TRUNCATED NONSMOOTH NEWTON MULTIGRID METHODS FOR BLOCK-SEPARABLE MINIMIZATION PROBLEMS

CARSTEN GRÄSER AND OLIVER SANDER

In memory of Elias Pipping (1986–2017)

Abstract. The Truncated Nonsmooth Newton Multigrid (TNNMG) method is a robust and efficient solution method for a wide range of block-separable convex minimization problems, typically stemming from discretizations of nonlinear and nonsmooth partial differential equations. This paper proves global convergence of the method under weak conditions both on the objective functional, and on the local inexact subproblem solvers that are part of the method. It also discusses a range of algorithmic choices that allows to customize the algorithm for many specific problems. Numerical examples are deliberately omitted, because many such examples have already been published elsewhere.

AMS classification: 65K15, 90C25, 49M20

Keywords: multigrid, convex minimization, global convergence, block separable

1. Introduction

We consider minimization problems for energy functionals having block-separable nonsmooth terms. Given an objective functional $\mathcal{J} : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$, we assume that it has the form

$$\mathcal{J} = \mathcal{J}_0 + \varphi, \tag{1}$$

where $\mathcal{J}_0 : \mathbb{R}^n \to \mathbb{R}$ is coercive and continuously differentiable, and $\varphi : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is block separable, i.e., there are functionals $\varphi_i : \mathbb{R}^{n_i} \to \mathbb{R} \cup \{\infty\}$, $i = 1, \dots, M$ that are convex, proper, lower semi-continuous, and continuous on their domains such that

$$\varphi(v) = \sum_{i=1}^{M} \varphi_i(v_i),$$

where $\sum_{i=1}^{M} n_i = n$ and we implicitly identify $\mathbb{R}^n$ with $\prod_{i=1}^{M} \mathbb{R}^{n_i}$.

Such minimization problems occur in many fields of science; however, we are mainly interested in the case where they result from discretized nonsmooth partial differential equations (PDEs). Examples of such PDEs include contact and friction models in solid mechanics [26, 33], certain models of porous media flow [3], but also many variants of Allen–Cahn-type phase-field models [25]. The block-structure is typically either induced by grouping all vector components of a given vector-valued degree of freedom, or by grouping the degrees of freedom of a grid element in a Discontinuous-Galerkin-type discretization. Nonsmoothness can originate from the

Parts of this work were motivated by enjoyable discussions with Elias Pipping.

PDE itself, or from local convex constraints to admissible sets K_i, which can be incorporated through local indicator functionals $\varphi_i = \chi_{K_i}$. We do not make use of the relationship of $\mathcal{J}$ to PDEs other than by implicitly assuming that the individual block sizes n_i are small compared to their number M, and by proposing to solve linear subproblems inexactly by multigrid methods.

The nondifferentiable terms φ_i make minimizing functionals of the type (1) notoriously difficult. Interior-point methods [4] are expensive, because they solve entire sequences of linear problems, and preserving robust global convergence is difficult, if these problems are solved only inexactly. The same issues also plague semismooth Newton [32] and classical active-set methods, which are closely related [2, 20, 19]. Such methods do not actively exploit convexity and block-separability of the objective functional. They also struggle with the unbounded gradients that appear in several practically relevant phase-field models, which lead to ill-conditioned linear systems.

Similarly, the very generic subgradient and bundle methods make no use of the block structure or convexity, either. They also disregard any second-order information, and they are therefore very slow, and convergence is not always guaranteed [7]. In contrast, local relaxation and operator splitting approaches use the specific problem structure to obtain global convergence, but are also very slow for problems resulting from discretized PDEs [5, 25].

In this paper we present the Truncated Nonsmooth Newton Multigrid (TNNMG) method, a nonsmooth multigrid method for block-separable nonsmooth minimization problems. While we have proposed TNNMG before for various more specialized problems [10, 14, 13, 15], we now present it as a general framework that includes all previous incarnations of TNNMG as special cases.

The TNNMG method combines local relaxation methods with a generalized Newton approach. It consists of a nonlinear pre-smoother followed by an inexact linear correction step, typically one multigrid iteration. Thus it can be interpreted alternatively as a multigrid method with a nonlinear fine grid smoother, or as an inexact Newton method with a nonlinear corrector step [34]. The method does not regularize the problem, and does not involve parameters that would need to be selected manually. Numerous numerical experiments have shown that it achieves multigrid-like convergence behavior (i.e., mesh-independent convergence rates and linear time complexity) on a wide range of difficult nonlinear problems if reasonable initial iterates are available. In case of discretized PDEs, such initial iterates can typically be obtained by *nested iteration*, i.e., by using inexact solutions for coarser discretizations [10]. We emphasize that good initial iterates are not mandatory for convergence itself, but only to obtain multigrid-like convergence speed.

In this paper we prove global convergence of the method under very weak conditions on the functional, the smoother, and the linear correction. The convergence result subsumes previous results for obstacle [10] and contact problems [13], as well as for problems with polyhedral nonsmoothness [14, 25]. Unlike those previous results, we allow search spaces that do not coincide with the block structure, and we even allow non-direct sums of search spaces. This in particular allows us to reprove the main results of [14] in a much simpler way. We show convergence of the presented method to stationary points, such that we get convergence to minimizers, e.g., if $\mathcal{J}_0$ is convex. While the convergence proof only requires continuous differentiability of $\mathcal{J}_0$, the method uses a Newton-type linearization in a substep

such that we need at least Lipschitz continuity of $\mathcal{J}_0'$. The strongest results are obtained for strictly convex $\mathcal{J}$, but biconvex and certain quasiconvex problems are covered by the theory as well.

The convergence proof relies mainly on the convergence of the Gauß–Seidel-type pre-smoother by itself. Weak conditions on such smoothers are stated that ensure global convergence of TNNMG. These weak conditions allow various cheap inexact solvers to be used for the local nonsmooth subproblems. We propose various such inexact solvers and prove that they fulfill the conditions.

On the other hand, while the linear correction step does not really interfere in the convergence proof at all, it is crucial for *fast* convergence of the TNNMG method. The step involves constructing a sufficiently large subspace on which a Newton-type correction problem is well defined, and then solving this problem in a suitably inexact way. The standard way is to use a single geometric or algebraic multigrid step here. Such a choice then leads to the interpretation of the overall method as a multigrid method. However, other choices of solvers are possible, and can be useful in certain situations.

This article is structured as follows: We begin by formally stating the problem in Chapter 2. In Chapter 3 we introduce the TNNMG method, and in Chapter 4 we prove its global convergence. Chapter 5 discusses ways to inexactly solve the local minimization problems that make up the nonlinear multigrid smoother. Chapter 6 introduces truncated linearized correction problems, and Chapter 7 discusses efficient inexact solvers for those problems, which are crucial to obtain the overall multigrid-like convergence speed of TNNMG. An appendix collects a few technical results.

The paper deliberately omits numerical benchmarks of the method, because many of these have already appeared in the literature. Among others, we refer the reader to [13] for contact problems, [13, 15] for phase-field problems, and [34] for the performance of TNNMG on small-strain plasticity problems.

2. Block-separable nonsmooth functionals

The TNNMG method is designed to solve nonsmooth minimization problems

$$u^* \in \mathbb{R}^n : \qquad \mathcal{J}(u^*) \leq \mathcal{J}(v) \qquad \forall v \in \mathbb{R}^n, \tag{2}$$

for possibly nondifferentiable functionals $\mathcal{J} : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$. Throughout the paper we will assume that $\mathcal{J}$ is proper, coercive, lower semi-continuous, continuous on its domain $\operatorname{dom} \mathcal{J} := \{x \in \mathbb{R}^n \mid \mathcal{J}(x) < \infty\}$, and that $\operatorname{dom} \mathcal{J}$ is convex, but not necessarily closed.

The algorithm is based on a subspace decomposition

$$\mathbb{R}^n = \sum_{k=1}^{m} V_k \tag{3}$$

into m subspaces V_k of dimensions $n_k := \dim V_k$, $k = 1, \dots, m$. The sum in (3) does not need to be direct. For practical purposes we will also assume coordinate systems for each subspace in the sense that we identify V_k with $\mathbb{R}^{n_k}$ using prolongation maps

$$P_k : \mathbb{R}^{n_k} \to \mathbb{R}^n, \qquad k = 1, \dots, m, \tag{4}$$

such that for each k, P_k is an isomorphism from $\mathbb{R}^{n_k}$ to V_k.

If $\mathcal{J}$ is differentiable, then the splitting (3) is sufficient to characterize minimizers of $\mathcal{J}$ in the sense that $u \in \mathbb{R}^n$ is a global minimizer of $\mathcal{J}$ if it is a global minimizer with respect to each V_k. For nondifferentiable $\mathcal{J}$ this is no longer true, unless the subspace splitting is compatible with the nondifferentiable structure of $\mathcal{J}$. For counterexamples we refer to [8]. Here, we will make compatibility an additional assumption.

Definition 2.1. *The decomposition* (3) *is called* compatible with $\mathcal{J}$ *if*

$$\mathcal{J}(u) \leq \mathcal{J}(u+v) \qquad \forall v \in V_k, \quad k = 1, \dots, m \tag{5}$$

implies global optimality

$$\mathcal{J}(u) \leq \mathcal{J}(u+v) \qquad \forall v \in \mathbb{R}^n.$$

An important example for this situation are block-separable nonsmooth problems with a corresponding subspace decomposition. We consider problems with a nonsmooth, convex, and separable part, and a smooth but possibly nonconvex part.

Definition 2.2. *Let $R : \mathbb{R}^n \to \prod_{k=1}^M \mathbb{R}^{N_k}$ be an isomorphism with $R = (R_1, \dots, R_M)$ and surjective linear maps $R_k : \mathbb{R}^n \to \mathbb{R}^{N_k}$. We say that a function $\mathcal{J} : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is* block-separable nonsmooth with respect to R *if there is a continuously differentiable function $\mathcal{J}_0 : \mathbb{R}^n \to \mathbb{R}$, and convex, proper, lower-semicontinuous functions $\varphi_k : \mathbb{R}^{N_k} \to \mathbb{R} \cup \{\infty\}$ such that*

$$\mathcal{J}(v) = \mathcal{J}_0(v) + \underbrace{\sum_{k=1}^M \varphi_k(R_k v)}_{=:\varphi(v)}.$$

We note that for any such isomorphism $R : \mathbb{R}^n \to \prod_{k=1}^M \mathbb{R}^{N_k}$, the maps $R_k : (\ker R_k)^\perp \to \mathbb{R}^{N_k}$ are isomorphisms themselves, which together with $n = \sum_{k=1}^M N_k$ gives rise to the direct sum representation

$$\mathbb{R}^n = \bigoplus_{k=1}^M (\ker R_k)^\perp. \tag{6}$$

Using the subspaces induced by the block-separability leads to a compatible decomposition. However, the following result shows that more general decompositions are also possible.

