Computational Optimization and Applications (2024) 87:1–37
https://doi.org/10.1007/s10589-023-00515-x

Inexact proximal Newton methods in Hilbert spaces

Bastian Pötzl[1] · **Anton Schiela**[1] · **Patrick Jaap**[2]

Received: 23 September 2022 / Accepted: 26 July 2023 / Published online: 16 August 2023

Abstract
We consider proximal Newton methods with an inexact computation of update steps. To this end, we introduce two inexactness criteria which characterize sufficient accuracy of these update step and with the aid of these investigate global convergence and local acceleration of our method. The inexactness criteria are designed to be adequate for the Hilbert space framework we find ourselves in while traditional inexactness criteria from smooth Newton or finite dimensional proximal Newton methods appear to be inefficient in this scenario. The performance of the method and its gain in effectiveness in contrast to the exact case are showcased considering a simple model problem in function space.

Keywords Non-smooth optimization · Optimization in Hilbert space · Proximal Newton · Inexactness

Mathematics Subject Classification 49M15 · 49M37 · 65K10

1 Introduction

In the present work we extend the idea of Proximal Newton methods in Hilbert spaces as presented in [15] to admit an inexact computation of update steps by solving the respective subproblem only up to prescribed accuracy. We consider the composite

✉ Bastian Pötzl
bastian.poetzl@uni-bayreuth.de

Anton Schiela
anton.schiela@uni-bayreuth.de

Patrick Jaap
patrick.jaap@tu-dresden.de

[1] University of Bayreuth, Chair of Applied Mathematics, Universitätsstraße 30, 95440 Bayreuth, Germany

[2] Technische Universität Dresden, Institut für Numerische Mathematik, Zellescher Weg 12-14, 01069 Dresden, Germany

minimization problem

$$\min_{x \in X} F(x) := f(x) + g(x) \tag{1}$$

on some real Hilbert space $(X, \langle \cdot, \cdot \rangle_X)$ where $f : X \to \mathbb{R}$ is assumed to be smooth in some adequate sense and $g : X \to]-\infty, \infty]$ is possibly not. We pay particular attention to the infinite-dimensionality of the underlying Hilbert spaces and thus develop inexactness criteria for update step computation that are sufficiently easy to evaluate, help us preserve convergence properties of the exact case as considered in [15] and reduce the computational effort significantly.

For an overview of the development of Proximal Newton methods themselves consider [15]. Here, we want to focus on the realization of the inexactness aspect and consider corresponding most recent literature in this introductory section. The use of gradient-like inexactness criteria which can be seen as the direct generalization of the one for classical smooth Newton methods in [5] is quite common, cf. [3, 9, 11, 14, 24].

In [11] additional knowledge of bounds on the second-order bilinear forms as well as the Lipschitz constant of f' is necessary and only local convergence has been investigated in the inexact case. Globalization of the ensuing method has been achieved in [9] by using a Proximal Gradient substitute step in case the inexactly computed second order step does not suffice a sufficient decrease criterion or the step computation subproblem is ill-formed due to non-convexity which thus can be overcome as well. In [3] the particular case of L_1-regularization for machine learning applications has been considered. Thus, the inexactness criterion has further been specified and also here enhanced with a decrease criterion in the quadratic approximation of the composite objective function. The latter has then been tightened in order to achieve local acceleration. A similar route has been taken in the development of a globally convergent methods in [14] and [24].

Another approach to inexactness criteria is measuring the residual within the step computation subproblem. In [12], where objective functions consisting of the sum of a thrice continuously differentiable smooth part and a self-concordant non-smooth part have been considered, the residual vector within optimality conditions for update computation is supposed to be bounded in norm with respect to the already computed inexact step. However, the residual can also be measured via functional descent in the quadratic approximation of the composite objective F, cf. [10, 18]. While in [10] the second order model decrease bound against its optimal value has not directly been tested but simply assumed to hold after a finite (and fixed) number of subproblem solver iterations, the authors in [18] have taken the structure of their randomized coordinate descent subproblem solver into account and also have given quadratic bounds for the prefactor constant within their model descent estimate in order to obtain sufficient convergence results.

All of the above works have in common that they depend on the finite dimensional structure of the underlying Euclidean space. In particular, the efficient computation of proximal gradients, required for the evaluation of inexactness criteria, relies on the diagonal structure of the underlying scalar product $\langle \cdot, \cdot \rangle_X$, which is usually not present

in (discretized) function spaces, as for example, Sobolev spaces. Moreover, all current approaches consider fixed search directions which are then scaled by some step length parameter.

Our contributions beyond the above work can be summarized as follows: Most importantly, we replace the Euclidean space setting with a Hilbert space one in order to rigorously allow function space applications of our method. In particular, we are interested in the important case where X is a Sobolev space. Then, a diagonal approximation of $\langle \cdot, \cdot \rangle_X$ after discretization would lead to proximal operators that suffer from mesh-dependent condition numbers. For the efficient computation of proximal steps we thus take advantage of a non-smooth multigrid method. Specifically, we use a Truncated Non-smooth Newton Multigrid (TNNMG) method, cf. [7], in our numerical implementation. Consequently, our inexactness criteria need to be constructed in such a way that their evaluation is efficient in this context. Existing criteria can only be employed efficiently, if $\langle \cdot, \cdot \rangle_X$ enjoys a diagonal structure.

Additionally, ellipticity of the bilinear forms for forming quadratic approximations of our objective functional as well as convexity of the non-smooth part g has often been crucial in the literature. We drop these prerequisites and use a less restrictive framework of convexity assumptions for the composite objective function F. Finally, we do not demand second order differentiability with Lipschitz-continuous second order derivative of the smooth part f but instead settle for adequate semi-smoothness assumptions.

Let us now give the precise set of assumptions in which we will discuss the convergence properties of inexact Proximal Newton methods. As pointed out beforehand, we find ourselves in a real Hilbert space $(X, \langle \cdot, \cdot \rangle_X)$ with corresponding norm $\|v\|_X = \sqrt{\langle v, v \rangle_X}$ and dual space X^*. This choice of X also provides us with the Riesz-Isomorphism $\mathcal{R} : X \to X^*$, defined by $\mathcal{R}x = \langle x, \cdot \rangle_X$, which satisfies $\|\mathcal{R}x\|_{X^*} = \|x\|_X$ for every $x \in X$. Since $\mathcal{R}$ is non-trivial in general, we will not identify X and X^*.

Assumption 1 The smooth part of our objective functional $f : X \to \mathbb{R}$ is assumed to be continuously differentiable with Lipschitz-continuous derivative $f' : X \to X^*$, i.e., we can find some constant $L_f > 0$ such that for every $x, y \in X$ we obtain the estimate

$$\|f'(x) - f'(y)\|_{X^*} \leq L_f \|x - y\|_X . \tag{2}$$

As mentioned beforehand, we will use the base algorithm from [15] as our point of departure. This means that we consider a variation of the Proximal Newton method which is globalized by an additional norm term within the subproblem for step computation. As a consequence, the latter reads

$$\Delta x(\omega) := \operatorname*{argmin}_{\delta x \in X} \lambda_{x,\omega}(\delta x) \tag{3}$$

where the regularized second order decrease model $\lambda_{x,\omega} : X \to \mathbb{R}$ is given by

$$\lambda_{x,\omega}(\delta x) := f'(x)\delta x + \frac{1}{2} H_x(\delta x, \delta x) + \frac{\omega}{2} \|\delta x\|_X^2 + g(x + \delta x) - g(x) .$$

The updated iterate then takes the form $x_+(\omega) := x + \Delta x(\omega)$.

The second order model of the smooth part f from above also has to be endowed with adequate prerequisites. Notationally identifying the linear operators $H_x \in \mathcal{L}(X, X^*)$ with the corresponding symmetric bilinear forms $H_x : X \times X \to \mathbb{R}$, we write $(H_x v)(w) = H_x(v, w)$ and abbreviate $H_x(v)^2 = H_x(v, v)$.

Assumption 2 Uniform boundedness of the H_x along the sequence of iterates in the form

$$\exists\, M > 0 \,\forall\, x \in X : \|H_x\|_{\mathcal{L}(X,X^*)} \leq M \tag{4}$$

will also be of importance in what follows.

Assumption 3 Furthermore, we assume the existence of a mapping $\kappa_1 : X \to \mathbb{R}$ which is bounded from below such that the bound

$$\forall\, x \in X \;\forall v \in X : H_x(v)^2 := H_x(v, v) \geq \kappa_1(x) \|v\|_X^2 \tag{5}$$

holds, which can be interpreted as an ellipticity assumption on H_x if $\kappa_1(x)$ is positive. In this case, when considering exact (and smooth) Proximal Newton methods, where H_x is given by the Hessian of f at some point $x \in X$, (5) is equivalent to $\kappa_1(x)$-strong convexity of f. When considering general bilinear forms H without dependence on some x, we refer to (5) in the sense of $H(v)^2 \geq \kappa_1$ for some constant $\kappa_1 \in \mathbb{R}$ and all $v \in X$.

As a simple example for such an operator, one could imagine the mapping

$$H_x : H^1(\Omega) \to H^1(\Omega)^* , \quad H(u) := \left(v \mapsto \int_\Omega \nabla u(\xi) \cdot \nabla v(\xi) + x(\xi) u(\xi) v(\xi) \, \mathrm{d}\xi \right)$$

with $\Omega \in \mathbb{R}^n$ an open, bounded set and $x \in L^\infty(\Omega)$. Then, the operator H_x is satisfies both (5) with $\kappa_1 = -\|x\|_\infty$ and (4) with $M = \max\{\|x\|_\infty, 1\}$.

While in a sufficiently smooth setting $H_x := f''(x)$ is common, for most of the paper we may choose H_x freely in the above framework. For fast local convergence, however, we will impose a semi-smoothness assumption, cf. (17). Semi-smooth Newton methods in function space have been discussed, for example, in [8, 19–21].

Assumption 4 As far as the non-smooth part $g : X \to]-\infty, \infty]$ is concerned, we require lower semi-continuity as well as the existence of some $\kappa_2 \in \mathbb{R}$ such that the bound

$$g(sx + (1-s)y) \leq s g(x) + (1-s) g(y) - \frac{\kappa_2}{2} s(1-s) \|x - y\|_X^2 \tag{6}$$

holds for all $x, y \in X$ and $s \in [0, 1]$. While for $\kappa_2 \leq 0$ this estimate is often referred to as weak convexity, the case of $\kappa_2 > 0$ reduces (6) to κ_2-strong convexity of g. In the latter case we can then conclude that g is bounded from below, its level-sets $L_\alpha g$ are bounded for all $\alpha \in \mathbb{R}$ and that their diameter shrinks to 0 in the limit of $\alpha \to \inf_{x \in X} g$. Non-positivity of κ_2 allows g to be non-convex in a limited way. This assumption is often summarized to *weak κ_2-convexity* of g.

The theory behind Proximal Newton methods and the respective convergence properties evolve around the convexity estimates stated in (5) and (6). We will assign particular importance to the interplay of the convexity properties of f and g, i.e., the sum $\kappa_1(x) + \kappa_2$ will continue to play an important role over the course of the present treatise. In particular, the convexity of f only enters this quantity in a local way, depending on the current iterate x. Apparently, the update step in (3) is well defined for every $\omega > 0$ if $\kappa_1(x) + \kappa_2 > 0$. This holds also in the case of $\kappa_1(x) + \kappa_2 \leq 0$ for every $\omega > -(\kappa_1(x) + \kappa_2)$ due to the bounds stated in (5), (6) and the strong convexity of the norm term. For this reason, we will assume $\omega > -(\kappa_1(x) + \kappa_2)$ wherever it appears.

The above demands on f, g, H_x and ω constitute the standing assumptions for the further investigation which we impose for the entirety of the paper.

Let us now briefly outline the structure of our work: In Sect. 2 we introduce the notion of composite gradient mappings and consider some of their basic properties. Afterwards, in Sect. 3, we take advantage of the acquired knowledge and introduce the first inexactness criterion in order to investigate local convergence of our method as well as the influence of both damping and inexactness. Section 4 then considers the globalization phase of our inexact Proximal Newton method and for this reason introduces a second inexactness criterion which compares the functional decrease of inexact updates with steps originating from a simpler subproblem. Thus, we also achieve sufficient global convergence results. In order to then benefit from local acceleration, we investigate the transition to local convergence in Sect. 5. To this end, we need to ensure that close to optimal solutions also arbitrarily weakly damped update steps yield sufficient decrease. Lastly, we put our method to the test in Sect. 5.1 and display global convergence as well as local acceleration considering a simple model problem in function space. Concluding remarks can be found in Sect. 6.

2 Composite gradient mappings and their properties

The main goal to keep in mind is not only to introduce the concept of inexactness to the computation of update steps of the Proximal Newton method from [15] but also quantify the influence of damping update steps to the local convergence rate of our algorithm.

2.1 Definition and representation via proximal mappings

For this cause, we take advantage of the notion of regularized composite gradient mappings $G_\tau^\Phi : X \to X$ for some composite functional $\Phi : X \to \mathbb{R}$ of the form

$\Phi(x) := \phi(x) + \psi(x)$ with smooth part $\phi : X \to \mathbb{R}$ and non-smooth part $\psi : X \to \mathbb{R}$. More precisely, ϕ has to be at least continuously differentiable and ψ should satisfy Assumption 4 on g from before. Then, the aforementioned gradient mapping is defined via

$$G_\tau^\Phi(y) := -\tau\left[\operatorname*{argmin}_{\delta y \in X} \phi'(y)\delta y + \frac{\tau}{2}\|\delta y\|_X^2 + \psi(y+\delta y) - \psi(y)\right] \tag{7}$$

for $y \in X$ and some regularization parameter $\tau > -\kappa_2$ the assumptions on which we will specify further over the course of the current section. For the derivation of useful estimates for composite gradient mappings, the so-called scaled dual proximal mapping $\mathcal{P}_\psi^H : X^* \to X$, defined via

$$\mathcal{P}_\psi^H(\ell) := \operatorname*{argmin}_{z \in X} \psi(z) + \frac{1}{2}H(z,z) - \ell(z)$$

for arbitrary $\ell \in X^*$ and some symmetric bilinear form H sufficing (5) as well as some real valued function ψ satisfying (6) for constants $\kappa_1, \kappa_2 \in \mathbb{R}$ with $\kappa_1 + \kappa_2 > 0$, will play an important role.

In what is to come, we will take advantage of the following two crucial results concerning dual proximal mappings which have been stated and proven in [15]. The first one is a general estimate for the image of such operators which generalizes the assertions of the so called second prox theorem, cf. e.g. [2, Chapter 6.5]. The second one is a Lipschitz-continuity result.

