QUASI-SOLUTION OF LINEAR INVERSE PROBLEMS IN NON-REFLEXIVE BANACH SPACES

Christian Clason[*]   Andrej Klassen[*]

June 26, 2018

Abstract   We consider the method of quasi-solutions (also referred to as Ivanov regularization) for the regularization of linear ill-posed problems in non-reflexive Banach spaces. Using the equivalence to a metric projection onto the image of the forward operator, it is possible to show regularization properties and to characterize parameter choice rules that lead to a convergent regularization method, which includes the Morozov discrepancy principle. Convergence rates in a suitably chosen Bregman distance can be obtained as well. We also address the numerical computation of quasi-solutions to inverse source problems for partial differential equations in $L^\infty(\Omega)$ using a semismooth Newton method and a backtracking line search for the parameter choice according to the discrepancy principle. Numerical examples illustrate the behavior of quasi-solutions in this setting.

1 INTRODUCTION

We consider the operator equation $Au = y$ for a linear injective operator $A : U \to Y$ with dense range, where Y is a uniformly convex Banach space and U the – not necessarily reflexive – dual of a separable normed vector space. Instead of the exact data $y \in R(A)$ (the range of A), usually only a noisy measurement $y^\delta \in Y \setminus R(A)$ with $\|y^\delta - y\|_Y \le \delta$ for some noise level δ is available, and regularization methods have to be applied to obtain a stable (approximate) solution. Three related strategies that can be found in the literature are:

- *Tikhonov regularization* [23], consisting in solving for some $\alpha > 0$ the minimization problem

$$\text{(T)} \qquad \min_{u \in U} \frac{1}{2}\|Au - y\|_Y^2 + \frac{\alpha}{2}\|u\|_U^2;$$

- *Morozov regularization* (also known as the *method of residuals*) [18], consisting in solving the constrained minimization problem

$$\text{(M)} \qquad \min_{u \in U} \|u\|_U \quad \text{subject to} \quad \|Au - y\|_Y \le \delta;$$

[*]University Duisburg-Essen, Faculty of Mathematics, Thea-Leymann-Str. 9, 45127 Essen, Germany (christian.clason@uni-due.de, andrej.klassen@uni-due.de)

- *Ivanov regularization* (also known as the *method quasi-solutions*) [12], consisting in solving for some $\rho > 0$ the constrained minimization problem

$$\text{(I)} \qquad \min_{u \in U} \|Au - y\|_Y \quad \text{subject to} \quad \|u\|_U \leq \rho.$$

It is the last one that is the focus of this work. Let us first briefly comment on the relation of these approaches. Since A is linear, all three minimization problems are convex, and hence their solutions coincide for a suitable choice of α and ρ, respectively; see, e.g., [15, Sec. 3.5]. However, this is no longer the case for nonlinear operators, as was shown by a counterexample in [17] (which also, to the best of our knowledge, coined the term *Ivanov regularization*); see also [16, Ex. 1]. Furthermore, if U is not a Hilbert space, (T) is no longer necessarily easier to solve than (M) or (I). In fact, for $U = L^\infty(\Omega)$ (which motivates the problem setting stated in the beginning), [2] proposed solving (T) by transforming it into the form (I) with an additional penalty $\frac{\alpha}{2}\rho^2$ for (the in this case free variable) ρ. This motivates directly considering (I) as a regularization strategy.

A further motivation is parameter identification in partial differential equations, where the domain of the (nonlinear) forward mapping is given by pointwise almost everywhere restrictions on the parameter. If such restrictions already yield a regularization of the ill-posed parameter identification problem, unwanted smoothing from the introduction of a Tikhonov penalty can be avoided.

In this work, we consider as a first step admissible sets of the form

$$\text{(1.1)} \qquad M_\rho := \{u \in U : \|u\|_U \leq \rho\},$$

where we are particularly interested in the case $U = L^\infty(\Omega)$ for a bounded domain $\Omega \subset \mathbb{R}^d$. We then take as the *quasi-solution* of $Au = y$ the minimizer $u_\rho \in M_\rho \subset U$ of

$$\text{(1.2)} \qquad \inf_{u \in M_\rho} \|Au - y\|_Y,$$

where the radius $\rho \geq 0$ plays the role of regularization parameter. Its proper choice is one of the main issues in showing that the method of quasi-solutions can be considered a (convergent) regularization method, and the characterization of such choice rules – especially of a posteriori-type – is one of the main contributions of this work. The main difficulty here is the fact that U is a non-reflexive Banach space, which requires working in the weak* topology. This is addressed by considering (1.2) as a projection problem in Y, allowing exploiting the strong topology of this space. We also discuss the numerical solution of (1.2) for the case that A is the solution operator to an elliptic partial differential equation and $U = L^\infty(\Omega)$.

Let us remark on related literature. Quasi-solutions were introduced by Ivanov [12, 13, 15] as the stabilization of linear ill-posed problems in normed spaces through a compact restriction of the solution domain based on Tikhonov's theorem [23]. Further works [4, 14] weakened the requirement to weakly compact norm balls in uniformly convex Banach spaces; see also [15, 26, 27]. More general compact solution domains of monotone or convex and bounded functions were considered in [28, 29]. More recently, convergence rates for quasi-solutions to nonlinear inverse problems with norm balls in Hilbert scales were proved in [20]. However, the (a posteriori)

parameter choice for the radius was not discussed. Finally, in parallel to this work, convergence and convergence rates for both Morozov and Ivanov regularization under variational source conditions were derived in a more abstract setting in [16]. In contrast to this, the current work also treats the numerical computation of quasi-solutions in a non-reflexive Banach space, which to the best of our knowledge has not been done so far.

This paper is organized as follows. Section 2 collects results on metric projections onto image sets under A that will be needed in the following. In Section 3, we show that the method of quasi-solutions is indeed a regularization strategy, characterize parameter choice rules leading to a convergent regularization method, and derive convergence rates in a suitably chosen Bregman distance. Section 4 is concerned with the numerical solution of an inverse source problem for a partial differential equation with $U = L^\infty(\Omega)$ based on a semi-smooth Newton method and a line search for the a posteriori choice according to the discrepancy principle. Finally, Section 5 contains numerical examples illustrating the behavior of quasi-solutions in this setting.

2 QUASI-SOLUTION AS METRIC PROJECTION IN IMAGE SPACE

To avoid having to work with the weak or weak* topology in U, we consider the problem of finding a quasi-solution as a metric projection onto the range of A. While existence of a quasi-solution can also be shown by standard methods, the analysis in image space will be helpful for analyzing parameter choice rules.

For this purpose, we make the following assumption on the forward operator.

Assumption 2.1. The linear operator $A : U \to Y$

(i) is injective,

(ii) has dense range,

(iii) is bounded in the strong topologies of U and Y,

(iv) is weak*-to-strong closed, i.e., for any sequence $\{u_n\}_{n \in \mathbb{N}}$ with $u_n \rightharpoonup^* u$ in U and $Au_n \to y$ in Y we have that $Au = y$.

From the definition of M_ρ and these assumptions on A, we immediately obtain the following useful property.

Lemma 2.2. *For every $\rho \geq 0$, the image set $Q_\rho := A(M_\rho) \subset Y$ is non-empty, closed, convex, and bounded.*

Proof. Non-emptiness, convexity, and boundedness are obvious since M_ρ is a ball and A is linear and bounded. To see the closedness, let $\{y_n\}_{n \in \mathbb{N}}$ be a sequence in Q_ρ converging strongly to some $y \in Y$. By definition of Q_ρ, there thus exists for every $n \in \mathbb{N}$ a $u_n \in M_\rho$ such that $Au_n = y_n$. Since $\{u_n\}_{n \in \mathbb{N}} \subset M_\rho$ is bounded and U is the dual of a separable vector space, we can apply the Banach–Alaoglu theorem to extract a subsequence, still denoted by $\{u_n\}_{n \in \mathbb{N}}$, with $u_n \rightharpoonup^* u \in M_\rho$ (since the unit ball in U is weakly* sequentially closed). By the closedness Assumption 2.1 (iv) on A, we thus have $y = Au \in Q_\rho$. □

3

Since Y was assumed to be a uniformly convex Banach space, we can apply the following result from [15, Rem. 1.6.2, Thm. 1.6.4].

