

# Adaptive Regularization Minimization Algorithms with Non-Smooth Norms

S. Gratton\* and Ph. L. Toint†

27 V 2021

## Abstract

An adaptive regularization algorithm (AR1 $p$ GN) for unconstrained nonlinear minimization is considered, which uses a model consisting of a Taylor expansion of arbitrary degree and regularization term involving a possibly non-smooth norm. It is shown that the non-smoothness of the norm does not affect the  $\mathcal{O}(\epsilon_1^{-(p+1)/p})$  upper bound on evaluation complexity for finding first-order  $\epsilon_1$ -approximate minimizers using  $p$  derivatives, and that this result does not hinge on the equivalence of norms in  $\mathbb{R}^n$ . It is also shown that, if  $p = 2$ , the bound of  $\mathcal{O}(\epsilon_2^{-3})$  evaluations for finding second-order  $\epsilon_2$ -approximate minimizers still holds for a variant of AR1 $p$ GN named AR2GN, despite the possibly non-smooth nature of the regularization term. Moreover, the adaptation of the existing theory for handling the non-smoothness results in an interesting modification of the subproblem termination rules, leading to an even more compact complexity analysis. In particular, it is shown when the Newton's step is acceptable for an adaptive regularization method. The approximate minimization of quadratic polynomials regularized with non-smooth norms is then discussed, and a new approximate second-order necessary optimality condition is derived for this case. A specialized algorithm is then proposed to enforce first- and second-order conditions that are strong enough to ensure the existence of a suitable step in AR1 $p$ GN (when  $p = 2$ ) and in AR2GN, and its iteration complexity is analyzed. A final section discusses how practical approximate curvature measures may lead to weaker second-order optimality guarantees.

**Keywords:** nonlinear optimization, adaptive regularization, evaluation complexity, non-smooth norms, second-order minimizers.

## 1 Introduction

This paper is concerned with the derivation of upper bounds on the evaluation complexity of adaptive regularization algorithms for the solution of the smooth unconstrained nonconvex optimization problem

$$\min_{x \in \mathbb{R}^n} f(x). \quad (1.1)$$

This research area has been remarkably active in recent years (see, for instance, [27, 34, 8, 10, 12, 4, 22, 5, 6, 31, 23, 3, 2, 15]). Adaptive regularization algorithms, the class of methods

---

\*Université de Toulouse, INP, IRIT, Toulouse, France. Email: serge.gratton@enseeiht.fr Work partially supported by 3IA Artificial and Natural Intelligence Toulouse Institute, French "Investing for the Future - PIA3" program under the Grant agreement ANR-19-PI3A-0004"

†NAXYS, University of Namur, Namur, Belgium. Email: philippe.toint@unamur.be

considered here, compute steps from one iterate to the next by building and (often approximately) minimizing a model consisting of a truncated Taylor expansion of  $f$ , which is then “regularized” by adding a suitable power of the norm of the putative step. Several authors have considered various smooth norms for this regularization term [34, 10, 20, 6, 18, 17], showing that, under suitable assumptions, the resulting adaptive regularization method must find a first-order  $\epsilon_1$ -approximate minimizer for problem (1.1) (that is an iterate  $x_k$  such  $\|\nabla_x^1 f(x_k)\| \leq \epsilon_1$ ) in at most  $\mathcal{O}(\epsilon_1^{-(p+1)/p})$  evaluations of the objective function and its derivatives. In addition, second-order variants of this algorithm are bound to find a second-order  $\epsilon_2$ -approximate minimizer (that is an iterate  $x_k$  such the smallest eigenvalue of  $\nabla_x^2 f(x_k)$  exceeds  $-\epsilon_2$ ) in at most  $\mathcal{O}(\epsilon_2^{-(p+1)/(p-1)})$  such evaluations. The detailed algorithms considered in these contributions all depend on the central tenet that the regularized model (whose approximate minimization yields the step from one iterate to the next) is smooth, and thus that this approximate minimization can be carried out using algorithms for smooth functions and can be terminated using approximate optimality conditions for smooth problems. We show in this paper that the same evaluation complexity bounds still holds for first-order approximate minimizers in the case where non-smooth norms (such as  $\ell_1$  or  $\ell_\infty$ ) are considered, provided the algorithm is suitably modified. We also show that, when  $p = 2$ , the evaluation complexity bound in  $\mathcal{O}(\epsilon_2^{-3})$  is also maintained in the same context for a variant of the algorithm. Unsurprisingly, both results require redefining the termination conditions for model minimization, which makes them more flexible. As it turns out, the resulting modifications of the standard adaptive regularization method are extremely simple and their use in the complexity theory results in a remarkably compact formulation. They also shed a new light on the acceptability of the Newton step within adaptive regularization methods.

One may argue that, since all norms are equivalent in finite dimensional spaces, the stated complexity bound can be derived for any norm from known results in Euclidean norm [6, 14]. While this is true if one focuses on the order in  $\epsilon_1$  and  $\epsilon_2$  only, this ignores the influence of the norm equivalence constants, whose size can be significant when  $n$ , the dimension of the problem, grows. For instance the equivalence constant between the Euclidean and infinity norm is proportional to the square root of the problem’s dimension. Thus obtaining a given accuracy on the gradient norm in the infinity norm by simply applying the norm equivalence principle may require  $n^{(p+1)/2p}$  times more evaluations of the objective function and its derivatives than in the Euclidean one. The approach presented here attempts to avoid this potentially problematic increase in cost.

Of course, for the new algorithms to be practical, one needs to show that the model minimization subproblems are solvable by implementable methods. Focusing again on the case where  $p = 2$  and the model is a regularized quadratic, we derive a specialized second-order necessary optimality condition for the approximate minimization of such non-smooth functions. We then propose a new algorithm which is able to achieve first- and second-order approximate optimality for this problem and evaluate its iteration complexity. We finally discuss relaxed variants of the new algorithms that are sufficient for solving the subproblems of interest in algorithms for general functions, as well as their iteration complexity.

It is interesting to note that several authors [11, 16, 25] have proposed algorithms and associated complexity theory for the problem of minimizing a composite function of the form  $f(x) + h(c(x))$ , where  $f$  and  $c$  are smooth but  $h$  is potentially non-smooth. These approaches differ from our present objective in that, although the objective function may be non-smooth, the regularization term remains smooth. Moreover, the definition of the acceptable approxi-

mate minimizers circumvents the model’s non-smoothness by using the unmodified function  $h$  and, at variance with the approach described here, makes no attempt to exploit its generalized derivatives. We also note the contribution [1], where a first-order trust-region method (using a general norm) is proposed for the composite problem with weak assumptions on  $h$  (allowing non-smoothness and non-convexity), and for which an  $\mathcal{O}(\epsilon_1^{-2})$  evaluation complexity bound is shown. This bound is consistent with standard complexity bounds for trust-region methods (see [24] for a first proof on smooth problems). The use of models of degree exceeding one is considered, but, as is typically the case for trust-region methods, does not improve the complexity bound.

Our exposition is organized as follows. We present the problem and the first-order algorithm in Section 2 and derive its evaluation complexity theory in Section 3. Section 4 discusses the new approximate second-order necessary condition for global minimizers and establishes the upper bound on evaluation complexity for an adapted variant of the algorithm. A method for approximately minimizing regularized quadratics (enough for solving the subproblems arising in Sections 2 and 4) is then presented and analyzed in Section 5. A discussion of alternatives for the measure of curvature, a key ingredient in second-order conditions, is proposed in Section 6. Finally, some brief conclusions are outlined in Section 7.