Proposition 1 ([15], *Proposition 2 and Corollary 1) Let H and ψ satisfy the assumptions* (5) *and* (6) *with $\kappa_1 + \kappa_2 > 0$. Then for any $\ell \in X^*$ the image of the corresponding proximal mapping $u := \mathcal{P}_\psi^H(\ell)$ satisfies the estimate*

$$\big[\ell - H(u)\big](\xi - u) \le \psi(\xi) - \psi(u) - \frac{\kappa_2}{2}\|\xi - u\|_X^2$$

for all $\xi \in X$. Additionally, for all $\ell_1, \ell_2 \in X^*$ the following inequality holds:

$$\|\mathcal{P}_\psi^H(\ell_1) - \mathcal{P}_\psi^H(\ell_2)\|_X \le \frac{1}{\kappa_1 + \kappa_2}\|\ell_1 - \ell_2\|_{X^*}.$$

With the aid of scaled proximal mappings, we can express the composite gradient mapping as

$$G_\tau^\Phi(y) = \tau\big[y - \mathcal{P}_\psi^{\tau\mathcal{R}}\big(\tau\mathcal{R}y - \phi'(y)\big)\big], \tag{8}$$

where $\mathcal{R} : X \to X^*$ again denotes the Riesz-Isomorphism. Let us now justify the designation of G_τ^Φ as a regularized composite gradient mapping. If we consider the smooth case of $\psi = 0$, the proximal mapping takes the form $\mathcal{P}_\psi^H(\ell) = H^{-1}\ell$. This

fact carries over to the definition of the gradient mapping via

$$G_\tau^\Phi(y) = \tau\big[y - (\tau\mathcal{R})^{-1}\big(\tau\mathcal{R}y - \phi'(y)\big)\big] = \mathcal{R}^{-1}\phi'(y)$$

which resembles the infinite dimensional counterpart of the gradient $\nabla\phi$ in Euclidean space. Note that this consistency result holds for all $\tau > 0$.

Another consideration which expresses the consistency between G_τ^F and some actual 'smooth' gradient of $F = f + g$ with respect to our minimization problem (1) is the following: Let then $G_\tau^F(x_*) = 0$ hold for some $x_* \in X$ and $\tau \geq 0$. This is equivalent to the fixed point equation $x_* = \mathcal{P}_g^{\tau\mathcal{R}}\big(\tau\mathcal{R}x_* - f'(x_*)\big)$ which can then again be transformed to $-f'(x_*) \in \partial_F g(x_*) \subset X^*$ for the Fréchet subdifferential ∂_F. Consequently, we recognize that the composite gradient mapping is zero if and only if we evaluate it at stationary points of the underlying minimization problem (1).

2.2 Key properties and auxiliary estimates

For now, let us derive some key properties of the composite gradient mappings which will be crucial as we quantify the influence of both inexactness and damping to local convergence rates of our algorithm.

Before departing on this endeavor we introduce the modified quadratic model $\hat{F}_{x,\omega} : X \to \mathbb{R}$ of the composite objective functional F around $x \in X$ with regularization parameter ω via

$$\begin{aligned}\hat{F}_{x,\omega}(y) := F(x) + \lambda_{x,\omega}(y - x) \quad &= f(x) + f'(x)(y - x) + \frac{1}{2}H_x(y - x)^2 + g(y) \\ &+ \frac{\omega}{2}\|y - x\|_X^2 .\end{aligned} \tag{9}$$

The corresponding composite gradient mapping $G_\tau^{\hat{F}_{x,\omega}}$ will play an important role. In that regard, we note that in the framework of the definition of the gradient mapping in (7) we thus have $\Phi = \hat{F}_{x,\omega} = \hat{\phi} + \hat{\psi}$ with

$$\hat{\phi}(y) = f(x) + f'(x)(y - x) + \frac{1}{2}\big(H_x + \omega\mathcal{R}\big)(y - x)^2 \quad , \quad \hat{\psi}(y) = g(y) \tag{10}$$

and thereby $\hat{\phi}'(y) = f'(x) + \big(H_x + \omega\mathcal{R}\big)(y - x)$ for any $y \in X$. The following lemma provides us with helpful estimates for the norm difference of composite gradient mappings both from above and below.

Lemma 1 *For every $x, y, z \in X$ and with $\tau := \omega + \frac{1}{2}\big(\|H_x\|_{\mathcal{L}(X,X^*)} + \kappa_1(x)\big)$, the regularized composite gradient mapping suffices the estimate*

$$\tau\big(1 - \mathcal{H}\big)\|y - z\|_X \leq \|G_\tau^{\hat{F}_{x,\omega}}(y) - G_\tau^{\hat{F}_{x,\omega}}(z)\|_X \leq \tau\big(1 + \mathcal{H}\big)\|y - z\|_X \tag{11}$$

where we abbreviated $\mathcal{H} := \frac{\|H_x\|_{\mathcal{L}(X,X^*)} - \kappa_1(x)}{2(\tau+\kappa_2)}$.

Proof As we insert the characterizations of the respective regularized composite gradient mappings as in (8), we perceive that we can represent their norm difference via

$$\big\|G_\tau^{\hat{F}_{x,\omega}}(y) - G_\tau^{\hat{F}_{x,\omega}}(z)\big\|_X = \tau\big\|(y-z) - \big(\mathcal{P}_y - \mathcal{P}_z\big)\big\|_X$$

with abbreviations $\mathcal{P}_\xi := \mathcal{P}_g^{\tau\mathcal{R}}\big(\tau\mathcal{R}\xi - \big[f'(x) + \big(H_x + \omega\mathcal{R}\big)(\xi - x)\big]\big)$ for $\xi \in \{y, z\}$. This provides us with the bounds

$$\tau\big(\|y-z\|_X - \|\mathcal{P}_y - \mathcal{P}_z\|_X\big) \le \big\|G_\tau^{\hat{F}_{x,\omega}}(y) - G_\tau^{\hat{F}_{x,\omega}}(z)\big\|_X \le \tau\big(\|y-z\|_X + \|\mathcal{P}_y - \mathcal{P}_z\|_X\big)$$

from above and below for the norm difference of gradient mappings. This shows that for the proof of (11) it suffices to verify

$$\|\mathcal{P}_y - \mathcal{P}_z\|_X \le \mathcal{H}\|y-z\|_X = \frac{\|H_x\|_{\mathcal{L}(X,X^*)} - \kappa_1(x)}{2(\tau+\kappa_2)}\|y-z\|_X\,. \tag{12}$$

The Lipschitz result from Proposition 1 allows us to establish the following estimate for the norm difference of proximal mapping images in relation to their arguments:

$$\begin{aligned}\|\mathcal{P}_y - \mathcal{P}_z\|_X &\le \frac{1}{\tau+\kappa_2}\big\|\tau\mathcal{R}y - \big(H_x + \omega\mathcal{R}\big)(y-x) - \big(\tau\mathcal{R}z - \big(H_x + \omega\mathcal{R}\big)(z-x)\big)\big\|_{X^*} \\ &= \frac{1}{\tau+\kappa_2}\big\|\big((\tau-\omega)\mathcal{R} - H_x\big)(y-z)\big\|_{X^*} \le \frac{\|(\tau-\omega)\mathcal{R} - H_x\|_{\mathcal{L}(X,X^*)}}{\tau+\kappa_2}\|y-z\|_X\,.\end{aligned} \tag{13}$$

Let us now pay particular attention to the $\mathcal{L}(X, X^*)$-norm difference in the prefactor above. On the one hand, for any $\tau > -\kappa_2$, we can estimate it by

$$\|(\tau-\omega)\mathcal{R} - H_x\|_{\mathcal{L}(X,X^*)} \le |\tau-\omega| + \|H_x\|_{\mathcal{L}(X,X^*)}\,.$$

Nevertheless, with further assumptions on the gradient mapping regularization parameter τ we can deduce a better bound. To this end, we define $\lambda := \tau - \omega$ and choose λ_{opt} such that $\|\lambda\mathcal{R} - H_x\|_{\mathcal{L}(X,X^*)}$ is minimal. It is easy to see that the eigenvalues of the self-adjoint operator $H_x^\tau := \mathcal{R}^{-1}(\lambda\mathcal{R} - H_x)$ lie in the interval $\big[\lambda - \|H_x\|_{\mathcal{L}(X,X^*)}, \lambda - \kappa_1(x)\big]$.

In order to now minimize the norm of H_x^τ, we recognize that it equals the spectral radius of H_x^τ and thus want to establish a symmetrical interval where eigenvalues can be located. This yields the choice $\lambda_{\text{opt}} := \frac{1}{2}\big(\|H_x\|_{\mathcal{L}(X,X^*)} + \kappa_1(x)\big)$. In particular, this implies

$$\tau := \omega + \lambda_{\text{opt}} = \omega + \frac{1}{2}\big(\|H_x\|_{\mathcal{L}(X,X^*)} + \kappa_1(x)\big) \geq \omega + \frac{|\kappa_1(x)| + \kappa_1(x)}{2} \geq \omega + \kappa_1(x) > -\kappa_2$$

by our choice of ω and consequently

$$\begin{aligned}&\|(\tau-\omega)\mathcal{R} - H_x\|_{\mathcal{L}(X,X^*)} \\ &\quad = \|H_x^\tau\|_{\mathcal{L}(X,X)} = \|H_x\|_{\mathcal{L}(X,X^*)} - \lambda_{\text{opt}} = \frac{1}{2}\big(\|H_x\|_{\mathcal{L}(X,X^*)} - \kappa_1(x)\big)\,.\end{aligned}$$

Inserting this into the above estimate (13), we obtain (12) which completes the proof. □

For the next result, we take advantage of the solution property of exactly computed update steps from (3).

Proposition 2 *Let $\Delta x(\omega)$ be an exactly computed update step as in (3) at some $x \in X$. Then, for any $\tau > -\kappa_2$ the following identity holds:*

$$G_\tau^{\hat{F}_{x,\omega}}\big(x + \Delta x(\omega)\big) = 0\,. \tag{14}$$

Proof We consider the minimization problem within brackets in the definition of the regularized composite gradient mapping in (7). Here, we have to insert the derivative ϕ' of the smooth part of the regularized model $\hat{F}_{x,\omega}$ as in (10) evaluated at $y = x + \Delta x(\omega)$ which yields

$$\begin{aligned}&\operatorname*{argmin}_{\delta x \in X} \big[f'(x) + \big(H_x + \omega\mathcal{R}\big)\Delta x(\omega)\big]\delta x \\ &\quad + \frac{\tau}{2}\|\delta x\|_X^2 + g\big(x + \Delta x(\omega) + \delta x\big) - g\big(x + \Delta x(\omega)\big)\,.\end{aligned} \tag{15}$$

By strong convexity of the objective function for $\tau > -\kappa_2$, the above minimization problem has a unique solution $\delta\bar{x} \in X$. By first order optimality conditions, this solution then satisfies the dual space inclusion

$$0 \in f'(x) + \big(H_x + \omega\mathcal{R}\big)\Delta x(\omega) + \partial_F g\big(x + \Delta x(\omega) + \delta\bar{x}\big) + \tau\mathcal{R}\delta\bar{x} \tag{16}$$

for the Fréchet-subdifferential $\partial_F g$. Note here that the exactly computed update step $\Delta x(\omega)$ as a solution of (3) suffices

$$0 \in f'(x) + \big(H_x + \omega\mathcal{R}\big)\Delta x(\omega) + \partial_F g\big(x + \Delta x(\omega)\big)$$

which directly yields that $\delta\bar{x} = 0$ satisfies (16) and is thereby the unique solution of (15). This completes the proof of (14). □

Assumption 5 For the following, we require an approximation property of f' at stationary points x_* of our problem (1) with respect to the second order bilinear forms H_x which often appears in connection with semi-smoothness of f':

$$\left\| f'(x_*) - f'(x) - H_x(x_* - x) \right\|_{X^*} = o\left(\left\| x - x_* \right\|_X \right). \tag{17}$$

Adequate definitions of H_x can be given via a so-called Newton derivative from $\partial_N f'(x)$, also known as the generalized differential $\partial^* f'(x)$ for Lipschitz-continuous operators in finite dimensions, and for corresponding superposition operators, cf. [21, Chapter 3.2].

While finite-dimensional semi-smoothness of real-valued functions has been developed by Mifflin in [13], an extension to mappings between two finite dimensional spaces has then been given by Qi [16] as well as Qi and Sun [17]. The motivation for the concept is to develop locally q-superlinearly convergent Newton methods which are applicable despite the general non-smoothness of the underlying mapping.

This finite-dimensional notion of semi-smoothness can also be characterized by directional differentiability together with an approximation property of the above form. For the generalization to infinite-dimensional domain and image spaces, however, we do not require the directional differentiability of the corresponding mapping any more. In that regard, continuity near the point at which it is supposed to be semi-smooth together with the approximation property from (17) is sufficient, cf. [21] for an elaborate introduction to the concept of semi-smoothness for general operators between Banach spaces.

Let us now consider the difference of gradient mappings of the objective function F and its modified second order model $\hat{F}_{x,\omega}$ at stationary points x_* of problem (1).

Lemma 2 *Let the semi-smoothness assumption (17) hold at a stationary point $x_* \in X$. Then, the regularized composite gradient mapping satisfies the following estimate for each $\tau > -\kappa_2$ and $x \in X$ sufficiently close to x_*:*

$$\left\| G_\tau^F(x_*) - G_\tau^{\hat{F}_{x,\omega}}(x_*) \right\|_X \leq o\left(\left\| x_* - x \right\|_X \right) + \frac{\tau\, \omega}{\tau + \kappa_2} \left\| x_* - x \right\|_X .$$

Proof The proof here follows immediately by the characterization of the regularized composite gradient mapping as in (8) and the semi-smoothness of f' according to (17). To go into detail, by Proposition 1 we have

$$\begin{aligned}
&\left\| G_\tau^F(x_*) - G_\tau^{\hat{F}_{x,\omega}}(x_*) \right\|_X \\
&\quad = \tau \left\| \mathcal{P}_g^{\tau\mathcal{R}}\big(\tau\mathcal{R}x_* - f'(x_*)\big) - \mathcal{P}_g^{\tau\mathcal{R}}\big(\tau\mathcal{R}x_* - \big[f'(x) + \big(H_x + \omega\mathcal{R}\big)(x_* - x)\big]\big) \right\|_X \\
&\quad \leq \frac{\tau}{\tau + \kappa_2} \left\| \big(\tau\mathcal{R}x_* - f'(x_*)\big) - \big(\tau\mathcal{R}x_* - \big[f'(x) + \big(H_x + \omega\mathcal{R}\big)(x_* - x)\big] \right\|_{X^*} \\
&\quad \leq \frac{\tau}{\tau + \kappa_2} \left\| f'(x_*) - \big(f'(x) + (H_x + \omega\mathcal{R})(x_* - x)\big) \right\|_{X^*} \\
&\quad = o\left(\left\| x_* - x \right\|_X \right) + \frac{\tau\, \omega}{\tau + \kappa_2} \left\| x_* - x \right\|_X
\end{aligned}$$

the last identity of which follows by the aforementioned definition of $H_x \in \partial_N f'(x)$ as a Newton-derivative together with (17). □

2.3 An existing inexactness criterion

In the literature composite gradient mappings have been used in order to derive an inexactness criterion for update step computation within Proximal Newton methods. Based on an approach from the smooth case, cf. e.g. [5], the authors in [9, 11] took advantage of the composite gradient mapping G_τ^F to postulate the corresponding estimate which their inexact update steps have to satisfy. In a similar fashion, transferring the criterion from the smooth case to our globalization scheme using the damped update steps $\Delta s(\omega)$ from (3) yields

$$\left\| G_\tau^{\hat{F}_{x,\omega}}\big(x + \Delta s(\omega)\big) \right\|_X \leq \eta \left\| G_\tau^F(x) \right\|_X \tag{18}$$

for some yet to be specified forcing term $\eta > 0$. Here, $\hat{F}_{x,\omega}$ denotes the modified quadratic model from (9) above. This requirement can be understood as a relative error criterion for the composite gradient mapping in norm due to the optimality of exactly computed update steps as formulated in Proposition 2.