Proposition 2.3. *For every $y \in Y$, there exists a unique metric projection $y_\rho \in Q_\rho$ such that*

$$(2.1) \qquad \|y_\rho - y\|_Y = \inf_{q \in Q_\rho} \|q - y\|_Y.$$

Furthermore, the mapping $P_{Q_\rho} : y \mapsto y_\rho$ is continuous.

By the definition of Q_ρ, this yields the existence of a $u_\rho \in M_\rho$ with $y_\rho = Au_\rho$ that attains the minimum in (1.2). Together with the injectivity of A, we thus immediately obtain the existence of a unique quasi-solution.

Theorem 2.4. *For every $y \in Y$ and $\rho \geq 0$ there exists a unique solution $u_\rho \in M_\rho$ to (1.2).*

We can therefore introduce the *distance function*

$$(2.2) \qquad d : [0, \infty) \times Y \to [0, \infty) \qquad (\rho, y) \mapsto \|Au_\rho - y\|_Y = \min_{u \in M_\rho} \|Au - y\|_Y.$$

The following property of the distance function will be crucial in the following. It was first proved in [4] for the case that M_ρ is compact, and our extension to weakly* compact and convex subsets follows the line of their proof. (Here we point out that convex and compact sets are in general *not* weakly* compact.)

Proposition 2.5. *For every $y \in Y$, the mapping $\rho \mapsto d(\rho, y)$ is surjective on $(0, \|y\|_Y)$.*

Proof. Let $y \in Y$ and $\sigma \in (0, \|y\|_Y)$ be given. Since the range of A is dense in Y, there exists a sufficiently large ρ such that $Q_\rho \cap B_\sigma(y)$ is non-empty. The set of all such $\rho > 0$ is thus non-empty and bounded from below and therefore admits an infimum

$$\tilde{\rho} := \inf\{\rho > 0 : Q_\rho \cap B_\sigma(y) \neq \varnothing\}.$$

By the definition of the infimum, there thus exists a sequence $\{\rho_n\}_{n \in \mathbb{N}} \subset (0, \infty)$ with

$$S_n := Q_{\rho_n} \cap B_\sigma(y) \neq \emptyset$$

and $\rho_n \to \tilde{\rho}$. Without loss of generality, we can assume that this sequence is monotonically decreasing, which implies that $Q_{\rho_{n+1}} \subseteq Q_{\rho_n}$ and hence that $S_{n+1} \subseteq S_n$.

We now show that the intersection $S := \bigcap_{n \in \mathbb{N}} S_n$ is non-empty. First, since every $S_n \subset Y$ is bounded, convex, and closed (as the intersection of such sets) as well as non-empty, we can apply Proposition 2.3 to obtain for each $n \in \mathbb{N}$ an $s_n := P_{S_n}(0)$. Furthermore, the monotonicity of $\{\rho_n\}_{n \in \mathbb{N}}$ implies that $\|s_n\|_Y \leq \rho_1$. Hence, $\{s_n\}_{n \in \mathbb{N}}$ is bounded and, since uniformly convex Banach spaces are reflexive, thus contains a subsequence weakly converging to some $s \in Y$. As S is the intersection of convex and closed sets and therefore weakly closed, we obtain that $s \in S$.

We next show that $s \in Q_{\tilde{\rho}}$. First, by construction we have that $s \in S_n \subset Q_{\rho_n}$ for every $n \in \mathbb{N}$, i.e., there exists a $u \in M_{\rho_n}$ such that $Au = s$. (Since A is injective, u cannot depend on n.) We can thus pass to the limit to obtain that

$$\|u\|_U \leq \rho_n \to \tilde{\rho}.$$

4

In fact, we even have that $\|u\|_U = \tilde{\rho}$, because if the inequality were strict, we could find a $\hat{\rho} < \tilde{\rho}$ such that $s \in Q_{\hat{\rho}} \cap B_\sigma(y)$, in contradiction to the definition of $\tilde{\rho}$. This shows the claim.

Similarly, since $s \in S_n$ for all $n \in \mathbb{N}$, we have that $s \in B_\sigma(y)$. Assume now that $\|s - y\|_Y < \sigma$, i.e., s is an interior point of $B_\sigma(y)$. Then there exists for $h := -s$ an $\varepsilon > 0$ small enough such that $\bar{s} := s + \varepsilon h = (1 - \varepsilon)s \in B_\sigma(y)$. Since A is linear, we have that $0 \in Q_{\tilde{\rho}}$ and hence by Lemma 2.2 also $\bar{s} = \varepsilon 0 + (1 - \varepsilon)s \in Q_{\tilde{\rho}}$. (Note that the assumption $0 < \sigma < \|y\|_Y$ implies that $0 \notin B_\sigma(y)$ and thus in particular $s \neq 0$.) This implies that there exists a $\bar{u} \in M_{\tilde{\rho}}$ such that $A\bar{u} = \bar{s}$. But from the linearity and injectivity of A it then follows that $\bar{u} = (1 - \varepsilon)u$ and hence that

$$\|\bar{u}\|_U = (1 - \varepsilon)\|u\|_U = (1 - \varepsilon)\tilde{\rho} < \tilde{\rho},$$

again in contradiction to the definition of $\tilde{\rho}$. This implies that $\|s - y\|_Y = \sigma$.

It remains to show that $\|s - y\|_Y = \inf_{u \in M_{\tilde{\rho}}} \|Au - y\|_Y$. For this, we argue as in the last step: If there exists a $\hat{u} \in M_{\tilde{\rho}}$ such that $\|A\hat{u} - y\|_Y < \sigma$, we can again find a $\bar{u} \in M_{\tilde{\rho}}$ such that $A\bar{u} \in B_\sigma(y)$ but $\|\bar{u}\|_U < \tilde{\rho}$ to contradict the definition of $\tilde{\rho}$. Hence,

$$d(\tilde{\rho}, y) = \|s - y\|_Y = \sigma$$

as claimed. □

The proof of Proposition 2.5 shows in particular that unless $d(\rho, y) = 0$ – i.e., $y \in Q_\rho \subset R(A)$ – the unique quasi-solution $u_\rho \in M_\rho$ from Theorem 2.4 always lies on the boundary of M_ρ, a result that was already shown in [20, Prop. 2.2] in the case that U and Y are Hilbert spaces.

Corollary 2.6. *For every $\rho \geq 0$ and $y \in Y \setminus Q_\rho$, the solution to (1.2) satisfies $\|u_\rho\|_U = \rho$.*

This result also implies a strict monotonicity of the distance function.

Lemma 2.7. *For every $y \in Y \setminus R(A)$, the mapping $\rho \mapsto d(\rho, y)$ is strictly monotonically decreasing and continuous.*

Proof. If $\rho_1 \leq \rho_2$, the definition of the distance function and the fact that $M_{\rho_1} \subset M_{\rho_2}$ immediately imply that
$$d(\rho_2, y) = \inf_{u \in M_{\rho_2}} \|Au - y\|_Y \leq \inf_{u \in M_{\rho_1}} \|Au - y\|_Y = d(\rho_1, y).$$
Assume now that $\rho_1 < \rho_2$ but $d(\rho_1, y) = d(\rho_2, y)$. By Corollary 2.6, we have that $\|u_{\rho_1}\|_U = \rho_1 < \rho_2 = \|u_{\rho_2}\|_U$ and in particular $u_{\rho_1} \neq u_{\rho_2}$. But by injectivity of A, we have that $Au_{\rho_1} \neq Au_{\rho_2}$ as well, and thus the uniform convexity of Y implies that $\bar{u} := \frac{1}{2}(u_{\rho_1} + u_{\rho_2}) \in M_{\rho_2}$ satisfies

$$\|A\bar{u} - y\| = \|\tfrac{1}{2}(Au_{\rho_1} + Au_{\rho_2}) - y\|_Y < \frac{1}{2}\left(\|Au_{\rho_1} - y\|_Y + \|Au_{\rho_2} - y\|_Y\right) = d(\rho_2, y),$$

in contradiction to the definition of the distance function.