## 2 An first-order adaptive regularization in general norms

We consider the unconstrained nonlinear optimization problem (1.1) where  $f$  is a (potentially nonconvex)  $p$  times continuously differentiable function from  $\mathbb{R}^n$  to  $\mathbb{R}$ , for some integer  $p \geq 1$ . We define

$$T_{f,p}(x, s) \stackrel{\text{def}}{=} f(x) + \sum_{\ell=1}^p \frac{1}{\ell!} \nabla_x^\ell f(x)[s]^\ell,$$

the  $p$ -th order Taylor expansion of  $f$  at  $x$ , where the notation  $\nabla_x^\ell f(x)[s]^\ell$  denotes the symmetric  $\ell$ -dimensional tensor  $\nabla_x^\ell f(x)$  applied on  $\ell$  copies of the vector  $s$ .

As outlined in the introduction, adaptive regularization methods are iterative schemes that compute a step from an iterate  $x_k$  by constructing a regularized model  $m_k(s)$  of  $f(x_k + s)$  as

$$m_k(s) \stackrel{\text{def}}{=} T_{f,p}(x_k, s) + \frac{\sigma_k}{(p+1)!} \|s\|_r^{p+1}, \quad (2.1)$$

where the  $p$ -th order Taylor series is “regularized” by adding the term  $\sigma_k \|s\|_r^{p+1}/(p+1)!$  ( $\sigma_k$  is known as the “regularization parameter”) and where we allow  $\|\cdot\|_r$  to be a general possibly non-smooth norm. This implies that  $\|\cdot\|_r$  is convex and Lipschitz continuous with global Lipschitz constant equal to one. Given the  $\|\cdot\|_r$  norm and defining

$$\|S_j\|_{r,j} = \max_{\|s\|_r=1} |S_j[s]^j| \quad (2.2)$$

to be the norm of the  $j$ -dimensional symmetric (for  $j > 1$ ) tensor  $S_j$  induced by  $\|\cdot\|_r$ , we are now interested in finding, for some prespecified accuracy requirement  $\epsilon_1 \in (0, 1]$ , an  $\epsilon_1$ -approximate first-order critical point, that is a point  $x_{\epsilon_1}$  such that  $\|\nabla_x^1 f(x_{\epsilon_1})\|_{r,1} \leq \epsilon_1$ . Note that, because of (2.2),  $\|\cdot\|_{r,1}$  is the dual norm of  $\|\cdot\|_r$ .

The “regularization term” in (2.1) guarantees that  $m_k(s)$  is bounded below and thus makes the procedure of finding a step  $s_k$  by (approximately) minimizing  $m_k(s)$  well-defined.

However, at variance with the usual setting for adaptive regularization methods, *the model  $m_k(s)$  may no longer be smooth*. Once the step is computed, the value of the objective function at the trial point  $x_k + s_k$  is then computed. If the decrease in  $f$  from  $x_k$  to  $x_k + s_k$  is comparable to that predicted by the second-order Taylor series, the trial point is accepted as the new iterate and the regularization parameter is (possibly) reduced. If this is not the case, the trial point is rejected and the regularization parameter increased. The resulting algorithm is formally stated as the AR1pGN.

**Algorithm 2.1: First-Order Adaptive Regularization with General Norm (AR1pGN)**

**Step 0: Initialization.** An initial point  $x_0 \in \mathbb{R}^n$ , a regularization parameter  $\sigma_0$  and a desired final gradient accuracy  $\epsilon_1 \in (0, 1]$  are given. The constants  $\eta_1, \eta_2, \gamma_1, \gamma_2, \gamma_3, \theta_1$  and  $\sigma_{\min}$  are also given such that

$$\sigma_{\min} \in (0, \sigma_0], \quad 0 < \eta_1 \leq \eta_2 < 1, \quad \theta_1 > 1 \quad \text{and} \quad 0 < \gamma_1 < 1 < \gamma_2 < \gamma_3. \quad (2.3)$$

Compute  $f(x_0)$  and set  $k = 0$ .

**Step 1: Check for termination.** Terminate with  $x_{\epsilon_1} = x_k$  if

$$\|\nabla_x^1 f(x_k)\|_{r,1} \leq \epsilon_1. \quad (2.4)$$

**Step 2: Step calculation.** Compute a step  $s_k$  which sufficiently reduces the model  $m_k$  in the sense that

$$m_k(s_k) \leq m_k(0) \quad (2.5)$$

and

$$\|\nabla_s^1 T_{f,p}(x_k, s_k)\|_{r,1} \leq \theta_1 \frac{\sigma_k}{p!} \|s\|_r^p. \quad (2.6)$$

**Step 3: Acceptance of the trial point.** Compute  $f(x_k + s_k)$  and define

$$\rho_k = \frac{f(x_k) - f(x_k + s_k)}{T_{f,p}(x_k, 0) - T_{f,p}(x_k, s_k)}. \quad (2.7)$$

If  $\rho_k \geq \eta_1$ , then define  $x_{k+1} = x_k + s_k$ ; otherwise define  $x_{k+1} = x_k$ .

**Step 4: Regularization parameter update.** Set

$$\sigma_{k+1} \in \begin{cases} [\max(\sigma_{\min}, \gamma_1 \sigma_k), \sigma_k] & \text{if } \rho_k \geq \eta_2, \\ [\sigma_k, \gamma_2 \sigma_k] & \text{if } \rho_k \in [\eta_1, \eta_2), \\ [\gamma_2 \sigma_k, \gamma_3 \sigma_k] & \text{if } \rho_k < \eta_1. \end{cases} \quad (2.8)$$

Increment  $k$  by one and go to Step 1.

While the AR1pGN algorithm follows the main lines of existing adaptive regularization methods (see [10, 6] for example), we immediately note that the test (2.6) differs from the test

$\|\nabla_s^1 m_s(s_k)\|_2 \leq \theta_1 \|s_k\|_2^p$  which is used so far in the literature. Indeed, our framework no longer guarantees that  $\nabla_s^1 m_s(s)$  exists, due to the possible lack of smoothness of the regularization term. Note however that, if  $\|\cdot\|_r$  is differentiable everywhere except at the origin, then

$$\nabla_s^1 T_{f,p}(x_k, s_k) + \frac{\sigma_k}{p!} \|s\|_r^p \nabla_s^1 \|s\|_r = 0$$

at a nonzero first-order point of  $m_k(s)$ , and (2.6) holds at such a point since  $\|s\|_r$  is Lipschitz continuous with unit Lipschitz constant, and thus  $\|\nabla_s^1 \|s\|_r\| \leq 1$ . The condition (2.6) is therefore weaker than a more standard condition of the form  $\|\nabla_s^1 m_k(s_k)\| \leq \theta_1 \|s_k\|^p$ . This therefore suggests that the standard termination test may lead to subproblem oversolving, an important issue for numerical performance.