While in a finite dimensional Euclidean space setting the gradient mapping $G_\tau^{\hat{F}_{x,\omega}}\big(x + \Delta s(\omega)\big)$ can be evaluated efficiently due to the diagonal structure of the norm term, in an infinite dimensional setting the computation of it is quite demanding, even as expensive as computing the actual exact update step $\Delta x(\omega)$.

Consequently, evaluating (18) for every iteration within the subproblem solver becomes very costly and thereby immediately eclipses the savings we gain from inexactly computing the update steps. For this reason, we will resort to a different inexactness criterion.

3 First inexactness criterion and local convergence properties

As pointed out beforehand, we do not use an inexactness criterion of the form (18) due to its immense computational effort in function space. Instead, we exploit the advantageous properties of the TNNMG subproblem solver by resorting to an actual relative error estimate of the form

$$\left\| \Delta x(\omega) - \Delta s(\omega) \right\|_X \leq \eta \left\| \Delta x(\omega) \right\|_X \tag{19}$$

where $\Delta x(\omega)$ denotes the exact solution of the update step computation subproblem (3) and $\Delta s(\omega)$ is the corresponding inexact candidate. The influence of the forcing terms $\eta \geq 0$ on local convergence rates will be investigated in Theorem 1.

Before actually stating the local convergence results, let us remark that the inexactness criterion (19) is trivially satisfied by exactly computed update steps and η is a measure for the margin for error which we allow in the computation. Additionally, the fact that the inexactly computed update steps $\Delta s(\omega)$ are in our case iterates from the

convergent TNNMG subproblem solver implies that sooner or later within the solution process of (3) the requirement (19) will be satisfied.

Furthermore, let us comment on the efficient evaluation of this relative error estimate. At first sight, this is not completely obvious since apparently we do not have the exact solution $\Delta x(\omega)$ of the update computation subproblem (3) at hand. In order to deal with this issue, we take advantage of the multigrid structure of the iterative subproblem solver which we employ, i.e., the TNNMG method from [7]. By δ^j we denote TNNMG-corrections, let therefore $\Delta s^i(\omega) = \sum_{j=1}^{i} \delta^j$ be an iterate within the inner solver towards the exact solution $\Delta x(\omega)$ and θ the 'constant' multigrid convergence rate for $\|\delta^j\|_X \leq \theta \|\delta^{j-1}\|_X$. Simple triangle inequalities thus provide us with

$$\|\Delta x(\omega) - \Delta s^i(\omega)\|_X = \sum_{j=i+1}^{\infty} \|\delta^j\|_X \leq \|\delta^i\|_X \sum_{j=i+1}^{\infty} \theta^{j-i} = \frac{\theta}{1-\theta} \|\delta^i\|_X .$$

Similarly, for the norm of the exact solution we obtain

$$\begin{aligned} \|\Delta x(\omega)\|_X &= \Big\|\sum_{j=1}^{\infty} \delta^j\Big\|_X = \Big\|\Delta s^i(\omega) + \sum_{j=i+1}^{\infty} \delta^j\Big\|_X \geq \|\Delta s^i(\omega)\|_X - \Big\|\sum_{j=i+1}^{\infty} \delta^j\Big\|_X \\ &\geq \|\Delta s^i(\omega)\|_X - \frac{\theta}{1-\theta} \|\delta^i\|_X . \end{aligned}$$

Combining both of these estimates implies

$$\frac{\|\Delta x(\omega) - \Delta s^i(\omega)\|_X}{\|\Delta x(\omega)\|_X} \leq \frac{\frac{\theta}{1-\theta}\|\delta^i\|_X}{\|\Delta s^i(\omega)\|_X - \frac{\theta}{1-\theta}\|\delta^i\|_X} \overset{!}{\leq} \eta \tag{20}$$

as a sufficient and easy to evaluate alternative inexactness criterion for the relative error estimate (19). Numerical experiments, which we also incorporated to Sect. 5.1, clearly demonstrate that the performed triangle inequalities are sharper than one might have expected. Thus, the evaluation of the alternative criterion from (20) comes very close to using the actual relative error for our computations later on.

Another crucial auxiliary result for all of the present treatise is an equivalence estimate between exactly computed update steps which have been damped according to different regularization parameters. It generalizes [15, Lemma 6] insofar that this result is comprised here in the case of $\omega = 0$.

Lemma 3 *Let $\Delta x(\omega)$ and $\Delta x(\tilde{\omega})$ be exactly computed update steps at an iterate $x_k =: x$ according to (3) with regularization parameters satisfying $\omega > -(\kappa_1(x) + \kappa_2)$ and $\tilde{\omega} \geq \omega$. Then the following norm estimates hold:*

$$\|\Delta x(\omega) - \Delta x(\tilde{\omega})\|_X \leq \frac{\tilde{\omega} - \omega}{\omega + \kappa_1(x) + \kappa_2} \|\Delta x(\tilde{\omega})\|_X \tag{21}$$

$$\|\Delta x(\tilde{\omega})\|_X \leq \|\Delta x(\omega)\|_X \leq \frac{\tilde{\omega} + \kappa_1(x) + \kappa_2}{\omega + \kappa_1(x) + \kappa_2} \|\Delta x(\tilde{\omega})\|_X \tag{22}$$

Proof We consider the proximal representation of exactly computed update steps

$$x + \Delta x(\hat{\omega}) = x_+(\hat{\omega}) = \mathcal{P}_g^{H_x+\hat{\omega}\mathcal{R}}\big((H_x + \hat{\omega}\mathcal{R})x - f'(x)\big)$$

for $\hat{\omega} \in \{\omega, \tilde{\omega}\}$. Via Proposition 1, from these we can deduce the respective proximal inequalities

$$\begin{aligned}\big[\big(\hat{\omega}\mathcal{R} + H_x\big)x - f'(x) - \big(\hat{\omega}\mathcal{R} + H_x\big)x_+(\hat{\omega})\big]\big(\hat{\xi} - x_+(\hat{\omega})\big)& \\ \leq g(\hat{\xi}) - g\big(x_+(\hat{\omega})\big) - \frac{\kappa_2}{2}\big\|\hat{\xi} - x_+(\hat{\omega})\big\|_X^2&\end{aligned} \tag{23}$$

for any $\hat{\xi} \in X$ which we choose as $\hat{\xi} = x_+(\hat{\omega})$ for the respectively other $\hat{\omega} \in \{\omega, \tilde{\omega}\}$ and add the ensuing estimates in order to obtain

$$\big[\big(\omega\mathcal{R} + H_x\big)\Delta x(\omega) - \big(\tilde{\omega}\mathcal{R} + H_x\big)\Delta x(\tilde{\omega})\big]\big(\Delta x(\omega) - \Delta x(\tilde{\omega})\big) \leq -\kappa_2\big\|\Delta x(\omega) - \Delta x(\tilde{\omega})\big\|_X^2 .$$

We now insert a $(\omega\mathcal{R} + H_x)\Delta x(\tilde{\omega})$-term to the left-hand squared bracket and simplify which yields

$$\begin{aligned}\big(\omega\mathcal{R} + H_x\big)\big(\Delta x(\omega) - \Delta x(\tilde{\omega})\big)^2 + \kappa_2\big\|\Delta x(\omega) - \Delta x(\tilde{\omega})\big\|_X^2& \\ \leq (\tilde{\omega} - \omega)\mathcal{R}\big(\Delta x(\tilde{\omega}), \Delta x(\omega) - \Delta x(\tilde{\omega})\big)&\end{aligned}$$

where we can now additionally utilize (5) for the simpler form

$$(\omega + \kappa_1(x) + \kappa_2)\big\|\Delta x(\omega) - \Delta x(\tilde{\omega})\big\|_X^2 \leq (\tilde{\omega} - \omega)\mathcal{R}\big(\Delta x(\tilde{\omega}), \Delta x(\omega) - \Delta x(\tilde{\omega})\big) . \tag{24}$$

From here, we can take two paths both of which contribute to the completion of the proof. Firstly, we divide by $(\omega + \kappa_1(x) + \kappa_2) > 0$ and use the Cauchy-Schwarz-Inequality on the right-hand side which then implies

$$\big\|\Delta x(\omega) - \Delta x(\tilde{\omega})\big\|_X^2 \leq \frac{\tilde{\omega} - \omega}{\omega + \kappa_1(x) + \kappa_2}\big\|\Delta x(\tilde{\omega})\big\|_X\big\|\Delta x(\omega) - \Delta x(\tilde{\omega})\big\|_X ,$$

i.e., exactly (21) since the difference norm term can be assumed to be non-zero without loss of generality. Moving on, we take advantage of

$$\big\|\Delta x(\omega)\big\|_X \leq \big\|\Delta x(\omega) - \Delta x(\tilde{\omega})\big\|_X + \big\|\Delta x(\tilde{\omega})\big\|_X \leq \Big(1 + \frac{\tilde{\omega} - \omega}{\omega + \kappa_1(x) + \kappa_2}\Big)\big\|\Delta x(\tilde{\omega})\big\|_X$$

and thereby directly obtain the second inequality from (22). The other way to manipulate (24) is to simply drop the right-hand side due to $(\omega + \kappa_1 + \kappa_2) > 0$. This immediately yields

$$(\tilde{\omega} - \omega)\big\|\Delta x(\tilde{\omega})\big\|_X^2 \leq (\tilde{\omega} - \omega)\mathcal{R}\big(\Delta x(\tilde{\omega}), \Delta x(\omega)\big)$$

where we use the Cauchy-Schwarz-Inequality and divide by $(\tilde{\omega}-\omega)\|\Delta x(\tilde{\omega})\|_X$ which again can be assumed to be non-zero (and positive) without loss of generality. The ensuing estimate then constitutes the first part of (22), completing the proof. □

With the relative error inexactness criterion (19) as well as the auxiliary results concerning regularized composite gradient mappings from Sect. 2 and norm estimates from Lemma 3 at hand, we can now tackle the proof of the following local acceleration result.

Theorem 1 *Suppose that at a stationary point $x_* \in X$ of (1) the semi-smoothness assumption (17) holds together with $\kappa_1(x_k)+\kappa_2 > 0$ at iterates x_k close to x_*. Then, the inexact Proximal Newton method with update steps computed according to (3) at x_k with the inexactness criterion (19) for $\eta_k \geq 0$ exhibits the following local convergence behavior:*

(a) The sequence of iterates locally converges linearly if ω_k and η_k are sufficiently small, i.e., if there exists some constant $0 < \Theta < 1$ and $k_0 \in \mathbb{N}$ such that for all $k \geq k_0$ the following estimate holds:

$$\frac{1}{\omega_k + \kappa_1(x_k) + \kappa_2}\Big[\big(\omega_k + \|H_{x_k}\|_{\mathcal{L}(X,X^*)}+\kappa_2\big)\eta_k + \omega_k\Big] < \Theta\,. \tag{25}$$

(b) The sequence of iterates locally converges superlinearly in case both ω_k and η_k converge to zero.

Proof For the sake of simplicity, we will omit the sequence indices of all quantities here and denote $x = x_k$, $\omega = \omega_k$ and $\eta = \eta_k$ for the current iterate, regularization parameter and forcing term. For the next iterate, we write $x_+(\omega) = x_{k+1}(\omega)$ and $H_x = H_{x_k}$ stands for the current second order bilinear form.

For what follows, we fix $\tau := \omega + \frac{1}{2}\big(\|H_x\|_{\mathcal{L}(X,X^*)}+\kappa_1(x)\big)$ for the gradient mapping regularization parameter which allows us to take advantage of the auxiliary estimates deduced in Lemma 1. Under these circumstances, the first part of (11) from Lemma 1 provides us with

$$\begin{aligned}\|x_+(\omega) - x_*\|_X &\leq \frac{1}{\tau(1-\mathcal{H})}\|G_\tau^{\hat{F}_{x,\omega}}(x+\Delta s(\omega)) - G_\tau^{\hat{F}_{x,\omega}}(x_*)\|_X \\ &\leq \frac{1}{\tau(1-\mathcal{H})}\Big[\|G_\tau^{\hat{F}_{x,\omega}}(x+\Delta s(\omega))\|_X + \|G_\tau^{\hat{F}_{x,\omega}}(x_*)\|_X\Big]\end{aligned} \tag{26}$$

where we abbreviated the constant $\mathcal{H} := \frac{\|H_x\|_{\mathcal{L}(X,X^*)}-\kappa_1(x)}{2(\tau+\kappa_2)}$.