Together with the surjectivity from Proposition 2.5, the strict monotonicity implies the continuity. □

Finally, we will also require the continuity of the distance function with respect to its second argument.

Lemma 2.8. *For each $\rho \geq 0$, the mapping $y \mapsto d(\rho, y)$ is non-expansive (i.e., Lipschitz continuous with constant 1).*

Proof. Consider $y_1, y_2 \in Y$ and corresponding projections $q_1 := P_{Q_\rho} y_1$ and $q_2 := P_{Q_\rho} y_2$. The optimality property of the projections then imply that

$$d(\rho, y_1) = \|q_1 - y_1\|_Y \leq \|q_2 - y_1\|_Y \leq \|q_2 - y_2\|_Y + \|y_2 - y_1\|_Y,$$
$$d(\rho, y_2) = \|q_2 - y_2\|_Y \leq \|q_1 - y_2\|_Y \leq \|q_1 - y_1\|_Y + \|y_1 - y_2\|_Y,$$

which together yield that

$$|d(\rho, y_1) - d(\rho, y_2)| \leq \|y_1 - y_2\|_Y. \qquad \square$$

3 REGULARIZATION AND PARAMETER CHOICE

In this section, we show that the method of quasi-solutions is a regularization in the classical sense [5, 22] and leads – combined with an appropriate parameter choice rule for ρ – to a convergent regularization method. Since in this case the solution domain M_ρ is not compact and U is not assumed to be reflexive, we can only expect weak* convergence.

3.1 STABILITY AND CONVERGENCE

To show that quasi-solutions are a regularization of the ill-posed operator equation $Au = y$, we first need to show that for each $\rho \geq 0$, the solution to (1.2) depends continuously on $y \in Y$. We again exploit the equivalence of (1.2) with the metric projection onto $Q_\rho := A(M_\rho)$.

Theorem 3.1. *Let $\rho \geq 0$ and $y \in Y$ be given. Then for every sequence $\{y_n\}_{n\in\mathbb{N}} \subset Y$ with $y_n \to y$, the corresponding sequence of solutions $\{u_n\}_{n\in\mathbb{N}}$ to (1.2) for y_n in place of y converges weakly* to u_ρ.*

Proof. First, Proposition 2.3 yields for every $n \in \mathbb{N}$ a unique metric projection $q_n \in Q_\rho$ and hence, by definition of Q_ρ, a $u_n \in M_\rho$ such that $Au_n = q_n$. Furthermore, by continuity of the metric projection, we have that $q_n \to q := P_{Q_\rho} y$. As in the proof of Lemma 2.2, we can now extract a subsequence, still denoted by $\{u_n\}_{n\in\mathbb{N}}$, such that $u_n \rightharpoonup^* u \in M_\rho$ with $Au = q$, i.e., $u \in M_\rho$ minimizes $\|Au - y\|_Y$ over M_ρ. This shows that $u = u_\rho$ is the solution to (1.2). Since this solution is unique by Theorem 2.4, a subsequence–subsequence argument yields weak* convergence of the full sequence. $\qquad \square$

In the case of Ivanov regularization, convergence as $\rho \to \infty$ for fixed $y \in R(A)$ is obvious since for every $\rho \geq \|u^\dagger\|_U$ with $Au^\dagger = y$, the choice $u_\rho = u^\dagger \in M_\rho$ clearly solves (1.2). Nevertheless, we state the consequence as a theorem for the sake of completeness.

Theorem 3.2. *Let $y \in R(A)$ be given and $u^\dagger \in U$ such that $Au^\dagger = y$. Then for every sequence $\{\rho_n\}_{n\in\mathbb{N}}$ with $\rho_n \to \infty$, we have $u_{\rho_n} \to u^\dagger$.*

3.2 PARAMETER CHOICE

We now address the convergence of Ivanov regularization in combination with a parameter choice strategies for ρ. As we will see, there are fundamental differences to Tikhonov (and Morozov) regularization, and hence the results of this section are one of the main contributions for this work.

We begin by characterizing parameter choices that lead to a convergent regularization method. In the following, let $y \in R(A)$ be arbitrary, $\{\delta_n\}_{n \in \mathbb{N}}$ be a non-negative sequence with $\delta_n \to 0$, and $y^{\delta_n} \in Y$ with $\|y^{\delta_n} - y\|_Y \leq \delta_n$ for each $n \in \mathbb{N}$. Let further $u^\dagger \in U$ denote the (unique) solution to $Au = y$ and for given $\rho_n \geq 0$, let $u_{\rho_n}^{\delta_n} \in U$ denote the solution to

$$(3.1) \qquad \min_{u \in M_{\rho_n}} \|Au - y^{\delta_n}\|_Y.$$

Proposition 3.3. *If A is not continuously invertible, the sequence $\{u_{\rho_n}^{\delta_n}\}_{n \in \mathbb{N}}$ weakly* converges to $u^\dagger$ as $n \to \infty$ if and only if the following two conditions hold:*

  (i) *$\{\rho_n\}_{n \in \mathbb{N}}$ is bounded,*

  (ii) *$\liminf\limits_{n \to \infty} \rho_n \geq \|u^\dagger\|_U$.*

Proof. We first show that these conditions are sufficient by proceeding similarly to the proof of Theorem 3.1. First, condition (i) implies that $\{u_{\rho_n}^{\delta_n}\}_{n \in \mathbb{N}}$ is bounded, and hence we can extract a subsequence $u_n \rightharpoonup^* u$ for some $u \in U$. By passing to a further subsequence, we can in addition assume that $\rho_n \to \hat{\rho} \geq \|u^\dagger\|_U$ by condition (ii). This implies that

$$(3.2) \qquad \|P_{Q_{\rho_n}} y - y\|_Y = d(\rho_n, y) \to d(\hat{\rho}, y) = 0,$$

where we have used the continuity of the distance function from Lemma 2.8 and the fact that $Au^\dagger = y \in Q_{\hat{\rho}}$. Furthermore, for every such n, Proposition 2.3 yields the existence of a unique metric projection $P_{Q_{\rho_n}} y^{\delta_n} = Au_{\rho_n}^{\delta_n}$ which by definition satisfies

$$\|P_{Q_{\rho_n}} y^{\delta_n} - y^{\delta_n}\|_Y = \inf_{q \in Q_{\rho_n}} \|q - y^{\delta_n}\|_Y \leq \|P_{Q_{\rho_n}} y - y^{\delta_n}\|_Y.$$

Combining the above, we obtain that

$$\begin{aligned}
\|P_{Q_{\rho_n}} y^{\delta_n} - y\|_Y &\leq \|P_{Q_{\rho_n}} y^{\delta_n} - y^{\delta_n}\|_Y + \|y^{\delta_n} - y\|_Y \\
&\leq \|P_{Q_{\rho_n}} y - y^{\delta_n}\|_Y + \delta_n \\
&\leq \|P_{Q_{\rho_n}} y - y\|_Y + 2\delta_n.
\end{aligned}$$

Passing to the limit then yields that $P_{Q_{\rho_n}} y^{\delta_n} \to y$ and hence by the closedness that $Au = y$, i.e., $u = u^\dagger$. Convergence of the full sequence now follows again from a subsequence–subsequence argument.

