SpezialForschungsBereich F 32

Karl–Franzens Universität Graz
Technische Universität Graz
Medizinische Universität Graz

Minimization of Non-smooth, Non-convex Functionals by Iterative Thresholding

K. Bredies D. Lorenz S. Reiterer

SFB-Report No. 2014-015 November 2014

A–8010 GRAZ, HEINRICHSTRASSE 36, AUSTRIA

Supported by the
Austrian Science Fund (FWF)

Der Wissenschaftsfonds.

SFB sponsors:

- **Austrian Science Fund** (FWF)
- **University of Graz**
- **Graz University of Technology**
- **Medical University of Graz**
- **Government of Styria**
- **City of Graz**

SFB MOBIS Report 2014-015

Minimization of Non-smooth, Non-convex Functionals by Iterative Thresholding

Kristian Bredies · Dirk A. Lorenz · Stefan Reiterer

11 July 2014

Abstract Convergence analysis is carried out for a forward-backward splitting/generalized gradient projection method for the minimization of a special class of non-smooth and genuinely non-convex minimization problems in infinite dimensional Hilbert spaces. The functionals under consideration are the sum of a smooth, possibly non-convex and non-smooth, necessarily non-convex functional. For separable constraints in the sequence space, we show that generalized gradient projection method amounts to a discontinuous iterative thresholding procedure, which can easily be implemented. In this case we prove strong subsequential convergence, and moreover, show that the limit satisfies strengthened necessary conditions for a global minimizer, i.e., it avoids a certain set of non-global minimizers. Eventually, the method is applied to problems arising in the recovery of sparse data, where strong convergence of the whole sequence is shown, and numerical tests are presented.

Keywords Non-convex optimization · non-smooth optimization · gradient projection method · iterative thresholding

Mathematics Subject Classification (2000) 49M05 · 65K10

1 Introduction

In this article, we develop and analyze an iterative algorithm, which aims at minimizing non-smooth and non-convex functionals, covering the important special case of Tikhonov functionals for non-linear operators and non-convex penalty terms.

K. Bredies
Institute of Mathematics and Scientific Computing, University of Graz, Heinrichstraße 36, A-8010 Graz, Austria, E-mail: kristian.bredies@uni-graz.at

D. A. Lorenz (✉)
Institute for Analysis and Algebra, TU Braunschweig, D-38092 Braunschweig, Germany, E-mail: d.lorenz@tu-braunschweig.de

S. Reiterer
Institute of Mathematics and Scientific Computing, University of Graz, Heinrichstraße 36, A-8010 Graz, Austria, E-mail: stefan.harald.reiterer@gmail.com

The minimization of non-convex and non-smooth functionals is a delicate matter. On one hand, there is a class of popular generic algorithms such as simulated annealing, genetic algorithms and other derivative-free methods which can be performed with minimal assumptions on the objective functional. However, they tend to be impractical in higher-dimensional or infinite-dimensional spaces. On the other hand, many minimization algorithms for non-convex problems, which are based on derivatives, can be applied when the functional is sufficiently smooth. Convergence can, under suitable conditions, be established, also in infinite-dimensional spaces [1, 2]. When it comes to non-convex and non-smooth optimization problems in high (or infinite) dimensions, as it is typical for inverse problems in mathematical image processing or partial differential equations, there are only a few algorithms available, for instance the graduated non-convexity algorithm, which works for finite-dimensional, large-scale problems [3–5]. In the present paper, an alternative approach for the numerical minimization of such functionals is proposed. More precisely, we introduce a generalized gradient projection method (GGPM), which involves proximity mappings for non-convex functionals and show convergence.

The class of problems the algorithms addresses is finding a minimizer of functionals on real Hilbert spaces, that are the sum of a smooth part (with Lipschitz continuous derivative) and a non-smooth, non-convex part. The algorithm is similar in spirit to the well known *proximal gradient methods*, also known as *forward-backward methods* (cf., e.g., [6, 7]) or *generalized gradient methods* [8]: One takes forward gradient steps with respect to the smooth part, and then a backward (proximal) step with respect to the non-smooth part. One difficulty here is, that, due to non-convexity, the proximal step may be multivalued and not non-expansive. Recently (and after this paper was published as a preprint), similar non-convex problems have been treated with similar ideas, e.g., in [9, 10] and [11, 12] in which the non-smooth part is assumed to be convex, while the smooth part may be non-convex. In some of these works, the so-called Kurdyka-Łojasiewicz (KŁ) inequality is assumed to hold for the objective function. This is a powerful but weak assumption that allows in this cases to conclude the convergence of the whole sequence of iterates. In this work we do not use the KŁ inequalitym but use tools from algebraic geometry to prove convergence of the sequence of iterates. More related work on non-convex problems exists, e.g., for proximal-point methods [13] or for methods involving non-convex projections [14]. While proximity operators are well-studied for convex functionals [15], the non-convex case has been of interest to researchers only recently [16, 17].

In particular, we discuss, as an important example, separable non-smooth functionals that are *necessarily* non-convex, where a complete characterization of the proximity mapping will be derived. Under general assumptions, the GGPM can then be shown to converge subsequentially to a stationary point if the step-sizes are bounded away from a critical constant. However, we will see that these step-size can, in general, not lead to global minimizers in the limit and that step-sizes approaching this critical constant are beneficial. We therefore also consider a variant of the GGPM, where subsequential convergence to so-called quasi-global minimizers, i.e., stationary points which share some necessary properties with the global minimizers, can be ensured. Finally, we apply this framework to Tikhonov regularization for solving the ill-posed linear inverse problem with non-convex penalties in sequence spaces. If the exponent of the sequence space is also rational, the quasi-global minimizers are shown to be discrete, leading to the convergence of GGPM for the whole sequence (a result which has been shown previously using semi-algebraic functions and the Kurdyka-Łojasewiec inequality, [9]).

The paper is organized as follows: Section 2 presents the considered framework, the GGPM for this case and derives some basic properties. In Section 3, we investigate the proximity operator for certain one-dimensional non-convex functions. In Section 4, we combine the results of the previous sections to treat the case of separable non-smooth functionals in the infinite dimensional Hilbert space of square summable sequences. There we show subsequential convergence of GGPM to stationary points for step-sizes bounded away from a critical constant and investigate quasi-global minimizers. We show that letting the step-sizes in GGPM approach the critical constant, it is possible to obtain subsequential convergence to such quasi-global minimizers. In Section 5, the results are specialized further to sequence spaces with non-convex quasi-norms in sequence spaces and, in particular, to Tikhonov regularization for linear inverse problems. Using tools from algebraic geometry, it is shown that in this setting and for rational exponent of the sequence space, the whole sequence converges to a quasi-global minimizer. Section 6 presents numerical experiments and Section 7 summarizes and concludes the paper.

2 The Generalized Gradient Projection Method

The class of problems the algorithm addresses is finding a minimizer of the functional

$$\min_{u \in X} T_\alpha(u) \quad \text{with} \quad T_\alpha(u) := S(u) + \alpha R(u)$$

in a real Hilbert space X. Here, S is assumed to be smooth, i.e., differentiable with Lipschitz-continuous derivative. In contrast to that, the functional R may be non-convex and non-smooth. The proposed method, GGPM, then consists of iteratively computing

$$u^{n+1} \in \operatorname*{argmin}_{v \in X} \frac{\left\| v - u^n + sS'(u^n) \right\|^2}{2} + s\alpha R(v).$$

Here, we assume, that S', as well as the proximity mapping with respect to R, i.e., global minimizers of $v \mapsto \frac{1}{2} \|v - w\|^2 + s\alpha R(v)$, can efficiently be computed.

We start our analysis by first considering the general case of minimizing the functional

$$T_\alpha := S + \alpha R. \tag{1}$$

Throughout this paper, we assume the following on the involved functionals.

Assumption 2.1 (General setup) Let X be a real Hilbert space and $\alpha > 0$. For the functionals $S : X \to [0, \infty[$ and $R : X \to [0, \infty]$ we assume

(a) S is weakly sequentially lower-semicontinuous and continuously differentiable, the derivative S' is Lipschitz continuous with constant

$$L := \sup_{x \neq y} \frac{\left\| S'(x) - S'(y) \right\|}{\|x - y\|} > 0.$$

(b) R is proper, weakly sequentially lower-semicontinuous and coercive; in other words, $\mathrm{dom}(R) \neq \emptyset$, $R(u) \leq \liminf_{n\to\infty} R(u^n)$ whenever $u^n \rightharpoonup u$, and finally $\|u^n\| \to \infty$ implies $R(u^n) \to \infty$.

Observe that this implies, by the direct method, the existence of a minimizer for T_α. Moreover, minimizing T_α does not reduce to the minimization of R as the requirement that the Lipschitz constant of S' is positive implies that S is non-constant (if S is constant, then $S' = 0$ and $L = 0$).

The generalized gradient projection algorithm we are proposing builds on the gradient projection algorithm for constrained minimization problems, i.e., for a smooth S and a non-empty, convex and closed set C one considers the problem $\min_{u\in C} \; S(u)$. Recall that in this case, the method performs a step along the negative gradient $-S'(u^n)$ with a suitable step-size s_n, and by P_C projects back onto the feasible set: $u^{n+1} := P_C(u^n - s_n S'(u^n))$ (cf. [18,19]). The main idea of the GGPM for the solution of (1) is to replace the convex constraint C by a general functional R, and to replace the projection P_C by the associated proximity operator, and due to the similarity with the projection, we use the same notation, i.e.,

$$\mathrm{prox}_R : u \mapsto \operatorname*{argmin}_{v\in X} \frac{\|u-v\|^2}{2} + R(v). \tag{2}$$

Hence, the GGPM reads as

$$\begin{cases} u^0 \in \mathrm{dom}(R) \quad , \quad u^{n+1} \in G_{\alpha,s_n}(u^n), \\ G_{\alpha,s_n}(u) := \operatorname*{argmin}_{v\in X} \dfrac{\left\|u - s_n S'(u) - v\right\|^2}{2} + s_n \alpha R(v). \end{cases} \tag{3}$$

It can easily be seen that this iteration and in particular, the proximity operator, is well-defined if Assumption 2.1 is satisfied. The motivation for the consideration of the minimization problem (2) is that, in practice and as we will see later, it may be much easier to solve than the original problem since it only involves the regularization term R.

In the non-convex case, the operator prox_R may be set-valued because there may be several global minima—moreover local minima may exist. However, if one is able to evaluate prox_R (or $\mathrm{prox}_{s\alpha R}$, respectively), i.e., to calculate global minimizers of $\|u-v\|^2/2 + s\alpha R(v)$ with respect to v, then descent of the functional value is guaranteed, as shown in the following proposition, which is proven along the lines of [9, Section 5.1].

Proposition 2.1 (Basic descent) *Let Assumption 2.1 hold true. Then,* $\mathrm{prox}_{s\alpha R}$ *is non-empty for each* $s > 0$ *and it holds for every* $v \in G_{\alpha,s}(u)$ *that*

$$T_\alpha(v) \leq T_\alpha(u) - \tfrac{1}{2}\left(\tfrac{1}{s} - L\right)\|v-u\|^2. \tag{4}$$

From this proposition we conclude that a step-size $0 \leq s \leq L^{-1}$ does not increase the objective functional T_α. However, since T_α is bounded from below we get that the sequence $(T_\alpha(u^n))$ converges as soon as $0 \leq s_n \leq L^{-1}$ for each n. As a direct consequence, we moreover have the following corollary.

Corollary 2.1 (Decreasing difference of iterates) *In the situation of Proposition 2.1 and with a step-size sequence* (s_n) *satisfying* $0 < \underline{s} \leq s_n \leq \overline{s} < L^{-1}$ *for each* n*, the sequence* (u^n) *generated by the iteration* (3) *obeys*

$$\left\| u^{n+1} - u^n \right\| \to 0 \quad \text{as} \quad n \to \infty. \tag{5}$$

Proof The assertion follows from the estimate

$$\tfrac{1}{2}\big(\tfrac{1}{s_n} - L\big) \left\| u^{n+1} - u^n \right\|^2 \leq T_\alpha(u^n) - T_\alpha(u^{n+1})$$

and the observation that $(T_\alpha(u^n))$ is a converging sequence. □

In general, this does not lead to convergence, but if R is, for example, some power of a norm in a space which is compactly embedded in X, i.e.,

$$R(u) := \begin{cases} \|u\|_Z^p & u \in Z \\ \infty & u \notin Z \end{cases}, \qquad Z \hookrightarrow X, \quad p > 0, \tag{6}$$

then, from the boundedness of the sequence $(T_\alpha(u^n))$, it follows that $(R(u^n))$ is bounded and hence (u^n) is relatively compact in X, admitting a convergent subsequence $u^{n_j} \to u^*$ as $j \to \infty$. It is however, not clear whether the limit is a global solution (which is rather unlikely in general) or at least a point which is stationary in some sense. At the current stage, however, we can state the following.

Proposition 2.2 (Global minimizers are fixed points) *Let Assumption 2.1 hold true. For each global minimizer* u^* *of* T_α *and any* $0 < s < L^{-1}$*, we have*

$$G_{\alpha,s}(u^*) = \{u^*\}.$$

In other words: Each global minimizer of T_α *is a fixed point of the GGPM* (3)*.*

Moreover, the graph G_α *and the fixed points* $U_{\alpha,s}$*, according to*

$$G_\alpha := \{(u, v, s) \in X \times X \times [0, \infty[\ : \ v \in G_{\alpha,s}(u)\},$$
$$U_{\alpha,s} := \{u \in X \ : \ u \in G_{\alpha,s}(u)\},$$

are closed sets.

Proof Choosing $v \in G_{\alpha,s}(u^*)$ and utilizing (4) implies

$$T_\alpha(u^*) \leq T_\alpha(v) \leq T_\alpha(u^*) - \Big(\frac{1}{2s} - \frac{L}{2}\Big) \left\| v - u^* \right\|^2,$$

hence $\|v - u^*\|^2 \leq 0$, and consequently, $v = u^*$.

The closedness of G_α follows from continuity of S' and the sequential lower-semicontinuity of R since this shows that defining inequality for $(u, v, s) \in G_\alpha$ stays true under limits. The closedness of $U_{\alpha,s}$ follows from $i_s : u \mapsto (u, u, s)$ being continuous between $X \to X \times X \times [0, \infty[$ and $U_{\alpha,s} = i_s^{-1}(G_\alpha)$. □

Proposition 2.3 (Accumulation points are fixed points) *If* $((u^n, s_n))$ *is a sequence generated by* (3) *with step-sizes satisfying* $0 < \underline{s} \leq s_n \leq \overline{s} < L^{-1}$*, then each accumulation point of* (u^n) *is in* $U_{\alpha,s}$ *for some* $s \in [\underline{s}, \overline{s}]$*.*

If (u^n) *is relatively compact and possesses an isolated accumulation point, then the whole sequence converges to some* $u^* \in U_{\alpha,s}$*.*

Proof Taking a sequence of iterates and step-sizes $((u^n, s_n))$ in (3), and assuming that the subsequence (u^{n_j}) satisfies $u^{n_j} \to u^*$, one can extract another subsequence (not relabeled) such that, additionally, $s_{n_j} \to s$ for some $s \in [\underline{s}, \overline{s}]$. Now for the sequence $v^n := u^{n+1} \in G_{\alpha,s_n}(u^n)$, Corollary 2.1 implies that $v^{n_j} \to u^*$, hence $(u^*, u^*, s) \in G_\alpha$ by closedness, meaning $u^* \in U_{\alpha,s}$.