In particular, and at variance with other adaptive regularization methods, *the AR1pGN algorithm allows the Newton step*  $s_k = -\nabla_x^2 f(x_k)^{-1} \nabla_x^1 f(x_k)$  when  $p = 2$  and the Hessian  $\nabla_x^2 f(x_k)$  is positive definite, *provided the regularized model has not increased*, that is provided (2.5) holds. Indeed this step automatically ensures (2.6) since then  $\nabla_s^1 T_{f,2}(x_k, s_k) = 0$ . The condition (2.5) however avoids situations where the model decrease  $m_k(0) - m_k(s_k)$  is tiny but  $\|s_k\|_r$  is large, which is exactly what happens in the example [13] showing convergence of Newton's method to a first-order  $\epsilon_1$ -approximate minimizer in  $\mathcal{O}(\epsilon_1^{-2})$  evaluations.

We also note that we could use an iteration-dependent  $\theta_{1,k}$  in (2.6), provided it is strictly bounded below by one and bounded above by a constant. We have ignored this possibility for the sake of simplicity.

Having modified the requirements on the step, we now need to verify that the new conditions (2.5) and (2.6) are compatible. We start by deriving an expression for the subdifferential  $\partial(\|\cdot\|_r^{p+1})(s)$ .

**Lemma 2.1** We have that

$$\partial(\|\cdot\|_r)(s) = \{v \in \mathbb{R}^n \mid v^T s = \|s\|_r \quad \text{and} \quad \|v\|_{r,1} = 1\} \quad (2.9)$$

and

$$\partial_C(\|\cdot\|_r^{p+1})(s) = \partial(\|\cdot\|_r^{p+1})(s) = (p+1) \|s\|_r^p \partial(\|\cdot\|_r)(s), \quad (2.10)$$

where  $\partial_C$  denotes the Clarke subdifferential.

**Proof.** The identity (2.9) is standard (see [28, Example 3.1] for instance). By composition of the norm with the increasing convex differentiable function  $\phi(t) = t^{p+1}$  (on  $\mathbb{R}^+$ ), we obtain from [28, Theorem 4.3.1] that

$$\partial(\|\cdot\|_r^{p+1})(s) = \{\alpha s \in \mathbb{R}^n \mid \alpha \in \phi'(\|s\|_r) \quad \text{and} \quad s \in \partial(\|\cdot\|_r)(s)\},$$

which is the second equality in (2.10). Since  $\|\cdot\|_r^{p+1}$  is also Lipschitz continuous, it is Clarke regular and thus the Clarke subdifferential and the standard one coincide (see [19, Proposition 4.3]), giving the first equality in (2.10).  $\square$

This allows us to derive the following characterization of a minimizer of  $m_k$ .

**Lemma 2.2** Let  $s_k^*$  be a local minimizer of  $m_k$ . Then

$$\|\nabla_s^1 T_{f,p}(x_k, s_k^*)\|_{r,1} = \frac{\sigma_k}{p!} \|s_k^*\|_r^p. \quad (2.11)$$

**Proof.** Since  $m_k$  is Lipschitz continuous, the Clarke criticality of  $s_k^*$  implies that

$$0 \in \partial_C m_k(s_k^*) = \{\nabla_x^1 T_{f,p}(x_k, s_k^*)\} + \frac{\sigma_k}{(p+1)!} \partial_C(\|s_k^*\|_r^{p+1}), \quad (2.12)$$

where we have used the property of the Clarke subdifferential of the sum of two locally Lipschitz functions [19, Exercice 1.4] and the fact that, since  $T_{f,p}(x_k, s)$  is continuously differentiable as a function of  $s$ ,  $\partial_C T_{f,p}(x_k, \cdot)(s) = \{\nabla_s^1 T_{f,p}(x_k, s)\}$ . Using now (2.10), we deduce from this identity and (2.12) that there exists a vector  $\xi \in \partial(\|\cdot\|_r)(s_k^*)$  such that

$$\nabla_s^1 T_{f,p}(x_k, s_k^*) = -\frac{\sigma_k}{p!} \|s_k^*\|_r^p \xi. \quad (2.13)$$

Moreover, (2.9) implies that  $\|\xi\|_{r,1} = 1$ . Taking norms in (2.13) gives (2.11).  $\square$

The (scalar) necessary condition (2.11) is clearly weaker than the (vector) identity (2.12), but is nevertheless sufficient to derive the following crucial result.

**Corollary 2.3** A step satisfying both (2.5) and (2.6) always exists.

**Proof.** From

$$m_k(s) \geq \frac{\sigma_k}{(p+1)!} \|s\|_r^{p+1} - |f(x)| - \sum_{\ell=1}^p \frac{1}{\ell!} \|\nabla_x^\ell f(x_k)\|_{r,\ell} \|s\|_r^\ell,$$

we obtain  $\lim_{\|s\|_r \rightarrow +\infty} m_k(s) = +\infty$  which, together with the continuity of  $m_k(s)$ , implies that  $m_k$  admits at least one minimizer  $s_k^*$  over  $\mathbb{R}^n$ , satisfying  $m_k(s_k^*) \leq m_k(0)$ . Applying Lemma 2.2 then gives that (2.6) holds at  $s_k^*$  for any  $\theta_1 \geq 1$ .  $\square$

An important comment is in order at this point. Because the Clarke subdifferential of the norm is not necessarily continuous in our context, it may seem at first sight that obtaining a step satisfying the conditions (2.5) and (2.6) may require the computation of an exact minimizer  $s_k^*$  of the model, which is potentially costly. Fortunately, this fear is unfounded because both the left- and the right-hand sides of (2.6) are continuous functions of  $s$  and the inequality therefore also holds in a neighbourhood of  $s_k^*$  provided  $\theta_1 > 1$ . Any convergent minimization algorithm (such as those proposed, for instance, in [21, 33, 30, 11, 16, 25] or, more generally, in [29], or in Section 5) applied on the model is therefore bound to produce a suitable step  $s_k$  in a finite number of iterations.

Following well-established practice, we now define

$$\mathcal{S} \stackrel{\text{def}}{=} \{k \geq 0 \mid x_{k+1} = x_k + s_k\} = \{k \geq 0 \mid \rho_k \geq \eta_1\} \quad \text{and} \quad \mathcal{S}_k \stackrel{\text{def}}{=} \mathcal{S} \cap \{1, \dots, k\},$$

the set of indices of “successful iterations”, and the set of indices of successful iterations up to iteration  $k$ , respectively. We also recall a well-known result bounding the total number of iterations of an adaptive regularization method in terms of the number of successful ones.

**Lemma 2.4** [6, Theorem 2.4] Suppose that the AR1pGN algorithm is used and that  $\sigma_k \leq \sigma_{\max}$  for some  $\sigma_{\max} > 0$ . Then

$$k \leq |\mathcal{S}_k| \left( 1 + \frac{|\log \gamma_1|}{\log \gamma_2} \right) + \frac{1}{\log \gamma_2} \log \left( \frac{\sigma_{\max}}{\sigma_0} \right). \quad (2.14)$$

### 3 Evaluation complexity for the AR1pGN algorithm

Before discussing our analysis of evaluation complexity, we first formalize our assumptions on problem (1.1).

**AS.1**  $f$  is  $p$  times differentiable and its  $p$ -th derivative  $\nabla_x^p f(x)$  is globally Lipschitz continuous in the  $\|\cdot\|_r$  norm, that is there exists  $L_{f,p} \geq 0$  such that

$$\|\nabla_x^p f(x) - \nabla_x^p f(y)\|_{r,p} \leq L_{f,p} \|x - y\|_r \quad \text{for all } x, y \in \mathbb{R}^n,$$

where the  $\|\cdot\|_r$  norm in the left-hand side is defined by (2.2).