To show the necessity, first note that by assumption $R(A)$ is not closed and hence has no interior points (otherwise the algebraic interior of $R(A)$ would not be empty either, and the linearity of A would lead to the contradiction $R(A) = Y$.) For any $n \in \mathbb{N}$, we can thus find a

7

$y^{\delta_n} \in B_{\delta_n}(y) \setminus R(A)$. Corollary 2.6 then implies that the corresponding quasi-solution satisfies $\|u_{\rho_n}^{\delta_n}\|_U = \rho_n$. Hence, if $\{\rho_n\}_{n \in \mathbb{N}}$ is unbounded, $\{u_{\rho_n}^{\delta_n}\}_{n \in \mathbb{N}}$ is unbounded as well and therefore cannot weakly* converge (since weakly* convergent sequences are bounded by the Banach–Steinhaus Theorem). On the other hand, if $u_{\rho_n}^{\delta_n} \rightharpoonup^* u^\dagger$, then the weak* lower semi-continuity of the norm immediately implies that

$$\|u^\dagger\|_U \le \liminf_{n \to \infty} \|u_{\rho_n}^{\delta_n}\|_U = \liminf_{n \to \infty} \rho_n. \qquad \square$$

Proposition 3.3 shows in particular that Ivanov regularization converges under the constant parameter choice $\rho_n \equiv \|u^\dagger\|_U$. However, this is not a valid parameter choice in the strict sense, which may only depend on δ and y^δ, but not on the exact solution $u^\dagger$. In fact, Proposition 3.3 implies that there can be no convergent a priori parameter choice $\rho = \rho(\delta)$ unless the inverse problem is well-posed.

Corollary 3.4. *Let $\rho(\delta)$ be an a priori parameter choice. If $u_{\rho_n}^{\delta_n} \rightharpoonup^* u^\dagger$ for $\rho_n = \rho(\delta_n)$ with $\delta_n \to 0$ as $n \to \infty$, then A is continuously invertible.*

Proof. Since the claim must hold independently of $u^\dagger$, condition (ii) of Proposition 3.3 has to be satisfied for any $u^\dagger \in U$. This implies a fortiori that $\{\rho_n\}_{n \in \mathbb{N}}$ is unbounded, in contradiction to condition (i). Hence, A must be continuously invertible. $\qquad \square$

Of course, under the additional a priori information $u^\dagger \in M_{\rho^\dagger}$ for some $\rho^\dagger \ge 0$, the choice $\rho_n \equiv \rho^\dagger$ leads to a convergent regularization method (cf. [20]).

This leaves a posteriori choice rules $\rho = \rho(\delta, y^\delta)$. Here we consider the Morozov discrepancy principle [19], which for some $\tau > 1$ chooses $\rho(\delta, y^\delta)$ such that

$$(3.3) \qquad \delta \le \|Au_{\rho(\delta, y^\delta)}^\delta - y^\delta\|_Y \le \tau\delta.$$

This choice indeed leads to a convergent method.

Theorem 3.5. *Let $\rho_n := \rho(\delta_n, y^{\delta_n})$ be chosen according to (3.3) for some $\tau > 1$. Then $u_{\rho_n}^{\delta_n} \rightharpoonup^* u^\dagger$ as $n \to \infty$.*

Proof. We make use of the properties of the distance function (2.2) from Section 2. First note that for any $0 < \delta_n < \|y^{\delta_n}\|_Y$ (which holds for n sufficiently large provided $\|y\|_Y > 0$), Proposition 2.5 yields the existence of a ρ_n satisfying (3.3). It remains to verify the conditions of Proposition 3.3.

We first show that $\{\rho_n\}_{n \in \mathbb{N}}$ is bounded by $\|u^\dagger\|_U$. Assume that there exists an $n \in \mathbb{N}$ such that $\rho_n > \|u^\dagger\|_U$, which implies that $u^\dagger \in M_{\rho_n}$. Since $\|u_{\rho_n}^{\delta_n}\|_U = \rho_n$ for $\delta_n > 0$ by Corollary 2.6, we further have that $u_{\rho_n}^{\delta_n} \ne u^\dagger$. Hence it follows from the definition and uniqueness of the quasi-solution that

$$\|Au_{\rho(\delta, y^\delta)}^\delta - y^\delta\|_Y = \min_{u \in M_{\rho_n}} \|Au - y^{\delta_n}\|_Y < \|Au^\dagger - y^{\delta_n}\|_Y \le \delta_n,$$

in contradiction to the choice of ρ_n.

Assume now that

$$\liminf_{n \to \infty} \rho_n < \|u^\dagger\|_U =: \rho^\dagger.$$

8

Since the inequality is strict, there exists – after passing to a subsequence – for any $\varepsilon > 0$ an $N \in \mathbb{N}$ such that $\rho_n < \tilde{\rho} := \|u^\dagger\|_U - \varepsilon$ for all $n > N$. From Lemma 2.7 we thus obtain that $d(\rho_n, y^{\delta_n}) > d(\tilde{\rho}, y^{\delta_n})$ for all $n > N$. The parameter choice rule (3.3) and Lemma 2.8 then imply that

$$\lim_{n\to\infty} \tau\delta_n \geq \lim_{n\to\infty} d(\rho_n, y^{\delta_n}) \geq \lim_{n\to\infty} d(\tilde{\rho}, y^{\delta_n}) = d(\tilde{\rho}, y) > 0$$

since $u^\dagger \notin M_{\tilde{\rho}}$, in contradiction to $\delta_n \to 0$. □

The proof of Theorem 3.5 in fact shows that for $\rho_n := \rho(\delta_n, y^{\delta_n})$ chosen according to the discrepancy, we have $\rho_n \to \|u^\dagger\|_U$ from below.

3.3 CONVERGENCE RATES

Since we have only shown weak* convergence of u_ρ^δ to $u^\dagger$ as $\delta \to 0$, we cannot expect convergence rates for the strong error $\|u_\rho^\delta - u^\dagger\|_U$. As usual for inverse problems in Banach spaces, we thus consider rates for the error measured by the Bregman distance; see, e.g., [1, 6, 7, 10, 21]. We recall that for a convex functional $J : U \to \mathbb{R} \cup \{\infty\}$, the convex subdifferential of J at $u \in U$ with $J(u) < \infty$ is given by

$$(3.4) \qquad \partial J(u) = \{\xi \in U^* : \langle \xi, v - u \rangle_U \leq J(v) - J(u) \quad \text{for all } v \in U\}.$$

Note that the subdifferential can be empty unless $J(u) < \infty$ for all $u \in U$. For given $u \in U$ and $\xi \in \partial J(u)$, we can now define the Bregman distance

$$(3.5) \qquad D_J^\xi(v; u) = J(v) - J(u) - \langle \xi, v - u \rangle_U \qquad \text{for all } v \in U.$$

By the definitions, we have $D_J^\xi(v; u) \geq 0$ for any $v \in U$ as well $D_J^\xi(u; u) = 0$, which justifies using the Bregman distance to measure the error. Here, we choose $J(u) := \|u\|_U$ and point out that this choice is different from the classical one as suggested in [1]; to fit into their framework, we would have to take $J(u) = \delta_{M_\rho}(u)$ with $\delta_{M_\rho}(u) = 0$ if $u \in M_\rho$ and ∞ else, which however would make the Bregman distance uninformative for $v \notin M_\rho$.

Together with a classical source condition, we can then derive the expected convergence rate under the a posteriori choice rule (3.3).