Now suppose that u^* is an isolated accumulation point for (u^n), i.e., there is a $\delta > 0$ such that for each accumulation point u with $\|u^* - u\| < \delta$ it holds that $u = u^*$. Assume that not the whole sequence converges to u^*, i.e., there exists a $0 < \varepsilon < \delta$ and infinitely many n with $\|u^n - u^*\| \geq \varepsilon$. Denote by N the set of these n. Since there exists a subsequence of (u^n) with limit u^*, we can find infinitely many n with $n \in N$ and $n + 1 \notin N$. Denote by (u^{n_j}) the subsequence associated with these n. By construction ,$\|u^{n_j+1} - u^*\| < \varepsilon < \delta$, and by relative compactness, one can choose another subsequence (still denoted by (u^{n_j+1})) with $u^{n_j+1} \to u$ for some u. It follows that $\|u^* - u\| < \delta$, hence $u^{n_j+1} \to u^*$, and since $\|u^{n_j} - u^{n_j+1}\| \to 0$, also $u^{n_j} \to u^*$. Because of $n_j \in N$, this is a contradiction, hence the assumption that not the whole sequence converged must have been wrong. ◻

The sets $U_{\alpha,s}$ in Proposition 2.2 describe, in some sense, the fixed points of the iteration, hence Proposition 2.3 says that the GGPM converges subsequentially, if it converges, to a fixed point. This may not be a satisfactory result at the first glance, but one shall keep in mind that no convexity of R is assumed and hence, one can not expect strong results in this general setting.

Remark 2.1 A crucial condition in this section is the assumption that the step-size are strictly smaller than L^{-1}. We will refer to L^{-1} as *critical constant*, as we will observe in Section 4 that it plays a special role.

3 The One-dimensional Proximity Operator

In order to apply the algorithm (3) in infinite-dimensional spaces for separable constraints R in the next section, we study one-dimensional penalty functions ϕ and, in particular, the properties of the associated proximity operators. I.e., for $\phi : [0, \infty[\to [0, \infty[$ we investigate the multivalued function $\text{prox}_{s\alpha\phi}$, i.e.,

$$\text{prox}_{s\alpha\phi}(x) := \underset{y \in \mathbb{R}}{\text{argmin}} \ \frac{(y - x)^2}{2} + s\alpha\phi(|y|). \tag{7}$$

In the following, we identify multivalued functions $f : \mathbb{R} \rightrightarrows \mathbb{R}$ with subsets of $\mathbb{R} \times \mathbb{R}$ by its graph G_f, i.e., $f(x) = \{y \in \mathbb{R} \ : \ (x, y) \in G_f\}$. All pointwise operations on multivalued functions like addition, multiplication then have to be understood in an respective elementwise sense. This also applies to relations like $\leq$ and $\geq$.

As already mentioned, we are interested in the case where ϕ is necessarily non-convex. As we will see, the multivaluedness of $\text{prox}_{s\alpha\phi}$ is often not as severe as it may seem. Under suitable assumptions on ϕ it will be mainly single-valued except for two points, where is admits exactly two values. Hence, we will drop the term “multivalued” for the function $\text{prox}_{s\alpha\phi}$ in the following.

Throughout this paper, we usually assume the following on ϕ.

Assumption 3.1 (One-dimensional penalty function) Let $\phi : [0, \infty[\to [0, \infty[$ be continuous and non-decreasing with $\phi(0) = 0$.

(a) ϕ is differentiable on $]0,\infty[$ with $\phi'(x) \to \infty$ for $x \to 0$.
(b) The derivative ϕ' is strictly convex on $]0,\infty[$ with $\phi'(x)/x \to 0$ for $x \to \infty$.
(c) The function $x \mapsto \phi''(x)x$ is locally integrable on $[0,\infty[$.
(d) It holds that $\phi(x) \to \infty$ whenever $x \to \infty$.
(e) For each $b > 0$ there exists an $a > 0$ such that $\phi(x) \geq ax^2$ for $x \in [0,b]$.

Note that (a) already implies that ϕ cannot be convex. Moreover, from (b) follows that ϕ' is locally Lipschitz continuous and hence, ϕ'' is almost everywhere defined, so the requirement (c) indeed makes sense.

We start with the observation that it suffices to know $\operatorname{prox}_{s\alpha\phi}(x)$ for $x \geq 0$ as one can easily see that

$$\operatorname{prox}_{s\alpha\phi}(-x) = -\operatorname{prox}_{s\alpha\phi}(x) \quad \text{for all } x \in \mathbb{R}, \qquad \operatorname{prox}_{s\alpha\phi}(x) \geq 0 \quad \text{for all } x \geq 0.$$

So, without any loss of generality, we may assume $x, y \geq 0$. It is moreover useful to observe a necessary condition for $y \in \operatorname{prox}_{s\alpha\phi}(x)$:

$$y = 0 \quad \text{or} \quad x = (\mathrm{id} + s\alpha\phi')(y).$$

The following property on $\operatorname{prox}_{s\alpha\phi}$ resembles a result from [3], namely that Assumption 3.1(a) guarantees that $\operatorname{prox}_{s\alpha\phi}$ has a "dead-zone".

Lemma 3.1 (Deadzone of $\operatorname{prox}_{s\alpha\phi}$) *Let ϕ fulfill Assumption 3.1(a). Then, there exists a $\kappa_s > 0$ such that*

$$|x| < \kappa_s \implies \operatorname{prox}_{s\alpha\phi}(x) = 0.$$

Moreover, κ_s depends monotonically increasing on s.

Proof By definition of $\operatorname{prox}_{s\alpha\phi}$ we have, for $\operatorname{prox}_{s\alpha\phi}(x) \neq 0$,

$$x = \operatorname{prox}_{s\alpha\phi}(x) + s\alpha\phi'\big(\operatorname{prox}_{s\alpha\phi}(x)\big) = (\mathrm{id} + s\alpha\phi')\big(\operatorname{prox}_{s\alpha\phi}(x)\big).$$

Without any loss of generality, consider $x > 0$. Since $\phi' \geq 0$ and $\phi'(y)$ tends to infinity for $y \to 0$, the value $\kappa_s := \inf_{y>0}\ y + s\alpha\phi'(y)$ is positive, meaning that, for all $y > 0$ we have $(\mathrm{id} + s\alpha\phi')(y) \geq \kappa_s$, which proves the claim. Moreover, from the definition of κ_s it is obvious that it depends monotonically increasing on s. □

We need a preparatory lemma for proving continuity of the $\operatorname{prox}_{s\alpha\phi}$, the proof of which is given in the Appendix. It utilizes the following notions: For $f : I \to \mathbb{R}$ convex on some open interval I, define the monotone limits

$$f'_+(x) := \inf_{h>0} \frac{f(x+h) - f(x)}{h}, \qquad f'_-(x) := \sup_{h>0} \frac{f(x) - f(x-h)}{h} \qquad \text{for} \quad x \in I.$$

We have $f'_- \leq f'_+$. For $x \in I$, $f'_+(x) = f'_-(x)$ if and only if f is differentiable in x. In this case $f'_-(x) = f'_+(x) = f'(x)$, which holds for almost every $x \in I$. Both f'_- and f'_+ are monotonically increasing, strictly if f is strictly convex.

Lemma 3.2 *Assumption 3.1(a)–(c) implies the following properties:*

(a) for each $s > 0$, the function $\rho_s : y \mapsto y + s\alpha\phi'(y)$ exists on $]0,\infty[$, is strictly convex and attains a minimum at some $y_s > 0$,

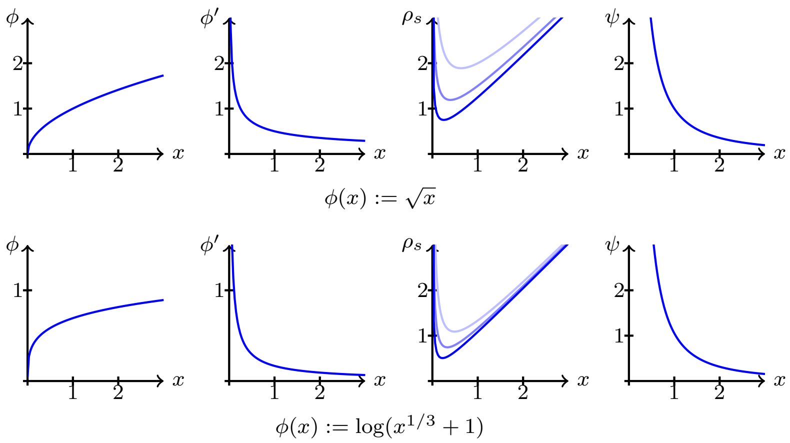


Fig. 1 Illustration of some ϕ satisfying Assumption 3.1, ϕ', ρ_s ($\alpha = 1$, $s \in \{1/2, 1, 2\}$) and ψ.

(b) the function $\psi : y \mapsto 2\bigl(\phi(y) - y\phi'(y)\bigr)/y^2$ *is strictly decreasing and onto as a mapping* $]0,\infty[\, \to \,]0,\infty[$,

(c) we have, for $y > 0$ *and any* $z \geq (\phi')'_-(y)$, *that* $\psi(y) > -z$.

Example 3.1 For $p \in \,]0,1[$, it can be verified that the functions $\phi(x) := x^p$, as well as $\phi(x) := \log(x^p + 1)$, satisfy Assumption 3.1. In Figure 1, you can see an illustration of the constructions in Lemma 3.2.

The continuity properties of $\mathrm{prox}_{s\alpha\phi}$ can now easily be deduced.

Lemma 3.3 (Exact formula and continuity of $\mathrm{prox}_{s\alpha\phi}$**)** *Assume that the Assumptions 3.1(a)–(c) are fulfilled and let* ρ_s *and* ψ *be defined as in Lemma 3.2. With*

$$\lambda_s := \psi^{-1}\bigl((s\alpha)^{-1}\bigr), \quad \tau_s := \rho_s \circ \psi^{-1}\bigl((s\alpha)^{-1}\bigr), \tag{8}$$

the functions $\mathrm{prox}_{s\alpha\phi}$ *obey*

$$\mathrm{prox}_{s\alpha\phi}(x) = \begin{cases} 0, & \textit{for } |x| \leq \tau_s \\ \mathrm{sign}(x)\rho_s^{-1}(|x|), & \textit{for } |x| \geq \tau_s, \end{cases}$$
$$\mathrm{prox}_{s\alpha\phi}(x) \in \,]-\infty, -\lambda_s] \cup \{0\} \cup [\lambda_s, \infty[.$$

In particular, $\mathrm{prox}_{s\alpha\phi}$ *takes exactly two values if* $|x| = \tau_s$, *is left- and right continuous there (in the sense that* $\mathrm{prox}_{s\alpha\phi}(\pm\tau_s)$ *are exactly the left and right limits). Moreover,* $\mathrm{prox}_{s\alpha\phi}$ *is continuous on* $\{|x| \neq \tau_s\}$.

Proof We fix $s > 0$ and denote $F_x(y) := \frac{1}{2}(y-x)^2 + s\alpha\phi(|y|)$. Furthermore, assume, without any loss of generality, $x \geq 0$ such that we only need to minimize F_x over $[0,\infty[$. Note the identities $F'_x = \rho_s - x$ and $(F'_x)'_+ = (\rho_s)'_+$. From Lemma 3.2 we know that there is a unique minimizer $y_s > 0$ of ρ_s. In accordance to the notation introduced in Lemma 3.1, we denote $\kappa_s := \rho_s(y_s)$.

We have already seen that $\mathrm{prox}_{s\alpha\phi}(x) = 0$ for $x < \kappa_s$. If $x \geq \kappa_s$, we have, for any $y < y_s$ with $\rho_s(y) = x$, that $(F'_x)'_+(y) = (\rho_s)'_+(y) < 0$ as otherwise, y_s would not be the minimizer. Consequently, there exists a $\varepsilon > 0$ such that it holds that $\rho_s(y+h) - x < \rho_s(y) - x = 0$ for $0 < h < \varepsilon$, so

$$F_x(y+h') - F_x(y) = \int_0^{h'} \rho_s(y+h) - x \,\mathrm{d}h < 0 \qquad \text{for each} \quad 0 < h' < \varepsilon.$$

Hence, y cannot be a local minimizer. As each global minimizer $y \in \operatorname{prox}_{s\alpha\phi}(x)$ of F_x is also a local one, it follows that either $y = 0$ or $y \geq y_s$. Now ρ_s is strictly monotonically increasing on $[y_s, \infty[$: For $y_s < y_1 < y_2$ it follows from $y_1 = \lambda y_s + (1-\lambda) y_2$ with $\lambda = (y_2 - y_1)/(y_2 - y_s) \in]0,1[$ and the strict convexity of ρ_s that

$$\rho_s(y_1) < \lambda \rho_s(y_s) + (1-\lambda)\rho_s(y_2) \leq \rho_s(y_2).$$

For $y_1 = y_s$ the inequality follows from the uniqueness of the minimizer y_s. Moreover, $\lim_{y\to\infty} \rho_s(y) = \infty$ since $\phi' \geq 0$. Hence, there is exactly one $y^* \geq y_s$ with $x = \rho_s(y^*)$ or, equivalently, $F_x'(y^*) = 0$. Thus, we just have to compare the values of $F_x(0)$ and $F_x(y^*)$ in order to determine the minimizers. It turns out that

$$\begin{aligned} F_x(y^*) - F_x(0) &= \frac{(s\alpha\phi'(y^*))^2}{2} + s\alpha\phi(y^*) - \frac{(y^* + s\alpha\phi'(y^*))^2}{2} \\ &= s\alpha\big(\phi(y^*) - y^*\phi'(y^*)\big) - \frac{(y^*)^2}{2}, \end{aligned}$$

which leads to

$$\begin{aligned} \operatorname{sign}\big(F_x(y^*) - F_x(0)\big) &= \operatorname{sign}\big(\psi(y^*) - (s\alpha)^{-1}\big) = \operatorname{sign}\big(\psi^{-1}((s\alpha)^{-1}) - y^*\big) \\ &= \operatorname{sign}(\lambda_s - y^*); \end{aligned}$$

the latter since ψ is invertible and strictly monotonically decreasing, see Lemma 3.2 (b). Finally, ρ_s is again strictly monotonically increasing on $\{y \geq y_s\}$ and $x = \rho_s(y^*)$, hence

$$\operatorname{sign}\big(F_x(y^*) - F_x(0)\big) = \operatorname{sign}\big(\rho_s(\lambda_s) - x\big) = \operatorname{sign}(\tau_s - x).$$

Note that necessarily $\tau_s \geq \rho_s(y_s)$ holds, and hence $\operatorname{prox}_{s\alpha\phi}(x) = 0$ for all $x < \tau_s$ and $\operatorname{prox}_{s\alpha\phi}(x) = \rho_s^{-1}(x)$ for $x > \tau_s$. For $x = \tau_s$, there are the two minimizers: 0 and λ_s.