Lemma 2.1. *Let $\mathcal{J} : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be convex and block-separable nonsmooth with respect to an isomorphism $R : \mathbb{R}^n \to \prod_{k=1}^M \mathbb{R}^{N_k}$. Assume that for any block $k = 1, \dots, M$ the decomposition contains a subspace $V_{k'}$ with $(\ker R_k)^\perp \subset V_{k'}$. Then the decomposition is compatible with $\mathcal{J}$.*

Proof. Let $u \in \mathbb{R}^n$ satisfy (5). To any $k = \{1, \dots, M\}$ we associate an index k' such that $(\ker R_k)^\perp \in V_{k'}$. The optimality of u in $u + V_{k'}$ implies the variational inequalities

$$\langle \mathcal{J}_0'(u), w_{k'} - u \rangle + \varphi(w_{k'}) - \varphi(u) \geq 0 \qquad \forall w_{k'} \in u + V_{k'} \tag{7}$$

for $k = 1, \dots, M$. Now let $w \in \mathbb{R}^n$ and $v = w - u$. Then, according to (6), there is a decomposition

$$v = \sum_{k=1}^{M} v_k, \qquad v_k \in (\ker R_k)^\perp \subset V_{k'}$$

with some k' for each k. Since the sum (6) is direct, we have $R_k v = R_k v_k$ and $R_i v_k = 0$ for $i \neq k$, which implies

$$\varphi(u + v_k) - \varphi(u) = \sum_{i=1}^{M} \big[\varphi_i(R_i(u + v_k)) - \varphi_i(R_i u)\big] = \varphi_k(R_k(w)) - \varphi_k(R_k u). \tag{8}$$

Inserting $w_{k'} = u + v_k$ into the variational inequalities (7), summing up for all $k = 1, \dots, M$, and using (8) finally gives

$$\langle \mathcal{J}_0'(u), w - u\rangle + \varphi(w) - \varphi(u) \geq 0.$$

Since $w \in \mathbb{R}^n$ was arbitrary and $\mathcal{J}$ is convex, this proves the assertion. □

Lemma 2.1 shows compatibility for general possibly non-direct subspace decompositions. Such decompositions occur, e.g., when using domain decomposition methods [18, 36, 37, 5] and nonlinear multi-level relaxation methods [29, 10, 1, 23, 24]. In the present paper we will concentrate on subspace decompositions where we either have

$$(\ker R_k)^\perp = V_k \qquad k = 1, \dots, m = M \tag{9}$$

or where each $(\ker R_k)^\perp$ can be written as a direct or non-direct sum of a subset of the subspaces $V_{k'}$. In the following we list a few examples of suitable energy functionals. For a set U, we denote by χ_U the indicator functional

$$\chi_U(z) := \begin{cases} 0 & \text{if } z \in U, \\ \infty & \text{otherwise.} \end{cases}$$

Example 2.1. Let $A \in \mathbb{R}^{n\times n}$ be symmetric and positive definite, $b \in \mathbb{R}^n$, and let $K_k \subset \mathbb{R}$, $k = 1, \dots, n$ be nonempty, closed, possibly unbounded intervals. Then

$$\mathcal{J}(v) := \underbrace{\frac{1}{2}\langle Av, v\rangle - \langle b, v\rangle}_{=:\mathcal{J}_0(v)} + \sum_{k=1}^{n} \underbrace{\chi_{K_k}(v_k)}_{=:\varphi_k(v_k)}$$

is block-separable nonsmooth, and the decomposition of $\mathbb{R}^n$ into $V_k = \{v \in \mathbb{R}^n \mid v_i = 0 \quad \forall i \neq k\}$ is compatible. This is the energy functional for the classic obstacle problem of minimizing $\mathcal{J}_0$ in the hypercube $K = \prod_{i=1}^n K_i$, see [10].

Example 2.2. Let $D_i \in \mathbb{R}^{d\times n}$, and $\gamma_i : \mathbb{R}^d \to \mathbb{R}$ be convex, continuously differentiable functions for $i = 1, \dots, l$. Then the functional

$$\mathcal{J}(v) := \underbrace{\sum_{i=1}^{l} \gamma_i(D_i v)}_{=:\mathcal{J}_0(v)} + \sum_{k=1}^{n} \underbrace{\chi_{[0,1]}(v_k)}_{=:\varphi_k(v_k)}$$

is block-separable nonsmooth, and the decomposition into $V_k = \{v \in \mathbb{R}^n \mid v_i = 0 \quad \forall i \neq k\}$ is compatible. Functionals of this form with $\gamma_i(z) = \omega_i \gamma(z)$ and a quadrature weight ω_i are obtained, e.g., for discretized minimal surface equations

with an obstacle [17], and shallow-ice glacier models [21, 22]. In both cases D_i represents the local evaluation of the gradient of a finite element function. For the minimal surface equation we have $\gamma(z) = \sqrt{1 + \|z\|_2^2}$, while the p-Laplace operator in shallow-ice models leads to $\gamma(z) = \|z\|_2^p$. For anisotropic phase-field models one obtains similar functions with an additional quadratic term [11].

Example 2.3. Let A and b be as above, $R : \mathbb{R}^n \to (\mathbb{R}^L)^m$ an isomorphism, and

$$G := \Big\{ v \in \mathbb{R}^L \mid \sum_{i=1}^{L} v_i = 1,\ v_i \geq 0 \quad \forall i = 1, \dots, L \Big\} \tag{10}$$

the L-dimensional Gibbs-simplex. Then the functional

$$\mathcal{J}(v) := \underbrace{\langle Av, v\rangle - \langle b, v\rangle}_{=:\mathcal{J}_0(v)} + \sum_{k=1}^{m} \underbrace{\chi_G(R_k v)}_{=:\varphi_k(R_k v)}$$

is block-separable nonsmooth, and the decomposition into $V_k = (\ker R_k)^\perp$ is compatible. Note that these subspaces take the form $V_k = e_k \otimes \mathbb{R}^L$ if we identify $\mathbb{R}^n = (\mathbb{R}^L)^m$, where e_k is the k-th canonical basis vector in $\mathbb{R}^m$. Functionals of this form occur when discretizing multi-component phase-field models [25]. It was shown in [14, 25] that another compatible subspace decomposition is given by $V_{k,i} = e_k \otimes \eta_i$, where $\eta_1, \dots, \eta_l$ denote the $l = \frac{1}{2}(L-1)L$ edge vectors of G.

Example 2.4. Let A, b, and R as above, $\|\cdot\| : \mathbb{R}^L \to \mathbb{R}$ a norm on $\mathbb{R}^L$, and $\omega_k > 0$ positive weights. Then the functional

$$\mathcal{J}(v) = \underbrace{\langle Av, v\rangle - \langle b, v\rangle}_{=:\mathcal{J}_0(v)} + \sum_{k=1}^{m} \underbrace{\omega_k \|R_k v\|}_{=:\varphi_k(R_k v)}$$

has the desired form with a compatible decomposition of $\mathbb{R}^n = (\mathbb{R}^L)^m$ into $V_k = e_k \otimes \mathbb{R}^L$. Functionals of this type are, e.g., obtained for certain friction laws, and in primal formulations of small-strain plasticity [16].

3. The Truncated Nonsmooth Newton Multigrid method

The *Truncated Nonsmooth Newton Multigrid Method* (TNNMG) was introduced in [10] for quadratic obstacle problems, and later generalized to variational inequalities of the second kind with separable nonsmooth nonlinearities [13, 9]. Similar to these cases, the extension to block-separable nonsmooth problems of the form (2) will be based on a nonlinear block Gauß–Seidel iteration and an additional linear correction step. For given initial iterate $u^0 \in \operatorname{dom} \mathcal{J}$ and iteration number $\nu \in \mathbb{N}_0$, one step of the TNNMG method is defined as follows:

Input: Given u^ν
begin Nonlinear pre-smoothing
  Set $w^{\nu,0} := u^\nu$
  for $k = 1, \dots, m$ **do**
    Compute $w^{\nu,k} \in w^{\nu,k-1} + V_k$:

$$w^{\nu,k} :\approx \operatorname*{arg\,min}_{v \in w^{\nu,k-1} + V_k} \mathcal{J}(v)$$

  end
  Set $u^{\nu+\frac{1}{2}} := w^{\nu,m}$
end

begin Truncated linear correction
  Determine large subspace $W_\nu \subset \mathbb{R}^n$ such that $\mathcal{J}|_{u^{\nu+\frac{1}{2}}+W_\nu}$ is C^2 near $u^{\nu+\frac{1}{2}}$.
  Compute $v^\nu \in W_\nu$ as

$$v^\nu :\approx -\Big(\mathcal{J}''(u^{\nu+\frac{1}{2}})|_{W_\nu\times W_\nu}\Big)^{-1}\Big(\mathcal{J}'(u^{\nu+\frac{1}{2}})|_{W_\nu}\Big) \tag{11}$$

end
begin Post-processing
  Compute the projection $\tilde{v}^\nu := \Pi_{\operatorname{dom}\mathcal{J}-u^{\nu+1/2}}(v^\nu)$
  Compute $\rho_\nu \in [0,\infty)$ such that $\mathcal{J}(u^{\nu+\frac{1}{2}}+\rho_\nu\tilde{v}^\nu) \le \mathcal{J}(u^{\nu+\frac{1}{2}})$
end
Output: Set $u^{\nu+1} := u^{\nu+\frac{1}{2}} + \rho_\nu\tilde{v}^\nu$

Using the prolongation operators P_k of (4), the incremental minimization problems in Line 5 for a given intermediate iterate $w = w^{\nu,k-1} \in \mathbb{R}^n$ can be written as minimization problems in $\mathbb{R}^{n_k}$

$$\operatorname*{arg\,min}_{v\in w+V_k}\mathcal{J}(v) = w + \operatorname*{arg\,min}_{v\in V_k}\mathcal{J}(w+v) = w + P_k\operatorname*{arg\,min}_{\xi\in\mathbb{R}^{n_k}}\mathcal{J}(w+P_k\xi). \tag{12}$$

We are mainly interested in block-separable problems originating from discretized PDEs where the block-sizes n_k are independent of the problem size n. If the subspaces V_k are chosen according to the blocks or even finer, then the size of the subproblems (12) is in $O(1)$, which allows to use more expensive schemes for their approximate solution.

In Line 11 we denote by $\mathcal{J}'(u^{\nu+\frac{1}{2}})|_{W_\nu}$ and $\mathcal{J}''(u^{\nu+\frac{1}{2}})|_{W_\nu\times W_\nu}$ the gradient and Hessian of the restriction of the objective functional $\mathcal{J}$ to the subspaces W_ν and $W_\nu \times W_\nu$, respectively. While $\mathcal{J}$ is in general not even differentiable, these are still well-defined expressions since $\mathcal{J}$ is twice continuously differentiable on W_ν by construction of W_ν. In implementations it is convenient to represent these restricted derivatives with respect to coordinates in $\mathbb{R}^n$ by extending them with zero to the orthogonal complement of W_ν. In such a situation, the restricted Hessian is not invertible as a map $\mathbb{R}^n \to \mathbb{R}^n$. However, if $\mathcal{J}$ is convex and $\mathcal{J}_0''(v)$ exists and is positive definite for any v then the correction problem (11) still has a unique solution in W_ν. In general this can be expressed by replacing $(\mathcal{J}''(u^{\nu+\frac{1}{2}})|_{W_\nu\times W_\nu})^{-1}$ by the Moore–Penrose pseudo-inverse of $\mathcal{J}''(u^{\nu+\frac{1}{2}})$ on $\mathbb{R}^n$. Similar arguments hold if $\mathcal{J}_0$ is not C^2, but has a Lipschitz-continuous first derivative. Then, $\mathcal{J}''$ needs to be replaced by a suitable generalized derivative of $\mathcal{J}'$.

Simply adding the linear coarse grid correction v^ν from Line 9 to $u^{\nu+\frac{1}{2}}$ to obtain the new iterate $u^{\nu+1}$ may lead to infeasibility of $u^{\nu+1}$, since Line 9 is not aware of the domain of $\mathcal{J}$. Adding a damped version of v^ν can ensure feasibility, but this may lead to very small damping parameters and thus to poor convergence. As an alternative we first compute a Euclidean projection $\tilde{v}^\nu$ of v^ν into the domain in Line 14. While this ensures feasibility of $u^{\nu+\frac{1}{2}} + \tilde{v}^\nu$ it may still increase energy, because the projection is only aware of the domain, but not of the values of $\mathcal{J}$. To ensure energy decrease damping is then applied to the projected correction.