Proposition 3.6. *Assume there exists a $w \in Y^*$ with $\xi := A^*w \in \partial J(u^\dagger)$. If $\rho = \rho(\delta, y^\delta)$ is chosen according to (3.3), then*

$$(3.6) \qquad D_J^\xi(u_\rho^\delta; u^\dagger) \leq C\delta.$$

Proof. First note that the proof of Theorem 3.5 together with Corollary 2.6 implies that

$$\|u_\rho\|_U = \rho \leq \|u^\dagger\|_U.$$

9

Hence, we can use the source condition together with the parameter choice rule to estimate

$$
\begin{aligned}
D_J^\xi(u_\rho; u^\dagger) &= \|u_\rho\|_U - \|u^\dagger\|_U - \langle A^* w, u_\rho - u^\dagger \rangle_U \\
&\le \left| \langle A^* w, u_\rho - u^\dagger \rangle_U \right| = \left| \langle w, A(u_\rho - u^\dagger) \rangle_Y \right| \\
&\le \|w\|_{Y^*} \|Au_\rho - Au^\dagger\|_Y \\
&\le \|w\|_{Y^*} \left(\|Au_\rho - y^\delta\|_Y + \|y^\delta - y\|_Y \right) \\
&\le \|w\|_{Y^*} (\tau + 1)\delta. \qquad \square
\end{aligned}
$$

Remark 3.7. We end this section by remarking on the interpretation of the convergence in Bregman distance in our context. Recall that the subgradient ξ in (3.6) satisfies by definition.

$$
(3.7) \qquad \xi \in \partial \| \cdot \|_{L^\infty(\Omega)}(u^\dagger) = \left\{ \xi \in L^\infty(\Omega)^* : \langle \xi, u^\dagger \rangle_{L^\infty(\Omega)} = \|u^\dagger\|_{L^\infty(\Omega)}, \|\xi\|_{L^\infty(\Omega)^*} = 1 \right\}.
$$

Assume now that the set $C := \{x \in \Omega : |u(x)| = \rho\}$ is measurable. In this case, it is straightforward to verify that one possible choice is $\xi \in L^1(\Omega) \subset (L^\infty)^*$ defined pointwise almost everywhere by

$$
(3.8) \qquad \xi(x) := \begin{cases} |C|^{-1} & \text{if } \left\{ x \in \Omega : u^\dagger(x) = \rho \right\} \\ 0 & \text{if } \left\{ x \in \Omega : |u^\dagger(x)| < \rho \right\}, \\ -|C|^{-1} & \text{if } \left\{ x \in \Omega : u^\dagger(x) = -\rho \right\} \end{cases}
$$

where $|C|$ denotes the Lebesgue measure of C. In this case, convergence in Bregman distance entails pointwise convergence $u_\rho^\delta(x) \to u^\dagger(x)$ on the active set where $|u^\dagger(x)| = \rho$. Of course, this choice does not satisfy the source condition $\xi \in R(A^*) \subset H^1(\Omega)$, so Proposition 3.6 does not imply a pointwise convergence on this set with rate $O(\delta)$.

4 NUMERICAL SOLUTION

We now address the numerical computation of the quasi-solution problem (1.2) for the case $U = L^\infty(\Omega)$ and $Y = L^2(\Omega)$ for a bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$, $d \in \{1, 2, 3\}$, using a combination of a semi-smooth Newton method and a backtracking procedure for the a posteriori parameter choice of ρ. As a model problem, we consider the identification of the source term $u \in L^\infty(\Omega)$ from observation of the state $y \in H^1(\Omega)$ given as the weak solution to

$$
(4.1) \qquad \begin{cases} -\Delta y + cy = u & \text{in } \Omega, \\ \partial_\nu y = 0 & \text{on } \partial\Omega, \end{cases}
$$

for given $c > 0$. Setting now

$$
A : L^\infty(\Omega) \to L^2(\Omega), \qquad u \mapsto y,
$$

it is easy to see that A is linear, injective, continuous, and weakly* closed. It is well-known from PDE-constrained optimization that in this case u_ρ is a solution to (1.2) (which is known in that context as a *bang-bang control problem*, see, e.g., [3]) if and only if

$$
(4.2) \qquad u_\rho = \operatorname{proj}_{M_\rho}(u_\rho - p_\rho),
$$

where

$$\text{proj}_{M_\rho} : L^2(\Omega) \to L^2(\Omega), \qquad [\text{proj}_{M_\rho}(v)](x) =: \begin{cases} \rho & \text{if} \quad v(x) > \rho \\ v(x) & \text{if} \quad |v(x)| \leq \rho \\ -\rho & \text{if} \quad v(x) < -\rho \end{cases}$$

and p_ρ is the solution to the *adjoint problem*

$$(4.3) \qquad \begin{cases} -\Delta p + cp = f & \text{in } \Omega, \\ \partial_\nu p = 0 & \text{on } \partial\Omega, \end{cases}$$

for $f = Au_\rho - y^\delta$; see, e.g., [24, Lem. 2.26]. (Here we rely on the fact that the adjoint state $p_\rho \in H^1(\Omega)$ to allow a pointwise almost everywhere evaluation.)

We now wish to solve (4.2) using a locally superlinearly convergent semi-smooth Newton method [8, 25]. However, it is known that the pointwise projection is semi-smooth from $L^p(\Omega) \to L^q(\Omega)$ if and only if $p > q$, and hence (4.2) is not semi-smooth with respect to u_ρ. Usually, this is addressed by including additional (small) Tikhonov regularization in (1.2), which allows canceling u_ρ inside the projection; see, e.g., [8, 9] as well as [24, Thm. 2.28, Ch. 2.12.4]. Since the focus of this work is Ivanov regularization, we do not follow this route here and instead consider a finite element discretization of (1.2).

Let $Y_h \subset L^2(\Omega)$ be a finite-dimensional space spanned by the usual continuous piecewise linear nodal basis ("hat") functions based on the vertices $\{x_j\}_{j=1}^N$ of a triangulation of $\overline{\Omega}$. Taking now $y_h, u_h, p_h \in Y_h$ (which we identify with the vector of their basis coefficients, which coincide with the values at the nodes in this case), noting that functions in Y_h attain their maximum and minimum at the nodes, and introducing the stiffness matrix $K_h \in \mathbb{R}^{N \times N}$ and the mass matrix $M_h \in \mathbb{R}^{N \times N}$, we obtain the discretized conditions

$$(4.4) \qquad \begin{cases} (K_h + cM_h)y_h = M_h u_h, \\ (K_h + cM_h)p_h = M_h(y_h - y_h^\delta), \\ \qquad\quad u_h = \text{proj}_{[-\rho,\rho]}(u_h - p_h), \end{cases}$$

where $y_h^\delta \in Y_h$ denotes the L^2 projection of y^δ and the projection is now understood as componentwise. This is now semi-smooth from $\mathbb{R}^N$ to $\mathbb{R}^N$, where a Newton derivative at $v \in \mathbb{R}^N$ direction $w \in \mathbb{R}^N$ is given componentwise by

$$(4.5) \qquad [D_N \text{proj}_{[-\rho,\rho]}(v)w]_i = \begin{cases} w_i & \text{if } |w_i| \leq \rho, \\ 0 & \text{else.} \end{cases}$$

Introducing the active and inactive sets as

$$(4.6) \qquad \begin{cases} \mathcal{A}_+ := \{i : u_i - p_i > \rho\} \\ \mathcal{A}_- := \{i : u_i - p_i < -\rho\} \\ \mathcal{I} := \{i : |u_i - p_i| \leq \rho\} \end{cases}$$

and the corresponding characteristic functions via $[\mathbb{1}_{\mathcal{A}}]_i = 1$ if $i \in \mathcal{A}$ and 0 else, we can write a Newton step as

$$
(4.7) \quad B_k := \begin{pmatrix} K_h + cM_h & O & -M_h \\ -M_h & K_h + cM_h & O \\ O & \mathbb{1}_{\mathcal{A}_+^k \cup \mathcal{A}_-^k} & \mathbb{1}_{\mathcal{I}^k} - \mathbb{1}_{\mathcal{A}_+^k \cup \mathcal{A}_-^k} \end{pmatrix} \begin{pmatrix} \delta y \\ \delta p \\ \delta u \end{pmatrix}
$$

$$
= - \begin{pmatrix} (K_h + cM_h)y_h^k - M_h u_h^k \\ (K_h + cM_h)p_h^k - M_h(y_h^k - y_h^\delta) \\ u_h - \operatorname{proj}_{[-\rho,\rho]}(u_h^k - p_h^k) \end{pmatrix} =: -b^k
$$

and setting $y_h^{k+1} = y_h^k + \delta y$, $p_h^{k+1} = p_h^k + \delta p$, and $u_h^{k+1} = u_h^k + \delta u$. The iteration is terminated if either a fixed number $k_{\max}$ of iterations is exceeded or if both $\mathcal{A}_+^k = \mathcal{A}_+^{k-1}$ and $\mathcal{A}_-^k = \mathcal{A}_-^{k-1}$. In the latter case, (p_h^k, y_h^k, u_h^k) is already a solution to (4.4); see [11, Rem. 7.1.1]. To improve the robustness of the procedure, we also include dampening, which requires an additional criterion that $\|b^k\|_2 \leq \text{tol}$ for some small tol > 0 for successful termination. The full procedure is given in Algorithm 1.