Finally observe that ρ_s can be continuously inverted on $[\lambda_s, \infty[$, hence $\operatorname{prox}_{s\alpha\phi}$ is continuous on $[\tau_s, \infty[$ with the left limit in τ_s being λ_s. The claimed continuity on $[0, \tau_s]$ is trivial, and it is easy to see that $\operatorname{range}(\operatorname{prox}_{s\alpha\phi}) = \{0\} \cup \{|x| \geq \lambda_s\}$. □

Remark 3.1 We remark that the threshold τ_s from (8) is always greater than the minimum of ρ_s: The minimizer y_s satisfies $1 + s\alpha(\phi')'_-(y_s) \leq 0$ and consequently, $(\phi')'_-(y_s) \leq -(s\alpha)^{-1}$. According to the definition, $(s\alpha)^{-1} = \psi(\lambda_s)$, so due to Lemma 3.2 (c) we have $\psi(y_s) > \psi(\lambda_s)$ and hence $y_s < \lambda_s$. Since ρ_s is strictly monotonically increasing on $[y_s, \infty[$, $\min_{y>0} \rho_s(y) = \rho_s(y_s) = \kappa_s < \tau_s$ follows. Moreover, note that, in particular, we have that $(\rho_s)'_-(\tau_s) > 0$.

You can find illustrations of the functions $\operatorname{prox}_{s\alpha\phi}$ for some particular ϕ in Figure 2. In Section 4.3 we will see, that it will be crucial for the algorithm (3) to consider non-constant step-sizes s_n increasing to the critical constant. The following lemma will be helpful in this context.

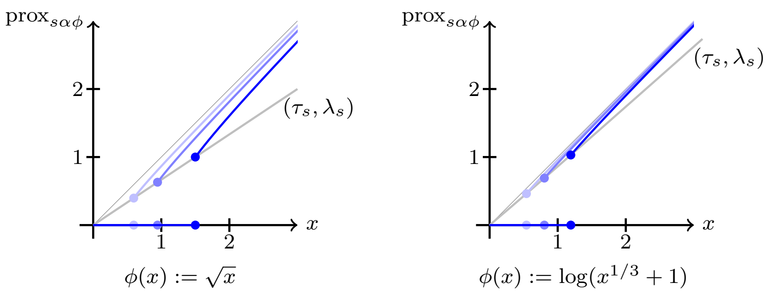


Fig. 2 Illustration of $\mathrm{prox}_{s\alpha\phi}$ for some particular ϕ and $s \in \{1/4, 1/2, 1\}$ ($\alpha = 1$). Additionally, the curve of breaking points (τ_s, λ_s) is depicted.

Lemma 3.4 (Joint continuity of $\mathrm{prox}_{s\alpha\phi}(x)$) *Let Assumption 3.1(a)–(c) be satisfied, λ_s and τ_s be as in (8), $s_n \to s^*$ monotonically increasing and $x_n \to x^*$. If $|x^*| \neq \tau_{s^*}$, then $\mathrm{prox}_{s_n\alpha\phi}(x_n) \to \mathrm{prox}_{s^*\alpha\phi}(x^*)$.*

Otherwise, the accumulation points of $(\mathrm{prox}_{s_n\alpha\phi}(x_n))$ correspond to one of the following three sets: $\{0\}, \{\mathrm{sign}(x^*)\lambda_{s^*}\}, \{0, \mathrm{sign}(x^*)\lambda_{s^*}\}$.

Proof Again, it suffices to assume that $x_n \geq 0$ for each n and $x^* \geq 0$. Then, note that the set $\{(x, s) \in \mathbb{R}^2 \ : \ s \in \,]0, \infty[, \ |x| \neq \tau_s\}$ is open as $s \mapsto \tau_s$ is continuous on $]0, \infty[$.

Suppose that $x^* < \tau_{s^*}$, implying that from some index on, $x_n < \tau_{s^*}$. Consequently, it holds that $\mathrm{prox}_{s_n\alpha\phi}(x_n) = 0$ from some index on, and this yields $\lim_{n\to\infty} \mathrm{prox}_{s_n\alpha\phi}(x_n) = 0 = \mathrm{prox}_{s^*\alpha\phi}(x^*)$. Analogously, $x^* > \tau_{s^*}$ implies that from some index on, $x_n > \tau_{s_n}$. Hence $\big(x^*, s^*, \mathrm{prox}_{s^*\alpha\phi}(x^*)\big)$ is in

$$M := \{(x, s, y) \in \mathbb{R}^3 \ : \ s \in \,]0, \infty[, \ x \geq \kappa_s, \ y \geq y_s\},$$

denoting by y_s and κ_s again the minimizing argument and minimum of ρ_s, respectively. It is easily seen that $F : M \to \mathbb{R}$, defined by

$$F(x, s, y) := y - x + s\alpha\phi'(y) = \rho_s(y) - x,$$

is locally Lipschitz continuous in M. Moreover, for $x^* \geq \tau_{s^*}$ we have that the generalized partial derivative satisfies $\frac{\partial F}{\partial y}\big(x^*, s^*, \mathrm{prox}_{s^*\alpha\phi}(x^*)\big) > 0$, the latter since $\tau_{s^*} > \rho_{s^*}(y_{s^*})$, see Remark 3.1. The implicit function theorem for Lipschitz functions [20] implies the Lipschitz continuity of the mapping that locally parameterizes the fiber $F^{-1}(0)$, i.e., $(s, x) \mapsto \rho_s^{-1}(x) = \mathrm{prox}_{s\alpha\phi}(x)$. We conclude that $\mathrm{prox}_{s_n\alpha\phi}(x_n) \to \mathrm{prox}_{s^*\alpha\phi}(x^*)$.

Finally, let $x^* := \tau_{s^*}$ and $y_n \in \mathrm{prox}_{s_n\alpha\phi}(x_n)$ for each n. Defining

$$I^- := \{n \in \mathbb{N} \ : \ x_n \leq \tau_{s_n}, \ y_n = 0\}, \qquad I^+ := \{n \in \mathbb{N} \ : \ x_n \geq \tau_{s_n}, \ y_n \neq 0\}$$

we get a disjoint partition of $\mathbb{N}$. Furthermore, provided that the respective sets are infinite,

$$\lim_{\substack{n\to\infty \\ n\in I^-}} y_n = 0, \qquad \lim_{\substack{n\to\infty \\ n\in I^+}} y_n = \rho_{s^*}^{-1}(x^*) = \lambda_{s^*},$$

the latter since necessarily, $y_n = \rho_{s_n}^{-1}(x_n)$ for $n \in I^+$, and $(s, x) \mapsto \rho_s^{-1}(x)$ is also locally Lipschitz continuous for $x^* \geq \tau_{s^*}$. The observation that at least of the sets I^-, I^+ is infinite now gives the claimed statement. □

Note that we did not utilize Assumption 3.1(d)–(e) yet. These conditions will play a role in the next section.

4 Application to Separable Constraints

The results of Section 3 will enable us to provide results on the GGPM beyond the elementary ones from Section 2. First observe that Assumption 3.1(d)–(e) on ϕ implies that the regularization functional R, according to

$$R(u) := \sum_k \phi(|u_k|), \tag{9}$$

is necessarily non-negative, coercive and weak sequentially lower semi-continuous in ℓ^2, see [21, Lemmas 3.2 and 3.4], consequently, Assumption 2.1(b) is satisfied. We therefore assume in the following, in addition to Assumption 3.1, that $X := \ell^2$ and $S : \ell^2 \to [0, \infty[$ fulfills Assumption 2.1 (a).

Therefore, we can postulate that Assumptions 2.1 and 3.1 are fulfilled in the following. Assuming that $S'(u)$ is computationally accessible, the main problem is the evaluation of the proximity operator. However, fortunately in the case of separable constraints, computing $\text{prox}_{s\alpha R}$ is reduced to the evaluation of $\text{prox}_{s\alpha\phi}$, according to (7). Knowing the (generally multi-valued) mapping $\text{prox}_{s\alpha\phi}$, iteration (3) amounts to

$$\begin{cases} u^0 \text{ such that } \sum_{k=1}^{\infty} \phi(|u_k^0|) < \infty, \\ u_k^{n+1} \in \text{prox}_{s_n\alpha R}\big(u^n - s_n S'(u^n)\big)_k = \text{prox}_{s_n\alpha\phi}\big(u_k^n - s_n S'(u^n)_k\big). \end{cases} \tag{10}$$

We carry out convergence analysis for this iteration and we recover the result from [9] that subsequential convergence to a stationary point holds if the step-sizes are bounded away from the critical constant L^{-1} and 0. However, as the problem is non-convex, there are usually stationary points, which are not global minimizers. Then, we derive further necessary properties of the sought global minimizers of T_α according to (1), which lead to the notion of *quasi-global minimizers*. This gives rise to a refined variant of the GGPM for separable constraints, and using this variant and letting the step-sizes increase to the critical constant L^{-1} in a suitable way then also yields subsequential convergence to quasi-global minimizers.

4.1 Step-sizes Bounded Away From the Critical Constant

We give properties of the iterates of the GGPM. Define the *support* of a $u \in \ell^2$ as the set of indices where u does not vanish, i.e., $\text{supp}(u) := \{k \in \mathbb{N} \ : \ u_k \neq 0\}$. Note that, since $\text{prox}_{s\alpha\phi}(x) = 0$ for all $|x| < \tau_s$, all iterates u^n, $n \geq 1$ have finite support. The crucial observation now is that a change of support always implies a "jump" of size κ_s and can hence occur only finitely many times.

Theorem 4.1 (Finite support changes and existence of strong accumulation points) *Let S, R and ϕ satisfy Assumption 2.1,* (9) *and Assumption 3.1. Consider* (u^n)*, generated by iteration* (10) *with step-size sequence* (s_n) *satisfying*

$$0 < \underline{s} \leq s_n \leq \overline{s} < L^{-1}.$$

Then, all iterates u^n $(n \geq 1)$ have a finite support, and the support only changes finitely many times.

Moreover, every subsequence of (u^n) has a strong accumulation point u^*, which is a fixed point in the sense $u^* \in G_{\alpha,s}(u^*)$ for some $s^* \in [\underline{s}, \overline{s}]$.

Proof Due to Lemma 3.1 there exists a $\kappa_{\underline{s}} > 0$ such that $|x| < \kappa_{\underline{s}}$ implies $\mathrm{prox}_{s_n\alpha\phi}(x) = 0$ for all n, the latter since $s_n \geq \underline{s} > 0$, and κ_s depends monotonically increasing on s. Hence, each non-zero component of u^n has magnitude greater or equal to $\kappa_{\underline{s}}$. We conclude that, if the support of u^{n+1} differs from u^n, we have $\|u^{n+1} - u^n\| \geq \kappa_{\underline{s}}$. However, from Proposition 2.1 we deduce that also

$$\frac{(\overline{s}^{-1} - L)\kappa_{\underline{s}}^2}{2} \leq \frac{(s_n^{-1} - L)\kappa_{\underline{s}}^2}{2} \leq T_\alpha(u^n) - T_\alpha(u^{n+1})$$

with the right-hand side going to zero as $n \to \infty$. Hence, a change of the support can only occur a finite number of times.

By assumption, T_α is coercive and hence, has a subsequence (u^{n_j}) with weak accumulation point u^*. Without any loss of generality, we can also assume that $\lim_{j\to\infty} s_{n_j} =: s^*$ with $s^* \in [\underline{s}, \overline{s}]$. By Theorem 4.1 there is an iteration index j_0 and a finite set $\Lambda \subset \mathbf{N}$ such that $u_k^{n_j} = 0$ for $j \geq j_0$ and $k \notin \Lambda$. Hence, we have for the finitely many $k \in \Lambda$ that $u_k^{n_j} \to u_k^*$ as $j \to \infty$, and infinitely often $u_k^{n_j} = u_k^* = 0$ (for $j > j_0$), which shows strong convergence of (u^{n_j}) to u^*. The above argumentation holds true for every subsequence of (u^n). Finally, due to Proposition 2.3, it holds that $u^* \in G_{\alpha,s^*}(u^*)$. □

Note that similar arguments have been used in [22] for $\phi := \chi_{\mathbb{R}\setminus\{0\}}$ in the finite dimensional case.

In general, we do not know whether this fixed point is a global minimizer. This is, of course, due to the non-convexity of T_α, where necessary conditions for optimality, which are based on differentiation, are in general not sufficient. However, the step-sizes s_n being bounded away from the critical constant L^{-1} encourage the iteration converging to local minima: Indeed, in Section 6.3 we show a (numerical) example in which non-critical step-sizes produce non-global solutions.

4.2 Necessary Conditions for Global Minimizers

Now we address the question to which extent global minimizers of T_α can be identified. We start with the following necessary condition.

Proposition 4.1 (Necessary conditions for global minimizers) *Let R be according to* (9) *and let ϕ fulfill Assumption 3.1. For $\lambda_{L^{-1}}$ and $\tau_{L^{-1}}$ according to* (8)*, each global minimizer u^* of T_α possesses the following properties:*

(a) The set $\Lambda := \{k \in \mathbb{N} \ : \ u_k^* \neq 0\}$ is finite,
(b) for all $k \in \Lambda$ it holds that $-S'(u^*)_k = \alpha \,\mathrm{sgn}(u_k^*)\phi'(|u_k^*|)$ and $|u_k^*| \geq \lambda_{L^{-1}}$,
(c) for all $k \notin \Lambda$ it holds that $|S'(u^*)_k| \leq L\tau_{L^{-1}}$.

Proof Suppose $u^* \in \ell^2$ is a global minimizer of T_α. Obviously, $-S'(u^*) \in \ell^2$. Exploiting the condition that $\lim_{x\downarrow 0} \phi'(x) = \infty$, we are able to find an $\varepsilon > 0$ such that $|\phi'(x)| \geq 1$ for $0 < x \leq \varepsilon$. Take a k for which $|u_k^*| \leq \varepsilon$ and $u_k^* \neq 0$, define $(T_\alpha)_k : \mathbb{R} \to \mathbb{R}$ by $(T_\alpha)_k(x) := T_\alpha(u^{k*}(x))$, where $((u^{k*}(x))_k := x$ and $(u^{k*}(x))_l := u_l^*$ if $l \neq k$, and differentiate to get

$$S'(u^*)_k + \alpha \operatorname{sgn}(u_k^*)\phi'(|u_k^*|) = 0 \quad \Rightarrow \quad |S'(u^*)|_k = \alpha\phi'(|u_k^*|) \geq \alpha.$$

Consequently, $\alpha^{-2}\left\|S'(u^*)\right\|^2 \geq \#\{k \in \mathbb{N} \mid |u_k^*| \leq \varepsilon \wedge u_k^* \neq 0\}$, and since the norm is finite as well as only finitely many u_k^* can satisfy $|u_k^*| > \varepsilon$, $\Lambda := \{k \in \mathbb{N} \; : \; u_k^* \neq 0\}$ has to be finite. This proves the first assertion as well as the first part of the second.

