomes a box constrained quadratic program for the dicretized version of x , which we solve with the method from [12] using the Matlab code `mkr_box` provided to us by Philipp Hungerländer, Alpen-Adria Universität Klagenfurt. All implementations have been done in Matlab.

We performed test computations on a 2-d domain $\omega_o = \omega_c = \Omega = (-1, 1)^2$, on a regular computational finite element grid consisting of $2 \cdot N \cdot N$ triangles, with $N = 32$. We first of all consider $\kappa = 1$ (below we will also show results with $\kappa = 100$) and the piecewise constant exact source function

$$x_{ex} = -10 + 20 \cdot \mathbb{1}_B, \quad (81)$$

where B is the ball of radius 0.2 around $(-0.4, -0.3)$ cf. Fig. 1, and correspondingly set $\rho = 10$. In order to avoid an inverse crime, we generated the synthetic data on a

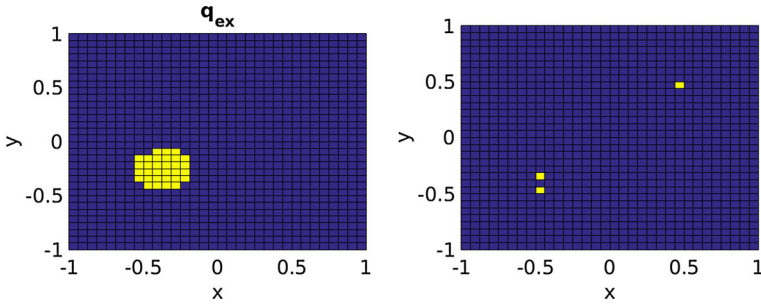


Fig. 1 Left: exact source x_{ex} ; right: locations of spots for testing weak $*$ L^∞ convergence

Table 1 Convergence as $\delta \rightarrow 0$: averaged errors of five test runs with uniform noise

δ	err_{spot_1}	err_{spot_2}	err_{spot_3}	$err_{L^1(\Omega)}$
0.1000	0	4.0818	8.0043	0.0627
0.0667	0.1558	3.6454	7.8451	0.0541
0.0333	0	3.0442	6.5726	0.0370
0.0100	0	0	3.9091	0.0188

finer grid and, after projection of u_{ex} onto the computational grid, we added normally distributed random noise of levels $\delta \in \{0.001, 0.01, 0.1\}$ to obtain synthetic data y^δ .

In all tests we start with the constant function with value zero for x_0 . Moreover, we always set $\tau = 1.1$. According to our convergence result Theorem 1 with $\mathcal{R} = \|\cdot\|_{L^\infty(\Omega)}$, we can expect weak $*$ convergence in $L^\infty(\Omega)$ here. Thus we computed the errors in certain spots within the two homogeneous regions and on their interface,

$$spot_1 = (0.5, 0.5), \quad spot_2 = (-0.4, -0.3), \quad spot_3 = (-0.4, -0.5),$$

cf. Fig. 1, more precisely, on $\frac{1}{N} \times \frac{1}{N}$ squares located at these spots, corresponding to the piecewise constant L^1 functions with these supports in order to exemplarily test weak $*$ L^∞ convergence. Additionally we computed L^1 errors.

Table 1 provides an illustration of convergence as δ decreases. For this purpose, we performed five runs on each noise level for each example and list the average errors.

In Fig. 2 we plot the reconstructions for $\kappa = 1$ and $\kappa = 100$. For $\kappa = 1$, the noise levels $\delta \in \{0.1, 0.667, 0.333, 0.01\}$ correspond to a percentage of $p \in \{5.6, 18.5, 37.1, 55.6\}$ of the L^2 deviation of the exact state from the background state $u_0 = -10^{1/3}$. In case of $\kappa = 100$, where the background state is $u_0 = -0.1^{1/3}$ the corresponding percentages are $p \in \{17.9, 59.7, 119.4, 179.2\}$. For an illustration of the noisy data as compared to the exact ones, see Figs. 3 and 4. Indeed, the box constraints enable to cope with relatively large noise levels, even in the rather nonlinear regime with $\kappa = 100$.