As a next step, we take a look at the first norm term in brackets in (26). We use (14) from Proposition 2 together with the second part of (11) from Lemma 1 for $y := x + \Delta s(\omega)$ and $z := x + \Delta x(\omega)$ in order to obtain the following estimate:

$$\begin{aligned}\|G_\tau^{\hat{F}_{x,\omega}}\big(x+\Delta s(\omega)\big)\|_X &= \|G_\tau^{\hat{F}_{x,\omega}}\big(x+\Delta s(\omega)\big) - G_\tau^{\hat{F}_{x,\omega}}\big(x+\Delta x(\omega)\big)\|_X \\ &\leq \tau\big(1+\mathcal{H}\big)\|\Delta x(\omega) - \Delta s(\omega)\|_X\,.\end{aligned}$$

For the ensuing norm difference we take advantage of the relative error estimate inexactness criterion (19) together with the monotonicity of update step norms concerning the damping parameter ω as in Lemma 3. Additionally, the superlinear convergence for full update steps $\Delta x{:=}\Delta x(0)$ close to optimal solutions (cf. [15, Theorem 1]) is important here:

$$\left\|\Delta x(\omega) - \Delta s(\omega)\right\|_X \leq \eta \left\|\Delta x(\omega)\right\|_X \leq \eta \left\|\Delta x\right\|_X \leq o\big(\left\|x - x_*\right\|_X\big) + \eta \left\|x - x_*\right\|_X . \tag{27}$$

By the stationarity of x_* together with Lemma 2, for the second term in brackets in (26) we have

$$\left\|G_\tau^{\hat{F}_{x,\omega}}(x_*)\right\|_X = \left\|G_\tau^{\hat{F}_{x,\omega}}(x_*) - G_\tau^F(x_*)\right\|_X \leq o\big(\left\|x - x_*\right\|_X\big) + \frac{\omega\tau}{\tau + \kappa_2}\left\|x - x_*\right\|_X . \tag{28}$$

The estimates (27) and (28) suffice to quantify the influence of either inexactness or damping on local convergence rates of our algorithm. Inserting both of them into (26) above yields

$$\left\|x_+(\omega) - x_*\right\|_X \leq \frac{(1 + \mathcal{H})\eta + \frac{\omega}{\tau+\kappa_2}}{1 - \mathcal{H}}\left\|x - x_*\right\|_X + o\big(\left\|x - x_*\right\|_X\big) . \tag{29}$$

All that remains to do now is simplify the rather complicated prefactor term within the estimate above. We expand the fraction by $2(\tau + \kappa_2)$ and use that by the definition of τ we have

$$2(\tau + \kappa_2) = 2(\omega + \kappa_2) + \left\|H_x\right\|_{\mathcal{L}(X,X^*)} + \kappa_1\kappa_1(x) + \kappa_2 .$$

This provides us with

$$\begin{aligned}\frac{(1 + \mathcal{H})\eta + \frac{\omega}{\tau+\kappa_2}}{1 - \mathcal{H}} &= \frac{\big(2(\tau + \kappa_2) + \left\|H_x\right\|_{\mathcal{L}(X,X^*)} - \kappa_1\big)\eta + 2\omega}{2(\tau + \kappa_2) - \left\|H_x\right\|_{\mathcal{L}(X,X^*)} + \kappa_1(x) + \kappa_2} \\ &= \frac{\big(\omega + \left\|H_x\right\|_{\mathcal{L}(X,X^*)} + \kappa_2\big)\eta + \omega}{\omega + \kappa_1(x) + \kappa_2} .\end{aligned}$$

Inserting this identity to (29) now directly yields

$$\left\|x_+(\omega) - x_*\right\|_X \leq \frac{1}{\omega + \kappa_1(x) + \kappa_2}\big[\big(\omega + \left\|H_x\right\|_{\mathcal{L}(X,X^*)} + \kappa_2\big)\eta + \omega\big]\left\|x - x_*\right\|_X + o\big(\left\|x - x_*\right\|_X\big) . \tag{30}$$

Now, both of the asserted cases for local convergence behavior are an immediate consequence of (30) and the uniform boundedness of the H_x stated in (4). □

Remark 1 The estimate (25) yields a couple of algorithmically relevant insights. First, the linear convergence factor Θ can only be small, if both ω_k and η_k are small. Hence, computing steps very accurately does only pay off if ω_k is very small. We will see in Sect. 5 that close to optimal solutions arbitrarily small regularization parameters $\omega_k \approx 0$ can indeed be used.

Second, if we neglect $\omega_k \approx 0$, then (25) simplifies to

$$\frac{\left\|H_{x_k}\right\|_{\mathcal{L}(X,X^*)}+\kappa_2}{\kappa_1(x)+\kappa_2}\eta_k \leq \Theta,$$

where the prefactor on the left hand side can be interpreted as a local condition number of the problem. Indeed, for $\kappa_2 = 0$ it coincides with the condition number of H_x relative to $\|\cdot\|_X$. Thus, to achieve a given rate of local convergence, η_k has to be chosen the tighter the higher the condition number. This underlines the necessity of an adequate choice of function space X and norm $\|\cdot\|_X$.

Additionally, we were able to extend the local convergence result from [15, Theorem 1] insofar that we quantified the influence of damping update steps on (local) convergence rates. We are now also aware of more insightful criteria for linear or superlinear convergence of our method respectively. This helps us understand the process of local convergence of the (inexact) Proximal Newton method to an even greater extent.

4 Global convergence properties

Now that we have clarified the local convergence properties of our inexact Proximal Newton method depending on the forcing terms in criterion (18), we want to take into consideration whether the globalization scheme via the additional norm term in (3) still fulfills its purpose and yields some global convergence results.

4.1 Cauchy decrease steps and the subgradient model

In order to achieve such a result, we will introduce a second crucial criterion which the inexactly computed update steps $\Delta s_k(\omega_k)$ have to satisfy in order to be admissible for our method. It can be viewed as an adopted strategy from smooth trust region methods where rather cheap so-called Cauchy decrease steps are used to measure functional value descent for the actual update steps, cf. e.g. [4, Chapter 6].

There are several conceivable ways to define and compute such comparative Cauchy decrease steps. A canonical choice would be a simple Proximal Gradient step, i.e., the minimizer of the regularized linear model

$$\lambda^C_{x,\hat{\omega}}(\delta x) := f'(x)\delta x + \frac{\hat{\omega}}{2}\left\|\delta x\right\|_X^2 + g(x+\delta x) - g(x)\,,\ \delta x \in X\,.$$

As was the problem with evaluating the gradient mapping for our first inexactness criterion, also this procedure is as expensive as computing the exact Proximal Newton step right away in our general Hilbert space setting. Thus, the idea arises to find some comparative update step which we can compute with marginal effort in order to measure its functional value descent and then compare it to our inexact update step.

To this end, we define the subgradient model descent of F around $x \in X$ with respect to $\mu \in \partial_F g(x)$ and regularization parameter $\hat{\omega} > 0$ by

$$\lambda^{\mu}_{x,\hat{\omega}}(\delta x) := f'(x)\delta x + \mu\, \delta x + \frac{\hat{\omega}}{2}\|\delta x\|_X^2\,,\ \delta x \in X\,, \tag{31}$$

and we refer to the respective minimizer

$$\Delta x^{\mu}(\hat{\omega}) := \operatorname*{argmin}_{\delta x \in X} \lambda^{\mu}_{x,\hat{\omega}}(\delta x) \tag{32}$$

as the corresponding subgradient step. Before introducing the second inexactness criterion which makes use of the above model and step, we will establish an analytical connection between (31) and our initially defined regularized second order decrease model $\lambda_{x,\omega}$ from (9). To this end, we remember that the regularization parameter $\omega \geq 0$ is generally chosen such that the modified non-smooth part

$$\tilde{g}\colon X \to \mathbb{R} \quad , \quad \tilde{g}(x) := g(x) + \frac{1}{2}\big(H_x + \omega\mathcal{R}\big)(x)^2$$

is convex and thus the subproblem (3) allows for a unique solution. Consequently, the characterization of the convex subdifferential $\partial\tilde{g}(x)$ yields that for any $\tilde{\mu} = \mu + (H_x + \omega\mathcal{R})x \in \partial\tilde{g}(x)$ with $\mu \in \partial_F g(x)$ we have that

$$\tilde{g}(x+\delta x) \geq \tilde{g}(x) + \tilde{\mu}\,\delta x \quad \text{and thus} \quad g(x+\delta x) - g(x) + \frac{1}{2}H_x(\delta x)^2 + \frac{\omega}{2}\|\delta x\|_X^2 \geq \mu\,\delta x$$

holds for any $\delta x \in X$ and $\mu \in \partial_F g(x)$. We immediately obtain that

$$\begin{aligned}\lambda^{\mu}_{x,\hat{\omega}}(\delta x) &= f'(x)\delta x + \frac{\hat{\omega}}{2}\|\delta x\|_X^2 + \mu\delta x \\ &\leq f'(x)\delta x + \frac{1}{2}H_x(\delta x)^2 + \frac{\hat{\omega}+\omega}{2}\|\delta x\|_X^2 + g(x+\delta x) - g(x) = \lambda_{x,\hat{\omega}+\omega}(\delta x)\end{aligned} \tag{33}$$

is true for any $\delta x \in X$. In particular, this estimate apparently also holds for the respective minima of the decrease models of the composite objective function. For that reason, from (33) we obtain

$$\lambda^{\mu}_{x,\hat{\omega}}\big(\Delta x^{\mu}(\hat{\omega})\big) \leq \lambda_{x,\hat{\omega}+\omega}\big(\Delta x(\hat{\omega}+\omega)\big) \leq -\frac{1}{2}\big(\hat{\omega}+\omega+\kappa_1(x)+\kappa_2\big)\|\Delta x(\hat{\omega}+\omega)\|_X^2 \tag{34}$$

for any $\hat{\omega} > 0$ where the last estimate constitutes a result from the exact case in [15, Eq.(19)] and will give us norm-like descent in the objective functional later on. Obviously, we now want to link this norm-like decrease within the subgradient model to the regularized second order decrease model $\lambda_{x,\omega}\big(\Delta s(\omega)\big)$ for our inexactly computed update step $\Delta s(\omega)$ and lastly to the direct descent within the objective functional F.

4.2 Second inexactness criterion and efficient evaluation

We will establish the first one of these connections via the actual second inexactness criterion which will thus also be checked within our algorithm and implementation. For this purpose, it is sufficient if an inexactly computed update step $\Delta s(\omega)$ satisfies the estimate

$$\lambda_{x,\omega}\big(\Delta s(\omega)\big) \leq \lambda^{\mu}_{x,\tilde{\omega}}\big(\Delta x^{\mu}(\tilde{\omega})\big) \quad \text{for some} \quad \tilde{\omega} < \tilde{\omega}_{\max} \tag{35}$$

where the upper bound $\tilde{\omega}_{\max} > 0$ is an algorithmic parameter yet to be specified. This inequality now constitutes our formal second inexactness criterion which we will also refer to as the *subgradient inexactness criterion*.

Let us shortly elaborate on the efficient evaluation of this estimate and from there derive the actual implementation of the criterion: The solution property of $\Delta x^{\mu}(\tilde{\omega})$ provides us with first order conditions for the corresponding minimization problem in the form of

$$0 = f'(x) + \mu + \hat{\omega}\mathcal{R}\Delta x^{\mu}(\tilde{\omega})$$

and thus $\Delta x^{\mu}(\tilde{\omega}) = -(\tilde{\omega}\mathcal{R})^{-1}\big(f'(x) + \mu\big)$. For a given value of $\lambda_{x,\omega}\big(\Delta s(\omega)\big)$, i.e., descent along an inexactly computed update step within the regularized second order model, we can thus theoretically determine $\tilde{\omega}$ such that (35) is satisfied with equality. This can be seen as follows:

$$\begin{aligned}
\lambda_{x,\omega}\big(\Delta s(\omega)\big) \overset{!}{=} \lambda^{\mu}_{x,\tilde{\omega}}\big(\Delta x^{\mu}(\tilde{\omega})\big) &= \big(f'(x) + \mu\big)\Delta x^{\mu}(\tilde{\omega}) + \frac{\tilde{\omega}}{2}\big\|\Delta x^{\mu}(\tilde{\omega})\big\|_X^2 \\
&= \big(f'(x) + \mu\big)\big[-(\tilde{\omega}\mathcal{R})^{-1}\big(f'(x) + \mu\big)\big] + \frac{\tilde{\omega}}{2}\big\|-(\tilde{\omega}\mathcal{R})^{-1}\big(f'(x) + \mu\big)\big\|_X^2 \\
&= -\frac{1}{2\tilde{\omega}}\big\|f'(x) + \mu\big\|_{X^*}^2
\end{aligned} \tag{36}$$

which provides us with the theoretical value

$$\tilde{\omega} = -\frac{\big\|f'(x) + \mu\big\|_{X^*}^2}{2\lambda_{x,\omega}\big(\Delta s(\omega)\big)} \overset{!}{<} \tilde{\omega}_{\max} \tag{37}$$

for the regularization parameter within the subgradient minimization problem (32). This quantity should remain bounded in order to enable the proof of global convergence

results later on. Thus, as also pointed out in (37), we establish a sufficient estimate for our subgradient inexactness criterion (35) by demanding boundedness of $\tilde{\omega}$ from above by $\tilde{\omega}_{\max}$. Note here that—as can be seen in (36)—the value for $\lambda^{\mu}_{x,\tilde{\omega}}\big(\Delta x^{\mu}(\tilde{\omega})\big)$ increases as $\tilde{\omega}$ does. Since globalization mechanisms in general should only provide worst case estimates and not slow down the convergence of our algorithm, we want the subgradient inexactness criterion to only interfere with update step computation on rare occasions and thus choose $\tilde{\omega}_{\max}$ very large.

The dual norm occurring in the numerator of (37) is computed as follows: we compute the minimizer of the linear subgradient model $\Delta x^{\mu}(1) \in X$ from (32) and afterwards evaluate the linear functional $f'(x) + \mu \in X^*$ there. Here, the Fréchet-subdifferential element $\mu \in \partial_F g(x)$ is chosen such that the norm $\big\| f'(x) + \mu \big\|_{X^*}$ is as small as possible. Obviously, this depends on the specific minimization problem at hand but due to the non-smooth nature of g it is often possible to exploit the set-valued subdifferential for this purpose.

Let us add some remarks concerning satisfiability of the subgradient inexactness criterion: As mentioned above, the freedom of choice of μ within $\partial_F g(x)$ opens up possibilities to decrease the value of $\big\| f'(x) + \mu \big\|_{X^*}$ right away. Additionally, considering the exact case for update step computation is very insightful in order to see that the criterion will be fulfilled by late iterations of the inner solver. For now, we interpret $\big\| f'(x) + \mu \big\|_{X^*} \approx \operatorname{dist}\big(\partial_F F(x), 0\big)$, i.e., we assume $\mu \in \partial_F g(x)$ to be chosen (nearly) optimally for our purpose of finding solutions of (1).

Proposition 3 *Assume that there exists some constant $C > 0$ such that*

$$\big\| f'(x) + \mu \big\|_{X^*} \leq C \operatorname{dist}\big(\partial_F F(x), 0\big)$$

holds at some iterate $x_k =: x$ for $\mu \in \partial_F g(x)$. Then, the subgradient inexactness criterion (35) is eventually satisfied by iterates $\Delta s(\omega)$ of convergent solvers for the subproblem (3) in case

$$\tilde{\omega}_{\max} > \frac{C^2\big(\omega + L_f + \big\| H_x \big\|_{\mathcal{L}(X,X^*)}\big)^2}{\omega + \kappa_1(x) + \kappa_2} \tag{38}$$

holds for the upper bound $\tilde{\omega}_{\max}$ from (35).