Next, we will show that, if $u_k^* \neq 0$ and $|u_k^*| < \lambda_{L^{-1}}$ for some k, setting u_k^* to zero strictly decreases the functional value of T_α and hence, u^* was not a global minimizer. Let $v_k^* := 0$ and $v_l^* := u_l^*$ for $l \neq k$ such that $v^* - u^* = -u_k^* e_k$. Expand S at u^* and compare $T_\alpha(v^*)$ with $T_\alpha(u^*)$:

$$\begin{aligned}
T_\alpha(u^*) - T_\alpha(v^*) &= \alpha \sum_{l=1}^{\infty} \phi(|u_l^*|) - \alpha \sum_{l \neq k} \phi(|u_l^*|) - \big\langle S'(u^*) \big| \, v^* - u^* \big\rangle \\
&\quad - \int_0^1 \big\langle S'\big(u^* + t(v^* - u^*)\big) - S'(u^*) \big| \, v^* - u^* \big\rangle \, \mathrm{d}t \\
&\geq u_k^* S'(u^*)_k - L\frac{(u_k^*)^2}{2} + \alpha\phi(|u_k^*|) \\
&= \frac{\alpha(u_k^*)^2}{2}\Big(\frac{2\big(\phi(|u_k^*|) - |u_k^*|\phi'(|u_k^*|)\big)}{(u_k^*)^2} - (L^{-1}\alpha)^{-1}\Big),
\end{aligned}$$

remembering that $u_k^* \neq 0$ implies $S'(u^*)_k = -\alpha \operatorname{sgn}(u_k^*)\phi'(|u_k^*|)$. The term on the right-hand side is positive if and only if $|u_k^*| < \psi^{-1}\big((L^{-1}\alpha)^{-1}\big) = \lambda_{L^{-1}}$ (also confer Lemma 3.2), thus $|u_k^*| < \lambda_{L^{-1}}$ implies $T_\alpha(v^*) < T_\alpha(u^*)$. Hence, for u^* being a global minimizer, it is necessary that $|u_k^*| \geq \lambda_{L^{-1}}$ for each $k \in \Lambda$, concluding the proof of the second assertion.

For the remainder, we will utilize a similar line of argumentation. Take a $k \notin \Lambda$, i.e. $u_k^* = 0$ and denote by $w_k^* := -S'(u^*)_k$. Our interest is in finding sufficient criteria which lead to a functional descent by just varying the k-th component. Let $v_l^* := u_l^*$ for $l \neq k$ and $v_k^* \in \mathbb{R}$. With this, we get, similarly to the above,

$$\begin{aligned}
T_\alpha(u^*) - T_\alpha(v^*) &\geq w_k^* v_k^* - L\frac{(v_k^*)^2}{2} - \alpha\phi(|v_k^*|) \\
&= -L\Big(\frac{1}{2}\Big(v_k^* - \frac{w_k^*}{L}\Big)^2 + \frac{\alpha}{L}\phi(|v_k^*|)\Big) + \frac{(w_k^*)^2}{2L}. \qquad (11)
\end{aligned}$$

Maximizing the right-hand side with respect to v_k^* gives $v_k^* = \operatorname{prox}_{L^{-1}\alpha\phi}\big(w_k^* L^{-1}\big)$ which is only non-zero if $|w_k^*| \geq L\tau_{L^{-1}}$, see Lemma 3.3. Moreover, the solutions of (7) are only non-unique in $\pm\tau_{L^{-1}}$, so

$$|w_k^*| > L\tau_{L^{-1}} \;\; \Rightarrow \;\; \frac{1}{2}\Big(v_k^* - \frac{w_k^*}{L}\Big)^2 + \frac{\alpha}{L}\phi(|v_k^*|) < \frac{(w_k^*)^2}{2L^2},$$

and plugging this into (11) yields $T_\alpha(u^*) - T_\alpha(v^*) > 0$. Consequently, if $|w_k^*| > L\tau_{L^{-1}}$ for a $u_k^* = 0$, then u^* was not a global minimizer what was to show. □

Note that the thresholds (8) for $s = L^{-1}$ play an important role for global optimality as they occur in the necessary conditions. The properties (a)–(c) moreover tell us what is necessary for the GGPM to be able to detect global minimizers.

Definition 4.1 (Quasi-global minimizers) A point $u^* \in \ell^2$, which fulfills the conditions (a)–(c) in Proposition 4.1, is called a *quasi-global minimizer* of T_α.

Remark 4.1 In order to get consistent results, we use, in the sequel, a single-valued selection of $G_{\alpha,s}$:

$$G^+_{\alpha,s}(u)_k := \begin{cases} 0, & \text{if } u_k \neq 0, |u_k - sS'(u)_k| < \tau_s \\ 0, & \text{if } u_k = 0, |u_k - sS'(u)_k| \leq \tau_s \\ \rho_s^{-1}(u_k - sS'(u)_k), & \text{if } u_k \neq 0, |u_k - sS'(u)_k| \geq \tau_s \\ \rho_s^{-1}(u_k - sS'(u)_k), & \text{if } u_k = 0, |u_k - sS'(u)_k| > \tau_s \end{cases}, \tag{12}$$

where we defined $\rho_s^{-1}(x) := -\rho_s^{-1}(-x)$ for $x \leq -\tau_s$ (and hence we have for all $|y| \geq \lambda_s$ that $\rho_s^{-1}\big(y + s\alpha\,\mathrm{sign}(y)\phi'(|y|)\big) = y$). The rule (12) basically tells in the ambiguous cases to select 0 whenever $u_k = 0$, and the non-zero value of $\mathrm{prox}_{s\alpha\phi}$ otherwise.

Proposition 4.2 (Global minimizers need critical step-sizes) *A global minimizer u^* for the Tikhonov functional T_α is a fixed point of $G^+_{\alpha,L^{-1}}$. Conversely, each fixed point of this mapping is a quasi-global minimizer of T_α.*

Proof Let u^* be a minimizer and denote again by Λ the set of indices, where $u^*_k \neq 0$. If $k \in \Lambda$, then $-L^{-1}S'(u^*)_k = L^{-1}\alpha\,\mathrm{sgn}(u^*_k)\phi'(|u^*_k|)$ as well as $|u^*_k| \geq \lambda_{L^{-1}}$; hence

$$\big|u^*_k - L^{-1}S'(u^*)_k\big| = \big|u^*_k + L^{-1}\alpha\,\mathrm{sgn}(u^*_k)\phi'(|u^*_k|)\big| = \rho_{\frac{1}{L}}\big(|u^*_k|\big) \geq \rho_{\frac{1}{L}}(\lambda_{\frac{1}{L}}) = \tau_{\frac{1}{L}}$$

as $\rho_{L^{-1}}$ is strictly monotonically increasing on $[\lambda_{L^{-1}}, \infty[$ (also see Remark 3.1). Consequently, by the particular single-valued selection $v^* := G^+_{\alpha,L^{-1}}(u^*)$, according to (12) always yields the non-zero value, i.e.,

$$v^*_k = \rho^{-1}_{L^{-1}}\big(u^*_k - L^{-1}S'(u^*)_k\big) = \rho^{-1}_{L^{-1}}\big(u^*_k + L^{-1}\alpha\,\mathrm{sign}(u^*_k)\phi'(|u^*_k|)\big) = u^*_k$$

by the definition of $\rho_{L^{-1}}$. Take a $k \notin \Lambda$ and observe that $|L^{-1}S'(u^*)_k| \leq \tau_{L^{-1}}$ implies $G^+_{\alpha,L^{-1}}\big(u^* - L^{-1}\alpha S'(u^*)\big)_k = 0$ again by the single-valued selection (12). Hence $v^*_k = u^*_k = 0$, and consequently, u^* is a fixed point.

Now suppose that u^* obeys $u^* = G^+_{\alpha,L^{-1}}(u^*)$. Obviously, u^* has only finitely many non-zero coefficients, meaning that property (a) of Proposition 4.1 is satisfied. For $u^*_k \neq 0$ we have

$$u^*_k - L^{-1}S'(u^*)_k = u^*_k + L^{-1}\alpha\,\mathrm{sign}(u^*_k)\phi'(|u^*_k|),$$

thus $S'(u^*)_k + \alpha\,\mathrm{sign}(u^*_k)\phi'(|u^*_k|) = 0$, and since the range of $\mathrm{prox}_{L^{-1}\alpha\phi}$ is the set $\{0\} \cup \{|x| \geq \lambda_{L^{-1}}\}$, see Proposition 3.3, we also have property (b). Finally, $u^*_k = 0$ means that $L^{-1}|S'(u^*)_k| \leq \tau_{L^{-1}}$, so $|S'(u^*)_k| \leq L\tau_{L^{-1}}$, implying that u^* also obeys property (c). □

These results suggest that the GGPM should actually look for fixed points of $G_{\alpha,L^{-1}}$. However, taking L^{-1} as a step-size does not give strict descent of T_α (cf. Prop. 4.3). However, with step-sizes carefully approaching the critical constant one can still subsequential convergence to a fixed point of $G^+_{\alpha,L^{-1}}$.

4.3 Step-sizes Approaching the Critical Constant

Motivated by the findings in the previous subsection, we now proceed with a case where the step-size sequence (s_n) approaches L^{-1}. Again, the main problem here is that $s = L^{-1}$ does not necessarily lead to functional descent in Proposition 2.1. Hence, we refine the Lipschitz continuity assumption for S' to obtain results in this setting.

Definition 4.2 (Refined Lipschitz continuity) Let $K : \ell^2 \to \ell^2$ be Lipschitz continuous with constant $L > 0$. Then, the Lipschitz constant is *sharp in only finitely many dimensions*, iff there exists an orthogonal projection P onto a finite-dimensional space and an $L^* < L$ such that, with $Q := \mathrm{id} - P$, the estimate

$$\|K(u) - K(v)\|^2 \leq L^2 \|P(u-v)\|^2 + (L^*)^2 \|Q(u-v)\|^2 \tag{13}$$

is satisfied for each $u, v \in \ell^2$.

For instance, compact linear operators have a Lipschitz constant which is sharp in only finitely many dimensions. For concrete examples, we ask for the reader's patience till Section 5. With the help of Assumption 4.2, we get the following refinement of Proposition 2.1.

Proposition 4.3 (Refined descent for critical step-sizes) *Let S, R satisfy Assumption 2.1 with the Lipschitz constant of S' being sharp in only finitely many dimensions. Then, for each $u \in \ell^2$, $0 < s \leq L^{-1}$ and $v \in G_{\alpha,s}(u)$, the functional descent of T_α can be estimated by*

$$T_\alpha(v) \leq T_\alpha(u) - \frac{1}{2}\Big(\frac{1}{s} - L\Big) \|P(v-u)\|^2 - \frac{1}{2}\frac{sL^*+1}{sL+1}\Big(\frac{1}{s} - L^*\Big) \|Q(v-u)\|^2 . \tag{14}$$

Proof Using that $v \in G_{\alpha,s}(u)$ solves the minimization problems in (3) and that S' satisfies (13), we obtain for $v \neq u$:

$$\begin{aligned} T_\alpha(v) - T_\alpha(u) &\leq \int_0^1 \big\langle S'\big(u + t(v-u)\big) - S'(u) \big| \, v-u \big\rangle \, \mathrm{d}t - \frac{\|v-u\|^2}{2s} \\ &\leq \int_0^1 t\big(L^2 \|P(v-u)\|^2 + (L^*)^2 \|Q(v-u)\|^2\big)^{1/2} \, \mathrm{d}t \, \|v-u\| - \frac{\|v-u\|^2}{2s} \\ &= \frac{1}{2}\Big(\big(L^2 \|P(v-u)\|^2 + (L^*)^2 \|Q(v-u)\|^2\big)^{1/2} - \frac{\|v-u\|}{s}\Big) \|v-u\| . \end{aligned} \tag{15}$$

Note that $L^* \leq L \leq 1/s$, so the difference on the right-hand side is actually non-positive. By writing $\sqrt{a} - \sqrt{b} = (a-b)/(\sqrt{a} + \sqrt{b})$ and estimating, it becomes

$$\begin{aligned} &\big(L^2 \|P(v-u)\|^2 + (L^*)^2 \|Q(v-u)\|^2\big)^{1/2} - \frac{\|v-u\|}{s} \\ &\quad \leq \Big(L + \frac{1}{s}\Big)^{-1} \frac{\big(L^2 - s^{-2}\big) \|P(v-u)\|^2 + \big((L^*)^2 - s^{-2}\big) \|Q(v-u)\|^2}{\|v-u\|} . \end{aligned} \tag{16}$$

Combining (15) and (16), rewriting $(L^*)^2 - s^{-2} = (L^* + s^{-1})(L^* - s^{-1})$ and expanding finally gives (14). □

Proceeding similarly to Section 4.1 we would like to show that the support of the iterates u^n stay in a finite set even for $\overline{s} = L^{-1}$.

Lemma 4.1 (Uniform finite support for critical step-sizes) *Let S, R and ϕ satisfy Assumption 2.1,* (9) *and Assumption 3.1 with the Lipschitz constant of S' being sharp in only finitely many dimensions. For* (u^n)*, generated by iteration* (10) *with step-sizes*

$$0 < \underline{s} \leq s_n \leq \overline{s} := L^{-1},$$

the supports of all u^n for $n \geq 1$ are still contained in a finite set.

Proof First recall that, by Proposition 2.1, $(T_\alpha(u^n))$ is a non-increasing sequence. Hence, as T_α is coercive (as a consequence of Assumption 3.1, see the beginning of this section), (u^n) is bounded. In the sequel, we examine the bounded sequence (v^n) given by $v^n := u^{n+1} - u^n$.

Let P and Q be chosen according to Assumption 4.2. First, we like to show that, for each $\varepsilon > 0$ there exists a $k_0 \in \mathbb{N}$ such that for the truncation operator M_{k_0} defined by

$$(M_{k_0} u)_k := \begin{cases} 0, & k \leq k_0 \\ u_k, & k > k_0 \end{cases},$$

it holds that $\|M_{k_0} P v^n\| \leq \varepsilon$ for all n. For this purpose, choose an orthonormal basis $z_1, \ldots, z_m$ for $\operatorname{range} P$ such that $Pu = \sum_{j=1}^m \langle z_j \,|\, u \rangle z_j$. Then, for each $u \in \ell^2$,

$$\|M_{k_0} P u\| \leq \Big(\sum_{j=1}^m \|M_{k_0} z_j\| \Big) \|u\| \,.$$

Choosing k_0 large enough, one can achieve that the sum becomes arbitrarily small. In particular, as (v^n) is bounded, we can find for any given $\varepsilon > 0$ a $k_0 \in \mathbb{N}$, such that $\|M_{k_0} P v^n\| \leq \varepsilon$ for each n.