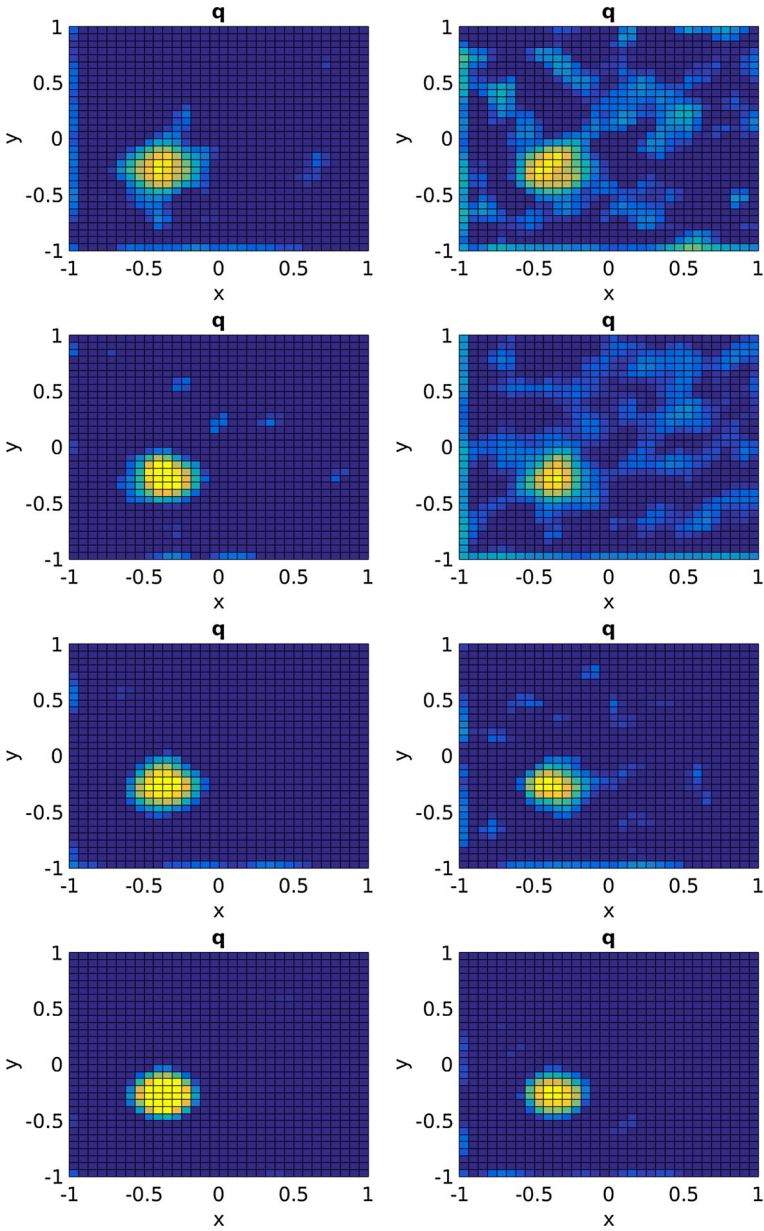


Fig. 2 Reconstructions from noisy data with $\delta \in \{0.1, 0.667, 0.333, 0.01\}$ (top to bottom) for $\kappa = 1$ (left) and $\kappa = 100$ (right)