**AS.2** There exists a constant  $f_{\text{low}}$  such that  $f(x) \geq f_{\text{low}}$  for all  $x \in \mathbb{R}^n$ .

Assumption AS.1 recasts the usual context for the analysis of complexity of adaptive regularization methods in the context of the general norms, while AS.2 ensures that problem (1.1) is well-defined. AS.1 yields the well-known Lipschitz error bounds.

**Lemma 3.1** Suppose that AS.1 holds and that  $k \in \mathcal{S}$ . Then

$$|f(x_{k+1}) - T_{f,p}(x_k, s_k)| \leq \frac{L_{f,p}}{(p+1)!} \|s_k\|_r^{p+1} \quad (3.1)$$

and

$$\|\nabla_x^j f(x_{k+1}) - \nabla_s^j T_{f,p}(x_k, s_k)\|_{r,j} \leq \frac{L_{f,p}}{(p-j+1)!} \|s_k\|_r^{p-j+1} \quad (3.2)$$

for  $j \in \{1, \dots, p\}$ .

**Proof.** The proof is a direct extension of [15, Lemma 2.1] with  $\beta = 1$  that now uses  $\|\cdot\|_r$  instead of  $\|\cdot\|_2$  and exploits (2.2). It is given in appendix for completeness.  $\square$

The analysis in the rest of this section follows that presented in [6] quite closely. We first state a simple lower bound on the decrease of the Taylor expansion.

**Lemma 3.2**

$$\Delta T_{f,p}(x_k, s_k) \stackrel{\text{def}}{=} T_{f,p}(x_k, 0) - T_{f,p}(x_k, s) \geq \frac{\sigma_k}{(p+1)!} \|s_k\|_r^{p+1}. \quad (3.3)$$

**Proof.** Direct from (2.5) and (2.1).  $\square$

We next derive an upper bound on the regularization parameter.

**Lemma 3.3** Suppose that AS.1 holds. Then, for all  $k \geq 0$ ,

$$\sigma_k \leq \sigma_{\max} \stackrel{\text{def}}{=} \gamma_3 \max \left[ \sigma_0, \frac{L_{f,p}}{(1-\eta_2)} \right]. \quad (3.4)$$

**Proof.** See [6, Lemma 2.2]. Using (2.7), (3.1), and (3.3), we obtain that

$$|\rho_k - 1| \leq \frac{(p+1)! |f(x_k + s_k) - T_{f,p}(x_k, s_k)|}{\sigma_k \|s_k\|_r^{p+1}} \leq \frac{L_{f,p}}{\sigma_k}.$$

Thus, if  $\sigma_k \geq L_{f,p}/(1-\eta_2)$ , then  $\rho_k \geq \eta_2$ , iteration  $k$  is successful and (2.8) implies that  $\sigma_{k+1} \leq \sigma_k$ . The mechanism of the algorithm then guarantees that (3.4) holds.  $\square$

The next lemma remains in the spirit of [6, Lemma 2.3], but now takes the new condition (2.6) into account, avoiding any reference to the model's derivative and resulting in a simpler proof.

**Lemma 3.4** Suppose that AS.1 holds and that  $k \in \mathcal{S}$  before termination. Then

$$\|s_k\|_r^p \geq \frac{p!}{L_{f,p} + \theta_1 \sigma_{\max}} \epsilon_1. \quad (3.5)$$

**Proof.** Successively using the fact that termination does not occur at iteration  $k$ , the triangle inequality, (3.2) for  $j = 1$ , condition (2.6) and (3.4), we deduce that

$$\begin{aligned} \epsilon_1 &< \|\nabla_x^1 f(x_{k+1})\|_{r,1} \\ &\leq \|\nabla_x^1 f(x_{k+1}) - \nabla_x^1 T_{f,p}(x_k, s_k)\|_{r,1} + \|\nabla_x^1 T_{f,p}(x_k, s_k)\|_{r,1} \\ &\leq \frac{L_{f,p}}{p!} \|s_k\|_r^p + \theta_1 \frac{\sigma_k}{p!} \|s_k\|_r^p. \end{aligned}$$

This in turn directly implies (3.5).  $\square$



We may now resort to the classical “telescoping sum” argument to obtain the desired complexity result.

**Theorem 3.5** Suppose that AS.1–AS.2 hold. Then the AR1 $p$ GN algorithm requires at most

$$\frac{(p+1)!}{\eta_1 \sigma_{\min}} \left( \frac{L_{f,p} + \theta_1 \sigma_{\max}}{p!} \right)^{\frac{p+1}{p}} \frac{f(x_0) - f_{\text{low}}}{\epsilon_1^{\frac{p+1}{p}}}$$

successful iterations and evaluations of  $\{\nabla_x^i f\}_{i=1,2}$  and at most

$$\frac{(p+1)!}{\eta_1 \sigma_{\min}} \left( \frac{L_{f,p} + \theta_1 \sigma_{\max}}{p!} \right)^{\frac{p+1}{p}} \frac{f(x_0) - f_{\text{low}}}{\epsilon_1^{\frac{p+1}{p}}} \left( 1 + \frac{|\log \gamma_1|}{\log \gamma_2} \right) + \frac{1}{\log \gamma_2} \log \left( \frac{\sigma_{\max}}{\sigma_0} \right)$$

evaluations of  $f$  to produce a vector  $x_{\epsilon_1} \in \mathbb{R}^n$  such that  $\|\nabla_x^1 f(x_{\epsilon_1})\|_{r,1} \leq \epsilon_1$ .

**Proof.** Let  $k$  be the index of an iteration before termination. Then, using AS.2, the definition of successful iterations, (3.3) and (3.5),

$$|\mathcal{S}_k| \leq \frac{(p+1)!}{\eta_1 \sigma_{\min}} \left( \frac{L_{f,p} + \theta_1 \sigma_{\max}}{p!} \right)^{-\frac{p+1}{p}} \frac{f(x_0) - f_{\text{low}}}{\epsilon_1^{\frac{p+1}{p}}}$$

for any  $k$  before termination, and the first conclusion follows since the derivatives are only evaluated once per successful iteration. Applying now Lemma 2.4 gives the second conclusion.  $\square$

## 4 Approximate second-order minimizers for $p = 2$

We now turn the second-order case and from now on, limit our analysis to the case where  $p = 2$ . We are thus interested in finding approximate second-order minimizers, that is, in line with our desire to measure distances in the primal space with the norm  $\|\cdot\|_r$ , points at which

$$\lambda_r[\nabla_x^2 f(x_k)] \geq -\epsilon_2 \quad \text{where} \quad \lambda_r[H] \stackrel{\text{def}}{=} \min_{v \neq 0} \frac{\langle Hv, v \rangle}{\|v\|_r^2}. \quad (4.1)$$

Guaranteeing this result may however be unreachable in practice, because the computation of  $\lambda_r[H]$  may be unrealistically expensive for some norms. For instance, if one consider the case where  $\|\cdot\|_r = \|\cdot\|_\infty$ , computing  $\lambda_r[H]$  for indefinite  $H$  amounts to solving an indefinite quadratic optimization problem with bound constraints, a problem which is known to be NP-complete [32, 35]. As a consequence, we may have to live with approximations. Observe that, if  $u_a$  is any vector with  $\|u_a\|_r = 1$ , then  $\lambda_r[H] \leq \langle Hu_a, u_a \rangle$ . Thus, in what follows, we assume that, for a symmetric  $H$ , we can compute a vector  $u_a$  and an associated scalar  $\lambda_a$  such that, for a given constant  $\tau \in (0, 1]$ ,

$$\|u_a\|_r = 1, \quad \lambda_a \stackrel{\text{def}}{=} \langle Hu_a, u_a \rangle \quad \text{and} \quad \left( \lambda_a \in \left[ \lambda_r[H], \tau \lambda_r[H] \right] \quad \text{if} \quad \lambda_r[H] < 0 \right), \quad (4.2)$$

a requirement potentially more affordable than (4.1). We comment in Section 6 on strategies related to this approximation. Meanwhile, and with this caveat in mind, we now establish a second-order necessary condition for a global minimizer of a regularized quadratic model  $m$ . As a first step, we derive a lower bound on the model decrease that can be obtained along a direction of sufficient negative curvature.