Next, we show that $Qv^n \to 0$ as $n \to \infty$. Assume the opposite, i.e., there exists a $c > 0$ such that $\|Q(u^{n+1} - u^n)\| \geq c$ for infinitely many n. For these n, Proposition 4.3 yields

$$\frac{c^2}{2} \frac{s_n L^* + 1}{s_n L + 1} \Big(\frac{1}{s_n} - L^* \Big) \leq T_\alpha(u^n) - T_\alpha(u^{n+1}),$$

which is a contradiction since the right-hand side tends to zero as $n \to 0$, while the left-hand side is bounded away from zero. Hence, $Qv^n \to 0$ as $n \to \infty$.

Setting $\varepsilon := \kappa_{\underline{s}}/2$, we are thus able to find a k_0 and n_0 such that

$$\|M_{k_0} v^n\| \leq \|M_{k_0} P v^n\| + \|Q v^n\| \leq \frac{\kappa_{\underline{s}}}{2} + \|Q v^n\| < \kappa_{\underline{s}}$$

for all $n \geq n_0$. Now, as in Theorem 4.1, if the supports of u^{n+1} and u^n differ for some index $k \geq k_0$, then $\|M_{k_0}(u^{n+1} - u^n)\| = \|M_{k_0} v^n\| \geq \kappa_{\underline{s}}$, hence $n < n_0$. Consequently, the support of u^n does only change finitely many times outside of $\{1, \ldots, k_0\}$. As $u^1, \ldots, u^{n_0}$ have finite support, it follows by induction that the support of all iterates u^n, $n \geq 1$ are contained in a finite set, which is the statement we wanted to show. □

The previous lemmas and propositions are the essential ingredients for proving subsequential convergence of the GGPM when (s_n) is monotonically increasing with limit L^{-1} in a suitable manner. This will be done in the following theorem.

Theorem 4.2 (Strong subsequential convergence to quasi-global minimizers) *Let S, R and ϕ satisfy Assumption 2.1,* (9) *and Assumption 3.1 with the Lipschitz constant of S' being sharp in only finitely many dimensions. If the GGPM is performed according to* (12) *with step-size choice*

$$s_n := \frac{n+1}{L(n+1)+1},$$

then (u^n) is relatively compact and possesses a strong accumulation point u, which is a quasi-global minimizer of T_α.

Proof Note that (u^n) is bounded and, by Lemma 4.1, contained in a finite-dimensional subspace, hence relatively compact. To show the existence of a strong accumulation point as stated, we claim that, for each $\varepsilon > 0$ and each $n_0 \in \mathbb{N}$, there exists a $n \geq n_0$ such that $\|u^{n+1} - u^n\| < \varepsilon$. Assume the opposite, which means that there is a $\varepsilon > 0$ and a n_0 such that $\|u^{n+1} - u^n\|^2 \geq \varepsilon^2$ for all $n \geq n_0$. According to Proposition 2.1,

$$\frac{\varepsilon^2}{2}\Big(\frac{1}{s_n} - L\Big) \leq T_\alpha(u^n) - T_\alpha(u^{n+1}),$$

which yields, by the particular choice of s_n, after summing up,

$$\sum_{n=n_0}^{n_1-1} \frac{1}{n+1} \leq \frac{2}{\varepsilon^2}\big(T_\alpha(u^{n_0}) - T_\alpha(u^{n_1})\big)$$

for any $n_1 > n_0$. Since $(T_\alpha(u^n))$ is a non-increasing sequence, the right-hand side is bounded as $n_1 \to \infty$, while the left-hand side is not, a contradiction.

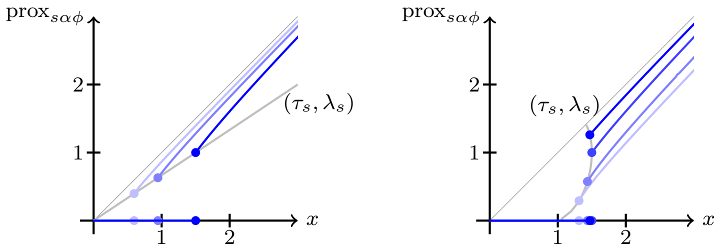
Hence, the above statement has to hold. Choosing $\varepsilon := \frac{1}{j}$ for each $j \in \mathbb{N}$, we can therefore construct a subsequence for which $\|u^{n_j+1} - u^{n_j}\| \to 0$ as $j \to \infty$. As (u^{n_j}) is still bounded and the supports of u^{n_j} are contained in a finite set (confer Lemma 4.1), there exists a strongly convergent subsequence. Hence, by further restricting (u^{n_j}), we can achieve that also $u^{n_j} \to u$ for some $u \in \ell^2$.

Next we show that $u^{n_j+1} = G^+_{\alpha,s_{n_j}}(u^{n_j}) \to G^+_{\alpha,L^{-1}}(u)$ as $j \to \infty$. From the above immediately follows that $G^+_{\alpha,s_{n_j}}(u^{n_j}) \to u$, hence it suffices to show that also $G^+_{\alpha,s_{n_j}}(u^{n_j})_k \to G^+_{\alpha,L^{-1}}(u)_k$ for each k. Since S' is Lipschitz continuous, it holds that $u^{n_j} - s_{n_j}S'(u^{n_j}) \to u - L^{-1}S'(u)$. In particular, $u^{n_j}_k - s_{n_j}S'(u^{n_j})_k \to u_k - L^{-1}S'(u)_k$ for each k. If, for a fixed k, it holds that $|u_k - L^{-1}S'(u)_k| \neq \tau_{L^{-1}}$, then

$$\begin{aligned} G^+_{\alpha,s_{n_j}}(u^{n_j})_k &= \operatorname{prox}_{s_{n_j}\alpha\phi}\big(u^{n_j}_k - s_{n_j}S'(u^{n_j})_k\big) \\ &\to \operatorname{prox}_{L^{-1}\alpha\phi}\big(u_k - L^{-1}S'(u)_k\big) = G^+_{\alpha,L^{-1}}(u)_k \end{aligned}$$

by Lemma 3.4. If $|u_k - L^{-1}S'(u)_k| = \tau_{L^{-1}}$, then Lemma 3.4 and the fact that $G^+_{\alpha,s_{n_j}}(u^{n_j})_k = u^{n_j+1}_k \to u_k$ give either $u_k = 0$ or $u_k = \rho^{-1}_{L^{-1}}\big(u^k - L^{-1}S'(u)_k\big) \neq 0$. In both cases, $u_k = G^+_{\alpha,L^{-1}}(u)_k$ by the definition (12).

Thus, $u = G^+_{\alpha,L^{-1}}(u)$ from which Proposition 4.2 gives the desired properties. □



$p = 1/2$ fixed, $s \in \{1/4,\ 1/2,\ 1\}$

$s = 1$ fixed, $p \in \{7/8,\ 3/4,\ 1/2,\ 1/4\}$

Fig. 3 Thresholding functions $\mathrm{prox}_{s\alpha\phi}$ for $\phi(x) := |x|^p$ according to Lemma 5.1 for $\alpha = 1$ and different values of s and p.

To summarize, this means that one can design an algorithm for which at least one subsequential limit u shares the same necessary conditions as the global minimizer u^*. Hence, if Theorem 4.2 is applicable, one can argue that the GGPM produces reasonable approximations to a minimizer of T_α.

5 Application to ℓ^p Penalties

This section deals with the special case $\phi(x) := |x|^p$ with $0 < p < 1$, which is clearly a non-convex penalty. First, it is analyzed how the proximity operator can be calculated (cf. [23, 24]). This result can be easily derived from the general statements of Lemma 3.3.

Lemma 5.1 (The proximal mapping for ℓ^p penalties) *Let $0 < p < 1$ and $\alpha > 0$. The function $\phi(x) := |x|^p$ satisfies Assumption 3.1. The corresponding $\mathrm{prox}_{s\alpha\phi}$ and thresholds, according to (8), can be expressed by (slightly abusing notation)*

$$\mathrm{prox}_{s\alpha\phi}(x) = \begin{cases} 0, & \text{for } |x| \le \tau_s \\ \big(\,\cdot\, + s\alpha p \,\mathrm{sign}(\,\cdot\,)|\,\cdot\,|^{p-1}\big)^{-1}(x), & \text{for } |x| \ge \tau_s \end{cases},$$

$$\tau_s := \frac{2-p}{2-2p}\big(2s\alpha(1-p)\big)^{\frac{1}{2-p}} \quad , \quad \lambda_s := \big(2s\alpha(1-p)\big)^{\frac{1}{2-p}}.$$

Proof All the statements can easily be verified by computation. Assumption 3.1(d)–(e) is almost trivially fulfilled. Regarding Assumption 3.1(a)–(c), note that ϕ is arbitrarily differentiable on $]0,\infty[$. The function ϕ' is strictly convex because of $\phi''' > 0$ for positive arguments. We have $\phi'(x) \to \infty$ as $x \to 0$, as well as $\phi'(x)/x = px^{p-2} \to 0$ as $x \to \infty$. Moreover, $x\phi''(x) = p(p-1)x^{p-1}$ is easily seen to be locally integrable on $[0,\infty[$. Clearly, $\psi(y) = 2(1-p)y^{p-2}$, so one computes $\lambda_s = \big(2s\alpha(1-p)\big)^{1/(2-p)}$, and consequently

$$\tau_s = \rho_s(\lambda_s) = \lambda_s + s\alpha p \lambda_s^{p-2}\lambda_s = \Big(1 + \frac{p}{2-2p}\Big)\lambda_s = \frac{2-p}{2-2p}\big(2s\alpha(1-p)\big)^{\frac{1}{2-p}}.$$

□

The functions $\operatorname{prox}_{s\alpha\phi}$ perform a thresholding, and for p varying from 1 to 0 they evolve from soft to hard thresholding as already noted in [23, 24], see Figure 3. With Proposition 4.1 in mind, we also get the following.

Remark 5.1 If $S : \ell^2 \to [0, \infty[$, one can immediately estimate the number of non-zero coefficients for a minimizer u^* of T_α:

$$S(0) = T_\alpha(0) \geq T_\alpha(u^*) \geq \alpha \sum_{k=1}^{\infty} |u_k^*|^p \geq \#\{u_k^* \neq 0\}\alpha\big(2L^{-1}\alpha(1-p)\big)^{p/(2-p)}$$

implies

$$\#\{u_k^* \neq 0\} \leq 2S(0)\big((2\alpha)^{2/p}L^{-1}(1-p)\big)^{p/(p-2)} .$$

The above shows that the GGPM is applicable for ℓ^p-penalties with S, according to Assumption 2.1 (a), and leads to convergence whenever one of the Theorems 4.1 or 4.2 is applicable. Roughly speaking, without additional assumptions we only get subsequential convergence to stationary points. In some situations, where more about the structure of the problem is known, one can actually see that the additional assumptions introduced in Section 4 are satisfied, and convergence can be established. One class for which this is possible is the usual setting of Tikhonov regularization for the solution of linear inverse problems. In the following, let, in addition to $0 < p < 1$, H be a Hilbert space, $K : \ell^2 \to H$ be linear and continuous, $f \in H$ be some given data, which is usually noise-perturbed, and $\alpha > 0$ a regularization parameter. Set

$$S(u) := \frac{\|Ku - f\|_H^2}{2}, \qquad R(u) := \|u\|_p^p = \sum_{k=1}^{\infty} |u_k|^p \tag{17}$$

and consider again

$$\min_{u \in \ell^2} \; T_\alpha(u) \quad \text{with} \quad T_\alpha(u) := S(u) + \alpha R(u),$$

which now amounts to Tikhonov regularization. Note that $S'(u) = K^*(Ku - f)$ is Lipschitz continuous. Clearly, the Lipschitz constant is given by $\|K^*K\|$, and Theorem 4.1 is applicable with $0 < \underline{s} \leq s_n \leq \overline{s} < \|K^*K\|$, yielding subsequential convergence to stationary points. We moreover observe the following:

Lemma 5.2 (Linear operators with sharp Lipschitz constant in only finitely many dimensions) *Consider S according to (17), with $K : \ell^2 \to H$ being bounded and continuous. If the space, for which $K^*Ku = \|K^*K\|\, u$ holds, is finite-dimensional and $\|K^*K\|$ is not an accumulation point of the eigenvalues of K^*K, then S', according to $S'(u) = K^*(Ku - f)$, has a Lipschitz constant which is sharp in only finitely many dimensions.*

Proof The statement follows immediately by setting P as the projection onto the eigenspace for $\|K^*K\|$ and noting that $K^*K = \|K^*K\|\, P + QK^*KQ$, where $\|QK^*KQ\| < \|K^*K\|$. □

Algorithm 1 GGPM for Tikhonov regularization with ℓ^p-constraints.

Input: $K : \ell^2 \to H$, $f \in H$, $\alpha > 0$, $0 < p < 1$.
Initialize u^0 with $\sum_k |u_k^0|^p < \infty$, set $L \leftarrow \|K^*K\|$, $n = 1$.
repeat
  Set $s_n \leftarrow (n+1)/(L(n+1)+1)$.
  Calculate threshold

$$\tau_n := \frac{2-p}{2-2p}\Big(2s_n\alpha(1-p)\Big)^{\frac{1}{2-p}}.$$

  Calculate $v^n = u^n - s_n K^*(Ku^n - f)$.
  for k **do**
    if $(|v_k^n| \leq \tau_n \wedge u_k^n = 0) \vee (|v_k^n| < \tau_n \wedge u_k^n \neq 0)$ **then**
      Set $u_k^{n+1} \leftarrow 0$.
    else
      Set $u_k^{n+1} \leftarrow \big(\,\cdot\, + s_n\alpha p \operatorname{sign}(\,\cdot\,)|\cdot|^{p-1}\big)^{-1}(v_k^n)$.
    end if
  end for
  Set $n \leftarrow n+1$.
until stopping criterion fulfilled

In particular, Theorem 4.2 is applicable for compact operators K and hence, we have strong subsequential convergence to a fixed point even in the case of $s_n := (n+1)/(L(n+1)+1)$. The GGPM for minimizing T_α with (17) in this situation is summarized as an abstract method in Algorithm 1.

In the remainder of this section, we show that, for rational $p \in \,]0,1[$ and K satisfying the so-called "finite-basis injectivity" (FBI) property [8], all accumulation points for the iterates generated by GGPM are isolated. This will imply the convergence of the whole sequence. The proof bases on algebraic considerations for the stationary points.

Proposition 5.1 (Discreteness of quasi-global minimizers for rational p) *Let R and S according to* (17) *with K satisfying the FBI property, i.e., K is injective whenever restricted to finitely many coefficients. Furthermore, let p be rational, i.e., $p := q/r$ with $q, r \in \mathbb{N}$ as well as $r > q \geq 1$.*

Then, the set of quasi-global minimizers of T_α is discrete.