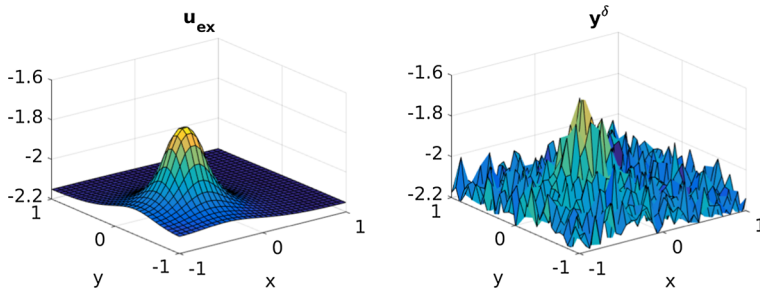


Fig. 3 Exact and noisy data ($\delta = 0.1$) for $\kappa = 1$

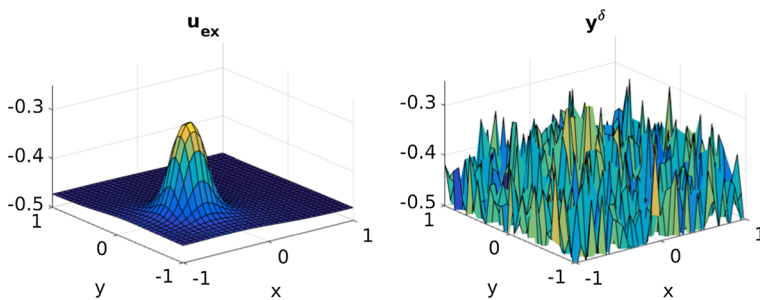


Fig. 4 Exact and noisy data ($\delta = 0.1$) for $\kappa = 100$

6 Conclusions and remarks

In this paper we have studied convergence of the Tikhonov type, the Ivanov type, and the Morozov type IRGNM with a stopping rule based on the discrepancy principle type. To the best of our knowledge, the Ivanov and Morozov IRGNMs have not been studied so far and in all three Tikhonov, Ivanov, and Morozov type IRGNMs, convergence results without source conditions so far use stronger assumptions than the tangential cone condition used here. We also consider discretized versions of the methods and provide discretization error bounds that still guarantee convergence. Moreover, we discuss goal oriented dual weighted residual error estimators that can be used in an adaptive discretization scheme for controlling these discretization error bounds. An inverse source problem for a nonlinear elliptic boundary value problems illustrates our theoretical findings in the special situations of measure valued and L^∞ sources. We also provide some computational results with the Ivanov IRGNM for the case of an L^∞ source. Numerical implementations and tests for a measure valued source, together with adaptive discretization is subject of ongoing work, based on the approaches from [4–6, 18, 19]. Future research in this context will be concerned with convergence rates results for the Ivanov and Morozov IRGNMs under source conditions.

Acknowledgements Open access funding provided by University of Klagenfurt. The authors wish to thank Philipp Hungerländer, Alpen-Adria Universität Klagenfurt, for providing us with the Matlab code based on the method from [12]. Moreover, the authors gratefully acknowledge financial support by the Austrian

Science Fund FWF under the grants I2271 “Regularization and Discretization of Inverse Problems for PDEs in Banach Spaces” and P30054 “Solving Inverse Problems without Forward Operators” as well as partial support by the Karl Popper Kolleg “Modeling-Simulation-Optimization”, funded by the Alpen-Adria-Universität Klagenfurt and by the Carinthian Economic Promotion Fund (KWF).

Moreover, we wish to thank both reviewers for fruitful comments leading to an improved version of the manuscript.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix

Lemma 1 For all $a, b > 0$, $p \geq 1$ and $\gamma, \epsilon \in (0, 1)$

$$(a + b)^p \leq (1 + \gamma)^{p-1} a^p + \left(\frac{1 + \gamma}{\gamma} \right)^{p-1} b^p \quad (82)$$

and, if additionally $a \geq b$, also

$$(a - b)^p \geq (1 - \epsilon)^{p-1} a^p - \left(\frac{1 - \epsilon}{\epsilon} \right)^{p-1} b^p. \quad (83)$$

Proof The estimate in (82) can be done by solving the following extremal value problems

$$C_\gamma = \max_{x>0} \phi(x), \quad C_\epsilon = \max_{x>0} \Phi(x),$$

where

$$\phi(x) := ((1 + x)^p - (1 + \gamma)^{p-1})x^{-p} \text{ and } \Phi(x) := ((1 - \epsilon)^{p-1} - (1 - x)^p)x^{-p},$$

since for any $\gamma, \epsilon \in (0, 1)$,

$$\phi(x) \leq C_\gamma \text{ and } \Phi(x) \leq C_\epsilon \text{ for all } x > 0$$

with $x := b/a$, $a, b > 0$ is equivalent to (24).