The post-processing steps are depicted in Figure 1. In particular, note that using a damped, projected correction may lead to larger steps compared to pure damping without preceding projection.

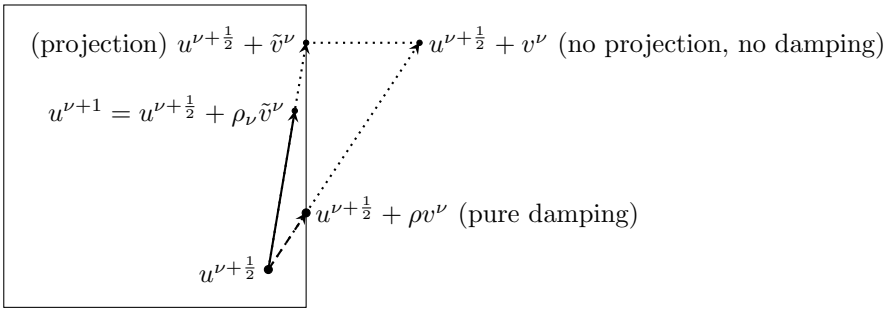

FIGURE 1. Illustration of the TNNMG post-processing steps. The square represents the admissible set $\operatorname{dom} \mathcal{J}$.

4. Global convergence

We now prove that the TNNMG iteration converges to a stationary point for any initial iterate u^0. For this we need some assumptions on the inexact solution operators that map $w^{\nu,k-1}$ to $w^{\nu,k}$ in Line 5 of the TNNMG algorithm. To this end we introduce operators

$$\mathcal{M}_k : \operatorname{dom} \mathcal{J} \to \operatorname{dom} \mathcal{J}, \qquad \mathcal{M}_k - \mathrm{Id} : \operatorname{dom} \mathcal{J} \to V_k \tag{13}$$

such that

$$w^{\nu,k} = \mathcal{M}_k(w^{\nu,k-1}), \qquad w^{\nu,0} = u^\nu. \tag{14}$$

These operators represent the inexact minimization of $\mathcal{J}$ in the affine subspaces $w^{\nu,k-1} + V_k$ by, e.g., some iterative algorithm. Using this notation the inexact block Gauß–Seidel loop in Lines 1–8 takes the form

$$u^{\nu+\frac{1}{2}} = \mathcal{M}(u^\nu), \qquad \mathcal{M} := \mathcal{M}_m \circ \cdots \circ \mathcal{M}_1. \tag{15}$$

The convergence proof hinges on the continuity of the compound operators $\mathcal{J} \circ \mathcal{M}_k$. In case of exact minimization, $\mathcal{M}_k$ is

$$\mathcal{M}_k(\cdot) = \operatorname*{arg\,min}_{v \in (\cdot)+V_k} \mathcal{J}(v),$$

and we get continuity of $\mathcal{J} \circ \mathcal{M}_k = \min_{v \in (\cdot)+V_k} \mathcal{J}(v)$ as a direct consequence of Lemma 5.1 below. However, this continuity may not hold for general inexact minimization operators $\mathcal{M}_k$. Therefore we introduce additional operators

$$\tilde{\mathcal{M}}_k : \operatorname{dom} \mathcal{J} \to \operatorname{dom} \mathcal{J}, \qquad \tilde{\mathcal{M}}_k - \mathrm{Id} : \operatorname{dom} \mathcal{J} \to V_k \tag{16}$$

with continuous $\mathcal{J} \circ \tilde{\mathcal{M}}_k$ that bound the energy decrease realized by $\mathcal{M}_k$ from above, i.e., we assume that

$$\mathcal{J}(\mathcal{M}_k(v)) \leq \mathcal{J}(\tilde{\mathcal{M}}_k(v)) \qquad \forall v \in \operatorname{dom} \mathcal{J}.$$

These operators do not appear in the implementation of the method but will only be used as a tool to prove convergence.

Finally, we introduce an operator $\mathcal{C} : \operatorname{dom} \mathcal{J} \to \operatorname{dom} \mathcal{J}$ that represents the truncated linear correction and the post-processing in Lines 9–17 in the sense that

$$u^{\nu+1} = \mathcal{C}(u^{\nu+\frac{1}{2}}). \tag{17}$$

However, this operator does not have to be continuous in any way. Note that (15) and (17) together formalize the TNNMG algorithm.

Theorem 4.1. *Let $\mathcal{J} : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be coercive, proper, lower-semicontinuous, and continuous on its domain, and assume that $\mathcal{J}(w + (\cdot))$ has a unique global minimizer in all V_k, $k = 1, \dots, m$ for each $w \in \operatorname{dom} \mathcal{J}$. For local operators $\mathcal{M}_k$ of the form* (13)*, a correction operator $\mathcal{C} : \operatorname{dom} \mathcal{J} \to \operatorname{dom} \mathcal{J}$, and an initial guess $u^0 \in \operatorname{dom} \mathcal{J}$ let (u^ν) be given by the algorithm* (15) *and* (17)*. Assume that there are operators $\tilde{\mathcal{M}}_k$ of the form* (16) *such that the following holds:*

1. *Monotonicity: $\mathcal{J}(\mathcal{M}_k(v)) \le \mathcal{J}(\tilde{\mathcal{M}}_k(v)) \le \mathcal{J}(v)$ and $\mathcal{J}(\mathcal{C}(v)) \le \mathcal{J}(v)$ for all $v \in \operatorname{dom} \mathcal{J}$.*
2. *Continuity: $\mathcal{J} \circ \tilde{\mathcal{M}}_k$ is continuous.*
3. *Stability: $\mathcal{J}(\tilde{\mathcal{M}}_k(v)) < \mathcal{J}(v)$ if $\mathcal{J}(v)$ is not minimal in $v + V_k$.*

Then any accumulation point u of (u^ν) is stationary in the sense that

$$\mathcal{J}(u) \le \mathcal{J}(u + v) \qquad \forall v \in V_k, \quad k = 1, \dots, m. \tag{18}$$

The proof uses the following direct consequence of lower semi-continuity.

Lemma 4.2. *If $F : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is lower semi-continuous, then the sub-level set $F^{-1}((-\infty, C])$ is closed for any $C \in \mathbb{R}$. In particular, this guarantees that $\lim x^\nu \in \operatorname{dom} F$ for any convergent sequence x^ν with bounded $F(x^\nu)$.*

Proof of Theorem 4.1. Throughout the proof we assume that the sequences $w^{\nu,k}$ are defined according to (14). Now let (u^{ν_l}) be any convergent subsequence of (u^ν) with $u^{\nu_l} \to u$. Then by Lemma 4.2 we have $u \in \operatorname{dom} \mathcal{J}$. We will show that u is stationary in the sense of (18).

As the subsequences $(w^{\nu_l,k})_{l \in \mathbb{N}}$ are also bounded for any k there are subsubsequences (w.l.o.g. also indexed by ν_l) and limits w^k such that $w^{\nu_l,k} \to w^k$ for $l \to \infty$ for all $k = 1, \dots, m$. In particular, we have $w^{\nu_m,0} = u^{\nu_m} \to u =: w^0$. For any $k = 1, \dots, m$ we then have by monotonicity of $\mathcal{J}$ with respect to $\mathcal{M}_k$, $\tilde{\mathcal{M}}_k$, and $\mathcal{C}$

$$\mathcal{J}(w^{\nu_{l+1},k-1}) \le \mathcal{J}(w^{\nu_l,k}) \le \mathcal{J}(\tilde{\mathcal{M}}_k(w^{\nu_l,k-1})) \le \mathcal{J}(w^{\nu_l,k-1}) \le \mathcal{J}(u^0) < \infty,$$

and thus $w^k \in \operatorname{dom} \mathcal{J}$ by Lemma 4.2. Taking the limit $l \to \infty$ and using continuity of $\mathcal{J}$ and $\mathcal{J} \circ \tilde{\mathcal{M}}_k$ we get

$$\mathcal{J}(w^{k-1}) \le \mathcal{J}(w^k) \le \mathcal{J}(\tilde{\mathcal{M}}_k(w^{k-1})) \le \mathcal{J}(w^{k-1}), \tag{19}$$

which, together with stability, implies that $\tilde{\mathcal{M}}_k(w^{k-1})$ is the unique minimizer of $\mathcal{J}$ in $w^{k-1} + V_k$ and thus $\tilde{\mathcal{M}}_k(w^{k-1}) = w^{k-1}$. From $w^{\nu_l,k} - w^{\nu_l,k-1} \in V_k$ we get for the limit $w^k - w^{k-1} \in V_k$ and thus $w^k \in w^{k-1} + V_k$. Hence $\mathcal{J}(w^k) \le \mathcal{J}(w^{k-1})$ (as shown in (19)) and uniqueness of the minimizer in $w^{k-1} + V_k$ show $w^k = w^{k-1}$.

We conclude that we have $w^m = w^{m-1} = \cdots = w^0 = u$, and thus $\tilde{\mathcal{M}}_k(u) = u$ for $k = 1, \dots, m$. Hence we get (18) from the stability assumption (3). □

Under the additional assumption of a compatible subspace decomposition we can show convergence to minimizers of $\mathcal{J}$.

Corollary 4.3. *Assume that, additionally to the assumptions of Theorem 4.1, the subspace decomposition $\mathbb{R}^n = \sum_{k=1}^m V_k$ is compatible with $\mathcal{J}$. Then any accumulation point of u^ν is a global minimizer of $\mathcal{J}$.*

Corollary 4.4. *Assume that, additionally to the assumptions of Corollary 4.3, $\mathcal{J}$ has a unique global minimizer u^*, then u^ν converges to u^*.*

Proof. Since $\mathcal{J}$ is coercive and $\mathcal{J}(u^\nu)$ is monotonically decreasing, the sequence u^ν is bounded. Hence, if u^ν does not converge to u^* it must have an accumulation point $u \neq u^*$, which contradicts Corollary 4.3. □

These convergence results cover many important cases; in particular, they imply global convergence to a minimizer for all example functionals listed in Chapter 2. For reference we state global convergence for strictly convex, block-separable nonsmooth functionals as a separate corollary.

Corollary 4.5. *Let $\mathcal{J} : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be strictly convex, coercive, proper, lower semi-continuous, and continuous on its domain. Assume that $\mathcal{J}$ is block-separable nonsmooth, and consider the induced subspaces $V_k := (\ker R_k)^\perp$. Then the TNNMG method with smoother $\mathcal{M} = \mathcal{M}_m \circ \cdots \circ \mathcal{M}_1$ and linear correction $\mathcal{C}$ satisfying the* monotonicity*, continuity, and* stability *assumptions of Theorem 4.1 converges globally to the unique minimizer of $\mathcal{J}$.*

This corollary subsumes previous results from [14, 25, 13] and others. However, the theory covers many more general situations as well, such as the biconvex energies in [31], and certain quasi-convex functionals.

5. Solution of local subproblems

In the following we will give several examples of local operators $\mathcal{M}_k$ that satisfy the assumptions of Theorem 4.1.[1] Evaluating such an operator is equivalent to the inexact minimization

$$\operatorname*{arg\,min}_{v \in w^{\nu,k-1} + V_k} \mathcal{J}(v) \tag{20}$$

in the k-th affine subspace $w^{\nu,k-1} + V_k$. These evaluations appear as local subproblems in Line 5 of the TNNMG algorithm, and form the nonlinear Gauß–Seidel smoother. In the following we only consider a single subspace $V := V_k$. Introducing $N = n_k$, and dropping all other indices we will write (20) as

$$\operatorname*{arg\,min}_{v \in w + V} \mathcal{J}(v) \tag{21}$$

for a given $w \in \mathbb{R}^n$.