Proof We split the proof of this proposition into several steps. First, denote by U the set of quasi-global minimizers according to Definition 4.1, and denote by $\operatorname{supp}(v) := \{k \in \mathbf{N} |\, v_k \neq 0\}$ for $v \in U$. The first claim is that there exists an $\varepsilon > 0$ such that for all $u, v \in U$ with $\|u - v\| < \varepsilon$, we have $\operatorname{supp}(u) = \operatorname{supp}(v)$. Suppose that for some $u, v \in U$ it holds that $\operatorname{supp}(v) \neq \operatorname{supp}(u)$. Then, there exists a $k \in \mathbf{N}$ such that either $u_k \neq 0$ and $v_k = 0$, or $u_k = 0$ and $v_k \neq 0$. Without any loss of generality, let $v_k = 0$. From condition (b) for quasi-global minimality (see Proposition 4.1) it follows that $|u_k - v_k| = |u_k| \geq \lambda_{L^{-1}}$. Therefore $\|u - v\| \geq \lambda_{L^{-1}}$ and the statement follows with $\varepsilon = \lambda_{L^{-1}}$.

Let now $I \subset \mathbf{N}$ be an arbitrary finite set. We show that $\{u \in U \ : \ \operatorname{supp}(u) = I\}$ is a finite set. Together with the above, the statements of the proposition will then follow. Since I is finite, $\operatorname{span}(I) := \operatorname{span}\{e_k \ : \ k \in I\} \cong \mathbf{R}^N$, where $N = \#I$. Therefore, there exists a matrix representation $K_I \in \mathbf{R}^{N \times N}$ of the operator K restricted to $\operatorname{span}(I)$. Due to the the FBI property, K is injective on $\operatorname{span}(I)$, and consequently, K_I is a regular matrix. This implies that $K_I^* K_I$ is a symmetric and positive definite matrix. Without any loss of generality we assume from now on that $I = \{1, \ldots, N\}$, $u = (u_1, \ldots u_N) \in \mathbf{R}^N$, $f = (f_1, \ldots f_N)$, and $K = K_I$. Again by condition (b) for quasi-global minimality, every $u \in U$ has to satisfy the system of equations

$$u_k \neq 0 \quad \text{and} \quad -S'(u)_k = \alpha \operatorname{sign}(u_k) p |u_k|^{p-1} \qquad \text{for } k \in I.$$

Since $S'(u) = K^*(Ku - f)$ we get the system of equations

$$u_k \neq 0 \quad \text{and} \quad \big(K^*(Ku - f)\big)_k + \alpha \operatorname{sign}(u_k) p |u_k|^{p-1} = 0, \tag{18}$$

for all $k \in I$. Now fix a sign pattern of u, i.e., let $\sigma \in \{-1, 1\}^N$ be given. Letting $v \in \mathbf{C}^N$,

$$t_k = \begin{cases} \alpha p, & \text{if } \sigma_k = 1 \\ -\alpha p \mathrm{e}^{-\frac{\mathrm{i}\pi(q-r)}{r}}, & \text{if } \sigma_k = -1, \end{cases}$$

and writing $v^r := (v_1^r, \ldots, v_N^r)$ as well as $t := (t_1, \ldots, t_N)$, we consider the system of complex equations

$$K^*(Kv^r - f) + \operatorname{diag}(t) v^{q-r} = 0, \qquad \text{for } v \in (\mathbf{C} \setminus \{0\})^N. \tag{19}$$

Each u solving (18) with $\operatorname{sign}(u) = \sigma$ gives a solution to (19) via

$$v_k = \begin{cases} u_k^{1/r}, & \text{for } u_k > 0 \\ (-u_k)^{1/r} \mathrm{e}^{\frac{\mathrm{i}\pi}{r}}, & \text{for } u_k < 0. \end{cases}$$

If there are only finitely many $v \in (\mathbf{C} \setminus \{0\})^N$ solving (19), then, by injectivity of $u \mapsto v$ on $(\mathbb{R}\backslash\{0\})^N \to (\mathbf{C}\backslash\{0\})^N$, there are only finitely many u solving (18) with $\operatorname{sign}(u) = \sigma$. Consequently, as there are only finitely many $\sigma \in \{-1, 1\}^N$, the solution set of (18) has to be finite.

In order to show that (19) has finitely many solutions, we like to employ Lemma A.2. For this purpose, we multiply with $\operatorname{diag}(v^r)$ from the left to get the new system

$$\operatorname{diag}(v^r) K^* K v^r - \operatorname{diag}(K^* f) v^r + \operatorname{diag}(t) v^q = 0, \quad v \in \mathbf{C}^N,$$

whose solutions are a superset of the solutions of (19). After substituting $X = v^r$ and $a = v^q$ and setting $A = K^*K$, $b = -\operatorname{diag}(K^*f)$, $c = 0$ and $d = t$, we arrive at the system

$$\begin{aligned} \operatorname{diag}(X) A X + \operatorname{diag}(X) b + \operatorname{diag}(X) \operatorname{diag}(a) + \operatorname{diag}(a) d = 0, \\ a^r = X^q, \end{aligned}$$

which fulfills all conditions of Lemma A.2, and hence has only finitely many solution vectors (X, a). Now observe that there are at most q^N possible $v \in \mathbf{C}^N$, which are mapped to one $(X, a) \in \mathbf{C}^{2N}$ by $X = v^r$ and $a = v^q$ (which is rather a rough estimate). Thus (19) has only finitely many solutions, and consequently, (18) has also only finitely many solutions. □

Collecting the results, we finally have the following convergence statement.

Theorem 5.1 (Strong convergence to quasi-global minimizers for rational p) *If $p \in]0,1[$ is rational and $K : \ell^2 \to H$ is linear, continuous and such that*

(a) the set $\{u \in \ell^2 \ : \ K^*Ku = \|K^*Ku\|\}$ is a finite-dimensional subspace,
(b) the eigenvalues of K^*K do not accumulate at $\|K^*K\|$ and
(c) K satisfies the FBI property,

then each sequence (u^n) generated by Algorithm 1 converges to a quasi-global minimizer of T_α.

Proof Relative compactness of (u^n) and subsequential convergence to a quasi-global minimizer follows immediately from Theorem 4.2. According to Proposition 5.1, the latter form a discrete set in ℓ^2, which implies that u^* is isolated. Therefore, the whole sequence converges to u^* by Proposition 2.3. □

Note that the restriction that p has to be rational might appear unsatisfactory at the first glance. From the practical point of view, however, machine numbers are most commonly represented with finite precision and rational by definition. Hence, the restriction does actually not matter in numerical computations.

Remark 5.2 (ℓ^0 penalties) We remark on the (not covered) limiting case $p = 0$, i.e.

$$\min_{u\in\ell^2} \frac{\|Ku-f\|^2}{2} + \alpha\#\{u_k \neq 0\}. \tag{20}$$

In fact, this problem is not well-posed for general $\alpha > 0$ due to the lack of coercivity of the functional $\#\{u_k \neq 0\}$. Indeed, one can construct a compact linear operator such that the minimization problem (20) admits no solution: For a given $\alpha > 0$ choose a $f \in \ell^2$ such that $\|f\|^2/2 > \alpha$, and the support of f is infinite. Denote by (f^n) the sequence according to

$$f_k^n := \begin{cases} f_k, & \text{if } k \leq n, \\ 0, & \text{else} \end{cases},$$

and define $K : \ell^2 \to \ell^2$ by $Ku := \sum_{n=1}^{\infty} \frac{u_n}{n} f^n$, which gives a linear, continuous and compact operator. With $u^n := ne_n$ we deduce

$$\inf_{u\in\ell^2} \frac{\|Ku-f\|^2}{2} + \alpha\#\{u_k \neq 0\} \leq \lim_{n\to\infty} \frac{\|Ku^n - f\|^2}{2} + \alpha = \alpha \ .$$

As the support of f is infinite, each $u = ve_k$ for some $v \neq 0$ and $k \in \mathbb{N}$ results in a $\|Ku-f\|^2/2 + \alpha > \alpha$. For $u = 0$ we get the functional value $\|f\|^2/2$, which is greater than α. This is obviously also true for each $u \in \ell^2$ with $\#\{u_k \neq 0\} \geq 2$. Hence, no minimizer exists, cf. [25] and [26].

Although the problem (20) generally admits no solution, one can try to apply the GGPM with quadratic fidelity and the separable penalty term

$$S(u) := \frac{\|Ku-f\|^2}{2} \quad , \quad R(u) := \alpha \sum_{k=1}^{\infty} \phi(|u_k|) \quad , \quad \phi(x) = \begin{cases} 0, & \text{if } x = 0 \\ 1, & \text{if } x \neq 0 \ . \end{cases}$$

One easily sees that the minimization problem for $\mathrm{prox}_{s\alpha R}$, i.e.,

$$\min_{v\in\ell^2} \sum_{k=1}^{\infty} \big(u_k - s\big(K^*(Ku-f)\big)_k\big)^2 + \alpha\phi(|u_k|),$$

can be solved by doing a hard-thresholding on $w = u - sK^*(Ku - f)$:

$$\operatorname{prox}_{s\alpha R}(w)_k = \begin{cases} 0, & \text{for } |w_k| \leq \sqrt{2s\alpha} \\ w_k, & \text{for } |w_k| \geq \sqrt{2s\alpha} \end{cases}.$$

This iteration has been studied in finite dimensions in [22].

Note that the structure of the iteration is the same as for the ℓ^p-penalties with $0 < p < 1$. Furthermore, with $0 < s < L^{-1}$, Proposition 2.1 is still applicable, and since hard-thresholding also induces a jump, the iterates (u^n) do not change the sign pattern from some index on, reducing the iteration to

$$u_k^{n+1} = u_k^n - s\big(K^*(Ku^n - f)\big)_k$$

for k being in some finite set. This amounts to a Landweber iteration in finite dimensions, which always converges.

Hence, the hard-thresholding operation of [22] still converges in the infinite-dimensional setting even though a global minimizer does not necessarily exist. The ℓ^p-setting where $0 < p < 1$ has the advantage that we can get both: well-posedness as well as convergence of the algorithm.

6 Numerical Examples

We illustrate the behavior of the proposed algorithm with two examples. The purpose of this section is to give an impression of the characteristics of the algorithm, especially since it behaves somehow discontinuous.

6.1 Reconstruction With a Partial DCT-Matrix

The purpose of this example is to demonstrate the typical behavior of the iterated thresholding algorithm on a simple example. We considered the finite dimensional problem of reconstruction of a spiky signal from partial discrete cosine transformed (DCT) measurements. We generated an operator by taking 64 random rows of a DCT matrix of size 256×256. We generated spiky data by randomly choosing ten entries to have normally distributed values, added 5% noise and chose $\alpha = 5 \cdot 10^{-4}$. Moreover, we chose to use the increasing step-size rule $s_n := (n+1)/(L(n+1)+1)$. Figure 4 illustrates the behavior of the iteration. We plotted the behavior of the functional value T_α, and also the norm difference between two iterates

$$r^n := \Big\| u^n - \operatorname{prox}_{s_n\alpha R}\big(u^n - s_n K^*(Ku^n - g^\delta)\big)\Big\| . \tag{21}$$

We observe that the functional value is monotonically decreasing, and from time to time it jumps down. This effect is due to the jump in the iteration mapping G_{α,s_n} and happens when one coefficient jumps out of or into the dead-zone of G_{α,s_n}. These jumps are also present in the plot of the residual r, and of course, they are smaller for larger p since the jump in G_{α,s_n} is smaller. Finally one observes that, from some point on, the residual decreases monotonically. This may be due to the fact that the support of the minimizer is identified and not changing anymore and the algorithm behaves like a usual gradient descent.

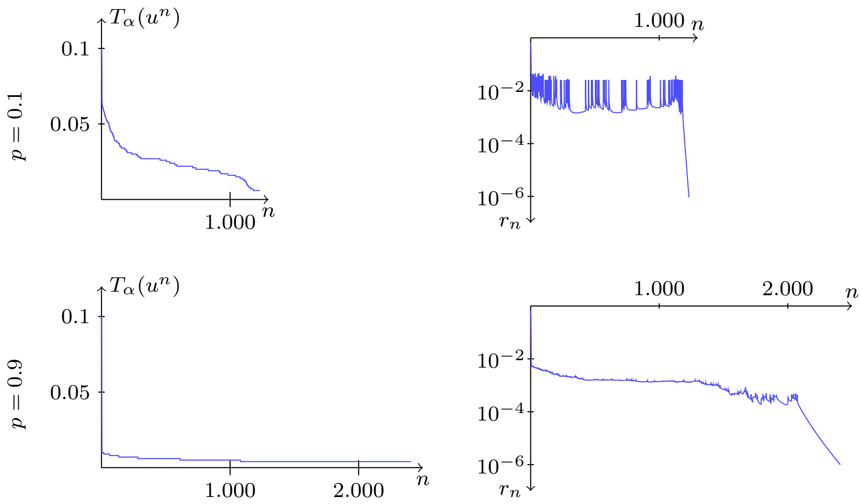


Fig. 4 Two typical runs for the partial-DCT example. Top row: $p = 0.1$, bottom row: $p = 0.9$. Left column: development of the objective value $T_\alpha(u^n)$, right column: development of the residual from (21).

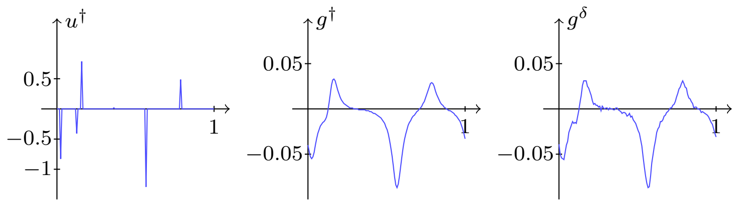


Fig. 5 Example with a smoothing operator. Left original solution $u^\dagger$, middle: original data $g^\dagger$, right: noisy data g^δ (5% noise).

6.2 Deblurring of Spikes

This purpose of the next example is to show the difference of the minimizers for different values of p. In this numerical example, we considered a discretized linear blurring operator F combined with a synthesis operator B associated with simple hat-functions. We generated a solution which consists of a few spikes and hence, has a sparse representation in hat functions. We generated noisy data with 5% noise, see Figure 5. Then, we chose $\alpha = 5 \cdot 10^{-5}$ and $p = 0.1$. Motivated by the previous example, we applied our algorithm until the norm of the residual r_n from (21) fell below the threshold $1 \cdot 10^{-9}$, leading to a reconstruction $u_p^{\alpha,\delta}$. To make a comparison with different values for p we calculated the discrepancy for $\left\| FBu_p^{\alpha,\delta} - g^\delta \right\|$ and chose α such that we obtained the same discrepancy for different values of p. The result is depicted in Figure 6. Concerning the properties of the solutions, one may note that smaller values of p lead to higher sparsity for the same discrepancy and to a more accurate reconstruction of the height of the peaks.