**Lemma 4.1** Let  $\phi(s) = f_0 + \langle g, s \rangle + \frac{1}{2} \langle Hs, s \rangle$  be a quadratic polynomial in  $s \in \mathbb{R}^n$ , and  $m(s) = \phi(s) + \frac{1}{6} \sigma \|s\|_r^3$ , where  $\sigma > 0$  is a constant and  $\|\cdot\|_r$  is a general norm. Consider  $s \neq 0$  and let  $\lambda_a$  and  $u_a$  satisfy (4.2), where we additionally choose the sign of  $u_a$  to ensure that  $\langle g + Hs, u_a \rangle \leq 0$  and assume that  $\lambda_a + \sigma \|s\|_r < 0$ . Then there exists an  $\alpha > 0$  such that

$$m(s) - m(s + \alpha \|s\|_r u_a) \geq \frac{3(\lambda_a + \sigma \|s\|_r)}{4\sigma^2} \left[ \psi(s) \sigma^2 \|s\|_r^2 - \frac{3}{4} (\lambda_a + \sigma \|s\|_r)^2 \right], \quad (4.3)$$

where

$$\psi(s) \stackrel{\text{def}}{=} \max \left[ 0, 1 + 2 \frac{\langle g + Hs, u_a \rangle}{\sigma \|s\|_r^2} \right]. \quad (4.4)$$

**Proof.** Setting  $d = \|s\|_r u_a$ , we have that, for  $\alpha > 0$ ,

$$\begin{aligned} m_k(s + \alpha d) &= m_k(s) + \alpha \langle g + Hs, d \rangle + \frac{1}{2} \alpha^2 \langle Hd, d \rangle + \frac{1}{6} \sigma \|s + \alpha d\|_r^3 - \frac{1}{6} \sigma \|s\|_r^3 \\ &\leq m_k(s) + \frac{1}{2} \alpha \sigma \|s\|_r^3 \left( 2 \frac{\langle g + Hs, u_a \rangle}{\sigma \|s\|_r^2} \right) + \frac{1}{2} \alpha^2 \lambda_a \|s\|_r^2 + \frac{1}{6} \sigma \|s + \alpha s\|_r^3 - \frac{1}{6} \sigma \|s\|_r^3. \end{aligned}$$

where we have used (4.2) and the fact that  $\|d\|_r = \|s\|_r$  implies the inequality  $\|s + \alpha d\|_r^3 \leq \|s + \alpha s\|_r^3$ . Moreover

$$\|s + \alpha s\|_r^3 - \|s\|_r^3 = [(1 + \alpha)^3 - 1] \|s\|_r^3 = \alpha [3 + 3\alpha + \alpha^2] \|s\|_r^3,$$

and hence, using (4.4),

$$m_k(s + \alpha d) \leq m_k(s) + \frac{1}{6} \sigma \|s\|_r^3 (3\alpha \psi(s) + 3\alpha^2 + \alpha^3) + \frac{1}{2} \alpha^2 \lambda_a \|s\|_r^2. \quad (4.5)$$

This in turn yields that, for  $\alpha > 0$ ,

$$m(s) - m(s + \alpha d) \geq -\frac{\alpha \|s\|_r^2}{2} \left[ \frac{\sigma \|s\|_r}{3} \alpha^2 + (\lambda_a + \sigma \|s\|_r) \alpha + \sigma \|s\|_r \psi(s) \right] \stackrel{\text{def}}{=} -\frac{\alpha \|s\|_r^2}{2} q_0(\alpha).$$

Now  $q_0 = a\alpha^2 + b\alpha + c$  is a convex quadratic in  $\alpha$  which admits a minimum for  $\alpha = -b/(2a)$  of value  $q(-b/(2a)) = c - b^2/(4a)$ . Since  $b = \lambda_a + \sigma \|s\|_r < 0$ ,

$$m(s) - m(s + \alpha d) \geq \left( \frac{3(\lambda_a + \sigma \|s\|_r)}{2\sigma \|s\|_r} \right) \frac{\|s\|_r^2}{2} \left[ \sigma \|s\|_r \psi(s) - \frac{3(\lambda_a + \sigma \|s\|_r)^2}{4\sigma \|s\|_r} \right]$$

for  $\alpha = -b/(2a) > 0$ , which implies (4.3).  $\square$

This leads to the following necessary optimality condition.

**Theorem 4.2** Let  $\phi(s) = f_0 + \langle g, s \rangle + \frac{1}{2}\langle Hs, s \rangle$  be a quadratic polynomial in  $s \in \mathbb{R}^n$ , and  $s_*$  be a global minimizer of  $m(s) = \phi(s) + \frac{1}{6}\sigma\|s\|_r^3$ , where  $\sigma > 0$  is a constant and  $\|\cdot\|_r$  is a general norm. Let  $\lambda_a$  and  $u_a$  satisfy (4.2), where we additionally choose the sign of  $u_a$  to ensure that  $\langle g + Hs, u_a \rangle \leq 0$ . Then

$$\lambda_a + \omega(s_*)\sigma\|s_*\|_r \geq 0, \quad (4.6)$$

where

$$\omega(s) \stackrel{\text{def}}{=} \begin{cases} 1 + \frac{2\sqrt{\psi(s)}}{\sqrt{3}} \leq 1 + \frac{2}{\sqrt{3}} \stackrel{\text{def}}{=} \kappa_\omega & \text{if } s \neq 0, \\ 1 & \text{otherwise,} \end{cases} \quad (4.7)$$

and  $\psi(s)$  is given by (4.4).

**Proof.** When  $H$  is positive-semidefinite, (4.6) follows trivially. Assume now that  $H$  admits at least one negative eigenvalue. Suppose first that  $s_* \neq 0$ . If  $\lambda_a + \sigma\|s_*\|_r \geq 0$ , (4.6) trivially follows. Suppose thus that  $\lambda_a + \sigma\|s_*\|_r < 0$ . Then (4.3) implies that there exists an  $\alpha > 0$  such that  $m(s_* + \alpha\|s_*\|_r u_a) < m(s_*)$  (which is impossible), unless

$$\psi(s_*)\sigma^2\|s_*\|_r^2 > \frac{3}{4}(\lambda_a + \sigma\|s_*\|_r)^2.$$

If  $\psi(s_*) = 0$ , this cannot happen. Otherwise, this last inequality requires that

$$\sqrt{\psi(s_*)}\sigma\|s_*\|_r > \sqrt{\frac{3}{4}}|\lambda_a + \sigma\|s_*\|_r| > -\frac{\sqrt{3}}{2}(\lambda_a + \sigma\|s_*\|_r),$$

which, given (4.7), yields (4.6).