5.1. Exact minimization.

The trivial choice for the local correction operator $\mathcal{M}$ is the exact minimization operator $\mathcal{M}^{\mathrm{ex}}$ given by

$$\mathcal{M}^{\mathrm{ex}}(w) := \operatorname*{arg\,min}_{v \in w + V} \mathcal{J}(v). \tag{22}$$

In the case where the minimizer is not unique, we assume that $\mathcal{M}^{\mathrm{ex}}(w)$ is any of the global minimizers in the affine subspace $w + V$. This is sufficient for the application of Theorem 4.1, because all assumptions there are stated in terms of $\mathcal{J} \circ \mathcal{M}$, which is invariant under the specific choice.

Lemma 5.1. *Let $\mathcal{J}$ be block-separable nonsmooth, and let the decomposition be induced by the block-structure, i.e.,* (9)*. Then the exact minimization operator $\mathcal{M} = \mathcal{M}^{ex}$ satisfies the assumptions of Theorem 4.1 with $\tilde{\mathcal{M}} = \mathcal{M}^{ex}$. In particular, the minimal energy $\mathcal{J} \circ \mathcal{M}^{ex}$ is continuous on* $\operatorname{dom} \mathcal{J}$.

[1] Note that "satisfy the assumptions" generally means that there exists a bounding minimization operator $\tilde{\mathcal{M}}_k$ with the necessary properties. In some cases this operator coincides with $\mathcal{M}_k$ itself.

Proof. The operator $\mathcal{M} = \tilde{\mathcal{M}} = \mathcal{M}^{\mathrm{ex}}$ is stable and monotone by construction. To show continuity of $\mathcal{J} \circ \mathcal{M}^{\mathrm{ex}}$, note that the block-separability of the functional

$$\mathcal{J}(v) = \mathcal{J}_0(v) + \sum_{k=1}^{M} \varphi_k(R_k v)$$

implies that its domain takes the form

$$\operatorname{dom} \mathcal{J} = \operatorname{dom} \varphi_k \times \Big(\prod_{i \neq k} \operatorname{dom} \varphi_i \Big).$$

The assertion then follows from Corollary A.2. □

If the subspace decomposition is not induced by the block-separable structure of $\mathcal{J}$, then the minimization (22) is in general not continuous. We refer to [14] for an example. The following result shows that continuity still holds if the domain is polyhedral. It is a generalization of a result shown in [14] for the case of a convex $\mathcal{J}$.

Lemma 5.2. *Let* $\operatorname{dom} \mathcal{J}$ *be a convex polyhedron. Then* $\mathcal{M} = \tilde{\mathcal{M}} = \mathcal{M}^{ex}$ *satisfies the assumptions of Theorem 4.1.*

Proof. Monotonicity and stability are again given by construction, while continuity for polyhedral domains is shown in Corollary A.3. □

We can also show that inexact versions of $\mathcal{M}^{\mathrm{ex}}$ satisfy these assumptions if they guarantee sufficient descent.

Lemma 5.3. *Let the functional* $\mathcal{J}$ *and the subspace decomposition satisfy either the assumptions of Lemma 5.1 or of Lemma 5.2. Assume that the operator*

$$\mathcal{M} : \operatorname{dom} \mathcal{J} \to \operatorname{dom} \mathcal{J}, \qquad \mathcal{M} - Id : \operatorname{dom} \mathcal{J} \to V$$

satisfies the sufficient descent condition

$$\mathcal{J}(w) - \mathcal{J}(\mathcal{M}(w)) \geq \varepsilon \big[\mathcal{J}(w) - \mathcal{J}(\mathcal{M}^{ex}(w)) \big]$$

for a fixed $\varepsilon > 0$*. Then* $\mathcal{M}$ *satisfies the assumptions of Theorem 4.1.*

Proof. By continuity of $\mathcal{J}$ on $\operatorname{dom} \mathcal{J}$, for any $w \in \operatorname{dom} \mathcal{J}$ there exists a $\tilde{w} \in \operatorname{dom} \mathcal{J} \cap w + V$ with

$$\mathcal{J}(\tilde{w}) = (1 - \varepsilon)\mathcal{J}(w) + \varepsilon \mathcal{J}(\mathcal{M}^{\mathrm{ex}}(w)).$$

Now we set $\tilde{\mathcal{M}}(w) = \tilde{w}$. Then we have

$$\mathcal{J}(\mathcal{M}(w)) \leq (1 - \varepsilon)\mathcal{J}(w) + \varepsilon \mathcal{J}(\mathcal{M}^{\mathrm{ex}}(w)) = \mathcal{J}(\tilde{\mathcal{M}}(w)) \leq \mathcal{J}(w),$$

which is the required monotonicity. Continuity of $\mathcal{J}$ and of $\mathcal{J} \circ \mathcal{M}^{\mathrm{ex}}$ (Lemma 5.1) imply continuity of $\mathcal{J} \circ \tilde{\mathcal{M}}$. Furthermore $\mathcal{J}(\tilde{\mathcal{M}}(w)) = \mathcal{J}(w)$ implies $\mathcal{J}(\mathcal{M}^{\mathrm{ex}}(w)) = \mathcal{J}(w)$ which shows stability. □

5.2. **Polyhedral Gauß–Seidel.** We now consider the case of a block-separable nonsmooth functional

$$\mathcal{J}(v) = \mathcal{J}_0(v) + \sum_{k=1}^{M} \varphi_k(R_k v)$$

where, additionally, each φ_k is piecewise smooth on a partition of its domain into a finite set a convex polyhedra. Furthermore we assume the induced subspace decomposition (9). Here, solving the piecewise smooth subproblems in each subspace V_k can be nontrivial problem.

For the case that $\operatorname{dom}\varphi_k$ is a simplex, [25] proposed to further split V_k into the one-dimensional spaces spanned by the simplex edges and to successively minimize in each of those spaces. This was generalized in [14] to general polyhedral partitions, by using spaces

$$V_k = \sum_{i=1}^{m_k} V_{k,i},$$

where the one-dimensional subspaces $V_{k,i}$ are aligned with the edges of the polyhedral partition of $\operatorname{dom}\varphi_k$. For the precise definition of these subspaces and the treatment of the case that the edges do not span the whole space we refer to [14].

Instead of showing for each V_k that this procedure takes the form of an inexact minimization operator $\mathcal{M}_k$ in V_k in the sense of Theorem 4.1, we can simply apply this theorem to the decomposition

$$\mathbb{R}^n = \sum_{k=1}^{m} \sum_{i=1}^{m_k} V_{k,i} \tag{23}$$

and use an exact or inexact minimization step in each one-dimensional subspace $V_{k,i}$. Such inexact minimization operators $\mathcal{M}_{k,i}$ for $V_{k,i}$ can be constructed by straightforward application of bisection with fully practical termination criteria, see, e.g., [14]. The resulting relaxation method is called the *Polyhedral Gauß–Seidel* method. A variant of the following theorem has been the main result of [14].

Theorem 5.4. *Let $\mathcal{J} : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be coercive, proper, lower-semicontinuous, and continuous on its domain. Suppose that there exists a finite partition of* $\operatorname{dom}\varphi$ *into nondegenerate, convex, closed polyhedra such that φ is piecewise smooth with respect to this partition. Consider the decomposition of $\mathbb{R}^n$ into one-dimensional subspaces $V_{k,i}$ such that any tangent cone of the partition is generated by vectors from the one-dimensional spaces $V_{k,i}$.*

On each such subspace, let $\mathcal{M}_{k,i}$ be the exact minimization operator (22) *or its inexact cousin in the sense of Lemma 5.3. Then all accumulation points of the TNNMG method are stationary. Additionally, if $\mathcal{J}$ is convex and has a unique minimizer, then the method will converge to that minimizer.*

Proof. The result follows directly from Theorem 4.1 and Corollary 4.4, if we can show continuity of $\mathcal{J} \circ \mathcal{M}_{k,i}$ and compatibility of the decomposition (23). The continuity of $\mathcal{J} \circ \mathcal{M}_{k,i}$ was shown in Lemmas 5.2 and 5.3 for the exact and the inexact case, respectively. The compatibility of the splitting with $\mathcal{J}$ was shown in Lemma 5.1 of [14]. □

5.3. **First-order models.** In this section we assume that $\mathcal{J}$ is block-separable nonsmooth and that the decomposition is induced by the block-structure, i.e., (9) holds. Then the local subproblems (21) take the form

$$v^* \in \mathbb{R}^N : \qquad f(v^*) \leq f(v) \qquad \forall v \in \mathbb{R}^N$$

where

$$f(v) = f_w(v) = \underbrace{\mathcal{J}_0(w + Pv)}_{=:f_0(v)} + \underbrace{\varphi_k(R_k w + v)}_{=:\psi(v)}, \tag{24}$$

and $P = P_k$ is the prolongation operator defined in (4). Under the assumption that f is convex we will now construct inexact solvers for these problems by solving approximate problems exactly. To this end we introduce the notion of first-order dominating models. These models differ from standard first-order models by only approximating the smooth part f_0 of the functional f, while the nonsmooth part ψ is treated exactly.

Since we only need these models for incremental problems we can simplify the situation without loss of generality by only considering models fitted to f in the origin.

Definition 5.1. *A functional $M : \mathbb{R}^N \to \mathbb{R} \cup \{\infty\}$ is called a* first-order model *for f at $0 \in \operatorname{dom} \psi$ if there is a continuously differentiable $M_0 : \mathbb{R}^N \to \mathbb{R}$ with*

$$M(v) = f_0(0) + \langle f_0'(0), v\rangle + M_0(v) + \psi(v)$$

such that $M_0(0) = 0$ and $M_0'(0) = 0$.

First-order models allow to detect minimizers, i.e., if the original functional is not minimal in 0, then neither is the model.

Lemma 5.5. *Let M be a first-order model of f at $0 \in \mathbb{R}^N$, and $v \in \mathbb{R}^N$ with $f(v) < f(0)$. Then there is a $t > 0$ with $M(tv) < M(0) = f(0)$.*

Proof. From convexity of f we get for any $t \in (0, 1)$ that

$$\begin{aligned} 0 > f(v) - f(0) &\geq \frac{1}{t}(f(tv) - f(0)) \\ &\geq \frac{1}{t}\big[\langle f_0'(0), tv\rangle + \psi(tv) - \psi(0)\big]. \end{aligned}$$

As M_0 is differentiable with $M_0'(0) = 0$ we get for sufficiently small $t > 0$ that

$$\frac{1}{2}\underbrace{(f(0) - f(v))}_{>0} \geq \frac{1}{t}\big[M_0(tv) + M_0(0)\big].$$

Adding both inequalities and multiplying by $t > 0$ yields

$$0 > \frac{t}{2}\big(f(v) - f(0)\big) \geq M(tv) - M(0). \qquad \square$$

Although a first-order model can only be minimal if the exact functional is minimal, minimizing the model does not necessarily decrease the value of the original functional. In order to obtain this property the model must also be *dominating*.

Definition 5.2. *A first-order model $M : \mathbb{R}^N \to \mathbb{R} \cup \{\infty\}$ for f at $0 \in \operatorname{dom} \psi$ is called* dominating *if it satisfies $M(v) \geq f(v)$ for all $v \in \mathbb{R}^N$.*

We will now show that exact minimization of a dominating first-order model leads to an inexact minimization operator $\mathcal{M}$ in the sense of Theorem 4.1. Here, the exact minimization of the model will play the role of the continuous minimization operator $\tilde{\mathcal{M}}$. We define

$$\mathcal{M}^{\mathrm{mod}}(w) := w + P \operatorname*{arg\,min}_{v \in \mathbb{R}^N} M(v). \tag{25}$$

We will only consider the case that M_0 does not depend on w. The following results can be generalized to w-dependent M_0 under strong technical continuity assumptions for M_0' with respect to w.