Algorithm 1 Damped semi-smooth Newton method

Input: y_h^δ, $\rho > 0$, $q \in (0,1)$, $i_{\max}$, $k_{\max}$, tol > 0, y_h^0, p_h^0, u_h^0

1: **for** $k = 1, \ldots, k_{\max}$ **do**
2:      compute active sets $\mathcal{A}_+^k, \mathcal{A}_-^k, \mathcal{I}^k$ from (4.6)
3:      compute Newton matrix B_k, right-hand side b^k from (4.7)
4:      **if** $\|b_k\|_2 < \text{tol}$ **then**
5:         set conv = true; **return**
6:      solve Newton step (4.7) for $(\delta y, \delta p, \delta u)$
7:      **for** $i = 1, \ldots, i_{\max}$ **do**
8:         set $y_h^i + q^{i-1}\delta y$, $p_h^i + q^{i-1}\delta p$, $u_h^i + q^{i-1}\delta u$
9:         compute b^i from (4.7)
10:        **if** $\|b^i\|_h < \|b^k\|_h$ **then**
11:          set $y_h^{k+1} = y_h^i$, $p_h^{k+1} = p_h^i$, $u_h^{k+1} = u_h^i$; **break**
12: set conv = false

Output: y_h^k, p_h^k, u_h^k, conv

In practice, we have to deal with the fact that Newton methods converge only locally. We therefore use a backtracking-type method for computing a regularization parameter ρ satisfying the discrepancy principle (3.3) as a continuation strategy. Our approach is based on the following two observations: First, since we are dealing with a discretized problem, the operator $A_h := (K_h + cM_h)^{-1}$ is invertible and thus any y_h^δ is in the range of A_h. There thus exists a $\hat{\rho}$ such that the constraint in the discrete version of (1.2) is inactive and the semi-smooth Newton method will therefore converge in a single step. Second, the solution for a given ρ_k will be a good starting point for $\rho_{k+1} < \rho_k$ provided the difference is not too large. We thus proceed in three phases:

I. Starting from some $\rho_0 \geq 0$, we quickly increase ρ_k until the Newton method converged in a fixed number of iterations and the residual is smaller than the noise level.

II. We then quickly decrease ρ_k until the Newton method fails to converge.

III. Starting from the last converged iteration, we adaptively decrease ρ_k further to ensure convergence of the Newton method, until the discrepancy principle (3.3) is satisfied.

The full procedure is given in Algorithm 2. Of course, we also terminate the iteration if the discrepancy principle is already satisfied in phase I or II. Furthermore, to reduce the number of computed quasi-solutions, the increase of ρ_k in phase I can be made adaptive as well.

Algorithm 2 Parameter choice

Input: $y^\delta, \delta, \tau > 1, \rho_0 \geq 0$
1: set $y_h^k = 0, p_h^k = 0, u_h^k = 0$
2: **for** $k = 0, \ldots$ **do**   ▷ Phase I
3:     compute y_h^k, p_h^k, u_h^k, conv using Algorithm 1 for $\rho_k, y_h^{k-1}, p_h^{k-1}, u_h^{k-1}$
4:     compute $d_k = \|y_k - y^\delta\|_2$
5:     **if** $d_k < \delta$ and conv = true **then**
6:        **break**
7:     set $\rho_{k+1} = \rho_k + \rho_0$
8: set $\rho_0 = \rho_k/2, y_h^0 = y_h^k, p_h^0 = p_h^k, u_h^0 = u_h^k$
9: **for** $k = 0, \ldots$ **do**   ▷ Phase II
10:     compute y_h^k, p_h^k, u_h^k, conv using Algorithm 1 for $\rho_k, y_h^{k-1}, p_h^{k-1}, u_h^{k-1}$
11:     compute $d_k = \|y_k - y^\delta\|_2$
12:     **if** $d_k > \delta$ or conv = false **then**
13:        **break**
14:     set $\rho_{k+1} = \rho_k/2$
15: set $\Delta\rho = \rho_{k-1}/2, \rho_0 = \rho_{k-1} - \Delta\rho, y_h^0 = y_h^k, p_h^0 = p_h^k, u_h^0 = u_h^k$
16: **for** $k = 0, \ldots$ **do**   ▷ Phase III
17:     compute y_h^k, p_h^k, u_h^k, conv using Algorithm 1 for $\rho_k, y_h^{k-1}, p_h^{k-1}, u_h^{k-1}$
18:     compute $d_k = \|y_k - y^\delta\|_Y$
19:     **if** $\delta \leq d_k \leq \tau\delta$ and conv = true **then**
20:        **break**
21:     **else if** $d_k > \tau\delta$ or conv = false **then**
22:        set $\Delta\rho \leftarrow \Delta\rho/2$
23:        set $\rho_{k+1} = \rho_k + \Delta\rho$
24:     **else**
25:        set $\rho_{k+1} = \rho_k - \Delta\rho$
Output: $y_h^k, p_h^k, u_h^k, \rho_k$

---

5 NUMERICAL EXAMPLE

We illustrate the regularization properties of the quasi-solution using a numerical example for the setting considered in Section 4, which is a moderately ill-posed problem. Specifically, we take $\Omega = [-1, 1]^2 \subset \mathbb{R}^2$ and discretize it using a uniform (Friedrichs–Keller) triangulation with $N = 128 \times 128$ vertices. The potential coefficient is fixed at $c = 1$. We then choose a piecewise constant true parameter $u^\dagger$ consisting of inclusions of different (positive and negative) heights with $\rho^\dagger := \|u^\dagger\|_{L^\infty(\Omega)} = 4$; see Figure 1a. (For the sake of conciseness, we omit the subscript h in the following.) From this, noisy data is generated by setting

$$y^\delta = y^\dagger + \frac{s}{100}\|y^\dagger\|_\infty \frac{\eta}{\|\eta\|_2},$$

where $y^\dagger = A_h u^\dagger$, $s > 0$ is a chosen relative noise percentage, $\eta \in \mathbb{R}^N$ is a random vector with $\eta_i, 1 \le i \le N$, independently normally distributed with mean 0 and standard deviation 1, and $\|\eta\|_2 := \eta^T M_h \eta$ is the discrete $L^2(\Omega)$-norm.

We then compute the quasi-solution u_ρ^δ using a Matlab implementation of the procedure discussed in Section 4, where the parameters are chosen as follows. In Algorithm 1, we set $q = 0.7$, $i_{\max} = 10$, $k_{\max} = 30$ and $\text{tol} = 10^{-9}$. In Algorithm 2, we set $\tau = 1.1$ and $\rho_0 = 10$.