Suppose now that  $s_* = 0$  and that  $\lambda_a < 0$ . It is then easy to verify that, if the sign of  $u_a$  is chosen to ensure that  $\langle g, u_a \rangle \leq 0$  and

$$\alpha \in \left[0, -\frac{3\lambda_a}{2\sigma}\right],$$

then, using (4.2),

$$m(\alpha u_a) = f_0 + \alpha\langle g, u_a \rangle + \frac{1}{2}\alpha^2\lambda_a + \frac{1}{6}\alpha^3\sigma \leq f_0 + \frac{9\lambda_a^3}{16\sigma^2} = m(0) + \frac{9\lambda_a^3}{16\sigma^2} < m(0), \quad (4.8)$$

which again contradicts the assumption that  $s_* = 0$  is a global minimum of  $m$ . Thus  $\lambda_a \geq 0$  and (4.6) also holds.  $\square$

It is remarkable that this lemma provides a “second-order” necessary condition for a global minimizer of quadratic polynomial regularized with a cubic term in a possibly non-smooth norm, despite the first and second derivatives of this objective function failing to exist.

Also note that, as long as  $0 < \|g + Hs\|_{r,1} \leq \frac{1}{2}\sigma\|s\|_r^2$ , our assumptions on  $u_a$  and the Cauchy-Schwarz inequality imply that

$$\begin{aligned} \psi(s) &= \max \left[ 0, 1 + \frac{\|g + Hs\|_{r,1}}{\frac{1}{2}\sigma\|s\|_r^2} \frac{\langle g + Hs, u_a \rangle}{\|g + Hs\|_{r,1}\|u_a\|_r} \right] \\ &\geq \max \left[ 0, 1 + \frac{\langle g + Hs, u_a \rangle}{\|g + Hs\|_{r,1}\|u_a\|_r} \right] \\ &= 1 + \frac{\langle g + Hs, u_a \rangle}{\|g + Hs\|_{r,1}\|u_a\|_r} \in [0, 1]. \end{aligned}$$

and the maximum with zero in (4.4) is unnecessary. This is particular the case if  $s$  is any local minimizer because Lemma 2.2.

It is interesting to pause at this point to stress that the necessary first- and second-order conditions (2.11) and (4.6), while sufficient for our purposes as we will see, are *merely necessary*, and by no means sufficient to guarantee a local minimizer. This is illustrated in Figure 4.1.

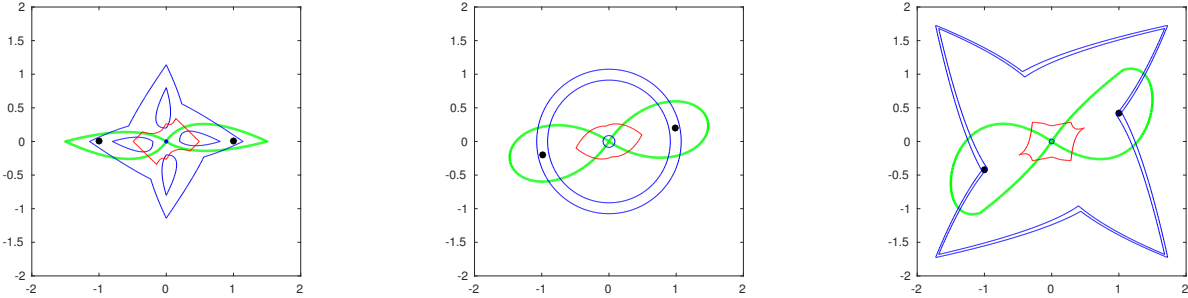


Figure 4.1: Admissible regions for the  $\ell_1$ - (left),  $\ell_2$ - (middle) and  $\ell_\infty$ - (right) norms

In this figure, a two-dimensional model is constructed with a zero gradient at the origin and an indefinite Hessian<sup>(1)</sup> scaled such that  $\lambda_r[H] = -1$ ,  $\tau = 1$  and with the regularization parameter  $\sigma$  is chosen equal to 6. The left picture corresponds to the choice  $\|\cdot\|_r = \|\cdot\|_1$ , the central one to  $\|\cdot\|_r = \|\cdot\|_2$  and the right one to  $\|\cdot\|_r = \|\cdot\|_\infty$ , all other parameters being identical. In each case, the region of the plane where (2.5) holds is the interior of the two green lobes and the regions where the deviation from (2.11) is bounded by  $0.01(\frac{1}{2})\sigma\|s\|_r^2$  are shown in blue, the first being the small region surrounding the origin (where the gradient is zero) and the second the zone between the two blue concentric curves. Finally, the region where the deviation from (4.6) does not exceed  $0.1\sigma\|s\|_r$  is the exterior of the region around the origin delineated in red. Note that the region around the origin which is admissible for (2.11) alone is excluded for (4.6). Thus in all cases, the admissible regions for (2.5), (2.11) and (4.6) consist of the regions limited by any of the shown curves and containing the minimizers marked with a black dot. We immediately notice that these regions are relatively large and may extend reasonably far from the minimizers. We also see that the geometry of these regions, while simple for the Euclidean norm, can be quite complicated for other norms.

Our algorithm for finding second-order  $\epsilon_2$ -approximate minimizers is described below.

<sup>(1)</sup>Chosen to be  $I - 2uu^T / \langle u, u \rangle$  with  $u^T = (5, 1)$ .

**Algorithm 4.1: Second-Order Adaptive Regularization with General Norm (AR2GN)**

**Step 0: Initialization.** An initial point  $x_0 \in \mathbb{R}^n$ , a regularization parameter  $\sigma_0$  a desired final gradient accuracy  $\epsilon, \epsilon_2 \in (0, 1]$  and a model degree  $p = 2$  are given. The constants  $\eta_1, \eta_2, \gamma_1, \gamma_2, \gamma_3, \theta_1 > 1, \theta_2 > 1$ , and  $\sigma_{\min}$  are also given such that

$$\sigma_{\min} \in (0, \sigma_0], \quad 0 < \eta_1 \leq \eta_2 < 1, \quad 0 < \gamma_1 < 1 < \gamma_2 < \gamma_3 \quad \text{and} \quad \tau \in (0, 1]. \quad (4.9)$$

Compute  $f(x_0)$  and set  $k = 0$ .

**Step 1: Check for termination.** Compute  $\lambda_{a,k}$  and  $u_{a,k}$  satisfying (4.2) with  $H = \nabla_x^2 f(x_k)$ , and terminate with  $x_\epsilon = x_k$  if

$$\|g_k\|_{r,1} \leq \epsilon_1 \quad \text{and} \quad \lambda_{a,k} \geq -\tau\epsilon_2. \quad (4.10)$$

**Step 2: Step calculation.** Compute a step  $s_k$  which sufficiently reduces the model  $m_k$  in the sense that (2.5) and (2.6) hold (for  $p = 2$ ) and, additionally,

$$\lambda_{a,k} + \theta_2 \omega(s_k) \sigma_k \|s_k\|_r \geq 0. \quad (4.11)$$

**Step 3: Acceptance of the trial point.** Compute  $f(x_k + s_k)$  and define  $\rho_k$  as in (2.7). If  $\rho_k \geq \eta_1$ , then define  $x_{k+1} = x_k + s_k$ ; otherwise define  $x_{k+1} = x_k$ .