Proof According to global convergence arguments in [15, Theorem 2] together with the assumed existence of $C > 0$ above, we can estimate

$$\big\| f'(x) + \mu \big\|_{X^*} \leq C \operatorname{dist}\big(\partial_F F(x), 0\big) \leq C\big(L_f + \big\| H_x \big\|_{\mathcal{L}(X,X^*)} + \omega\big) \big\| \Delta x(\omega) \big\|_X$$

for the exactly computed update step $\Delta x(\omega)$. Additionally, from [15, Eq.(19)] we infer that

$$\begin{aligned} \lambda_{x,\omega}(\Delta x(\omega)) &\leq -\frac{1}{2}\big(\omega + \kappa_1(x) + \kappa_2\big) \big\| \Delta x(\omega) \big\|_X^2 \;\Leftrightarrow\; -2\lambda_{x,\omega}(\Delta x(\omega)) \\ &\geq \big(\omega + \kappa_1(x) + \kappa_2\big) \big\| \Delta x(\omega) \big\|_X^2 \end{aligned}$$

is true in this scenario and we consequently obtain

$$\tilde{\omega} = -\frac{\left\| f'(x) + \mu \right\|_{X^*}^2}{2\lambda_{x,\omega}\big(\Delta x(\omega)\big)} \leq \frac{C^2\big(\omega + L_f + \left\| H_x \right\|_{\mathcal{L}(X,X^*)}\big)^2}{\omega + \kappa_1(x) + \kappa_2} < \infty\,. \tag{39}$$

Here, the convergence of the subproblem solver in the form that the respective objective value $\lambda_{x,\omega}\big(\Delta s(\omega)\big)$ tends to $\lambda_{x,\omega}\big(\Delta x(\omega)\big)$ from above comes into play. Thus, we can summarize

$$\tilde{\omega} = -\frac{\left\| f'(x) + \mu \right\|_{X^*}^2}{2\lambda_{x,\omega}\big(\Delta s(\omega)\big)} \underset{>}{\to} -\frac{\left\| f'(x) + \mu \right\|_{X^*}^2}{2\lambda_{x,\omega}\big(\Delta x(\omega)\big)} \leq \frac{C^2\big(\omega + L_f + \left\| H_x \right\|_{\mathcal{L}(X,X^*)}\big)^2}{\omega + \kappa_1(x) + \kappa_2}$$

for the theoretical value $\tilde{\omega}$ from (37). If now in particular the assumed estimate for the upper bound $\tilde{\omega}_{\max}$ holds, the assertion directly follows. □

Remark 2 The bound in (38) in particular remains finite in both limits $\omega \to 0$ and $x \to x_*$ for any stationary point $x_* \in X$ of problem (1) near which $\kappa_1(x) + \kappa_2 > 0$ uniformly holds.

The algorithmic strategy behind the subgradient inexactness criterion can now be summarized as follows: For the present iterate of the outer loop $x \in X$, we solve the linearized problem (32) for the computation of the dual norm $\left\| f'(x) + \mu \right\|_{X^*}$ and initiate the inner loop in order to determine the next inexact update step. At every iterate $\Delta s(\omega)$ of the inner solver for subproblem (3) we compute the corresponding subgradient regularization parameter $\tilde{\omega}$ from (37) and check $\tilde{\omega} < \tilde{\omega}_{\max}$. As a consequence of Proposition 3, either $\tilde{\omega}_{\max}$ is chosen large enough and we will eventually achieve $\tilde{\omega} < \tilde{\omega}_{\max}$ for some inexact step or we will compute an exact update step $\Delta x(\omega)$ which on its own provides us with global convergence of the sequence of iterates as presented in [15, Sect. 4].

4.3 Summary of inexactness criteria

With both of our inexactness criteria at hand, let us shortly reflect on their computational effort and compare it to possible alternatives: For the relative error criterion (19) in its form (20) only the evaluation of the fraction and its comparison to the forcing term is necessary since all occurring norms are already present within the subproblem solver. The subgradient inexactness criterion as described before requires the solution of the quadratic minimization problem (32) once per outer iteration of our method together with the evaluation of the quadratic model $\lambda_{x,\omega}\big(\Delta s(\omega)\big)$ at each inner iteration which is a cheap operation.

For comparative algorithms from literature, cf. [3, 9, 11], the gradient-like inexactness criterion (18) has to be assessed at every inner iteration together with one comparison of the second order decrease model value with its base value for $\delta x = 0$. As mentioned before, the former operation is very costly for non-diagonal function space norms, particularly in comparison to solving a linearized problem once per

outer iteration. This emphasizes both the necessity and the benefit of our adjustments to existing inexactness criteria. The summarized procedure can be retraced in the scheme of Algorithm 1.

4.4 Sufficient decrease criterion and global convergence

For global convergence in the case of inexactly computed update steps with the criteria introduced above we still have to carry out some more deliberations. The last missing ingredient in our recipe for norm-like descent within the composite objective functional is given by a sufficient decrease criterion which we have also used in the exact scenario in [15, Eq.(18)]. We say that an (inexactly computed) update step $\Delta s(\omega)$ is admissible for sufficient decrease if for some prescribed $\gamma \in]0, 1[$ the estimate

$$F\big(x + \Delta s(\omega)\big) - F(x) \leq \gamma \lambda_{x,\omega}\big(\Delta s(\omega)\big) \tag{40}$$

holds. Now, before justifying that (40) holds for sufficiently large values of the regularization parameter ω, let us combine estimates (40), (35), the monotonicity of $\lambda^{\mu}_{x,\tilde{\omega}}\big(\Delta x^{\mu}(\tilde{\omega})\big)$ with respect to $\tilde{\omega}$ as well as (34) from above and thus recognize that we obtain

$$\begin{aligned} F\big(x + \Delta s(\omega)\big) - F(x) &\leq \gamma \lambda_{x,\omega}\big(\Delta s(\omega)\big) = \gamma \lambda^{\mu}_{x,\tilde{\omega}}\big(\Delta x^{\mu}(\tilde{\omega})\big) \leq \gamma \lambda^{\mu}_{x,\tilde{\omega}+1}\big(\Delta x^{\mu}(\tilde{\omega}+1)\big) \\ &\leq -\frac{(\tilde{\omega} + \omega + 1 + \kappa_1(x) + \kappa_2)\gamma}{2} \big\|\Delta x(\tilde{\omega} + \omega + 1)\big\|_X^2 \\ &\leq -\frac{\gamma}{2} \big\|\Delta x(\tilde{\omega}_{\max} + \omega + 1)\big\|_X^2 . \end{aligned} \tag{41}$$

Note that we additionally used $\tilde{\omega} \geq 0$ and $\omega + \kappa_1(x) + \kappa_2 \geq 0$ as well as $\tilde{\omega} < \tilde{\omega}_{\max}$ together with the equivalence result from Lemma 3.

The following lemma ensures the satisfiability of the sufficient decrease criterion (40) as soon as ω is large enough.

Lemma 4 *The sufficient decrease criterion* (40) *is fulfilled by inexactly computed update steps* $\Delta s(\omega_+)$ *which additionally satisfy the inexactness criteria* (19) *and* (35) *if the regularization parameter* ω *satisfies the inequality*

$$\frac{1-\gamma}{(1+\eta)^2}(\omega + \kappa)^2 + \omega(\omega + \tilde{\omega}_{\max} + \kappa - L) \geq L(\tilde{\omega}_{\max} + \kappa)$$

where we abbreviated $\kappa := \kappa_1(x) + \kappa_2$ *and* $L := L_f - \kappa_1(x)$.

Proof The first inexactness criterion (19) provides us with the norm estimate

$$\big\|\Delta s(\omega)\big\|_X \leq \big\|\Delta s(\omega) - \Delta x(\omega)\big\|_X + \big\|\Delta x(\omega)\big\|_X \leq (1+\eta)\big\|\Delta x(\omega)\big\|_X . \tag{42}$$

With the aid of the second inexactness criterion (35), (34) and Lemma 3 we thus obtain

$$\begin{aligned}\lambda_{x,\omega}\big(\Delta s(\omega)\big) &= \lambda^{\mu}_{x,\tilde{\omega}}\big(\Delta x^{\mu}(\tilde{\omega})\big) \leq \lambda^{\mu}_{x,\tilde{\omega}_{\max}}\big(\Delta x^{\mu}(\tilde{\omega}_{\max})\big) \\ &\leq -\frac{1}{2}\big(\tilde{\omega}_{\max}+\omega+\kappa_1(x)+\kappa_2\big)\big\|\Delta x(\tilde{\omega}_{\max}+\omega)\big\|_X^2 \\ &\leq -\frac{\big(\omega+\kappa_1(x)+\kappa_1\big)^2}{2(1+\eta)^2\big(\tilde{\omega}_{\max}+\omega+\kappa_1(x)+\kappa_2\big)}\big\|\Delta s(\omega)\big\|_X^2 .\end{aligned} \tag{43}$$

Here, we recognize that the inequality from the assertion is equivalent to

$$\frac{L_f-\kappa_1-\omega}{2}\cdot\frac{2\big(\tilde{\omega}_{\max}+\omega+\kappa_1(x)+\kappa_2\big)(1+\eta)^2}{\big(\omega+\kappa_1(x)+\kappa_2\big)^2}\leq 1-\gamma$$

which together with (43) lets us estimate

$$\begin{aligned}F\big(x_+(\omega)\big)-F(x) &\leq f'(x)\Delta s(\omega)+\frac{L_f}{2}\big\|\Delta s(\omega)\big\|_X^2+g\big(x_+(\omega)\big)-g(x) \\ &\leq \lambda_{x,\omega}\big(\Delta s(\omega)\big)+\frac{1}{2}(L_f-\kappa_1(x)-\omega)\big\|\Delta s(\omega)\big\|_X^2 \\ &\leq \lambda_{x,\omega}\big(\Delta s(\omega)\big)-(1-\gamma)\lambda_{x,\omega}\big(\Delta s(\omega)\big)=\gamma\lambda_{x,\omega}\big(\Delta s(\omega)\big)\end{aligned}$$

and conclude that $\Delta s(\omega)$ yields sufficient decrease according to (40). □

Remark 3 The above result together with the assumption (5) and (6) on our objective functional also imply that the regularization parameter ω remains bounded over the course of the minimization process.

Let us now deduce the ensuing global convergence results for the inexact Proximal Newton method as presented in the scheme of Algorithm 1.

For this reason, we will first prove that the right-hand side of (41), i.e., the norm of exactly computed comparative steps $\Delta x(\tilde{\omega}_{\max}+\omega+1)$, converges to zero along the sequence of iterates generated by inexact updates. Here, it will come in handy to define $\omega^c:=\tilde{\omega}_{\max}+\omega+1$ for the regularization parameter of the comparative exact update steps. Note that this quantity is bounded both from above and below.

Lemma 5 *Let $(x_k)\subset X$ be the sequence generated by the inexact Proximal Newton method globalized via (3) starting at any $x_0\in X$. Additionally, suppose that the subgradient inexactness criterion (35) and the sufficient decrease criterion (40) are satisfied for all $k\in\mathbb{N}$. Then either $F(x_k)\to-\infty$ or $\big\|\Delta x_k(\omega^c)\big\|_X\to 0$ for $k\to\infty$.*

Proof By (41) the sequence $F(x_k)$ is monotonically decreasing. Thus, either $F(x_k)\to-\infty$ or $F(x_k)\to\underline{F}$ for some $\underline{F}\in\mathbb{R}$ and thereby in particular $F(x_k)-F(x_{k+1})\to 0$. As a consequence of (41), then also $\big\|\Delta x_k(\omega^c)\big\|_X\to 0$ holds. □

Note that the above result does not comprise the convergence of the sequence of iterates itself which is desirable in the context. In the exact case of update step computation

Data: Starting point $x_0 \in X$, sufficient decrease parameter $\gamma \in]0, 1[$, initial values ω_0 and $\eta_0, \varepsilon > 0$ for stopping criterion
Initialization: $k = 0$;
while $\frac{1+\omega_k}{1-\eta_k} \left\| \Delta s_k(\omega_k) \right\|_X \geq \varepsilon$ **do**
  Choose $\mu \in \partial_F g(x_k)$ and compute norm term for $\tilde{\omega}$ as in (37) via the linearized minimization problem (32);
  Compute a trial step $\Delta s_k(\omega_k)$ according to (3) which suffices the inexactness criteria (20) and (37);
  while *Sufficient decrease criterion* (40) *is not satisfied* **do**
    Increase ω_k appropriately;
    Recompute $\Delta s_k(\omega_k)$ as above;
  end
  Update current iterate to $x_{k+1} \leftarrow x_k + \Delta s_k(\omega_k)$;
  Decrease ω_k to some $\omega_{k+1} < \omega_k$ for next iteration;
  Decrease η_k to some $\eta_{k+1} < \eta_k$ for next iteration;
  Update $k \leftarrow k + 1$;
end

Algorithm 1: Inexact second order semi-smooth Proximal Newton Algorithm

it was possible to take advantage of first order optimality conditions of the exactly solved subproblem for the actual update steps and from there achieve a proper global convergence result at least in the strongly convex case, cf. [15, Theorem 3]. Due to the presence of inexactness in the update step computation this strategy has to be slightly adjusted in the current scenario, i.e., applied to the comparative update steps $\Delta x(\omega^c)$. To this end, for some $k \in \mathbb{N}$ and iterate $x_k \in X$ we introduce the so-called corresponding comparative iterate

$$y_k := x_k + \Delta x_k(\omega^c) = \mathcal{P}_g^{H_{x_k}+\omega^c\mathcal{R}}\left((H_{x_k} + \omega^c\mathcal{R})x_k - f'(x_k)\right). \tag{44}$$

Note here that the comparative iterate uses a theoretical exact update but origins at the iterate x_k which belongs to our inexact method. Also, for every $k \in \mathbb{N}$ the identity $y_k - x_k = \Delta x_k(\omega^c)$ holds by definition of y_k.

With this definition at hand, we are in the position to discuss at least subsequential convergence of our algorithm to a stationary point. In the following, we will assume throughout that the sequence of objective values $\left(F(x_k)\right)$ is bounded from below. We start with the case of convergence in norm:

Theorem 2 *Assume that the subgradient inexactness criterion* (35) *and the sufficient decrease criterion* (40) *are fulfilled. Then, all accumulation points* $\bar{x}$ *(in norm) of the sequence of iterates* (x_k) *generated by the inexact Proximal Newton method globalized via* (3) *are stationary points of problem* (1)*. Let now* $(x_{k_l}) \subset (x_k)$ *be the subsequence converging to* $\bar{x}$*. In particular, the corresponding comparative subsequence* (y_{k_l}) *defined via* (44) *satisfies*

$$\operatorname{dist}\left(\partial_F F(y_{k_l}), 0\right) \to 0 \quad \textit{and} \quad \left\| x_{k_l} - y_{k_l} \right\|_X \to 0,$$

i.e., also $y_{k_l} \to \bar{x}$ *for* $l \to \infty$.