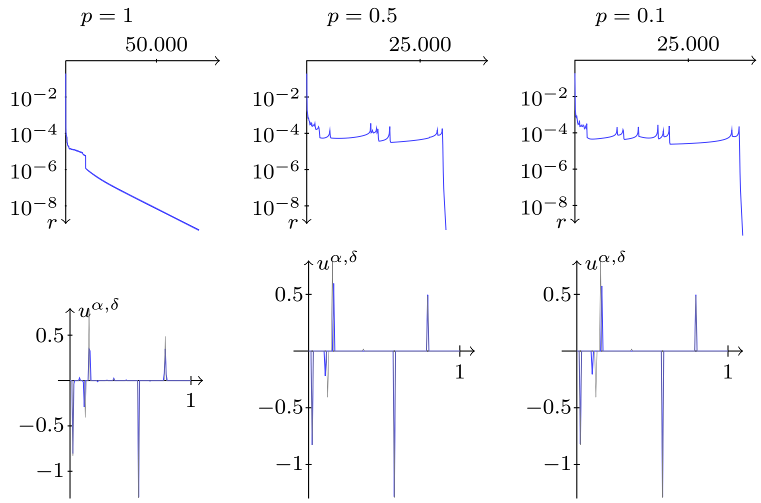


Fig. 6 Reconstructions for different values of p. The first row shows the norm of the residual, i.e. $\|u^n - \mathrm{prox}_{s\alpha R}(u^n - K^*(Ku^n - g^\delta)\|$. The second row shows the reconstructed $u^{\alpha,\delta}$ (and with slim lines the original $u^\dagger$).

6.3 Local Minimizers and Step-sizes

In this experiment, we used the same set-up as in Section 6.2. We created another spiky solution $u^\dagger$ and produced noisy data g^δ. Now we fixed $p = 0.2$, $\alpha = 2 \cdot 10^{-4}$ and ran GGPM with different step-sizes s_n. In the top-row of Figure 7 we show the progress of the objective function for stepsizes $s_n \equiv 0.2/L$, $s_n \equiv 0.5/L$ and $s_n := (n+1)/((n+1)L+1)$, respectively ($L = \|FB\|^2$). Note that the stepsizes that are bounded away from $1/L$ lead to larger objective values and hence, the method got stuck in a local minimizer. For the increasing stepsize, we obtain the lowest objective value and moreover, one that satisfies the additional necessary conditions for global minimizers (cf. Section 4.2). In the bottom row of Figure 7 we show the calculated minimizers $u_{0.2}$, $u_{0.5}$ and u_{inc}, corresponding to the different stepsizes. Note that indeed, the small stepsize misses several peaks and also that the minimizer for $s_n \equiv 0.5/L$ still misses one peak and also has several peaks at slightly shifted positions (compared to the peaks obtained with the increasing stepsize, cf. rightmost graph in Figure 7).

7 Concluding Remarks

We considered instances of non-smooth and non-convex minimization problems split into a smooth and a non-smooth, non-convex part with simple proximal mapping. Our analysis shows that, even in the general case of functionals, the proposed GGPM algorithm has convenient convergence properties. In the special case of separable constraints, our method amounts to an iterative thresholding procedure and is rather easy to implement. We remark that non-smooth and non-convex optimization problems are fairly hard to solve. Our algorithm gives strong subsequential convergence in the general separable case as well as strong

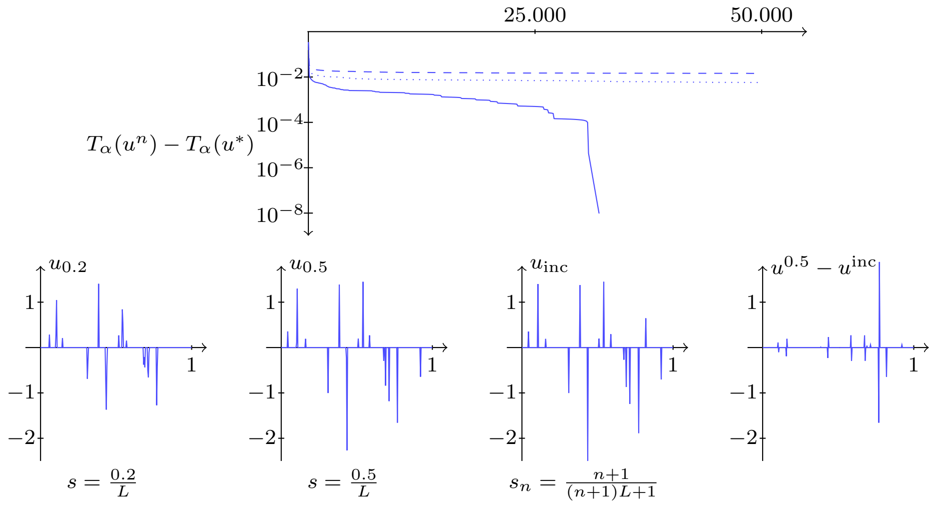


Fig. 7 Influence of the stepsize in GGPM. Top: Distance to the smallest objective value which could be produced for the different stepsizes ($s_n \equiv 0.2/L$ dashed, $s_n \equiv 0.5/L$ dotted, $s_n := (n+1)/((n+1)L+1)$, solid.) Bottom row: Corresponding local minimizers. Rightmost: Difference between the minimizers for $s_n := 0.5/L$ and $s_n := (n+1)/((n+1)L+1)$.

convergence if the smooth part is a squared discrepancy term with a compact operator satisfying the FBI property and non-convex sequence space quasi-norms penalties with rational exponent. Therefore, there is good reason to hope that it may reach a global minimizer in the special case of separable constraints. The question whether for irrational exponent it may happen that the minimizers are not discrete, is still open.

Finally, again for the latter special case, there is an approach based on iterative reweighting techniques [27] which also produces subsequences which converge to a critical point. However, one can show that these methods provably run into non-global minima if initialized wrong (e.g., if the operator equals the identity and too small or too large initialization) and no general strategy for the choice of initialization is available. This problem is partly circumvented in our method by step-sizes which approach the critical constant, cf. Section 4.3.

As possible generalizations of the GGPM one could imagine (in the spirit of [12]) to also allow for non-convex but smooth terms for the gradient step and it may even be possible to incorporate an inertial force like in the heavy ball method or an acceleration in Nesterov style (also, cf. [12]).

A Appendix

A.1 Proof of Lemma 3.2

As ϕ is non-decreasing, $\phi'(y) \geq 0$ for all $y > 0$, hence $\lim_{y\to\infty} \phi'(y) \geq 0$ (with ∞ allowed). Thus, each ρ_s is a strictly convex function with $\rho_s(y) \to \infty$ for $y \to 0$ and $y \to \infty$, which implies that a unique minimizer $y_s > 0$ exists.

Next, note that since ρ_s is strictly convex for some $s > 0$, it is locally Lipschitz. Consequently, ϕ' is locally Lipschitz and in particular absolutely continuous on compact intervals contained in $]0,\infty[$. This implies that ϕ' is differentiable almost everywhere in $]0,\infty[$, with $\phi'(y_1) - \phi'(y_0) = \int_{y_0}^{y_1} \phi''(t)\,\mathrm{d}t$ for each $y_1 > y_0 > 0$. Moreover, by assumption $t \mapsto t\phi''(t)$ is Lebesgue integrable on $[0,y]$ for each $y > 0$, hence Fubini's theorem is applicable. We compute

$$\phi(y) - y\phi'(y) = \phi(0) + \int_0^y (\phi'(\tau) - \phi'(y))\,\mathrm{d}\tau = -\int_0^y \int_\tau^y \phi''(t)\,\mathrm{d}t\,\mathrm{d}\tau = -\int_0^y t\phi''(t)\,\mathrm{d}t.$$

Since $(\phi')'_-$ is strictly monotonically increasing and coincides with ϕ'' where the latter exists, i.e., almost everywhere in $]0,\infty[$, we get for $y > 0$ and any $z \geq (\phi')'_-(y)$ that

$$-\int_0^y t\phi''(t)\,\mathrm{d}t = \int_0^y t(\phi')'_-(t)\,\mathrm{d}t > -z\int_0^y t\,\mathrm{d}t = -\frac{zy^2}{2}.$$

This already proves $\psi(y) > -z$. For ψ being strictly decreasing, consider $y > 0$ where $\phi''(y)$ exists and deduce

$$\psi'(y) = \frac{2}{y}\Big(\frac{2}{y^2}\int_0^y t\phi''(t)\,\mathrm{d}t - \phi''(y)\Big) < 0 \quad \Leftrightarrow \quad \psi(y) > -\phi''(y),$$

where the latter has already been established. Hence, $\psi' < 0$ almost everywhere in $]0,\infty[$ and as ψ is also locally Lipschitz in $]0,\infty[$, it is absolutely continuous in each compact interval contained in $]0,\infty[$. Consequently, one can conclude that ψ is strictly monotonically decreasing. Moreover, $\psi(y) > -(\phi')'_-(y)$ also implies $\psi(y) \to \infty$ as $y \to 0$ since $\phi'(y)$ is bounded around 0 whenever $(\phi')'_-(y)$ is bounded around 0. Finally, from the assumptions $\phi(y) \to \infty$ and $\phi'(y)/y \to 0$ as $y \to \infty$, it follows with L'Hôpital's rule that $\phi(y)/y^2 \to 0$. Consequently, $\psi(y) \to 0$ as $y \to \infty$ and, together with the above, $\psi :]0,\infty[\to]0,\infty[$ is onto. □

A.2 Tools from Algebraic Geometry

This section contains some technical lemmas for algebraic equations in order to prove Proposition 5.1. We make here use of the theory of Gröbner bases for solving algebraic equations.

We present the most important notations and definitions, which are needed for the theory here for the sake of convienience and in order to provide an uniform notation. A brief introduction to Gröbner bases can also be found in [28].

Definition A.1 (Some notations and definitions in polynomial rings)

- We denote by $\mathbf{P}$ the *polynomial ring* $\mathbf{C}[X_1,\ldots,X_n]$ in n variables with complex coefficients.
- A subgroup I of $\mathbf{P}$ with respect to $+$ is called an *ideal* of $\mathbf{P}$, iff for all $f \in \mathbf{P}$ and all $g \in I$ the product fg is also in I. We write $I \lhd \mathbf{P}$ if I is an ideal of $\mathbf{P}$.
- Let F be a subset of $\mathbf{P}$. The set
$$\langle F \rangle = \bigcap_{F \subset I \lhd \mathbf{P}} I$$
is called the *ideal generated by* F. It holds that
$$\langle f_1,\ldots,f_m \rangle = \Big\{\sum_{i=1}^m g_i f_i \ : \ f_i \in F,\ g_i \in \mathbf{P} \text{ for } i = 1,\ldots,m\Big\}.$$
- For $\alpha \in \mathbf{N}_0^n$ the power product X^α is defined by $X^\alpha := X_1^{\alpha_1}\cdots X_n^{\alpha_n}$.
- We denote with $\mathbf{T} \subset \mathbf{P}$ the set $\{X^\alpha \ : \ \alpha \in \mathbf{N}_0^n\}$ of *power products*. A power product X^α is *mixed* if $\alpha_i > 0$ for at least two i.
- The *lexicographical order* $<_{\text{lex}}$ of multindices $\alpha,\beta \in \mathbf{N}_0^n$ is defined as
$$\alpha <_{\text{lex}} \beta \quad \Leftrightarrow \quad \exists m > 0 : (\forall i < m : \alpha_i = \beta_i) \wedge \alpha_m < \beta_m.$$

- The *total degree lexicographical order* $\prec$ on the set $\mathbf{T}$ of power products is defined by

$$t_1 \prec t_2 \quad \Leftrightarrow \quad t_1 := X^\alpha, t_2 := X^\beta, \alpha, \beta \in \mathbf{N}_0^n, \\ |\alpha| < |\beta| \vee (|\alpha| = |\beta| \wedge \alpha <_{\text{lex}} \beta)$$

  with $|\alpha| = \sum_{j=1}^n \alpha_j$.
- For a polynomial $f \in \mathbf{P}$ we define the *coefficient* of the power product t of f as $C(f,t)$.
- We define the *support* of the polynomial f as $\text{supp}(f) := \{t \in \mathbf{T} \ : \ C(f,t) \neq 0\}$.
- The *degree* of the polynomial f is defined by $\deg(f) := \max\{|\alpha| \ : \ X^\alpha \in \text{supp}(f)\}$ while the *degree with respect to* X_i is given by $\deg_{X_i}(f) := \max\{\alpha_i \ : \ X^\alpha \in \text{supp}(f)\}$.
- The *leading powerproduct* of f is defined by $\text{LPP}(f) := \max_\prec\{t \ : \ t \in \text{supp}(f)\}$.
- The *trailing term* of f is the smallest powerproduct in $\text{supp}(f)$ with respect to $\prec$.

Definition A.2 (Varieties) Let $I \subset \mathbf{P}$ be a set of polynomials. Then the *variety* of I is defined as $V(I) := \{c \in \mathbf{C}^n \ : \ f(c) = 0 \text{ for all } f \in I\}$.

Definition A.3 Let $F \subset \mathbf{P}$. The *initial ideal* with respect to the order $\prec$ is defined by $\text{In}_\prec(F) := \langle\{\text{LPP}(f) \ : \ f \in F\}\rangle$.

Definition A.4 (Gröbner basis) Let $I \lhd \mathbf{P}$ be an ideal and $\prec$ the lexicographical total degree order. A finite $G \subset I$ such that $\text{In}_\prec(I) = \text{In}_\prec(G)$ is called a *Gröbner basis* for I with respect to $\prec$.

We call G a *reduced Gröbner basis* iff additionally

(i) each $g \in G$ is *normed*, i.e., the leading coefficient, $C(g, \text{LPP}(g))$, is 1,
(ii) the set $\{\text{LPP}(g) \ : \ g \in G\}$ minimally generates $\text{In}_\prec(I)$, i.e., for all $g \in G$ it holds that $\text{In}_\prec(G \setminus \{g\}) \neq \text{In}_\prec(G)$, and
(iii) no trailing term of any $g \in G$ lies in $\text{In}_\prec(I)$.

One can show that a reduced Gröbner basis exists for each ideal I and that it is unique. We will use the following characterization of Gröbner bases (the proof can be found in [29, Section 3.5] or [30, Theorem 2.4.1]).

Theorem A.1 *Let $I \lhd \mathbf{P}$ be an ideal and let $G \subset I$ be a finite set of normed polynomials. Then, the following statements are equivalent:*
(i) G is a Gröbner basis for I with respect to $\prec$.
(ii) It holds that $\{\text{LPP}(f) \ : \ f \in I\} = \{t \,\text{LPP}(g) \ : \ g \in G, \ t \in \mathbf{T}\}$.