**Step 4: Regularization parameter update.** Set  $\sigma_{k+1}$  according to (2.8), increment  $k$  by one and go to Step 1.

As is the case for  $\theta_1$  in the AR1pGN algorithm, choosing an iteration dependent  $\theta_{2,k}$  is possible provided it is strictly bounded below by 1 and bounded above by a constant.

The existence of a suitable step in the AR2GN algorithm directly hinges on Theorem 4.2.

**Lemma 4.3** A step satisfying (2.5), (2.6) (for  $p = 2$ ) and (4.11) always exists.

**Proof.** It follows from Corollary 2.3 and Theorem 4.2 (with  $\phi(s) = T_{f,2}(x_k, s)$ ,  $m = m_k$  and  $\sigma = \sigma_k$ ) and the bound  $\theta_2 > 1$  that the required conditions are satisfied at every global minimizer of the model  $m_k$ .  $\square$

As for the first-order case, continuity of  $\|s\|_r$  and of  $T_{f,2}(x, s)$  with respect to  $s$  implies that conditions (2.5), (2.6) and (4.11) also hold in a neighbourhood of a global minimizer whenever  $\theta_1 > 1$  and  $\theta_2 > 1$ . Such a neighbourhood can be reached for instance by using the algorithm discussed in Section 5.

Noting that Lemmas 2.4, 3.1, 3.2 and 3.3 remain valid for the AR2GN algorithm, we now provide a lower bound on the length of the step, which simplifies that of [14, Lemma 3.4].

**Lemma 4.4** Suppose that AS.1 holds for  $p = 2$  and that, for  $k \in \mathcal{S}$  before termination,  $\lambda_{a,k+1} < -\tau\epsilon_2$ . Then

$$\|s_k\|_r \geq \frac{\tau}{L_{f,2} + \tau^{-1}\theta_2\kappa_\omega\sigma_{\max}} \epsilon_2.$$

**Proof.** Let  $k \in \mathcal{S}$  such that  $\lambda_{a,k+1} < -\tau\epsilon_2$ . Since  $\min_z [a(z) + b(z)] \geq \min_z a(z) + \min_z b(z)$ , we deduce that

$$\begin{aligned} \lambda_{a,k+1} &\geq \lambda_r[\nabla_x^2 f(x_{k+1})] = \min_{\|d\|_r=1} \langle \nabla_x^2 f(x_{k+1})d, d \rangle \\ &= \min_{\|d\|_r=1} [\langle \nabla_x^2 f(x_{k+1})d, d \rangle - \langle \nabla_x^2 f(x_k)d, d \rangle + \langle \nabla_x^2 f(x_k)d, d \rangle] \\ &\geq \min_{\|d\|_r=1} [\langle \nabla_x^2 f(x_{k+1})d, d \rangle - \langle \nabla_x^2 f(x_k)d, d \rangle] + \min_{\|d\|_r=1} \langle \nabla_x^2 f(x_k)d, d \rangle \\ &= \min_{\|d\|_r=1} \langle (\nabla_x^2 f(x_{k+1}) - \nabla_x^2 f(x_k))d, d \rangle + \min_{\|d\|_r=1} \langle \nabla_x^2 f(x_k)d, d \rangle \\ &\geq -\|\nabla_x^2 f(x_{k+1}) - \nabla_x^2 f(x_k)\|_{r,2} + \lambda_r[\nabla_x^2 f(x_k)] \\ &\geq -\|\nabla_x^2 f(x_{k+1}) - \nabla_x^2 f(x_k)\|_{r,2} + \tau^{-1}\lambda_{a,k} \\ &\geq -(L_{f,2}\|s_k\|_r + \tau^{-1}\theta_2\omega(s_k)\sigma_k\|s_k\|_r), \end{aligned}$$

where we also used (4.2) (twice), (3.2) for  $p = j = 2$ , (4.11) and (4.7). The conclusion of the lemma then follows from Lemma 3.3, (4.7) and the fact that  $\lambda_{a,k+1} < -\tau\epsilon_2$ .  $\square$

This lemma and its proof show the potentially severe effect of the approximation bound  $\tau$  whose inverse occurs squared as a factor of  $\theta_2\kappa_\omega\sigma_{\max}$ . This is not surprising because the result is based on the closeness of  $\lambda_r[\nabla_x^2 f(x_k)]$  and  $\lambda_r[\nabla_x^2 f(x_{k+1})]$ , both of which are approximated within that factor.

We conclude our analysis by stating our final evaluation complexity bound for finding second-order  $\epsilon_2$ -approximate minimizers.

**Theorem 4.5** Suppose that AS.1–AS.2 hold for  $p = 2$  and let

$$\kappa_{\text{AR2GN}} \stackrel{\text{def}}{=} \max \left\{ \left[ \frac{1}{2} (L_{f,2} + \theta_1 \sigma_{\max}) \right]^{3/2}, \left[ \tau^{-1} (L_{f,2} + \tau^{-1} \theta_2 \kappa_\omega \sigma_{\max}) \right]^3 \right\}.$$

Then the AR2GN algorithm requires at most

$$\left( \frac{6 \kappa_{\text{AR2GN}}}{\eta_1 \sigma_{\min}} \right) \frac{f(x_0) - f_{\text{low}}}{\min \left[ \epsilon_1^{3/2}, \epsilon_2^3 \right]} \quad (4.12)$$

successful iterations and evaluations of  $g$  and  $H$  and at most

$$\left( \frac{6 \kappa_{\text{AR2GN}}}{\eta_1 \sigma_{\min}} \right) \frac{f(x_0) - f_{\text{low}}}{\min \left[ \epsilon_1^{3/2}, \epsilon_2^3 \right]} \left( 1 + \frac{|\log \gamma_1|}{\log \gamma_2} \right) + \frac{1}{\log \gamma_2} \log \left( \frac{\sigma_{\max}}{\sigma_0} \right) \quad (4.13)$$

evaluations of  $f$  to produce a vector  $x_\epsilon \in \mathbb{R}^n$  such that

$$\|\nabla_x^1 f(x_\epsilon)\|_{r,1} \leq \epsilon_1 \quad \text{and} \quad \lambda_r[\nabla_x^2 f(x_\epsilon)] \geq -\epsilon_2. \quad (4.14)$$

**Proof.** We prove the upper bounds (4.12) and (4.13) on the number of evaluations requested to produce an iterate  $x_\epsilon$  at iteration  $k_\epsilon$  such that

$$\|\nabla_x^1(x_\epsilon)\|_{r,1} \leq \epsilon_1 \quad \text{and} \quad \lambda_{a,k_\epsilon} \geq -\tau \epsilon_2$$

is identical to that of Theorem 3.5, except that  $p = 2$  and the decrease

$$\left( \frac{p!}{L_{f,p} + \theta_1 \sigma_{\max}} \right)^{\frac{p+1}{p}} \epsilon_1^{\frac{p+1}{p}}$$

is now replaced, using Lemmas 3.4 and 4.4, by

$$\min \left\{ \left( \frac{2}{L_{f,2} + \theta_1 \sigma_{\max}} \right)^{\frac{3}{2}} \epsilon_1^{\frac{3}{2}}, \left( \frac{1}{\tau^{-1} (L_{f,2} + \tau^{-1} \theta_2 \kappa_\omega \sigma_{\max})} \right)^3 \epsilon_2^3 \right\}.$$

yielding the constant  $\kappa_{\text{AR2GN}}$ . The desired conclusion then follows from (4.2).  $\square$

We emphasize again that a small value of  $\tau^{-1}$  could lead to disastrous growth of the constant  $\kappa_{\text{AR2GN}}$  in the above theorem, indicating potential difficulties. Consequences and alternative are discussed in Section 6.