Proof We simplify notation by referring to subsequence indices k_l as k. As mentioned beforehand, for the corresponding comparative sequence (y_k) we have $y_k - x_k = \Delta x_k(\omega^c)$ and consequently also $y_k \to \bar{x}$ holds by $\|\Delta x_k(\omega^c)\|_X \to 0$ due to Lemma 5. The proximal representation of y_k in (44) is equivalent to the minimization problem

$$y_k = \operatorname*{argmin}_{y \in X} g(y) + \frac{1}{2}\big(H_{x_k} + \omega^c \mathcal{R}\big)(y)^2 - \big((H_{x_k} + \omega^c \mathcal{R})x_k - f'(x_k)\big)y$$

which yields the first order optimality conditions given by the dual space inclusion

$$0 \in \partial_F g(y_k) + f'(x_k) + \big(H_{x_k} + \omega^c \mathcal{R}\big)(y_k - x_k)\,.$$

This, on the other hand, is equivalent to

$$\big(H_{x_k} + \omega^c \mathcal{R}\big)(x_k - y_k) + f'(y_k) - f'(x_k) \in \partial_F g(y_k) + f'(y_k) = \partial_F F(y_k) \quad (45)$$

the remainder term on the left-hand side of which we can estimate via

$$\begin{aligned}\big\|\big(H_{x_k} + \omega^c \mathcal{R}\big)(x_k - y_k) + f'(y_k) - f'(x_k)\big\|_{X^*} &\le \big(M + \omega^c + L_f\big)\|x_k - y_k\|_X \\ &= \big(M + \omega^c + L_f\big)\|\Delta x_k(\omega^c)\|_X \to 0\end{aligned}$$

for $k \to \infty$ where M denotes the uniform bound on the second order bilinear form norms from assumption (4).

In order to now achieve the optimality assertion of the accumulation point $\bar{x}$, we have to slightly adjust (45) for the use of the convex subdifferential and its direct characterization. To this end, we consider a bilinear form $Q : X \times X \to \mathbb{R}$ such that the function $\tilde{g} : X \to \mathbb{R}$ defined via $\tilde{g}(x) := g(x) + \frac{1}{2}Q(x)^2$, $x \in X$, is convex. As above, $Q := H_{x_k} + \omega_k \mathcal{R}$ is a reasonable choice. Inserting a $Q(y_k)$-term into (45) thus yields

$$\omega^c \mathcal{R}(x_k - y_k) + f'(y_k) - f'(x_k) \in \partial \tilde{g}(y_k) + \{f'(y_k) - Q(y_k)\}$$

for the convex subdifferential of $\tilde{g}$. The left-hand side now as before converges to zero in X^* and consequently, we know that for every $k \in \mathbb{N}$ there exists some $\tilde{\rho}_k \in \partial \tilde{g}(y_k)$ such that we can define $\tilde{\rho} := \lim_{k\to\infty} \tilde{\rho}_k = -f'(\bar{x}) + Q\bar{x}$ by the convergence of also y_k to $\bar{x}$. The lower semi-continuity of $\tilde{g}$ together with the definition of the convex subdifferential $\partial \tilde{g}$ directly yields

$$\begin{aligned}\tilde{g}(u) - \tilde{g}(\bar{x}) &= \tilde{g}(u) - g(\bar{x}) - \frac{1}{2}Q(\bar{x})^2 \ge \tilde{g}(u) - \liminf_{k\to\infty} g(y_k) - \lim_{k\to\infty} \frac{1}{2}Q(y_k)^2 \\ &= \liminf_{k\to\infty} \tilde{g}(u) - \tilde{g}(y_k) \ge \liminf_{k\to\infty} \tilde{\rho}_k(u - y_k) = \lim_{k\to\infty} \tilde{\rho}_k(u - y_k) = \tilde{\rho}(u - \bar{x})\end{aligned}$$

for any $u \in X$ which proves the inclusion $\tilde{\rho} \in \partial \tilde{g}(\bar{x})$. The evaluation of the latter limit expression can easily be retraced by splitting

$$\tilde{\rho}_k(u - y_k) = \tilde{\rho}_k(u - \bar{x}) + (\tilde{\rho}_k - \tilde{\rho})(\bar{x} - y_k) + \tilde{\rho}(\bar{x} - y_k)\,. \quad (46)$$

In particular, we recognize $\tilde{\rho} \in \partial\tilde{g}(\bar{x})$ as $-f'(\bar{x}) + Q\bar{x} \in \partial\tilde{g}(\bar{x})$ and equivalently $-f'(\bar{x}) \in \partial_F g(\bar{x})$ for the Frechét-subdifferential ∂_F. This implies $0 \in \partial_F F(\bar{x})$, i.e., the stationarity of our accumulation point $\bar{x}$. □

In [15] the criterion $\left\|\Delta x_k(\omega_k)\right\|_X \leq \varepsilon$ for some small $\varepsilon > 0$ has been used as a condition for the optimality of the current iterate up to some prescribed accuracy. Estimate (42) from above thus yields that also the norm of the inexactly computed update steps can be used as an optimality measure for the current iterate within our method.

However, small step norms $\left\|\Delta s_k(\omega_k)\right\|_X$ can also occur due to very large values of the damping parameter ω_k as a consequence of which the algorithm would stop even though the sequence of iterates is not even close to an optimal solution of the problem. In order to rule out this inconvenient case and incorporate the influence of inexactness, we consider the scaled version $\frac{1+\omega_k}{1-\eta_k}\left\|\Delta s_k(\omega_k)\right\|_X$ as the stopping criterion in the later implementations of our algorithm.

Let us now proceed to generalizing the convergence result from Theorem 2: While bounded sequences in finite dimensional spaces always have convergent subsequences, we can only expect *weak subsequential convergence* in general Hilbert spaces in this case. As one consequence, existence of minimizers of non-convex functions on Hilbert spaces can usually only be established in the presence of some compactness. On this count, we note that in (46) even weak convergence of $x_k \rightharpoonup \bar{x}$ would be sufficient. Unfortunately, in the latter case we cannot evaluate $f'(x_k) \to f'(\bar{x})$. In order to extend our proof to this situation, we require some more structure for both of the parts of our composite objective functional. The proof is completely analogous to the one of [15, Theorem 3].

Theorem 3 *Let f be of the form $f(x) = \hat{f}(x) + \check{f}(Kx)$ where K is a compact operator. Additionally, assume that $g + \hat{f}$ is convex and weakly lower semi-continuous in a neighborhood of stationary points of (1). Suppose that $\check{f}$ satisfies the assumptions made on f beforehand. Then weak convergence of the sequence of iterates $x_k \rightharpoonup \bar{x}$ suffices for $\bar{x}$ to be a stationary point of (1).*

If F is strictly convex and radially unbounded, the whole sequence (x_k) converges weakly to the unique minimizer x_* of F. If F is κ-strongly convex, with $\kappa > 0$, then $x_k \to x_*$ in norm.

5 Transition to local convergence

In order to now benefit from the local acceleration result in Theorem 1, we have to manage the transition from the globalization phase above to the local convergence phase described beforehand. To this end, we have to make sure that (at least close to stationary points of (1)) arbitrarily small regularization parameters $\omega \geq 0$ yield update steps that give us sufficient decrease in F according to the criterion formulated in (40). This endeavor has also been part of the investigation of the exact case in [15, Section 6] but as for all aspects of our convergence analysis has to be slightly adapted here.

As a starting point, a rather technical auxiliary result is required. It sets the limit behavior of inexact update steps in relation with the distance of consecutive iterates to the minimizer of (1).

Lemma 6 *Let x and $x_+(\omega) = x + \Delta s(\omega)$ be two consecutive iterates with update step $\Delta s(\omega)$ sufficing (19) for some $0 \leq \eta < 1$. Furthermore, consider a stationary point x_* of (1). Then the following estimates eventually hold for $\kappa_1 + \kappa_2 > 0$:*

$$\|x_+(\omega) - x_*\|_X \leq (3 + 2\eta)\|x - x_*\|_X \quad , \quad \|x - x_*\|_X \leq \frac{2}{1-\eta}\big(1 + \frac{\omega}{\kappa_1 + \kappa_2}\big)\|\Delta s(\omega)\|_X .$$

Proof Our proof here mainly exploits the local superlinear convergence of exactly computed and undamped update steps $\Delta x{:=}\Delta x(0)$ from [15, Theorem 1] and then uses the respective estimates in order to introduce the influences of both damping and inexactness. For the first asserted estimate, we take a look at

$$\begin{aligned}\|x_+(\omega) - x_*\|_X &\leq \|x - x_*\|_X + \|\Delta s(\omega)\|_X \leq \|x - x_*\|_X + (1+\eta)\|\Delta x\|_X \\ &\leq (2+\eta)\|x - x_*\|_X + (1+\eta)\|x + \Delta x - x_*\|_X\end{aligned}$$

where the second step involved (42) together with $\|\Delta x(\omega)\|_X \leq \|\Delta x\|_X$ as proven in Lemma 3. From here, we use the superlinear convergence of exact updates in the form of the existence of some function $\psi : [0, \infty[\to [0, \infty[$ with $\psi(t) \to 0$ for $t \to 0$ such that

$$\|x + \Delta x - x_*\|_X = \psi\big(\|x - x_*\|_X\big)\|x - x_*\|_X$$

holds in the limit of $x \to x_*$. Thus, we obtain

$$\|x_+(\omega) - x_*\|_X \leq \big[(2+\eta) + (1+\eta)\psi\big(\|x - x_*\|_X\big)\big]\|x - x_*\|_X \leq (3 + 2\eta)\|x - x_*\|_X$$

since eventually we can assume the ψ-term to be smaller than one. This completes the proof of the first asserted estimate.

For the second one we take advantage of

$$\|\Delta x\|_X \leq \big(1 + \frac{\omega}{\kappa_1(x) + \kappa_2}\big)\|\Delta x(\omega)\|_X$$

from Lemma 3 together with again the superlinear convergence as above and find that

$$\begin{aligned}\|x - x_*\|_X &\leq \|x + \Delta x - x_*\|_X + \|\Delta x\|_X \leq \psi\big(\|x - x_*\|_X\big)\|x - x_*\|_X \\ &\quad + \big(1 + \frac{\omega}{\kappa_1(x) + \kappa_2}\big)\|\Delta x(\omega)\|_X\end{aligned}$$

holds. Since the ψ-term eventually will be smaller than $\frac{1}{2}$, from here we infer

$$\|x - x_*\|_X \leq \frac{1 + \frac{\omega}{\kappa_1(x) + \kappa_2}}{1 - \psi\big(\|x - x_*\|_X\big)}\|\Delta x(\omega)\|_X \leq 2\big(1 + \frac{\omega}{\kappa_1(x) + \kappa_2}\big)\|\Delta x(\omega)\|_X .$$

The inexactness of update step computation now enters the above estimate using the inequality $\|\Delta x(\omega)\|_X \leq \frac{1}{1-\eta}\|\Delta s(\omega)\|_X$ which can easily be retraced via

$$\begin{aligned}(1-\eta)\|\Delta x(\omega)\|_X &\leq \|\Delta x(\omega)\|_X - \|\Delta x(\omega) - \Delta s(\omega)\|_X \\ &\leq \|\Delta x(\omega) - \big(\Delta x(\omega) - \Delta s(\omega)\big)\|_X = \|\Delta s(\omega)\|_X\end{aligned}$$

with the inexactness criterion (19). This completes the proof of the lemma. □

Remark 4 In particular, these eventual norm estimates have implications on the limit behavior of the respective terms. If we now have $\xi = o\big(\|x_+(\omega) - x_*\|_X\big)$ for some $\xi \in X, \xi = o\big(\|x - x_*\|_X\big)$ immediately holds and from there we obtain $\xi = o\big(\|\Delta s(\omega)\|_X\big)$ in the same way.

In what follows, it will be important several times that the second order bilinear forms H_x satisfy a bound of the form

$$\big(H_{x_+(\omega)} - H_x\big)\big(x_+(\omega) - x_*\big)^2 = o\big(\|x - x_*\|_X^2\big) \text{ for } x \to x_*\,. \tag{47}$$

It is easy to see that the bound holds if either we have uniform boundedness of the second order bilinear forms together with superlinear convergence of the iterates or if we have continuity of the mapping $x \mapsto H_x$ together with mere convergence of the iterates to x_*. Note here that the same assumption has been made in the exact case in [15] for the admissibility of undamped and arbitrarily weakly damped update steps. In our scenario, we conclude that according to Theorem 1 it is sufficient that both the regularization parameters $\omega_k \geq 0$ and the forcing terms $\eta_k \geq 0$ converge to zero as we approach a stationary point $x_* \in X$ of (1) together with assumption (4) from the introductory section. We will later on establish this convergence of (ω_k) and (η_k) in the specific implementation of our algorithm.

With the auxiliary estimates from Lemma 6 and Lemma 3 together with the thoroughly discussed additional assumption from (47) at hand, we can now turn our attention to the actual admissibility of arbitrarily small update steps close to stationary points of (1).

Assumption 6 For that matter, we furthermore suppose f to be second order semi-smooth at stationary points x_* of (1) with respect to the mapping $H : X \to \mathcal{L}(X, X^*), x \mapsto H_x$, which expresses itself via the estimate

$$f(x_* + \xi) = f(x_*) + f'(x_*)\xi + \frac{1}{2}H_{x_*+\xi}(\xi, \xi) + o\big(\|\xi\|_X^2\big) \quad \text{for} \quad \|\xi\|_X \to 0\,. \tag{48}$$

This notion generalizes second order differentiability in our setting but its definition slightly differs from semi-smoothness of f' as qualified in (17). For further elaborations on this concept of differentiability, consider [15, Sect. 5].

Proposition 4 *Suppose that the additional assumptions* (47) *and* (48) *hold. Furthermore, assume that the update steps* $\Delta s(\omega)$ *computed as inexact solutions of* (3) *at*

iterates $x \in X$ for some $\omega \geq 0$ satisfy the inexactness criteria (19) for $\eta \geq 0$ and (35). Then, $\Delta s(\omega)$ is admissible for sufficient decrease according to (40) for any $\gamma < 1$ if x is sufficiently close to a stationary point $x_* \in X$ of (1) near which $\kappa_1(x) + \kappa_2 > 0$ holds.