Theorem A.2 is also from the theory of Gröbner bases, and gives a connection between the finiteness of varieties and the structure of the reduced Gröbner basis (the proof can be found in [29, Section 3.6], [30, Proposition 3.7.1]).

Theorem A.2 *Let $I \lhd \mathbf{P}$ be an ideal, and G the reduced Gröbner basis with respect to $\prec$. Then, the following statements are equivalent:*
(i) $V(I)$ is finite.
(ii) There exists $g_1, \ldots, g_n \in G$, and $m_1, \ldots, m_n \in \mathbf{N}_0$, such that $\text{LPP}(g_i) = X_i^{m_i}$.
(iii) The quotient space $\mathbf{P}/I$ is a finite dimensional $\mathbf{C}$ vector space.

Finally, we have all the tools to obtain the desired results about the solutions of algebraic equations. They are formulated in the following two lemmas.

Lemma A.1 *Let $n \in \mathbf{N}$, $A \in \mathbf{C}^{n\times n}$ be a Hermitian positive definite matrix. The system of quadratic equations* $\text{diag}(X)AX = 0$, *has only the trivial solution* $X = 0$.

Proof Let X be a solution of $\text{diag}(X)AX = 0$. Then, for each i, we have $X_i = 0$ or $(AX)_i = 0$. Hence, we can rearrange the indices $1, \ldots, n$ in such a way that for a $1 \leq k \leq n$, that $X_1, \ldots, X_k \neq 0$, and $X_{k+1}, \ldots, X_n = 0$. In order to exchange the indices i and j, we have to exchange the i-th and the j-th rows as well as the i-th and j-th columns of A. Since the eigenvalues of A are invariant under those operations, the resulting new matrix $\tilde{A}$ is also positive definite and Hermitian. Since $X_{k+1}, \ldots, X_n = 0$ we can eliminate the corresponding rows and columns of the system. Denoting the resulting matrix by $\tilde{\tilde{A}}$ and $\tilde{X} = (X_1, \ldots, X_k)$, it follows that

$$\tilde{\tilde{A}}\tilde{X} = 0.$$

From the Sylvester criterion we know that $\tilde{\tilde{A}}$ is positive definite as it is a submatrix of $\tilde{A}$. Therefore, $\tilde{X} = 0$, and hence, $X = 0$. □

Lemma A.2 *Let $A \in \mathbf{C}^{n\times n}$ be a positive definite, Hermitian matrix and $b, c, d \in \mathbf{C}^n$. Then, the system*

$$\operatorname{diag}(X)AX + \operatorname{diag}(X)b + \operatorname{diag}(X)\operatorname{diag}(a)c + \operatorname{diag}(a)d = 0,$$
$$a_i^r = X_i^q, \qquad i = 1, \ldots, n,$$

with $r > q$ has only finitely many solutions.

Proof Define the polynomials

$$f_i := f(X_1, \ldots, X_n) = X_i\Big(\sum_{j=1}^{n} A_{ij}X_j\Big) + X_i b_i + X_i a_i c_i + a_i d_i = 0$$

and

$$\tilde{f}_i = X_i\Big(\sum_{j=1}^{n} A_{ij}X_j\Big).$$

Furthermore set $\tilde{I} = \langle \tilde{f}_1, \ldots, \tilde{f}_n\rangle$.

Let $\tilde{G}$ be the reduced Gröbner basis for $\tilde{I}$ with respect to the total degree lexicographical order $\prec$. From Lemma A.1 and Theorem A.2 we know that there are $m_1, \ldots, m_n \in \mathbf{N}$ such that $\mathrm{LPP}(\tilde{g}_i) = X_i^{m_i}$ for $i = 1, \ldots, n$. Since $\tilde{G} \subseteq \tilde{I}$ there are $r_{ij} \in \mathbf{P}$, such that

$$\tilde{g}_i = \sum_{j=1}^{n} \tilde{f}_j r_{ij} \tag{22}$$

for all $i = 1, \ldots, n$. Let us now fix some $i \in \{1, \ldots, n\}$.

Next note that $\tilde{f}_i$ only contains quadratic power products and that $X_i^2 \in \operatorname{supp}(\tilde{f}_i)$ since $A_{ii} > 0$ by positive definiteness of A. Therefore, $\deg(\tilde{f}_i) = 2$. Now split each r_{ij} according to

$$s_{ij} := \sum_{\substack{t\in\operatorname{supp}(r_{ij}),\\ \deg(t)\le m_i-2}} C(r_{ij}, t)t \qquad\qquad t_{ij} := \sum_{\substack{t\in\operatorname{supp}(r_{ij}),\\ \deg(t)> m_i-2}} C(r_{ij}, t)t.$$

We claim that $\sum_{j=1}^{n} t_{ij}\tilde{f}_j = 0$. Assume the opposite and observe that for each $t \in \mathbf{T}$ with $\deg(t) > m_i - 2$ and each $j = 1, \ldots, n$ it follows that $C(t\tilde{f}_j, s) = 0$ for each $s \in \mathbf{T}$ with $\deg(s) \le m_i$. Consequently, $C(\sum_{j=1}^{n} t_{ij}\tilde{f}_j, s) = 0$ for each $s \in \mathbf{T}$ with $\deg(s) \le m_i$ implying that $\deg(\sum_{j=1}^{n} t_{ij}\tilde{f}_j) > m_i$ since $\sum_{j=1}^{n} t_{ij}\tilde{f}_j \neq 0$. But we also have that

$$m_i = \deg(\tilde{g}_i) = \deg\Big(\sum_j s_{ij}\tilde{f}_j + \sum_j t_{ij}\tilde{f}_j\Big),$$

and $\deg\Big(\sum_j s_{ij}\tilde{f}_j\Big) \le m_i$, from which the contradiction $\deg(\sum_{j=1}^{n} t_{ij}\tilde{f}_j) \le m_i$ follows. Hence, $\sum_{j=1}^{n} t_{ij}\tilde{f}_j = 0$. In particular, $\tilde{g}_i = \sum_{j=1}^{n} s_{ij}\tilde{f}_j$, so, by possibly replacing r_{ij} with s_{ij}, we can achieve in (22) that $\deg(r_{ij}) \le m_i - 2$ for each $j = 1, \ldots, n$.

We define the polynomials $h_i := \sum_j r_{ij}f_j$ and consider them elements of the polynomial ring $\tilde{\mathbf{P}} := \mathbf{C}[a_1, \ldots, a_n, X_1, \ldots, X_n]$ in $2n$ variables, and extend the order $\prec$ to $\tilde{\mathbf{P}}$, i.e., we consider $a_1 \prec \ldots \prec a_n \prec X_1 \prec \ldots \prec X_n$. Additionally, we denote by $\tilde{\mathbf{T}}$ the set of power products of $\tilde{\mathbf{P}}$. It follows that

$$\begin{aligned} h_i = \sum_j r_{ij}f_j &= \sum_j r_{ij}(\tilde{f}_j + b_jX_j + c_jX_ja_j + d_ja_j) \\ &= \sum_j r_{ij}\tilde{f}_j + \sum_j r_{ij}(b_jX_j + c_jX_ja_j + d_ja_j) \\ &= \tilde{g}_i + \sum_j r_{ij}(b_jX_j + c_jX_ja_j + d_ja_j). \end{aligned}$$

Since we assumed that the b_j, c_j and d_j are constants, $\deg(r_{ij}) \le m_i - 2$ implies that

$$\deg\Big(\sum_j r_{ij}(b_j X_j + d_j a_j)\Big) \le m_i - 1, \qquad \deg\Big(\sum_j r_{ij} c_j X_j a_j\Big) \le m_i$$

and since $a_j \prec X_i$ for all j we have

$$\mathrm{LPP}(h_i) = \mathrm{LPP}(\tilde{g}_i) = X_i^{m_i}.$$

Now we define the ideal $J \lhd \tilde{\mathbf{P}}$ generated by $f_1, \ldots, f_n$ and the polynomials $k_j = a_j^r - X_j^q$, $j = 1, \ldots, n$. Clearly, $h_1, \ldots, h_n \in J$. Also, $\mathrm{LPP}(h_i) = X_i^{m_i}$ and $\mathrm{LPP}(k_i) = a_i^r$ for $i = 1, \ldots, n$, the latter since $r > q$.

In order to prove that $V(J)$ is finite we consider the reduced Gröbner basis G for J with respect to $\prec$. By the above, we have $X_i^{m_i}, a_i^r \in \{\mathrm{LPP}(f) \mid f \in J\}$, and by Theorem A.1 $\{\mathrm{LPP}(f) \ : \ f \in J\} = \{tg \ : \ g \in G,\ t \in \tilde{\mathbf{T}}\}$. Hence, for each $i = 1, \ldots, n$, $X_i^{m_i} = t_i \, \mathrm{LPP}(g_i)$ and $a_i^r = t_{i-n} \, \mathrm{LPP}(g_{i-n})$ for some $g_1, \ldots, g_{2n} \in G$, and $t_1, \ldots, t_{2n} \in \tilde{\mathbf{T}}$. This is only possible if there exist $\hat{m}_1, \ldots, \hat{m}_{2n}$ such that $\mathrm{LPP}(g_i) = X_i^{\hat{m}_i}$ and $\mathrm{LPP}(g_{i+n}) = a_i^{\hat{m}_{i+n}}$ for $i = 1, \ldots, n$. This implies by Theorem A.2 that $V(J) = V(\{f_1, \ldots, f_n, k_1, \ldots, k_n\})$ is finite, which is what we wanted to show. □

Acknowledgements The work of Kristian Bredies is supported by the Austrian Science Fund (FWF) under grant SFB32 (SFB “Mathematical Optimization and Applications in the Biomedical Sciences”). Dirk Lorenz acknowledges support by the DFG under grant LO 1436/2-1.

References

1. Nocedal, J., Wright, S.: Numerical Optimization. Springer Series in Operations Research and Financial Engineering. Springer, 2nd edition (2006)
2. Pytlak, R.: Conjugate Gradient Algorithms in Nonconvex Optimization, Nonconvex Optimization and Its Applications, volume 89. Springer (2009)
3. Nikolova, M.: Analysis of the recovery of edges in images and signals by minimizing nonconvex regularized least-squares. Multiscale Modelling and Simulation, 4(3), 960–991 (2005)
4. Nikolova, M.: Markovian reconstruction using a GNC approach. IEEE Transactions on Image Processing, 8(9), 1204–1220 (1999)
5. Blake, A., Zisserman, A.: Visual Reconstruction. MIT Press (1987)
6. Lions, P.L., Mercier, B.: Splitting algorithms for the sum of two nonlinear operators. SIAM Journal on Numerical Analysis, 16(6), 964–979 (1979)
7. Combettes, P.L., Wajs, V.R.: Signal recovery by proximal forward-backward splitting. Multiscale Model. Simul., 4(4), 1168–1200 (2005)
8. Bredies, K., Lorenz, D.A.: Linear convergence of iterative soft-thresholding. Journal of Fourier Analysis and Applications, 14(5–6), 813–837 (2008). doi:10.1007/s00041-008-9041-1
9. Attouch, H., Bolte, J., Svaiter, B.F.: Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods. Mathematical Programming Series B, 39pp (2011). doi:10.1007/s10107-011-0484-9
10. Ito, K., Kunisch, K.: A variational approach to sparsity optimization based on Lagrange multiplier theory. MOBIS SFB-Report 2011-014 (2011). Http://math.uni-graz.at/mobis/publications/SFB-Report-2011-014.pdf
11. Chouzenoux, E., Pesquet, J.C., Repetti, A.: Variable metric forward-backward algorithm for minimizing the sum of a differentiable function and a convex function. to appear in Journal of Optimization Theory and Applications, http://www.optimization-online.org/DB_HTML/2013/01/3749.html (2013)
12. Ochs, P., Chen, Y., Brox, T., Pock, T.: iPiano: Inertial proximal algorithm for non-convex optimization. http://lmb.informatik.uni-freiburg.de/Publications/2013/OB13 (2013)

13. Kaplan, A., Tichatschke, R.: Proximal point methods and nonconvex optimization. Journal of Global Optimization, 13, 389–406 (1998)
14. Lewis, A.S., Luke, D.R., Malick, J.: Local linear convergence for alternating and averaged nonconvex projections. Foundations of Computational Mathematics, 9, 485–513 (2009)
15. Rockafellar, R.T., Wets, R.J.B.: Variational Analysis, Grundlehren der mathematischen Wissenschaften, volume 317. Springer (1998)
16. Penot, J.P.: Proximal mappings. Journal of Approximation Theory, 94, 203–221 (1998)
17. Hare, W., Sagastizábal, C.: Computing proximal points of nonconvex functions. Mathematical Programming, Series B, 116(1–2), 221–258 (2009)
18. Demyanov, V.F., Rubinov, A.M.: Approximate Methods in Optimization Problems. Number 32 in Modern Analytic and Computational Methods in Science and Mathematics. American Elsevier (1970)
19. Dunn, J.C.: Global and asymptotic convergence rate estimates for a class of projected gradient processes. SIAM Journal on Control and Optimization, 19(3), 368–400 (1981)
20. Clarke, F.H.: Optimization and Nonsmooth Analysis. CRM Université de Montréal (1989)
21. Bredies, K., Lorenz, D.A.: Regularization with non-convex separable constraints. Inverse Problems, 25(8), 085011 (14pp) (2009). doi:10.1088/0266-5611/25/8/085011
22. Blumensath, T., Yaghoobi, M., Davies, M.: Iterative hard thresholding and l^0 regularisation. In: IEEE International Conference on Acoustics, Speech and Signal Processing (2007)
23. Moulin, P., Liu, J.: Analysis of multiresolution image denoising schemes using generalized Gaussian and complexity priors. IEEE Transactions on Information Theory, 45(3), 909–919 (1999)
24. Lorenz, D.A.: Non-convex variational denoising of images: Interpolation between hard and soft wavelet shrinkage. Current Development in Theory and Application of Wavelets, 1(1), 31–56 (2007)
25. Lorenz, D.A.: Convergence rates and source conditions for Tikhonov regularization with sparsity constraints. Journal of Inverse and Ill-Posed Problems, 16(5), 463–478 (2008). doi:10.1515/JIIP.2008.025
26. Grasmair, M.: Well-posedness classes for sparse regularisation. Reports of FSP S092 - "Industrial Geometry", University of Vienna, Austria (2010)
27. Lai, M.J., Wang, J.: An unconstrained ℓ_q minimization with $0 < q \leq 1$ for sparse solution of underdetermined linear systems. SIAM Journal on Optimization, 21(1), 82–101 (2011)
28. Sturmfels, B.: What is . . . a Gröbner basis? Notices Amer. Math. Soc., 52(10), 1199–1200 (2005)
29. Buchberger, B.: Introduction to Gröbner bases. In: Gröbner Basis and Applications, London Mathematical Society Lecture Notes Series, volume 251 (B. Buchberger, F. Winkler, eds.), pp. 3–31. Cambridge University Press (1998)
30. Kreuzer, M., Robbiano, L.: Computational commutative algebra, volume 1. Springer (2000)