Theorem 5.6. *For any $w \in \operatorname{dom} \mathcal{J}$ let f_w be as defined in (24). Set*

$$M_w(v) := f_{w,0}(0) + \langle f_{w,0}'(0), v\rangle + M_0(v) + \psi(v)$$

for some fixed function M_0, and assume that M_w is a first-order dominating model for f_w in each $w \in \operatorname{dom} \mathcal{J}$. Furthermore assume that M_0' is uniformly monotone, i.e., there are constants $\alpha > 0$ and $p > 1$ such that

$$\alpha \|u - v\|^p \le \langle M_0'(u) - M_0'(v), u - v\rangle \qquad \forall u, v \in \mathbb{R}^N.$$

Then $\mathcal{M}^{mod}$ as defined in (25) is an inexact minimization operator in the sense of Theorem 4.1.

Proof. By construction any dominating model leads to a monotone $\mathcal{M}^{\mathrm{mod}}$, and Lemma 5.5 shows that first-order models lead to stable $\mathcal{M}^{\mathrm{mod}}$. Now we show continuity of $\mathcal{M}^{\mathrm{mod}}$ and hence of $\mathcal{J} \circ \mathcal{M}^{\mathrm{mod}}$.

For $i = 1, 2$ let $w_i \in \operatorname{dom} \mathcal{J}$ and $v_i = \arg\min_{v \in \mathbb{R}^N} M_{w_i}(v)$. Then we have

$$\left\langle M_0'(v_i) + f_{w_i,0}'(0), v - v_i \right\rangle + \psi(v) - \psi(v_i) \ge 0.$$

Testing this variational inequality for $i = 1, 2$ with v_j, $j \ne i$, adding the results, and using monotonicity gives

$$\begin{aligned}\alpha \|v_1 - v_2\|^p \le \left\langle M_0'(v_1) - M_0'(v_2), v_1 - v_2 \right\rangle &\le \|f_{w_1,0}'(0) - f_{w_2,0}'(0)\| \|v_1 - v_2\| \\ &\le \|P\| \|\mathcal{J}_0'(w_1) - \mathcal{J}_0'(w_2)\| \|v_1 - v_2\|.\end{aligned}$$

Dividing by $\|v_1 - v_2\|$ and exploiting continuity of $\mathcal{J}_0'$ we find that $\arg\min_{v \in \mathbb{R}^N} M_w(v)$ and thus $\mathcal{M}$ depends continuously on w. □

The theory of first-order dominating models is useful because it is frequently easy to construct models that are much easier to minimize than the actual functional, while at the same time producing sufficient descent to act as a nonlinear smoother for a fast multigrid method. In the following we briefly discuss how such models can be constructed.

The proof of Lemma 5.8 below requires the following sufficient condition, which is of independent interest.

Lemma 5.7. *Let M be a first-order model of f at $0 \in \mathbb{R}^N$, and*

$$\langle f_0'(v) - f_0'(0), v\rangle \le \langle M_0'(v) - M_0'(0), v\rangle \qquad \forall v \in \mathbb{R}^N.$$

Then M is also a dominating model.

Proof. For $v \in \mathbb{R}^N$ we have

$$\begin{aligned} f(v) &= \psi(v) + f_0(0) + \langle f_0'(0), v\rangle + \int_0^1 \frac{1}{t}\langle f_0'(tv) - f_0'(0), tv\rangle \, dt \\ &\leq \psi(v) + f_0(0) + \langle f_0'(0), v\rangle + \int_0^1 \langle M_0'(tv), v\rangle - \langle M_0'(0), v\rangle \, dt = M(v), \end{aligned}$$

which is the assertion. □

With this result we can construct models with a quadratic smooth part.

Lemma 5.8. *Let $B \in \mathbb{R}^{N\times N}$ be a symmetric positive semi-definite matrix. Then*

$$M(v) = f_0(0) + \langle f_0'(0), v\rangle + \frac{1}{2}\langle Bv, v\rangle + \psi(v)$$

is a first-order model for f at 0. If B satisfies

$$\langle f_0'(v) - f_0'(0), v\rangle \leq \langle Bv, v\rangle \qquad \forall v \in \mathbb{R}^N, \tag{26}$$

then M is dominating.

Proof. The fact that M is a first-order model is trivial. If B satisfies (26), then $M_0'(v) - M_0'(0) = Bv$, and Lemma 5.7 implies that M is dominating. □

Example 5.1. Let $f(v) = \frac{1}{2}\langle Av, v\rangle - \langle b, v\rangle + \psi(v)$ and $\langle Av, v\rangle \leq \langle Bv, v\rangle$ for all $v \in \mathbb{R}^N$ for symmetric positive definite matrices A, B. Then

$$M(v) = \frac{1}{2}\langle Bv, v\rangle - \langle b, v\rangle + \psi(v)$$

is a dominating first-order model for f at 0.

Models of this type are interesting, because approximating matrices B can be easy to construct as long as A is small. The following radical choice can be applied to plasticity and friction problems, which are represented by Example 2.4.

Example 5.2. Let f and M be as in Example 5.1 with $\psi(v) = h(\|v_0 + v\|)$, and $B = \alpha I$ with $\alpha \geq \lambda_{\max}(A)$. Then M can be written as

$$\begin{aligned} M(v) &= \frac{\alpha}{2}\|v\|^2 - \langle b, v\rangle + \psi(v) \\ &= \frac{\alpha}{2}\|v_0 + v\|^2 - \langle b + \alpha v_0, v_0 + v\rangle + \frac{\alpha}{2}\|v_0\|^2 + \langle b, v_0\rangle + \psi(v) \\ &= \tilde{h}(\|v_0 + v\|) - \langle \tilde{r}, v_0 + v\rangle + \text{const}, \end{aligned}$$

with $\tilde{h}(t) = \frac{\alpha}{2}t^2 + h(t)$ and $\tilde{r} = b + \alpha v_0$.

In this example M is rotationally symmetric with respect to $v - v_0$. Hence the minimizer u of M can be computed by solving a scalar minimization problem on the line $\{t\tilde{r} - v_0 \mid t \in \mathbb{R}\} \subset \mathbb{R}^N$.

Lemma 5.9. *Let $h : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$, $r \in \mathbb{R}^N \setminus \{0\}$ and $v_0 \in \mathbb{R}^N$. Then any minimizer u of*

$$g : \mathbb{R}^N \to \mathbb{R} \cup \{\infty\}, \qquad g(v) = h(\|v_0 + v\|) - \langle r, v_0 + v\rangle$$

satisfies $u = tr - v_0$ for $t > 0$.

Proof. Let u be a minimizer and assume that u is not of the form $u = tr - v_0$. Then $|\langle r, v_0 + u\rangle| < \|r\|\|v_0 + u\|$. Now let $\tilde{u} = \frac{\|v_0+u\|}{\|r\|} r - v_0$. Then we have

$$\begin{aligned} g(\tilde{u}) &= h(\|v_0 + u\|) - \|r\|\|v_0 + u\| \\ &< h(\|v_0 + u\|) - |\langle r, v_0 + u\rangle| \le h(\|v_0 + u\|) - \langle r, v_0 + u\rangle = g(u), \end{aligned}$$

which contradicts the assumption. □

Another example where the simplification proposed in Example 5.2 can be helpful are problems with local simplex constraints as introduced in Example 2.3.

Example 5.3. Let f and M be as in Example 5.1 with $B = \alpha I$, $\alpha \ge \lambda_{\max}(A)$, and $\psi(v) = \chi_G(v_0 + v)$ where χ_G is the indicator function of the Gibbs simplex G defined in (10). Similar to Example 5.2 the dominating first-order model M can now be written as

$$M(v) = \frac{\alpha}{2}\|v_0 + v - \tilde{r}\|^2 + \chi_G(v_0 + v) + \text{const}$$

with $\tilde{r} = \frac{1}{\alpha} b + v_0$. Hence the minimizer u of M can be written as $u = u_0 - v_0$ where $u_0 = \arg\min_{z\in G} \|z - \tilde{r}\|^2$ is the Euclidean projection of $\tilde{r}$ into G. This is an important simplification, because this projection can be computed exactly with a simple algorithm in $O(N \log(N))$ time [15, 38].

The previous examples assumed that the smooth part $\mathcal{J}_0$ of the functional is quadratic, and replaced it by a simpler quadratic functional. For problems like Example 2.2 where this is not the case, using local first-order models can still be beneficial. The construction of such models is based on the following direct consequence of the chain rule.

Lemma 5.10. *Let the smooth part $\mathcal{J}_0$ of the global functional take the form*

$$\mathcal{J}_0(v) = \sum_{i=1}^{l} \gamma_i(D_i v)$$

with $D_i \in \mathbb{R}^{d\times n}$ and $\gamma_i : \mathbb{R}^d \to \mathbb{R}$ convex and continuously differentiable for $i = 1, \dots, l$. Assume that each γ_i' is Lipschitz continuous with a Lipschitz constant L_i. Then

$$\langle \mathcal{J}_0'(x) - \mathcal{J}_0'(y), x - y\rangle \le \langle \mathcal{B}(x - y), x - y\rangle \qquad \forall x, y \in \mathbb{R}^n$$

for the symmetric positive semi-definite matrix $\mathcal{B} = \sum_{i=1}^{l} L_i D_i^T D_i \in \mathbb{R}^{n\times n}$.

Example 5.4. Let the global problem take the form given in Example 2.2 with $\mathcal{J}_0$ as in Lemma 5.10. For the local problem consider the first-order model given in Lemma 5.8 with $B = P^T \mathcal{B} P$. Then Lemma 5.10 together with the chain rule shows that

$$\langle f_0'(x) - f_0'(y), x - y\rangle \le \langle B(x - y), x - y\rangle \qquad \forall x, y \in \mathbb{R}^N.$$

Hence, by Lemma 5.8 the model is also dominating. In contrast to f_0 the functional M_0 is quadratic such that the minimizer of M can be directly computed. For problems with more general nonsmooth terms φ_k where this is not the case, using an iterative method to minimize M will often be much faster because M_0 is cheaper to evaluate.

When using the first-order models presented here as the basis for an iterative method to actually solve the local minimization problem for f, the resulting algorithms are variants of so-called proximal methods. In case of the Examples 5.2 and 5.3 these are proximal gradient type methods, while the model in Example 5.4 would lead to a proximal Newton-type method. In the case that the nonsmooth part ψ of f is the indicator function of a convex set while the model M_0 for the smooth part is quadratic with a matrix $B = \alpha I$, such methods can also be viewed as projected gradient-type methods.

We emphasize that we do not propose and analyze the resulting proximal methods here, but only discuss their use as inexact local solvers inside of a nonlinear block Gauß–Seidel smoother. For a discussion of proximal gradient and proximal Newton-type iterations we refer to [27] and the references therein.

6. Truncated Newton corrections

6.1. **Constructing suitable correction spaces.** The second step of the TNNMG iteration consists of an inexact Newton step in an iteration-dependent subspace W_ν of $\mathbb{R}^n$ that is constructed such that all necessary derivatives of $\mathcal{J}$ are well-defined. In principle W_ν can be defined as the largest subspace of $\mathbb{R}^n$ such that there is an open ball $B_\varepsilon(u^{\nu+\frac{1}{2}})$ of radius $\varepsilon > 0$ around $u^{\nu+\frac{1}{2}}$ such that $\mathcal{J}$ is twice continuously differentiable on

$$(u^{\nu+\frac{1}{2}} + W_\nu) \cap B_\varepsilon(u^{\nu+\frac{1}{2}}). \tag{27}$$

However, depending on the problem, the practical construction of this subspace can be technical.