Proof We take a look back at the proof of [15, Proposition 8] and employ the same telescoping strategy in order to obtain

$$\begin{aligned} &f(x_+(\omega)) - f(x) - f'(x)\Delta s(\omega) - \frac{1}{2}H_x(\Delta s(\omega))^2 \\ &= \left[f(x_+(\omega)) - f(x_*) - f'(x_*)(x_+(\omega) - x_*) - \frac{1}{2}H_{x_+(\omega)}(x_+(\omega) - x_*)^2\right] \\ &\quad - \left[f(x) - f(x_*) - f'(x_*)(x - x_*) - \frac{1}{2}H_x(x - x_*)^2\right] \\ &\quad - \left[(f'(x) - f'(x_*))\Delta s(\omega) - H_x(x - x_*, \Delta s(\omega))\right] + \frac{1}{2}(H_{x_+(\omega)} - H_x)(x_+(\omega) - x_*)^2 \end{aligned}$$

where again we can use the second order semi-smoothness of f according to (48) for the first two terms as well as the semi-smoothness of f' as in (17) for the third one. This implies

$$\begin{aligned} f(x_+(\omega)) - f(x) - f'(x)\Delta s(\omega) - \frac{1}{2}H_x(\Delta s(\omega))^2 = o(\|x_+(\omega) - x_*\|^2) + o(\|x - x_*\|_X^2) \\ + o(\|x - x_*\|_X)\|\Delta s(\omega)\|_X + \rho(x, \omega) \end{aligned}$$

where we denoted $\rho(x, \omega) := \frac{1}{2}(H_{x_+(\omega)} - H_x)(x_+(\omega) - x_*)^2$. Due to the limit behavior of inexact update step norms investigated over the course of Lemma 6 this yields

$$f(x + \Delta s(\omega)) - f(x) - f'(x)\Delta s(\omega) - \frac{1}{2}H_x\big(\Delta s(\omega)\big)^2 = \rho(x, \omega) + o\big(\|\Delta s(\omega)\|_X^2\big). \tag{49}$$

As the next step towards the admissibility result, we define the prefactor function

$$\gamma(x, \omega) := \frac{F\big(x + \Delta s(\omega)\big) - F(x)}{\lambda_{x,\omega}\big(\Delta s(\omega)\big)}$$

which should be larger than some $\tilde{\gamma} \in]0, 1[$ for $\Delta s(\omega)$ to yield sufficient decrease according to (40). Thus, it suffices to show the convergence of $\gamma(x, \omega)$ to anything greater equal than one for any $\omega \geq 0$ in the limit of $x \to x_*$. The identity (49) from above now provides us with

$$F\big(x + \Delta s(\omega)\big) - F(x) = \lambda_{x,\omega}(\Delta s(\omega)) - \frac{\omega}{2}\|\Delta s(\omega)\|_X^2 + \rho(x, \omega) + o\big(\|\Delta s(\omega)\|_X^2\big)$$

which we insert into the prefactor function from above and estimate

$$\begin{aligned}\gamma(x,\omega) &= 1 + \frac{-\frac{\omega}{2}\|\Delta s(\omega)\|_X^2 + \rho(x,\omega) + o(\|\Delta s(\omega)\|_X^2)}{\lambda_{x,\omega}\big(\Delta s(\omega)\big)} \\ &= 1 + \frac{\frac{\omega}{2}\|\Delta s(\omega)\|_X^2 + o(\|\Delta s(\omega)\|_X^2) - \rho(x,\omega)}{|\lambda_{x,\omega}\big(\Delta s(\omega)\big)|}\end{aligned} \tag{50}$$

since from the computation strategy for $\Delta s(\omega)$ we in particular have

$$\lambda_{x,\omega}\big(\Delta s(\omega)\big) \leq \lambda^{\mu}_{x,\tilde{\omega}}\big(\Delta x^{\mu}(\tilde{\omega})\big) \leq -\frac{1}{2}\big\|\Delta x(\omega^c)\big\|_X^2 \leq 0 \tag{51}$$

following the later steps of (41). For the absolute value of the second order decrease model we can use (51) together with Lemma 3 and (42) to obtain

$$\begin{aligned}|\lambda_{x,\omega}\big(\Delta s(\omega)\big)| \geq |\lambda^{\mu}_{x,\tilde{\omega}}\big(\Delta x^{\mu}(\tilde{\omega})\big)| \geq \frac{1}{2}\big\|\Delta x(\omega^c)\big\|_X^2 \geq \frac{1}{2}\Big(\frac{\omega+\kappa_1(x)+\kappa_2}{\omega^c+\kappa_1(x)+\kappa_2}\Big)^2\big\|\Delta x(\omega)\big\|_X^2 \\ \geq \frac{1}{2}\Big(\frac{\omega+\kappa_1(x)+\kappa_2}{(1+\eta)(\omega^c+\kappa_1(x)+\kappa_2)}\Big)^2\big\|\Delta s(\omega)\big\|_X^2 =: C\big\|\Delta s(\omega)\big\|_X^2\end{aligned} \tag{52}$$

where $C = C(\omega, \omega^c, \kappa_1(x)+\kappa_2, \eta) > 0$ denotes the constant from above. In particular, note that C remains bounded in the limit of $\omega \to 0$ and is also well-defined in the limit case of $\omega = 0$ close to stationary points x_* with $\kappa_1(x) + \kappa_2 > 0$ for x near x_*.

We may assume that the numerator of the latter expression in (50) is non-positive, otherwise the desired inequality for $\gamma(x, \omega)$ is trivially fulfilled. Thus, we take advantage of (52) in order to decrease the positive denominator to achieve

$$\gamma(x,\omega) \geq 1 + \frac{\omega}{2C} - \varepsilon - \frac{\rho(x,\omega)}{C\big\|\Delta s(\omega)\big\|_X^2}$$

where for any $\varepsilon > 0$ there exists a neighborhood of the optimal solution x_* such that the above estimate holds.

Now, the assumption (47) for the ρ-term immediately implies the eventual admissibility of $\Delta s(\omega)$ for sufficient decrease according to (40). □

5.1 Numerical results

Let us now showcase the functionality of our inexact Proximal Newton method and also compare its performance to the case of exact computation of update steps which have been investigated in [15]. In order to make the influence of inexactness more clearly visible, we have decided to enhance the function space problem from there such that update step subproblems are harder to solve and thus it takes more TNNMG steps in order to find an exact solution.

The Objective Functional. To this end, we now consider the following function space problem on $\Omega := [0,1]^3 \subset \mathbb{R}^3$: Instead of finding a scalar function, we expanded the problem to finding a vector field

$$u \in H^1_{\Gamma_D}(\Omega, \mathbb{R}^3) := \left\{ v \in H^1(\Omega, \mathbb{R}^3) \,\middle|\, v = 0 \text{ on } \Gamma_D \right\}$$

where the Dirichlet boundary is given by $\Gamma_D := \{0\} \times [0,1] \times [0,1]$. The solution which we are looking for minimizes the composite objective functional F defined via

$$F(u) := f(u) + \int_\Omega c \|u\|_2 \, \mathrm{d}x \tag{53}$$

for again some parameter $c > 0$ as a weight of the Euclidean L_2-norm term where the smooth part $f\colon H^1_{\Gamma_D}(\Omega, \mathbb{R}^3) \to \mathbb{R}$ is now given by

$$f(u) := \int_\Omega \frac{1}{2} \|\nabla u\|_F^2 + \alpha \max\left(\|\nabla u\|_F - 1, 0 \right)^2 + \beta \frac{u_1^3 u_2^2 u_3}{1 + u_1^2 + u_2^2 + u_3^2} + \rho \cdot u \, \mathrm{d}x$$

with parameters $\alpha, \beta \in \mathbb{R}$ as well as a force field $\rho\colon \Omega \to \mathbb{R}^3$. The norm $\|\cdot\|_F$ denotes the Frobenius norm of the respective Jacobian matrices ∇u.

We note that f technically does not satisfy the assumptions made on the smooth part of the composite objective functional specified above in the case $\alpha \neq 0$ due to the lack of semi-smoothness of the corresponding squared max-term with gradient arguments. However, we think that slightly going beyond the framework of theoretical results for numerical investigations can be instructive.

We will choose the force-field ρ to be constant on Ω and to this end introduce the so-called load factor $\tilde{\rho} > 0$ which then determines $\rho = \tilde{\rho}(1,1,1)^T$. Again, for the sake of simplicity, we will refer to this load factor as ρ. Now that we have fully prescribed the composite objective functional F, we recognize that its non-smooth part g is again merely given by the integrated Euclidean L_2-norm term with constant prefactor $c > 0$.

Specifics on the Implementation. We use automatic differentiation by adol-C in order to establish the second order model and TNNMG to solve update step computation subproblems, cf. [22]. Additionally, the subproblem solver is provided with stopping criteria in the form of our inexactness criteria (20) and (37) with corresponding parameters $\eta_k \in [0, 1[$ for each iteration and global $\tilde{\omega}_{\max} > 0$.

Another topic of interest concerning the implementation of our algorithm is the choice of the aforementioned parameters ω, η and $\tilde{\omega}_{\max}$ governing the convergence behavior of our method. While – as discussed in its introduction in (37) – $\tilde{\omega}_{\max}$ can be chosen constant and is supposed to be very large, this is not the case for the regularization parameters ω and the forcing terms η. For ω, we use the heuristic approach of doubling it in case the sufficient decrease criterion is not fulfilled and multiplying it by $\frac{1}{2^n}$ if the update has been accepted. Here, $n \in \mathbb{N}$ denotes the number of subsequent successful update steps

Similarly, we multiply the forcing term η by 0.6 for accepted updates and leave it as it is in case the increment has been rejected by the sufficient decrease criterion. These

rather simple strategies for the choice of parameters ensure the convergence of both η and ω to zero along the sequence of iterates and thus also from a theoretical standpoint enable superlinear convergence as formulated in Theorem 1. For the constant determining the subgradient inexactness criterion, we decided to choose $\tilde{\omega}_{\max} = 10^{10}$.

The stopping criterion for our algorithm takes into account both the regularization parameter and the forcing term. The respective threshold value is given by $\varepsilon = 10^{-10}$.

Test scenarios and test machine. Firstly, we will demonstrate the consistency between results of the inexact method and the exact version the functionality of which has been thoroughly investigated in [15]. There, also the superiority of Proximal Newton approaches in comparison with common first order methods like Proximal Gradient and FISTA has been emphasized. In our case, exactly computing update steps means neglecting the additionally introduced inexactness criteria, and computing steps up to numerical accuracy in TNNMG. We remember that there a relative norm threshold for increments is considered as a stopping criterion.

Afterwards, we exhibit the gains in effectiveness by enhancing the exact algorithm with the inexactness criteria introduced above. Lastly, we analyze the implementation of the latter criteria and try to get a grasp on how they affect the process of solving the subproblem for update step computation. All results within the current section have been computed after conducting three uniform grid refinements of the cubical domain Ω which results in $8^4 = 4096$ grid elements.

Furthermore, all tests are executed single-threaded on a Intel(R) Core(TM) i5-8265U CPU with clock frequency fixed to 1600 Mhz in order to avoid overheating and to ensure comparability of all test runs. The test machine runs the current snapshot of Debian 12, including updates as of January 30, 2023. The C++ Codes are compiled with the flags -03 -DNDEBUG using the gcc compiler in version 12.2.0.

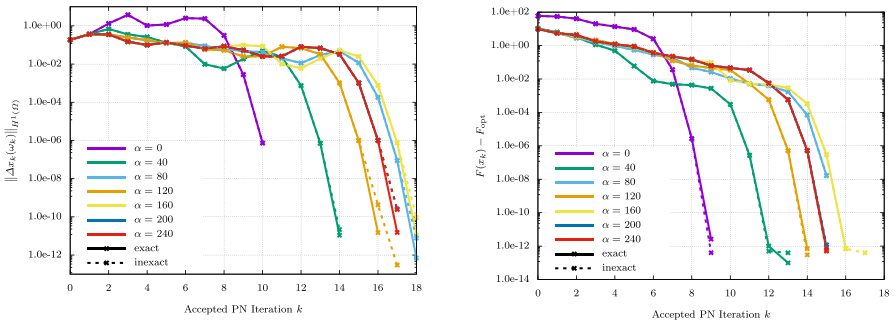
All in all, we use (53) with fixed parameters $c = 10$, $\beta = 10$, $\rho = -20$ and let $\alpha \geq 0$ vary. Thereby, increasing α magnifies the influence of the squared max-term in (53) and thus makes the corresponding minimization problem harder to solve.

Equivalence of Computed Solutions. This effect already becomes apparent in Fig. 1a where update step norms for accepted iterates are depicted for both the exact and inexact version of our method. Together with the plot of energy differences to the optimal value from Fig. 1b, this in particular suggests the equivalence of results achieved by both variants of the Proximal Newton algorithm. This expectation is validated by the computation of the relative error across all grid points y^i of our discretization via the straight-forward formula

$$\mathrm{err}_{\mathrm{rel}}(y^i) := \frac{\left\| u_{\mathrm{ex}}(y^i) - u_{\mathrm{inex}}(y^i) \right\|_2}{\left\| u_{\mathrm{ex}}(y^i) \right\|_2}$$

where we denoted by u_{ex} or u_{inex} the respective results of the exact or inexact method. In our simulations, this relative error expression reveals that the maximal discrepancy between the solutions found by the respective methods is even below numerical accuracy and yields zero in computational evaluation.

Improvements in Computational Efficiency. With the validation that the inexact variant of our Proximal Newton method achieves the same solution and general



(a) Correction Norms $\|\Delta x_k(\omega_k)\|_{H^1(\Omega)}$.

(b) Energy Differences $F(x_k) - F_{\text{opt}}$.

Fig. 1 Graphs of correction norms and energy differences to the optimal value for $c = 10$, $\beta = 10$, $\rho = -20$ and $\alpha \in \{0, 40, ..., 240\}$ for the exact and inexact Proximal Newton method

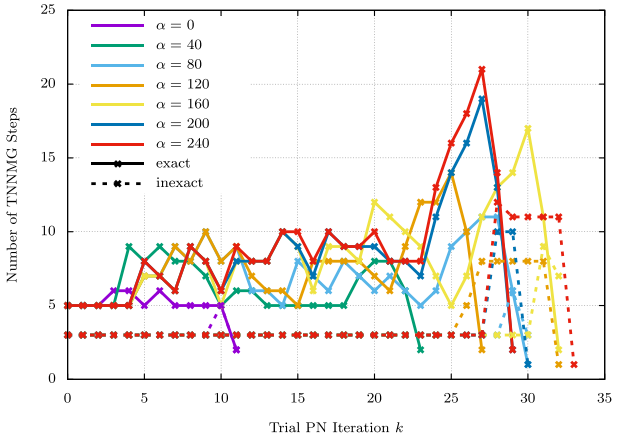


Fig. 2 Number of TNNMG iterates required for update step computation in every trial Proximal Newton step for $c = 10$, $\beta = 10$, $\rho = -20$ and $\alpha \in \{0, 40, ..., 240\}$

convergence behavior as the exact method at hand, we can now turn our attention to the actual reason for which we have made the deliberations considering inexact computation of update steps: computational efficiency.

The gain in efficiency already becomes apparent as we take a look at the plot from Fig. 2, where the number of required TNNMG iterations for computing the respective Proximal Newton trial update step is depicted. In particular, the Proximal Newton steps incorporate both accepted and declined iterates. Furthermore, we can recognize that the decrease of the forcing term η from the relative error criterion forces also the inexact version of our method to compute rather accurate solutions to the subproblems in the later stages of algorithm. This in particular enables the local superlinear convergence as we have verified in Theorem 1. In the globalization phase, however, it is easy to see that we spare many (apparently unnecessary) subproblem solver iterations and thus also save valuable computational time.