Solving for C_γ , we have

$$\phi'(x) = px^{-(p+1)}((1 + \gamma)^{p-1} - (1 + x)^{p-1}) \begin{cases} = 0 & \iff x = \gamma, \\ < 0 & \text{for } x > \gamma, \\ > 0 & \text{for } x < \gamma, \end{cases}$$

which means that

$$\max \phi(x) = \phi(\gamma) = \left(\frac{1 + \gamma}{\gamma} \right)^{p-1},$$

so defining $C_\gamma := \left(\frac{1+\gamma}{\gamma}\right)^{p-1}$ and writing the resulting inequality in terms of a and b we have the desired formula.

The formula in (83) is derived analogously. □

Lemma 2 *Let $k \in \mathbb{N}$, $(d_j)_{1 \leq j \leq k+1} \subseteq [0, \infty)$, $(\mathcal{R}_j)_{1 \leq j \leq k+1} \subseteq [0, \infty)$, $\alpha_0, \mathcal{R}^\dagger, c, q \neq \theta \in (0, \infty)$. Then*

$$\forall j \in \{0, \dots, k\} : d_{j+1} + \alpha_0 \theta^j \mathcal{R}_{j+1} \leq q d_j + \alpha_0 \theta^j \mathcal{R}^\dagger + c, \tag{84}$$

implies that

$$d_{k+1} + \alpha_0 \theta^k \mathcal{R}_{k+1} < q^{k+1} d_0 + \left(\frac{1}{1 - \frac{q}{\theta}}\right) \alpha_0 \theta^k \mathcal{R}^\dagger + \left(\frac{1}{1 - q}\right) c. \tag{85}$$

Proof We first show by induction that for all $l \in \{0, \dots, k\}$

$$d_{k+1} + \alpha_0 \theta^k \mathcal{R}_{k+1} \leq q^{l+1} d_{k-l} + \left(1 + \frac{q}{\theta} + \dots + \left(\frac{q}{\theta}\right)^l\right) \alpha_0 \theta^k \mathcal{R}^\dagger + (1 + q + \dots + q^l) c. \tag{86}$$

Indeed, for $l = 0$, (86) is just (84) with $j = k$. Suppose that (86) holds for l , then using (84) with $j = k - (l + 1)$, we obtain the formula for $l + 1$

$$\begin{aligned} & d_{k+1} + \alpha_k \mathcal{R}_{k+1} \\ & \leq q^{l+1} \left(q d_{k-(l+1)} + \alpha_0 \theta^{k-(l+1)} \mathcal{R}^\dagger + c \right) \\ & \quad + \left(1 + \frac{q}{\theta} + \dots + \left(\frac{q}{\theta}\right)^l \right) \alpha_0 \theta^k \mathcal{R}^\dagger + (1 + q + \dots + q^l) c \\ & = q^{(l+1)+1} d_{k-(l+1)} + \left(1 + \frac{q}{\theta} + \dots + \left(\frac{q}{\theta}\right)^l + \left(\frac{q}{\theta}\right)^{l+1} \right) \alpha_0 \theta^k \mathcal{R}^\dagger \\ & \quad + (1 + q + \dots + q^l + q^{l+1}) c, \end{aligned}$$

and the induction proof is complete.

Hence, setting $l = k$ in (86) and using the geometric series formula, we get the assertion (85).

References

1. Bakushinsky, A.B.: The problem of the convergence of the iteratively regularized Gauss–Newton method. *Comput. Math. Math. Phys.* **32**, 1353–1359 (1992)
2. Bakushinsky, A.B., Kokurin, M.: *Iterative Methods for Approximate Solution of Inverse Problems*. Kluwer, Dordrecht (2004)
3. Becker, R., Vexler, B.: Mesh refinement and numerical sensitivity analysis for parameter calibration of partial differential equations. *J. Comput. Phys.* **206**, 95–110 (2005)
4. Clason, C., Kunisch, K.: A duality-based approach to elliptic control problems in non-reflexive Banach spaces. *ESAIM Control Optim. Calc. Var.* **17**(1), 243–266 (2011)