In many cases, the implementation can be simplified by using a correction space W_ν that is smaller than possible. The practical behavior of the TNNMG method may be unaffected when W_ν is replaced by a slightly smaller space, and the convergence results hold for all spaces W_ν that allow for a well-defined Newton problem.

Let the energy functional $\mathcal{J}$ be block-separable nonsmooth i.e.,

$$\mathcal{J}(v) = \mathcal{J}_0(v) + \sum_{k=1}^{M} \varphi_k(R_k v), \qquad v \in \mathbb{R}^n,$$

with a C^2 functional $\mathcal{J}_0$. Then it is clear, that the space W_ν can be defined as a product space

$$W_\nu := R^{-1} \prod_{k=1}^{M} W_{\nu,k} \tag{28}$$

where $W_{\nu,k} \subset \mathbb{R}^{N_k}$ is a local subspace such that φ_k is locally smooth near $R_k u^{\nu+\frac{1}{2}}$ in the above given sense. Here we make use of the isomorphism $R : \mathbb{R}^n \to \prod_{k=1}^{M} \mathbb{R}^{N_k}$ translating between $\mathbb{R}^n$ and the block representation according to Definition 2.2 such that the k-th block of $v \in \mathbb{R}^n$ is given by $R_k v \in \mathbb{R}^{N_k}$.

The easiest way to construct a suitable correction space W_ν is by disabling entire blocks. For a given pre-smoothed iterate $u^{\nu+\frac{1}{2}}$ define the set of "inactive" blocks

$$\mathcal{N}_\nu^\circ := \Big\{ k = 1, \dots, M \mid \varphi_k''(R_k(\cdot)) \text{ exists and is continuous on } B_\varepsilon(R_k u^{\nu+\frac{1}{2}}) \text{ for some } \epsilon > 0 \Big\}. \tag{29}$$

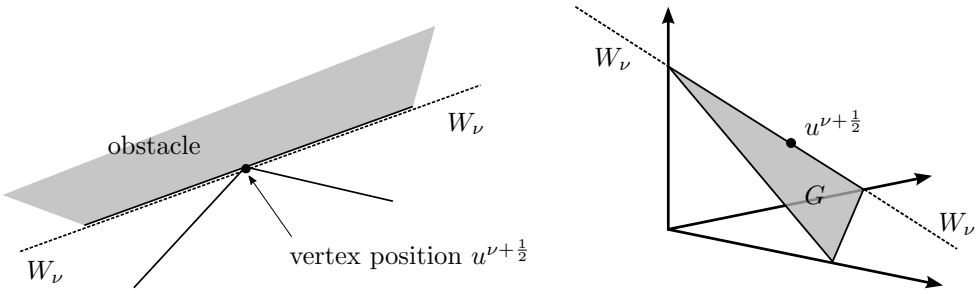

FIGURE 2. Constructing the correction spaces W_ν. Left: Contact problem. At a grid vertex in contact, W_ν can contain the tangential direction. Right: Simplex-constrained problem. The correction spaces can contain the affine span of simplex faces.

Then frequently a reasonable correction space is

$$W_\nu := R^{-1} \prod_{k=1}^{M} \begin{cases} \mathbb{R}^{N_k} & \text{if } k \in \mathcal{N}_\nu^\circ, \\ \{0\}^{N_k} & \text{otherwise.} \end{cases} \tag{30}$$

When all blocks are one-dimensional, this definition even yields the optimal space. It also produces the optimal space for functionals as in Example 2.4, where the nonsmoothness is a block-wise norm function

$$\varphi_k(v) := \|v\| \qquad \forall v \in \mathbb{R}^{N_k}.$$

In other cases $\mathcal{J}'$ and $\mathcal{J}''$ may exist in larger spaces than the one defined by (30). As an example, consider a contact problem in linear elastomechanics. There, it can be convenient to group the degrees of freedom in d-dimensional blocks, corresponding to the Lagrange nodes of the finite element space. For each $k = 1, \dots, M$, the nonsmooth function φ_k is then the indicator functional of a half-space in $\mathbb{R}^d$, which models the restriction on the normal displacement of the k-th Lagrange node. If the degrees of freedom in a block k are in contact, the naive approach (30) disables that block completely. On the other hand, φ_k is still C^2 in the tangential plane of the half-space, and a bigger space W_ν could therefore be constructed by including these tangent directions [13] (Figure 2, left). Such a space then includes sliding along the obstacle in the linear correction step.

Vector-valued phase-field models as described in Example 2.3 form a more extreme example. In this example the degrees of freedom are grouped in blocks of dimension $N^k = L$, and the nondifferentiability φ_k is the indicator functional χ_G of the Gibbs simplex G (10). Here, block-wise truncation as in (30) would truncate the whole space, since χ_G is nowhere differentiable. Hence the linear correction would entirely be removed and the method would degenerate to the pure nonlinear smoother. Truncating the block only if the solution is on the relative boundary of G and keeping the tangent space to G otherwise is still not a good choice, because solutions typically live on the boundary of G in almost every block $k = 1, \dots, M$. Again this would effectively result in an almost complete removal of the linear correction.

Fortunately, for this example the largest possible correction space W_ν is much larger compared to the resulting space in both of these simple constructions. To

really construct W_ν we consider the active sets

$$\mathcal{N}^\bullet_{\nu,k} := \Big\{ j = 1, \dots, L \mid (R_k u^{\nu+\frac{1}{2}})_j = 0 \Big\}$$

for each block $R_k u^{\nu+\frac{1}{2}} \in G$. Then $\mathcal{J}$ is differentiable near $u^{\nu+\frac{1}{2}}$ in the space

$$W_\nu := R^{-1} \prod_{k=1}^{M} \operatorname{span}\Big\{ \eta^{ij} \in \mathcal{E} \,\Big|\, i,j \notin \mathcal{N}^\bullet_{\nu,k} \Big\}, \tag{31}$$

where

$$\mathcal{E} = \Big\{ \eta^{ij} = e^i - e^j \in \mathbb{R}^L \,\Big|\, 1 \le i < j \le L \Big\}$$

is the set of edges of G. It is easy to see that this space is in fact the maximal subspace of $\mathbb{R}^n$ where $\mathcal{J}(u^{\nu+\frac{1}{2}} + \cdot)$ is differentiable near the origin: On the one hand restriction of the nonsmooth term $\varphi(u^{\nu+\frac{1}{2}} + \cdot)$ to W_ν is constant in a ball around the origin such that the functional is locally C^2. On the other hand $\varphi(u^{\nu+\frac{1}{2}} + \cdot)$ has a jump from 0 to ∞ at the origin along any other direction.

Remark 6.1. *Unfortunately, in an implementation, the decision of whether a degree of freedom is active or not can only be taken in an approximate sense, by replacing a condition like $R_k u = 0$ by $|R_k u| \le \varepsilon$ due to problems caused by finite-precision arithmetic. This effectively replaces the arbitrarily small positive ε that appears in* (27) *and* (29) *by a fixed one. The choice of ε can be important: if ε is too small, then the solver may become unstable. If it is too large, the convergence rates deteriorate. This effectively introduces a parameter into the algorithm; however, in a finite element context, while this parameter depends on the boundary value problem that is being solved, it does not depend on the discretization.*

For some applications continuous differentiability is not the only criterion that guides the construction of the correction spaces W_ν. For example, more realistic examples of phase-field models include a nonlinear energy whose derivatives become singular when approaching the boundary of the admissible set [12]. Using a linearization of $\mathcal{J}$ in a space constructed as discussed above may lead to unbounded derivatives as some u_k approach the singularity. As a consequence the condition number of the linear system may become arbitrarily large, independently of the spatial discretization parameter leading to numerical problems and slow convergence of linear solvers.

To avoid this the general truncation strategy can be extended by—additionally to twice continuous differentiability—requiring that the second derivatives do not become "too large" in the truncated ball (27). In practice this can be achieved by also checking the values of second-order derivatives or the distance to the singularity in the construction of the space. For example, in a multi-component phase-field model where φ_k takes the form

$$\varphi_k(z) = \chi_G(z) + \sum_{j=1}^{L} \varphi_{k,j}(z_j)$$

and each $\varphi'_{k,j}$ and $\varphi''_{k,j}$ is singular in zero, this leads to active sets defined by

$$\mathcal{N}^\bullet_{\nu,k} := \Big\{ j = 1, \dots, L \,\Big|\, (R_k u^{\nu+\frac{1}{2}})_j = 0 \text{ or } \varphi''_{k,j}((R_k u^{\nu+\frac{1}{2}})_j) > C \Big\}$$

for a large constant C. These active sets are then used in the definition (31) of the correction space W_ν. This modification can lead to a considerable increase in efficiency.

Certain energies, such as the anisotropic fracture energy in [31], have a smooth energy contribution $\mathcal{J}_0$ that is is not C^2 but only differentiable with a locally Lipschitz-continuous derivative (LC^1). For such cases, the definition of the space W_ν can be relaxed, by replacing second derivatives in the definition of W_ν by generalized derivatives of $\mathcal{J}'$. In this case the correction step (11) takes the form of a generalized Newton-step as proposed, e.g., in [32].

6.2. **Algebraic representations.** In practical implementations of the TNNMG method, matrix and vector representations of $\mathcal{J}''(u^{\nu+\frac{1}{2}})|_{W_\nu\times W_\nu}$ and $\mathcal{J}'(u^{\nu+\frac{1}{2}})|_{W_\nu}$ are needed, respectively. For simplicity, implementations are suggested to always work with matrices and vectors in $\mathbb{R}^n$, and extend $\mathcal{J}''(u^{\nu+\frac{1}{2}})|_{W_\nu\times W_\nu}$ and $\mathcal{J}'(u^{\nu+\frac{1}{2}})|_{W_\nu}$ by zero to the orthogonal complement of W_ν in $\mathbb{R}^n$.

This extension is simple if W_ν is constructed by the truncation of active blocks (30). In this case, the vector representation of $\mathcal{J}'(u^{\nu+\frac{1}{2}})|_{W_\nu}$ is simply the vector in $\mathbb{R}^n$ that contains the partial derivatives of $\mathcal{J}$ for all $k\in\mathcal{N}_\nu^\circ$ (which exist by construction of $\mathcal{N}_\nu^\circ$), and zero elsewhere. Similarly, the matrix representation of $\mathcal{J}''(u^{\nu+\frac{1}{2}})|_{W_\nu\times W_\nu}$ contains the second partial derivative in all blocks where neither the column nor the row index is active, and zero otherwise. This makes the matrix semi-definite.