Figures 1b to 1f show the quasi-solutions for $s = 1, 0.02, 0.01, 0.001, 0.0001$ with corresponding $\rho = \rho(\delta, y^\delta) \approx 2.5, 3.896, 3.952, 3.995, 3.999$ according to the discrepancy principle. First, we note that in all cases $\|u_\rho^\delta\|_{L^\infty(\Omega)} = \rho$ as expected from Corollary 2.6. For $s = 1$, we have in fact that $u_\rho^\delta(x) = \rho$ almost everywhere, which leads to a poor reconstruction quality. However, as s and therefore $\delta := \|y^\delta - y^\dagger\|_2$ decreases, the quality increases until the quasi-solution is virtually identical to $u^\dagger$ for $s = 0.0001$. While the reconstruction only becomes really acceptable for $s < 0.01$, two properties are of note: First, ignoring the noisy background, the location and shape of the inclusions is recovered well even for large s; in particular, no smoothing is visible in any of the reconstructions. Second, this is especially the case for the inclusions of largest magnitude for which $u_\rho^\dagger(x) = \rho^\dagger$, for which location and shape are reconstructed perfectly (as can be seen – with a bit of effort – even for $s = 1$ in Figure 1b). This ties in with Remark 3.7, which indicates improved convergence behavior at these points.
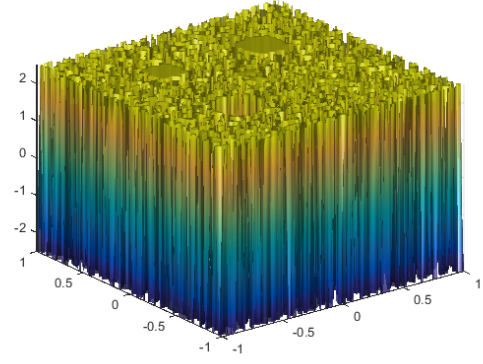
This is further illustrated by quantitative results for a larger selection of values of the noise percentage s in Table 1, where we give the actual noise level δ, the discrepancy $d(\rho, y^\delta)$ achieved by the parameter choice Algorithm 2, and the corresponding reconstruction error of the quasi-solution. For the sake of completeness, we give here both the $L^2(\Omega)$-error $\|u_\rho^\delta - u^\dagger\|_2$ and the $L^\infty(\Omega)$-error $\|u_\rho^\delta - u^\dagger\|_\infty$ as well as the error of the duality pairing $\langle \xi, u_\rho^\delta - u^\dagger \rangle$ in the discrete Bregman distance $D_J^\xi(u_\rho^\delta; u^\dagger)$ for $\xi \in \mathbb{R}^N$ constructed according to Remark 3.7, i.e.,

$$(5.1) \qquad \xi_i = \begin{cases} d & \text{if } u_i^\dagger = \rho^\dagger, \\ 0 & \text{if } |u_i^\dagger| < \rho^\dagger, \\ -d & \text{if } u_i^\dagger = -\rho^\dagger, \end{cases} \qquad \text{for} \qquad d = \left| \left\{ i : |u_i^\dagger| = \rho^\dagger \right\} \right|.$$
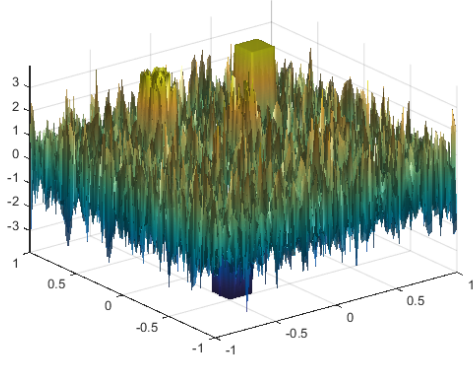
(Note that in the discrete setting, A_h is invertible and hence $\xi \in R(A_h^*)$.) It can be observed that at least for $s \le 0.01$, all these errors behave as $O(\delta)$, corroborating both Proposition 3.6 and the observed pointwise convergence.
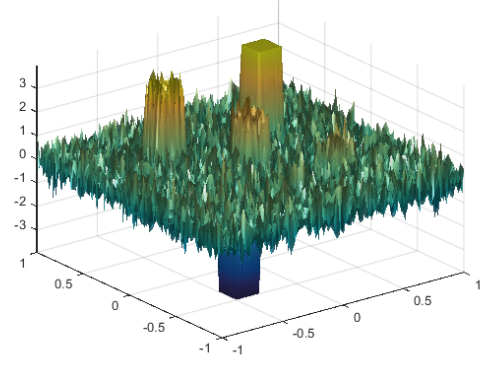
14

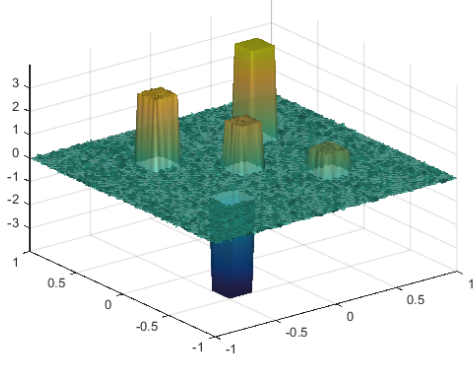(a) $u^\dagger$ with $\|u^\dagger\|_{L^\infty(\Omega)} = 4$
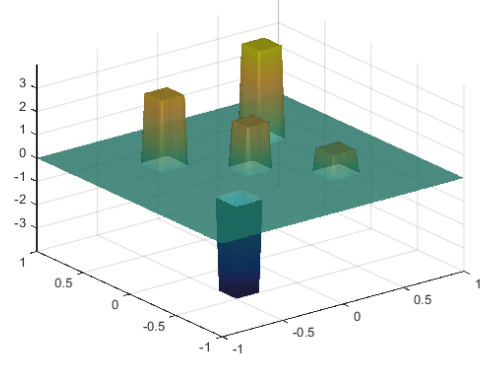
(b) u_ρ^δ for $s = 1\%$, $\rho \approx 2.5$
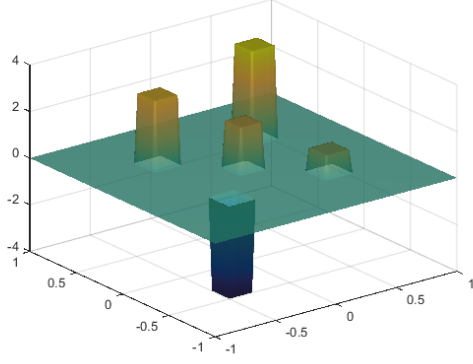
(c) u_ρ^δ for $s = 0.02$, $\rho \approx 3.896$

(d) u_ρ^δ for $s = 0.01$, $\rho \approx 3.952$

(e) u_ρ^δ for $s = 0.001$, $\rho \approx 3.995$

(f) u_ρ^δ for $s = 0.0001$, $\rho \approx 3.999$

Figure 1: true parameter $u^\dagger$ and quasi-solutions u_ρ^δ for different noise percentages

6 CONCLUSION

We have extended the method of quasi-solutions, or Ivanov regularization, to non-reflexive Banach spaces and shown weak* stability and convergence as well as convergence rates in Bregman distances. In particular, we have characterized parameter choice rules that lead to a convergent regularization method. While it turns out that in this setting, a true a priori choice is not feasible, the classical a posteriori choice according to the discrepancy principle is possible and can be exploited as a continuation strategy for the efficient numerical computation of quasi-solutions to inverse source problems using a Newton-type method. Numerical examples illustrate the regularization properties of this approach.

The results in Section 5 show that the method of quasi-solutions is indeed a very weak regularization. On the one hand, this is beneficial since it doesn't introduce additional smoothing as Tikhonov regularization with an $L^2(\Omega)$ penalty would. On the other hand, the reconstructions show large residual noise unless the data noise is relatively small. This indicates that at least for inverse source problems, Ivanov regularization is likely too weak for practical application in general; an exception would be problems where it is expected that $u^\dagger \in \{-\rho^\dagger, \rho^\dagger\}$ almost everywhere or where the main interest lies in the correct localization of strong inclusions.