5. Clason, C., Kunisch, K.: A measure space approach to optimal source placement. *Comput. Optim. Appl.* **53**(1), 155–171 (2012)
6. Casas, E., Clason, C., Kunisch, K.: Approximation of elliptic control problems in measure spaces with sparse solutions. *SIAM J. Control Optim.* **50**(4), 1735–1752 (2012)
7. Dombrowskaja, I., Ivanov, V.K.: On the theory of certain linear equations in abstract spaces. *Sib. Mat. Z.* **6**, 499–508 (1965)
8. Engl, H., Hanke, M., Neubauer, A.: *Regularization of Inverse Problems*. Kluwer, Dordrecht (1996)
9. Grasmair, M., Haltmeier, M., Scherzer, O.: The residual method for regularizing ill-posed problems. *Appl. Math. Comput.* **218**, 2693–710 (2011)
10. Hanke, M.: A regularization Levenberg–Marquardt scheme, with applications to inverse groundwater filtration problems. *Inverse Prob.* **13**, 79–95 (1997)
11. Hohage, T., Werner, F.: Iteratively regularized Newton-type methods for general data misfit functionals and applications to Poisson data. *Numer. Math.* **123**, 745–779 (2013)
12. Hungerländer, P., Rendl, F.: A feasible active set method for strictly convex problems with simple bounds. *SIAM J. Opt.* **25**, 1633–1659 (2015)
13. Ivanov, V.K.: On linear problems which are not well-posed. *Dokl. Akad. Nauk SSSR* **145**, 270–272 (1962)
14. Ivanov, V.K.: On ill-posed problems. *Mat. Sb. (N.S.)* **61**(103), 211–223 (1963)
15. Ivanov, V.K., Vasin, V.V., Tanana, V.P.: *Theory of Linear Ill-Posed Problems and Its Applications, Inverse and Ill-Posed Problems Series, VSP* (2002)
16. Jin, Q., Zhong, M.: On the iteratively regularized Gauss–Newton method in Banach spaces with applications to parameter identification problems. *Numer. Math.* **124**, 647–683 (2013)
17. Kaltenbacher, B., Hofmann, B.: Convergence rates for the iteratively regularized Gauss–Newton method in Banach spaces. *Inverse Prob.* **26**, 035007 (2010)
18. Kaltenbacher, B., Kirchner, A., Veljović, S.: Goal oriented adaptivity in the IRGNM for parameter identification in PDEs: I. Reduced formulation. *Inverse Prob.* **30**, 045001 (2014)
19. Kaltenbacher, B., Kirchner, A., Vexler, S.: Goal oriented adaptivity in the IRGNM for parameter identification in PDEs II: all-at once formulations. *Inverse Prob.* **30**, 045002 (2014)
20. Kaltenbacher, B., Neubauer, A., Scherzer, O.: *Iterative Regularization Methods for Nonlinear Ill-Posed Problems*. Walter de Gruyter, Berlin (2008)
21. Kaltenbacher, B., Schöpfer, F., Schuster, T.: Convergence of some iterative methods for the regularization of nonlinear ill-posed problems in Banach spaces. *Inverse Prob.* **25**, 065003 (2009)
22. Lorenz, D., Worliczek, N.: Necessary conditions for variational regularization schemes. *Inverse Prob.* **29**, 075016 (2013)
23. Morozov, V.A.: Choice of parameter for the solution of functional equations by the regularization method. *Dokl. Akad. Nauk SSSR* **175**, 1225–8 (1967)
24. Neubauer, A., Ramlau, R.: On convergence rates for quasi-solutions of ill-posed problems. *ETNA* **41**, 81–92 (2014)
25. Scherzer, O.: Convergence criteria of iterative methods based on Landweber iteration for solving nonlinear problems. *J. Math. Anal. Appl.* **194**, 911–933 (1995)
26. Seidman, T.I., Vogel, C.R.: Well posedness and convergence of some regularisation methods for nonlinear ill posed problems. *Inverse Prob.* **5**, 227–238 (1989)
27. Tikhonov, A.N., Arsenin, V.Y.: *Solutions of Ill-Posed Problems*. Wiley, New York (1977)
28. Werner, F.: On convergence rates for iteratively regularized Newton-type methods under a Lipschitz-type nonlinearity condition. *J. Inverse Ill Posed Probl.* **23**, 75–84 (2015)