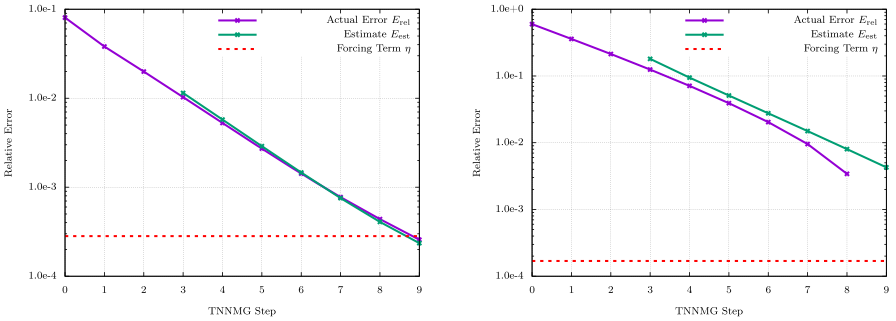
Table 1 Comparative statistics for the exact and inexact variant of our Proximal Newton method for $c = 10$, $\beta = 10$, $\rho = -20$ and $\alpha \in \{0, 40, \ldots, 240\}$

α	Variant	PN-Iterations			TNNMG-It.	Wall-Time in sec.		
		Acc	Decl	Total		TNNMG	Assembler	Total
0	Exact	11	3	14	60	13.99	89.02	117.40
	Inexact	11	3	14	37	9.37	88.81	112.67
40	Exact	15	9	24	147	30.74	109.68	166.94
	Inexact	15	9	24	72	15.43	109.52	152.12
80	Exact	19	12	31	214	44.79	139.86	218.99
	Inexact	19	12	31	96	20.67	139.45	195.11
120	Exact	17	11	28	211	43.97	124.78	201.52
	Inexact	18	15	33	124	26.49	134.69	198.99
160	Exact	19	14	33	271	56.59	139.11	232.25
	Inexact	19	14	33	109	23.41	140.70	201.61
200	Exact	17	13	30	254	52.99	131.44	217.46
	Inexact	18	13	31	105	22.60	138.89	196.66
240	Exact	18	12	30	268	56.01	131.81	220.97
	Inexact	18	16	34	141	30.27	142.25	211.08

This reduction of required TNNMG steps can be ascribed to the inexactness criteria which we have introduced over the course of the current chapter. Even though it has been a central concern of ours to also provide efficient ways for the evaluation these prerequisites for inexact update steps, this still might negate our abovementioned gains in efficiency. In order to dispel this worry, we have recorded the essential data concerning overall algorithmic efficiency for both the exact and inexact variation of our method across all test scenarios in Table 1.

While the number of accepted ("Acc."), declined ("Decl."), and total Proximal Newton iterations required for finding the solution of the minimization problem overall are the same for both alternatives, both the number of total TNNMG iterations and wall-time needed for this endeavor reveal the gains in efficiency of the inexact method. In particular, the evaluation of inexactness criteria is included in the TNNMG wall-time share. The advantageous properties of the modified algorithm become more and more apparent as α and thereby the complexity of the underlying minimization problem increases.

However, we have to note that across all numerical tests here the determining factor for the total wall-time of the respective run is the time required by the assembler, i.e., the time it takes to compute gradients and hessians, and to from there establish the respective second order problems which are then solved for the computation of Proximal Newton update steps. Furthermore, the wall-time shares of TNNMG and the assembler do not add up to the total time elapsed over one run of the algorithm since the latter additionally incorporates e.g. the evaluation of decrease criteria and update procedures of iterates and parameters.



(a) Proximal Newton trial step $k = 28$.

(b) Proximal Newton trial step $k = 29$.

Fig. 3 Comparison of the relative error as required in (19) and its estimator from (20) within the computation of Proximal Newton trial steps $k = 28$ and $k = 29$ while minimizing (53) for $c = 10$, $\beta = 10$, $\rho = -20$ and $\alpha = 200$

Having in mind the goal of the introduction of inexactness, on the other hand, we can still declare this endeavor as a success. As far as the time for solving the step computation subproblems is concerned, we have spared 51.7% across the above numerical tests which is a significant improvement. In particular for problems where first and second order models can be computed explicitly without depending on automatic differentiation software, this gain in effectiveness is crucial.

Investigation of Inexactness Criteria As mentioned beforehand, we also want to take a look at how the inexactness criteria affect the solution process of the step computation subproblems. To this end, we consider two aspects each of which covers one of our criteria based on exemplary computations of Proximal Newton steps: On the one hand, in order to investigate the relative error criterion (19), we compute every Proximal Newton step twice. Within the first computation, we neglect inexactness criteria which allows us to then compute the actual relative error E_{rel} of the TNNMG iterates in the second and actually inexact computation process. This makes it possible to compare the actual relative error to the estimate E_{est} which we use for easier evaluation, cf. (20).

As can be seen in the left-hand part of Table 2 and the plots in Fig. 3 for representative trial step computations, both of these quantities stay within the same order of magnitude. This lets us infer that the employed triangle inequality for the deduction of (20) is surprisingly sharp in practice. Note that the estimated error E_{est} is not assigned within the first two TNNMG iterations since we have to take more of these into consideration in order to obtain a valid estimate for multigrid convergence rates θ in (20). The respective column in Table 2 reveals that the estimated convergence rate then remains relatively constant over the minimization of the quadratic model which suggests it to be measured adequately by our procedure.

Furthermore, the graph for the computation of trial step $k = 29$ in Fig. 3b shows that the forcing term in this case was so small that the relative error criterion (19) could not be met by iterates of the subproblem solver before the latter stopped computation due to the default criterion from TNNMG. Thus, the relative error to the (numerically) exact solution of the subproblem is zero for the last data point in the actual relative

Table 2 Overview for inexactness criteria along TNNMG iterations i_k in Proximal Newton trial step $k = 28$ while minimizing (53) for $c = 10$, $\beta = 10$, $\rho = -20$ and $\alpha = 200$. Listed are the actual relative error E_{rel}, its estimate E_{est}, the forcing term η, the estimated TNNMG convergence rate θ, the subgradient regularization parameter $\tilde{\omega}$, and its upper bound $\tilde{\omega}_{\max}$

i_k	E_{rel}	E_{est}	η	θ	$\tilde{\omega}$	$\tilde{\omega}_{\max}$
1	0.0805495	Not assigned	0.000282111	1.03008e−06	8.71446	1e+10
2	0.0381125	Not assigned	0.000282111	0.0487216	8.81993	1e+10
3	0.0199241	0.0111777	0.000282111	0.372641	8.79332	1e+10
4	0.0102817	0.0114414	0.000282111	0.533085	8.53583	1e+10
5	0.00527283	0.0057486	0.000282111	0.524131	8.91433	1e+10
6	0.00271465	0.00289124	0.000282111	0.517709	8.41266	1e+10
7	0.00142414	0.00146153	0.000282111	0.513961	8.8467	1e+10
8	0.00077351	0.000754427	0.000282111	0.514819	8.57349	1e+10
9	0.000438303	0.000407447	0.000282111	0.522954	8.31665	1e+10
10	0.000257378	0.0002354	0.000282111	0.539877	8.20661	1e+10

error which also explains why it is missing in the corresponding logarithmic plot. In particular, the exact computation of update steps close to optimal solutions is crucial for the local acceleration of our method as shown in Theorem 1. All in all, we conclude that the estimate which implicitly uses the convergence rate of our multigrid subproblem solver constitutes an adequate and easy-to-evaluate alternative to the actual relative error.

On the other hand, we also consider the subgradient inexactness criterion (35). As mentioned beforehand, we have introduced this criterion for globalization purposes with the intention that it would not interfere with the minimization process, especially in the local acceleration phase close to optimal solutions. In fact, we have noticed that throughout our tests the determining quantity for further solving the subproblem was the relative error estimate and not that $\tilde{\omega}$ from (37) was too large. For example, over the TNNMG-iterations of the Proximal Newton trial step considered in Fig. 3a we had nearly constant $\tilde{\omega} \approx 8.5$, clearly remaining below our choice of $\tilde{\omega}_{\max} = 10^{10}$.

6 Conclusion

We have extended the globally convergent and locally accelerated Proximal Newton method in Hilbert spaces from [15] to inexact computation of update steps. Additionally, we have improved local convergence proofs by considering regularized gradient mappings and have thereby disclosed the influence of damping and inexactness to local convergence rates. We have found inexactness criteria that suit the general infinite-dimensional Hilbert space setting of the present treatise and can be evaluated cheaply within every iteration of the subproblem solver. Using these inexactness criteria, we have also been able to carry over all convergence results, local as well as global, from the exact case. The application of our method to actual function space problems is enabled by using an efficient solver for the step computation subproblem, the Trun-

cated Non-smooth Newton Multigrid Method. We have displayed functionality and efficiency of our algorithm by considering a simple model problem in function space.

Room for improvement is definitely present in the choice of both regularization parameters ω and forcing terms η. The former can be addressed by different approaches like estimates for residual terms of the quadratic model established in subproblem (3), cf. [23], or adapted strategies for controlling time step sizes in computing solutions of ordinary differential equations. For the forcing terms on the other hand, adaptive choices have already been studied for inexact Newton methods e.g. in [1, 6]. While these can be carried over to our non-smooth scenario, it also appears to be promising to tie the choice of regularization parameters and forcing terms together due to their similar convergence behavior. This idea both reduces the computational effort and better reflects the problem structure at hand.

Funding Open access funding enabled and organized by Projekt DEAL. This work was funded by the DFG SPP 1962: Non-smooth and Complementarity-based Distributed Parameter Systems – Simulation and Hierarchical Optimization; Project number: SCHI 1379/6-1.

Data availability The datasets generated and analysed during the current study are not publicly available due the fact that they constitute an excerpt of research in progress but are available from the corresponding author on reasonable request.

Declarations

Conflict of interest All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

References

1. An, H.B., Mo, Z.Y., Liu, X.P.: A choice of forcing terms in inexact newton method. J. Comput. Appl. Math. **200**(1), 47–60 (2007). https://doi.org/10.1016/j.cam.2005.12.030
2. Beck, A.: First-order methods in optimization. Soc. Ind. Appl. Math. (2017). https://doi.org/10.1137/1.9781611974997
3. Byrd, R.H., Nocedal, J., Oztoprak, F.: An inexact successive quadratic approximation method for l-1 regularized optimization. Math. Program. **157**(2), 375–396 (2015). https://doi.org/10.1007/s10107-015-0941-y
4. Conn, A.R., Gould, N.I.M., Toint, P.L.: Trust region methods. Soc. Ind. Appl. Math. (2000). https://doi.org/10.1137/1.9780898719857
5. Dembo, R.S., Eisenstat, S.C., Steihaug, T.: Inexact Newton methods. SIAM J. Num. Anal. **19**(2), 400–408 (1982). https://doi.org/10.1137/0719025
6. Deuflhard, P.: Newton methods for nonlinear problems. Affine invariance and adaptive algorithms, Series Computational Mathematics, vol. 35, 2nd edn. Springer (2006)

7. Gräser, C., Sander, O.: Truncated nonsmooth newton multigrid methods for block-separable minimization problems. IMA J. Num. Anal. **39**(1), 454–481 (2018). https://doi.org/10.1093/imanum/dry073
8. Hintermüller, M., Ito, K., Kunisch, K.: The primal-dual active set strategy as a semismooth newton method. SIAM J. Optim. **13**(3), 865–888 (2002). https://doi.org/10.1137/s1052623401383558
9. Kanzow, C., Lechner, T.: Globalized inexact proximal newton-type methods for nonconvex composite functions. Comput. Optim. Appl. (2020). https://doi.org/10.1007/s10589-020-00243-6
10. pei Lee, C., Wright, S.J.: Inexact successive quadratic approximation for regularized optimization. Comput. Optim. Appl. **72**(3), 641–674 (2019). https://doi.org/10.1007/s10589-019-00059-z
11. Lee, J.D., Sun, Y., Saunders, M.A.: Proximal newton-type methods for minimizing composite functions. SIAM J. Optim. **24**(3), 1420–1443 (2014). https://doi.org/10.1137/130921428
12. Li, J., Andersen, M.S., Vandenberghe, L.: Inexact proximal newton methods for self-concordant functions. Math. Meth. Oper. Res. **85**(1), 19–41 (2016). https://doi.org/10.1007/s00186-016-0566-9
13. Mifflin, R.: Semismooth and semiconvex functions in constrained optimization. SIAM J. Control Optim. **15**(6), 959–972 (1977). https://doi.org/10.1137/0315061
14. Mordukhovich, B.S., Yuan, X., Zeng, S., Zhang, J.: A globally convergent proximal newton-type method in nonsmooth convex optimization. Math. Program. (2022). https://doi.org/10.1007/s10107-022-01797-5
15. Ptözl, B., Schiela, A., Jaap, P.: Second order semi-smooth proximal newton methods in Hilbert spaces. Comput. Optim. Appl. **82**(2), 465–498 (2022). https://doi.org/10.1007/s10589-022-00369-9
16. Qi, L.: Convergence analysis of some algorithms for solving nonsmooth equations. Math. Oper. Res. **18**(1), 227–244 (1993). https://doi.org/10.1287/moor.18.1.227
17. Qi, L., Sun, J.: A nonsmooth version of Newton's method. Math. Program. **58**(1–3), 353–367 (1993). https://doi.org/10.1007/bf01581275
18. Scheinberg, K., Tang, X.: Practical inexact proximal quasi-newton method with global complexity analysis. Math. Program. **160**(1–2), 495–529 (2016). https://doi.org/10.1007/s10107-016-0997-3
19. Schiela, A.: A simplified approach to semismooth Newton methods in function space. SIAM J. Optim. **19**(3), 1417–1432 (2008). https://doi.org/10.1137/060674375
20. Ulbrich, M.: Nonsmooth newton-like methods for variational inequalities and constrained optimization problems in function spaces. Habilitation Thesis (2002)
21. Ulbrich, M.: Semismooth Newton methods for variational inequalities and constrained optimization problems in function spaces. Soc. Ind. Appl. Math. (2011). https://doi.org/10.1137/1.9781611970692
22. Walther, A., Griewank, A.: Getting started with ADOL-c. In: Combinatorial Scientific Computing, pp. 181–202. Chapman and Hall/CRC (2012). https://doi.org/10.1201/b11644-8
23. Weiser, M., Deuflhard, P., Erdmann, B.: Affine conjugate adaptive newton methods for nonlinear elastomechanics. Optim. Meth. Softw. **22**(3), 413–431 (2007). https://doi.org/10.1080/10556780600605129
24. Yue, M.C., Zhou, Z., So, A.M.C.: A family of inexact SQA methods for non-smooth convex minimization with provable convergence guarantees based on the luo–tseng error bound property. Mathematical Programming **174**(1–2), 327–358 (2018). https://doi.org/10.1007/s10107-018-1280-6