Further study of quasi-solutions in Banach spaces would therefore be justified. Besides the extension to nonlinear parameter identification problems for PDEs mentioned in the introduction, it would be interesting to consider more general choices of M_ρ such as, e.g., $M_\rho = \{u : 0 \le u(x) \le \rho\}$ or $M_\rho = \{u : \rho^{-1} \le u(x) \le \rho\}$; the former would be relevant for the problem of identifying a potential coefficient in an elliptic PDE (and furnish a link to sparsity regularization), while the latter would be appropriate for identification of a diffusion coefficient. Finally, it is an open question whether heuristic parameter choice rules for ρ are possible, e.g., based on a model function approach for the distance function $d(\rho, y^\delta)$.

Table 1: quantitative results: noise percentage s, effective noise level δ, reconstruction errors in $L^2(\Omega)$- and $L^\infty(\Omega)$-norms and related to Bregman distance D_J^ξ

s	δ	$d(\rho, y^\delta)$	$\rho(\delta, y^\delta)$	$\|u_\rho^\delta - u^\delta\|_\infty$	$\|u_\rho^\delta - u^\delta\|_2$	$\langle \xi, u_\rho^\delta - u^\dagger \rangle$
10^0	$1.310 \cdot 10^{-3}$	$1.321 \cdot 10^{-3}$	2.500	$4.500 \cdot 10^0$	$3.571 \cdot 10^0$	$1.500 \cdot 10^0$
10^{-1}	$1.310 \cdot 10^{-4}$	$1.400 \cdot 10^{-4}$	3.594	$5.594 \cdot 10^0$	$3.101 \cdot 10^0$	$4.063 \cdot 10^{-1}$
10^{-2}	$1.310 \cdot 10^{-5}$	$1.358 \cdot 10^{-5}$	3.952	$2.894 \cdot 10^0$	$5.620 \cdot 10^{-1}$	$4.836 \cdot 10^{-2}$
10^{-3}	$1.310 \cdot 10^{-6}$	$1.367 \cdot 10^{-6}$	3.995	$2.051 \cdot 10^{-1}$	$5.661 \cdot 10^{-2}$	$4.944 \cdot 10^{-3}$
10^{-4}	$1.310 \cdot 10^{-7}$	$1.391 \cdot 10^{-7}$	3.999	$2.317 \cdot 10^{-2}$	$5.726 \cdot 10^{-3}$	$4.987 \cdot 10^{-4}$
10^{-5}	$1.310 \cdot 10^{-8}$	$1.357 \cdot 10^{-8}$	4.000	$2.070 \cdot 10^{-3}$	$5.612 \cdot 10^{-4}$	$4.844 \cdot 10^{-5}$

REFERENCES

[1] BURGER & OSHER, Convergence rates of convex variational regularization, *Inverse Problems* 20 (2004), 1411, DOI: 10.1088/0266-5611/20/5/005.

[2] CLASON, ITO & KUNISCH, A minimum effort optimal control problem for elliptic PDEs, *ESAIM Math. Model. Numer. Anal.* 46 (2012), 911–927, DOI: 10.1051/m2an/2011074.

[3] DECKELNICK & HINZE, A note on the approximation of elliptic control problems with bang-bang controls, *Computational Optimization and Applications* 51 (2012), 931–939, DOI: 10.1007/s10589-010-9365-z.

[4] DOMBROVSKAJA & IVANOV, On the theory of certain linear equations in abstract spaces, *Sibirsk. Mat. Z* 6 N.3 (1965), 499–508.

[5] ENGL, HANKE & NEUBAUER, *Regularization of Inverse Problems*, Kluwer Academic Publishers, Dordrecht, 1996, DOI: 10.1007/978-94-009-1740-8.

[6] FLEMMING, Theory and examples of variational regularization with non-metric fitting functionals, *J. Inverse Ill-Posed Probl.* 18 (2010), 677–699, DOI: 10.1515/JIIP.2010.031.

[7] GRASMAIR, Generalized Bregman distances and convergence rates for non-convex regularization methods, *Inverse Problems* 26 (2010), 115014, 16, DOI: 10.1088/0266-5611/26/11/115014.

[8] HINTERMÜLLER, ITO & KUNISCH, The primal-dual active set strategy as a semismooth Newton method, *SIAM J. Optim.* 13 (2002), 865–888, DOI: 10.1137/S1052623401383558.

[9] HINZE, A variational discretization concept in control constrained optimization: The linear-quadratic case, *Computational Optimization and Applications* 30 (2005), 45–61, DOI: 10.1007/s10589-005-4559-5.

[10] HOFMANN, KALTENBACHER, PÖSCHL & SCHERZER, A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators, *Inverse Problems* 23 (2007), 987–1010, DOI: 10.1088/0266-5611/23/3/009.

[11] ITO & KUNISCH, *Lagrange Multiplier Approach to Variational Problems and Applications*, Society for Industrial & Applied Mathematics (SIAM), 2008, DOI: 10.1137/1.9780898718614.

[12] IVANOV, On linear problems which are not well-posed, *Dokl. Akad. Nauk SSSR* 145 N.2 (1962), 270–272.

[13] IVANOV, On ill-posed problems, *Mat. Sb. (N.S.)* 61(103) N.2 (1963), 211–223.

[14] IVANOV, Ill-posed problems in topological spaces, *Mat. Sb. (N.S.)* 10 N.5 (1969), 1065–1074.

[15] IVANOV, VASIN & TANANA, *Theory of Linear Ill-posed Problems and its Applications*, VSP, Utrecht, 2002.

[16] KALTENBACHER & KLASSEN, On convergence and convergence rates for Ivanov and Morozov regularization and application to some parameter identification problems in elliptic PDEs, *Inverse Problems* 34 (2018), 055008, DOI: 10.1088/1361-6420/aab739.

[17] LORENZ & WORLICZEK, Necessary conditions for variational regularization schemes, *Inverse Problems* 29 (2013), 075016, DOI: 10.1088/0266-5611/29/7/075016.

[18] Morozov, Choice of parameter in solving functional equations by the method of regularization, *Dokl. Akad. Nauk SSSR* 175 (1967), 1225–1228.

[19] Morozov, *Methods for Solving Incorrectly Posed Problems*, Springer-Verlag, New York, 1984.

[20] Neubauer & Ramlau, On Convergence rates for quasi-solutions of ill-posed problems, *Electron. Trans. Numer. Anal* 41 (2014), 81–92, URL: http://etna.math.kent.edu/volumes/2011-2020/vol41/abstract.php?vol=41&pages=81-92.

[21] Resmerita, Regularization of ill-posed problems in Banach spaces: convergence rates, *Inverse Problems* 21 (2005), 1303, DOI: 10.1088/0266-5611/21/4/007.

[22] Schuster, Kaltenbacher, Hofmann & Kazimierski, *Regularization Methods in Banach Spaces*, De Gruyter, Berlin, 2012, DOI: 10.1515/9783110255720.

[23] Tikhonov, On the solution of ill-posed problems and the method of regularization, *Dokl. Akad. Nauk SSSR* 151 (1963), 501–504.

[24] Tröltzsch, *Optimal Control of Partial Differential Equations*, American Mathematical Society, 2010, DOI: 10.1090/gsm/112.

[25] Ulbrich, *Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces*, SIAM, 2011, DOI: 10.1137/1.9781611970692.

[26] Vasin, The method of quasi-solutions by Ivanov is the effective method of solving ill-posed problems, *J. Inv. Ill-Posed Problems* 16 (2008), 537–552, DOI: 10.1515/JIIP.2008.028.

[27] Vasin & Ageev, *Ill-posed Problems with A Priori Information*, VSP, Utrecht, 1995, DOI: 10.1515/9783110900118.

[28] Yagola, Leonov & Titarenko, Data Errors and an Error Estimation for Ill-Posed Problems, *Inverse Problems in Engineering* 10 (2002), 117–129, DOI: 10.1080/10682760290031195.

[29] Yagola & Titarenko, A posteriori error estimation for ill-posed problems on some sourcewise represented or compact sets, in: *Ill-posed and inverse problems*, VSP, Zeist, 2002, 425–442.