Remark 6.2. *It would be easy to make the matrix positive definite by replacing the zeros on the diagonal by a fixed positive constant α. However, this constant, if chosen badly, may influence the condition number of the linear correction system. If possible we prefer to modify the linear solver to handle the semi-definite problem directly (see Chapter 7).*

For the general case, denote by Π_S the orthogonal projection onto a closed subspace S, and let $\mathcal{J}'(u^{\nu+\frac{1}{2}})$ and $\mathcal{J}''(u^{\nu+\frac{1}{2}})$ be a generalized gradient and Hessian of $\mathcal{J}$ at $u^{\nu+\frac{1}{2}}\in\mathbb{R}^n$, respectively. Here we understand "generalized" in the sense that the entries are the partial derivatives of $\mathcal{J}$ wherever they exist, and arbitrary elsewhere. If $\mathcal{J}$ is block-separable nonsmooth, then, by the block structure (28) of W_ν, the global projection $\Pi_{W_\nu}:\mathbb{R}^n\to W_\nu$ is (up to the isomorphism R) a block-diagonal matrix with diagonal blocks

$$(R\Pi_{W_\nu}R^{-1})_{kk}=R_k\Pi_{W_\nu}R_k^{-1}=\Pi_{W_{\nu,k}}\in\mathbb{R}^{N_k\times N_k},$$

where $\Pi_{W_{\nu,k}}$ is the projection onto $W_{\nu,k}=R_kW_\nu$. For example, for simplex-constrained problems we get [15]

$$\left(\Pi_{W_{\nu,k}}\right)_{ij}=\begin{cases}\delta_{ij}-\frac{1}{L-|\mathcal{N}_{\nu,k}^\circ(R_ku^{\nu+\frac{1}{2}})|} & \text{if } i,j\notin\mathcal{N}_{\nu,k}^\circ(w_k),\\ 0 & \text{otherwise.}\end{cases}$$

Using Π_{W_ν} we can write the representations of the truncated gradient and Hessian in terms of the coordinate system in $\mathbb{R}^n$ as

$$\begin{aligned}g_\nu&:=\mathcal{J}'(u^{\nu+\frac{1}{2}})|_{W_\nu}=\Pi_{W_\nu}^T\mathcal{J}'(u^{\nu+\frac{1}{2}}),\\ H_\nu&:=\mathcal{J}''(u^{\nu+\frac{1}{2}})|_{W_\nu\times W_\nu}=\Pi_{W_\nu}^T\mathcal{J}''(u^{\nu+\frac{1}{2}})\Pi_{W_\nu}.\end{aligned}$$

As Π_{W_ν} is block-diagonal $\mathcal{J}''(u^{\nu+\frac{1}{2}})|_{W_\nu \times W_\nu}$ inherits the sparsity pattern of $\mathcal{J}''(u^{\nu+\frac{1}{2}})$, and all block-entries can be computed independently according to

$$(R\mathcal{J}''(u^{\nu+\frac{1}{2}})|_{W_\nu \times W_\nu} R^{-1})_{ij} = \Pi_{W_{\nu,i}}^T \big(R\mathcal{J}''(u^{\nu+\frac{1}{2}})R^{-1}\big)_{ij} \Pi_{W_{\nu,j}}.$$

Using this matrix representation the solution of (11) is given by $v^\nu = -H_\nu^+ g_\nu$, where $(\cdot)^+$ is the Moore–Penrose pseudo-inverse, and this is independent of the entries selected for non-existing partial derivatives. See [15] for more details on TNNMG methods for vector-valued phase-field problems.

7. Inexact solution of linear subproblems

The final piece in the puzzle is how to produce suitable inexact solutions of the linear correction systems (11). The convergence result in Theorem 4.1 only requires that the linear correction shall not increase the energy, a feature that is always true due to the subsequent line search step. However, for fast error reduction rates you need something that causes global information exchange. Various alternatives exist; the best choice depends on the situation.

7.1. Geometric multigrid iterations.

In many cases, the minimization problem for $\mathcal{J}$ on $\mathbb{R}^n$ originates from the discretization of a partial differential equation on a function space $\mathcal{S}_J$. The default choice for an inexact solver for (11) is then a single V-cycle iteration of geometric multigrid. With this choice, the TNNMG algorithm can be interpreted as a multigrid method with a nonlinear pre-smoother, a truncated linear coarse grid correction, and a particular line search. One expects multigrid convergence rates at least asymptotically, and these are indeed frequently observed in practice [34, 13, 10]. Doing more than a single iteration, or a W-cycle iteration, may sometimes increase performance.

Implementations of geometric multigrid need minor modifications to cope with the restriction to the truncated correction space W_ν. Suppose we have a hierarchy $\mathcal{S}_0 \subset \cdots \subset \mathcal{S}_J$ of subspaces, with the natural embedding of $\mathcal{S}_{k-1}$ into $\mathcal{S}_k$ given by the prolongation matrix $\mathcal{P}_k \in \mathbb{R}^{\dim \mathcal{S}_k \times \dim \mathcal{S}_{k-1}}$. We set $A_J := \mathcal{J}''(u^{\nu+\frac{1}{2}})|_{W_\nu \times W_\nu}$ with the extension by zero to the orthogonal complement of W_ν as described in the previous chapter. From this we construct a hierarchy of stiffness matrices by setting $A_{k-1} := \mathcal{P}_k^T A_k \mathcal{P}_k$. We further assume the existence of linear smoothing operators $B_k : \mathbb{R}^{\dim \mathcal{S}_k} \to \mathbb{R}^{\dim \mathcal{S}_k}$. A linear multigrid step for (11) starting from a zero initial value takes the form

Input: Given $u^{\nu+\frac{1}{2}}$
Set $r^J = -\mathcal{J}'(u^{\nu+\frac{1}{2}})|_{W_\nu}$
for $k = J, \dots, 1$ **do**
  Compute $\hat{v}^k = B_k r^k$
  Set $r^{k-1} = \mathcal{P}_k^T (r^k - A_k \hat{v}^k)$
end
Compute $\hat{v}^0 = B_0 r^0$
Output: $v^\nu = \Pi_{W_\nu} \sum_{k=0}^{J} \Big(\prod_{l=k+1}^{J} \mathcal{P}_l \Big) \hat{v}^k$

In this simple form there is a single pre-smoothing step and no post-smoothing. The extension to a V-cycle with multiple pre- and post-smoothing steps is straightforward.

The first difference to a standard linear multigrid step is the projection Π_{W_ν} in the last line. It is needed since, while the method does in principle act in the quotient space $\mathbb{R}^n/\ker A_J = W_\nu$, it represents all corrections in the larger space $\mathbb{R}^n$. Since the prolongation operators are not aware of the kernel they may create spurious contributions in $\ker A_J$. However the corrections in the quotient space W_ν are invariant under those contributions and the latter can easily be removed by the additional projection onto W_ν.

The second difference is the choice of smoothers B_k. Since the matrices A_k are in general symmetric but positive semi-definite only, standard choices may not work as expected. Instead of approximating the inverse A_k^{-1} the smoothers B_k are intended to approximate the pseudo-inverse matrices A_k^+ such that they are invariant under any residual contribution in the kernel of the respective A_k. In the following, we drop the level index k for simpler notation and assume nested indices representing the block-structure of the matrices. A common choice is a block Gauß–Seidel method based on a block-triangular decomposition $A = D + L + L^T$, where $D_{ij} = \delta_{ij} A_{ij}$. However, $D + L$ may be singular because the diagonal entries D_{ii} are in general only positive semi-definite. Solving local systems involving the D_{ii} therefore requires a kernel-invariant local solver like, e.g., the CG method, or subspace correction methods as in [28].

Remark 7.1. *In certain situations, the following cheap modification can help to avoid numerical difficulties due to the infinite condition number of D_{ii}. For a small parameter $0 < \alpha \ll 1$ set*

$$\left(\tilde{D}_{ii}\right)_{lm} := \begin{cases} (D_{ii})_{lm} & \text{if } l \neq m, \\ \alpha & \text{if } l = m \text{ and } (D_{ii})_{lm} = 0, \\ \left((1+\alpha)D_{ii}\right)_{lm} & \text{if } l = m \text{ and } (D_{ii})_{lm} \neq 0. \end{cases}$$

Then $\tilde{D}_{ii}$ is symmetric positive definite, and the resulting correction obtained by a CG method for $\tilde{D}_{ii}$ can be viewed as a damped version of the one obtained for $D_{ii} + \Pi_{\ker D_{ii}}$. Numerical experiments in [15] *showed that the CG method was robust for $\alpha \geq 10^{-14}$, and the overall convergence rates were hardly influenced as long as $\alpha \leq 10^{-4}$.*

The modified linear multigrid algorithm becomes particularly simple if the restriction to the space W_ν is constructed by simply zeroing out certain matrix rows and columns. In this case, a suitable smoother is obtained by modifying a scalar Gauß–Seidel method to simply skip all rows with a zero matrix diagonal entry, and return a zero correction there.

7.2. **Algebraic multigrid iterations.** Geometric multigrid is difficult to apply if the PDE problem is posed on a given grid without a grid hierarchy. Luckily, in this situation algebraic multigrid (AMG) iterations can be used as well [6, 35]. In principle, any type of algebraic multigrid iteration can be used as part of the TNNMG method. The overall TNNMG convergence rates will depend on the quality of the AMG iteration as an independent solver. An AMG step that has merely value as a preconditioner is typically not sufficient because TNNMG expects suitably scaled inexact solutions to the linear system (11). However, this may be cured by using

an additional line search for the linear problem, i.e., by using single step of an AMG-preconditioned gradient method for (11).

If reduction to the truncated correction space W_ν makes the linear problems semi-definite, then the AMG step needs to be modified in the same way as the geometric multigrid step described above. As these modifications only concern the smoothers and additional projection operators, the techniques used for geometric multigrid work for AMG, too.

7.3. **Direct solvers and others.** If, for some reason, multigrid solvers are not available, then it is in principle possible to use any other solver for the correction problems (11). For example, one possible choice is (a number of steps of) a preconditioned CG method, which will even work out-of-the-box on semidefinite problems.

As a radical choice, it is even possible to use direct solvers, if the restriction to W_ν is implemented in a way that produces invertible correction matrices. TNNMG then stops being a multigrid method, and mutates into something related to nonsmooth Newton methods and active-set method. Certain predictor–corrector methods used for small-strain plasticity problems can be interpreted as variants of the TNNMG method with a direct solver for the correction problems [30, 34].

Appendix A. Continuity of minimization

In this section we present an auxiliary result on the continuity of minimization operators. More precisely we show for $F : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ and a family of subspaces $(V_x)_{x\in\mathbb{R}^n}$ of $\mathbb{R}^n$ that

$$m : \operatorname{dom} F \to \mathbb{R}, \qquad m(x) := \min_{v\in V_x} F(x+v)$$

is continuous if F and the family (V_x) are suitably well behaved. Only Corollaries A.2 and A.3 are used in this paper, in Section 5 on nonlinear smoothers. We nevertheless state the more general Theorem A.1 here for future reference.

Definition A.1. *Let $\mathcal{V} := (V_x)_{x\in\mathbb{R}^n}$ be a family of subspaces of $\mathbb{R}^n$. We call $K \subset \mathbb{R}^n$* stable with respect to translations in $\mathcal{V}$ *if for any sequence $x^\nu \in K$ with $x^\nu \to x \in K$ we have:*

1. *Any sequence $v^\nu \in V_{x^\nu}$ with $v^\nu \to v$ satisfies $v \in V_x$.*
2. *If $v \in V_x$ such that $x + v \in K$, then there is a sequence $v^\nu \in V_{x^\nu}$ such that $x^\nu + v^\nu \in K$ and $v^\nu \to v$.*

The main result shows continuity of m for suitable functions F whose domain $\operatorname{dom} F$ satisfies this stability assumption.

Theorem A.1. *Let $F : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ such that F is coercive, lower semi-continuous, and continuous on its domain. Furthermore, let $\mathcal{V} = (V_x)_{x\in\mathbb{R}^n}$ be a family of subspaces of $\mathbb{R}^n$ such that $K = \operatorname{dom} F$ is stable with respect to translations in $\mathcal{V}$. Then $x \mapsto m(x) = \min_{v\in V_x} F(x+v)$ is continuous on $\operatorname{dom} F$.*

Proof. First we note that coercivity and lower semi-continuity of F imply that for any $z \in K$ there is exists a minimizer $\mathcal{A}(z) := \arg\min_{v\in z+V_z} F(z+v) \in K$ (not necessarily unique) such that $m(z) = F(\mathcal{A}(z))$ is well-defined.

Now let $x^\nu \in K$ be a sequence with $x^\nu \to x \in K$ and assume that $F(\mathcal{A}(x^\nu)) \not\to F(\mathcal{A}(x))$. Continuity of F on K and coercivity imply

$$-\infty < \min_{z\in\mathbb{R}^n} F(z) \le F(\mathcal{A}(x^\nu)) \le F(x^\nu) \le C < \infty.$$

Hence $F(\mathcal{A}(x^\nu))$ is bounded and by coercivity $\mathcal{A}(x^\nu)$ is also bounded. Then there must be a subsequence (w.l.o.g. also indexed by ν) such that $\mathcal{A}(x^\nu) \to \hat{x}$ and Lemma 4.2 gives $\hat{x} \in K$. Hence we have

$$F(\mathcal{A}(x^\nu)) \to F(\hat{x}) \neq F(\mathcal{A}(x)) \tag{32}$$

by continuity of F on K. On the other hand the stability of K with respect to $\mathcal{V}$ together with $V_{x^\nu} \ni \mathcal{A}(x^\nu) - x^\nu \to \hat{x} - x$ gives $\hat{x} - x \in V_x$ and thus $F(\hat{x}) > F(\mathcal{A}(x))$.

Finally, the stability of K with respect to $\mathcal{V}$ also implies that there is a sequence $v^\nu \in V_{x^\nu}$ such that $x^\nu + v^\nu \in K$ and $v^\nu \to \mathcal{A}(x) - x \in V_x$. Then we have

$$F(\mathcal{A}(x^\nu)) \le F(x^\nu + v^\nu) \to F(\mathcal{A}(x)) < F(\hat{x}),$$

which contradicts (32). □

The first condition in Definition A.1 is trivially satisfied if V_x is a fixed subspace of $\mathbb{R}^n$. A simple example where the second condition also holds are domains with product structure.

Corollary A.2. *Let $F : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be coercive, lower semi-continuous, and continuous on its domain. Furthermore let $K := \operatorname{dom} F = K_1 \times K_2$ be a product of subsets $K_1 \subset \mathbb{R}^{n_1}$ and $K_2 \subset \mathbb{R}^{n_2}$ and let $V = \mathbb{R}^{n_1} \times \{0\}$. Then K is stable with respect to translations in V, and m is continuous on K.*

Proof. To show stability of K with respect to translations in V let $x^\nu \in K_1 \times K_2$ with $x^\nu \to x \in K_1 \times K_2$ and $v \in (K_1 \times K_2 - x) \cap V$. Then we have $v^\nu = (x_1 + v_1 - x_1^\nu, 0) \in (K - x^\nu) \cap V$ and $v^\nu \to v$. □

Another example of stability in the sense of Definition A.1 are convex polyhedral domains. A property similar to the second condition in Definition A.1 and related to the stability of tangent cones was shown in [14] for this case. We will give a direct proof for convex polyhedra without using tangent cones.

Corollary A.3. *Let $F : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be coercive, lower semi-continuous, and continuous on its domain. Furthermore let $K = \operatorname{dom} F$ be a (not necessarily closed) convex polyhedron, and V any fixed subspace of $\mathbb{R}^n$. Then K is stable with respect to translations in V, and m is continuous on K.*

Proof. First we note that K can be written as

$$K = \big\{x \in \mathbb{R}^n \mid \langle b_i, x\rangle \lhd_i c_i \text{ for } i = 1, \dots, k\big\},$$

where $\lhd_i$ is either "$<$" or "$\le$" for each i. To show stability of K with respect to translations in V let $x^\nu \in K$ with $x^\nu \to x \in K$, and $v \in V$ with $x + v \in K$. For each ν and i we define

$$\lambda_i^\nu := \begin{cases} 1 & \text{if } \langle b_i, v\rangle \le 0, \\ 1 & \text{if } \langle b_i, v\rangle > 0 \text{ and } \langle b_i, x^\nu - x\rangle \le 0, \\ \max\left\{0, 1 - \frac{\langle b_i, x^\nu - x\rangle}{\langle b_i, v\rangle}\right\} & \text{if } \langle b_i, v\rangle > 0 \text{ and } \langle b_i, x^\nu - x\rangle > 0, \end{cases}$$

and $\lambda^\nu = \min_i \lambda_i^\nu \geq 0$. Then we have $\langle x^\nu + \lambda v, b_i\rangle \lhd_i c_i$ for all $\lambda \in [0, \lambda_i^\nu]$ and thus $v^\nu = \lambda^\nu v \in (K - x^\nu) \cap V$. Furthermore $x^\nu \to x$ implies $\lambda^\nu \to 1$ and hence $v^\nu \to v$. □

References

[1] L. Badea. Convergence rate of a Schwarz multilevel method for the constrained minimization of nonquadratic functionals. *SIAM J. Numer. Anal.*, 44(2):449–477, 2006.

[2] M. Bergounioux, K. Ito, and K. Kunisch. Primal–dual strategy for constrained optimal control problems. *SIAM J. Control Optim.*, 37:1176–1194, 1999.

[3] H. Berninger, R. Kornhuber, and O. Sander. Fast and robust numerical solution of the Richards equation in homogeneous soil. *SIAM J. Numer. Anal.*, 49(6):2576–2597, 2011.

[4] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[5] C. Carstensen. Domain decomposition for a non-smooth convex minimization problem and its application to plasticity. *Numer. Linear Algebra Appl.*, 4(3):177–190, 1997.

[6] R. D. Falgout. An introduction to algebraic multigrid. In *Special Issue on Multigrid Computing*, volume 8 of *Computing in Science and Engineering*, pages 24–33. 2006.

[7] C. Geiger and C. Kanzow. *Theorie und Numerik restringierter Optimierungsaufgaben*. Springer, 2002.

[8] R. Glowinski, J. L. Lions, and R. Trémolières. *Numerical Analysis of Variational Inequalities*. Number 8 in Studies in Mathematics and its Applications. North-Holland Publishing Company, 1981.

[9] C. Gräser. *Convex Minimization and Phase Field Models*. PhD thesis, Freie Universität Berlin, 2011.

[10] C. Gräser and R. Kornhuber. Multigrid methods for obstacle problems. *J. Comp. Math.*, 27(1):1–44, 2009.

[11] C. Gräser, R. Kornhuber, and U. Sack. Time discretizations of anisotropic Allen–Cahn equations. *IMA J. Numer. Anal.*, 33(4):1226–1244, 2013.

[12] C. Gräser, R. Kornhuber, and U. Sack. Numerical simulation of coarsening in binary solder alloys. *Comp. Mater. Sci.*, 93:221–233, 2014.

[13] C. Gräser, U. Sack, and O. Sander. Truncated nonsmooth Newton multigrid methods for convex minimization problems. In M. Bercovier, M. Gander, R. Kornhuber, and O. Widlund, editors, *Domain Decomposition Methods in Science and Engineering XVIII*, volume 70 of *LNCSE*, pages 129–136. Springer, 2009.

[14] C. Gräser and O. Sander. Polyhedral Gauß–Seidel converges. *J. Numer. Math.*, 22(3):221–254, 2014.

[15] C. Gräser and O. Sander. Truncated nonsmooth Newton multigrid methods for simplex-constrained minimization problems. IGPM Preprint 384, RWTH Aachen, 2014.

[16] W. Han and B. D. Reddy. *Plasticity*. Springer, second edition, 2013.

[17] H. Hardering. Numerical approximation of capillary surfaces in a negative gravitational field. *Interfaces and Free Boundaries*, 15(3):263–280, 2013.

[18] B. Heimsund, X.-C. Tai, and J. Xu. Rate of convergence for parallel subspace correction methods for nonlinear variational inequalities. In N. Debit et al., editors, *Domain Decomposition Methods in Science and Engineering*, pages 127–138. CIMNE, 2002.

[19] M. Hintermüller, K. Ito, and K. Kunisch. The primal–dual active set strategy as a semismooth Newton method. *SIAM J. Optim.*, 13(3):865–888, 2002.

[20] R. H. W. Hoppe. Multigrid algorithms for variational inequalities. *SIAM J. Numer. Anal.*, 24:1046–1065, 1987.

[21] G. Jouvet, E. Bueler, C. Gräser, and R. Kornhuber. A nonsmooth Newton multigrid method for a hybrid, shallow model of marine ice sheets. In J. Li, H. Yang, and E. Machorro, editors, *Recent Advances in Scientific Computing and Applications*, volume 586 of *Contemporary Mathematics*, pages 197–205. American Mathematical Society, 2013.

[22] G. Jouvet and C. Gräser. An adaptive Newton multigrid method for a model of marine ice sheets. *J. Comp. Phys.*, 252:419–437, 2013.

[23] R. Kornhuber. Monotone multigrid methods for elliptic variational inequalities I. *Numer. Math.*, 69:167–184, 1994.

[24] R. Kornhuber. Monotone multigrid methods for elliptic variational inequalities II. *Numer. Math.*, 72:481–499, 1996.

[25] R. Kornhuber and R. Krause. Robust multigrid methods for vector-valued Allen–Cahn equations with logarithmic free energy. *Comput. Vis. Sci.*, 9(2):103–116, 2006.
[26] R. Krause and O. Sander. Fast solving of contact problems on complicated geometries. In R. Kornhuber, R. Hoppe, J. Périaux, O. Pironneau, O. Widlund, and J. Xu, editors, *Domain Decomposition Methods in Science and Engineering*, pages 495–502. Springer Verlag, 2005.
[27] J. D. Lee, Y. Sun, and M. Saunders. Proximal Newton-type methods for minimizing composite functions. *SIAM J. Optim.*, 24(3):1420–1443, 2014.
[28] Y.-L. Lee, J. Wu, J. Xu, and L. Zikatanov. A sharp convergence estimate for the method of subspace corrections for singular systems of equations. *Math. Comp.*, 77(262):831–850, 2008.
[29] J. Mandel. A multilevel iterative method for symmetric, positive definite linear complementarity problems. *Appl. Math. Optimization*, 11:77–95, 1984.
[30] J. Martin and S. Caddemi. Sufficient conditions for convergence of the Newton–Raphson iterative algorithm in incremental elastic–plastic analysis. *Eur. J. Mech. A Solids*, 13(3):351–365, 1994.
[31] C. Miehe, F. Welschinger, and M. Hofacker. Thermodynamically consistent phase-field models of fracture: Variational principles and multi-field FE implementations. *Internat. J. Numer. Methods Engrg.*, 83:1273–1311, 2010.
[32] J.-P. Penot. *Calculus Without Derivatives*. Springer, 2012.
[33] E. Pipping, O. Sander, and R. Kornhuber. Variational formulation of rate- and state-dependent friction problems. *Zeitschrift für Angewandte Mathematik und Mechanik (ZAMM)*, 95(4):377–395, 2015.
[34] O. Sander. Solving primal plasticity increment problems in the time of a single predictor–corrector iteration. arXiv: 1707.03733, 2017.
[35] K. Stüben. A review of algebraic multigrid. *J. Comput. Appl. Math.*, 128(1):281–309, 2001.
[36] X.-C. Tai. Rate of convergence for some constraint decomposition methods for nonlinear variational inequalities. *Numer. Math.*, 93(4):755–786, 2003.
[37] X.-C. T. Tai and J. Xu. Global and uniform convergence of subspace correction methods for some convex optimization problems. *Math. Comp.*, 71:105–124, 2001.
[38] W. Wang and M. Á. Carreira-Perpiñán. Projection onto the probability simplex: An efficient algorithm with a simple proof, and an application. arXiv: 1309.1541, 2013.

Carsten Gräser, Freie Universität Berlin, Institut für Mathematik, Arnimallee 6, 14195 Berlin, Germany
E-mail address: graeser@mi.fu-berlin.de

Oliver Sander, Technische Universität Dresden, Institut für Numerische Mathematik, Zellescher Weg 12–14, 01069 Dresden, Germany
E-mail address: oliver.sander@tu-dresden.de