## 5 An algorithm for approximate minimization of regularized quadratics

This section is devoted to the definition and analysis of a method whose purpose is to minimize a model of the form

$$m(s) = f_0 + \langle g, s \rangle + \frac{1}{2} \langle Hs, s \rangle + \frac{1}{6} \sigma \|s\|_r^3, \quad (5.1)$$

approximately, but enough for the conditions requested in Step 2 of the AR1pGN (for  $p = 2$ ) and AR2GN algorithms to hold for  $m = m_k$ . We first state a simple technical lemma.

**Lemma 5.1** Consider the quadratic polynomial  $q(t) = at^2 + bt + c$  with  $a \neq 0$  and  $c > 0$ . Then, for any  $\nu > 0$ ,

$$q(t_*) > \frac{1}{2}c \quad \text{for} \quad t_* = \min \left[ \frac{c}{\nu + 3|b|}, \frac{1}{3} \sqrt{\frac{c}{|a|}} \right]. \quad (5.2)$$

**Proof.** We immediately obtain that

$$q(t_*) \geq c - |b| \left( \frac{c}{\nu + 3|b|} \right) - |a| \left( \frac{1}{9} \frac{c}{|a|} \right) \geq c \left( 1 - \frac{1}{3} - \frac{1}{9} \right) > \frac{1}{2}c.$$

□

The constant  $\nu$  in (5.2) is introduced to safeguard against  $b = 0$  and its value can be chosen for convenience in what follows.

We may start building our specialized method (which we will call the RQMIN algorithm) for minimizing the regularized quadratic (5.1). The algorithm will unsurprisingly be iterative and we will denote its successive iterates by  $\{s_k\}_{k \geq 0}$  (the index  $k$  refers, for the rest of this section, to RQMIN iterations, and therefore  $g_k = g + Hs_k$ ). We will also make the choice to start from the origin, that is  $s_0 = 0$ . Moreover, we will construct the iterates  $s_k$  to ensure that the sequence  $\{m(s_k)\}_{k \geq 0}$  is monotonically decreasing from  $m(0)$ . To motivate the forthcoming detailed description of the algorithm ensuring this property, we now consider the magnitude of the model decrease which can be obtained at a given iterate  $s_k$ , if any. We know from Lemma 2.2 that, if  $s_k$  were a local minimizer of  $m$ , then

$$\|g_k\|_{r,1} = \frac{1}{2} \sigma \|s_k\|_r^2. \quad (5.3)$$

This is the condition that the RQMIN algorithm will strive to achieve. If (5.3) fails, we will now show that taking a step from  $s_k$  along a well chosen direction  $d_k$  does produce a model decrease

$$\Delta m(\alpha) \stackrel{\text{def}}{=} m(s_k) - m(s_k + \alpha d_k) = -\alpha \langle g_k, d_k \rangle - \frac{1}{2} \alpha^2 \langle H d_k, d_k \rangle - \frac{1}{6} \sigma \|s_k + \alpha d_k\|_r^3 + \frac{1}{6} \sigma \|s_k\|_r^3 \quad (5.4)$$

which is suitably large. We start by analyzing the case where the step is too short (in view of (5.3)), in which case a generalized ‘‘Cauchy point’’ will provide adequate descent.



**Lemma 5.2** Let  $s_k \in \mathbb{R}^n$  such that  $m(s_k) \leq m(0)$  and

$$\|g_k\|_{r,1} \geq \frac{1}{2}\sigma\|s_k\|_r^2. \quad (5.5)$$

Then

$$m(s_k) - m(s_k^C) \geq \frac{1}{2} \min \left[ \frac{\|\|g_k\|_{r,1} - \frac{1}{2}\sigma\|s_k\|_r^2\|^2}{1 + \frac{3}{2}(\|H\|_{r,2} + \sigma\|s_k\|_r)}, \frac{\|\|g_k\|_{r,1} - \frac{1}{2}\sigma\|s_k\|_r^2\|^{\frac{3}{2}}}{3\sqrt{\sigma}} \right], \quad (5.6)$$

where  $g_k = g + Hs_k$  and

$$s_k^C = s_k + \alpha_k^C d_k, \quad \text{with } d_k = \arg \min_{\|v\|_r=1} \langle g_k, v \rangle \quad \text{and} \quad \alpha_k^C = \arg \min_{\alpha>0} m(s_k + \alpha d_k). \quad (5.7)$$

**Proof.** If  $\|g_k\|_{r,1} = \frac{1}{2}\sigma\|s_k\|_r^2$ , the definition of  $s_k^C$  implies that  $m(s_k) - m(s_k^C) \geq 0$  and (5.6) trivially follows. Suppose therefore that the inequality in (5.5) is strict, and consider the unidimensional minimization of  $m(s_k + \alpha d_k)$  as a function of the scalar  $\alpha$ , where  $d_k$  is given by (5.7). Hence (5.4) holds.

Suppose first that  $s_k = 0$  and thus that  $g_k = g$ . Then

$$\Delta m(\alpha) = \alpha q_0(\alpha) \quad \text{where} \quad q_0(\alpha) = \|g_k\|_{r,1} - \frac{1}{2}\alpha \langle H d_k, d_k \rangle - \frac{1}{6}\sigma\alpha^2.$$

We have that  $q_0(0) = \|g_k\|_{r,1} > 0$  and  $q_0(\alpha)$  is a strictly concave quadratic. Hence the equation  $q_0(\alpha) = 0$  has a positive real root and we may apply Lemma 5.1 with  $\nu = 1$  to deduce that

$$q_0(\alpha_0) > \frac{1}{2}\|g_k\|_{r,1} \quad \text{where} \quad \alpha_0 = \min \left[ \frac{\|g_k\|_{r,1}}{1 + \frac{3}{2}|\langle H d_k, d_k \rangle|}, \frac{1}{3} \sqrt{\frac{\|g_k\|_{r,1}}{\frac{1}{6}\sigma}} \right],$$

and thus that

$$\Delta m(\alpha_0) \geq \alpha_0 q_0(\alpha_0) > \frac{1}{2}\|g_k\|_{r,1} \min \left[ \frac{\|g_k\|_{r,1}}{1 + \frac{3}{2}\|H\|_{r,2}}, \frac{1}{3} \sqrt{\frac{\|g_k\|_{r,1}}{\sigma}} \right]. \quad (5.8)$$

Suppose now that  $\|s_k\|_r > 0$  and define  $v_k = \|s_k\|_r d_k$ . Then, because  $\|v_k\|_r = \|s_k\|_r$ , we have that  $\|s_k + \alpha s_k\|_r \geq \|s_k + \alpha v_k\|_r$  and hence, from (5.4),

$$\Delta m(\alpha) \geq \alpha \|g_k\|_{r,1} \|s_k\|_r - \frac{1}{2}\alpha^2 \langle H v_k, v_k \rangle - \frac{1}{6}\sigma \|s_k + \alpha s_k\|_r^3 + \frac{1}{6}\sigma \|s_k\|_r^3. \quad (5.9)$$

Observe now that

$$\|s_k + \alpha s_k\|_r^3 - \|s_k\|_r^3 = [(1 + \alpha)^3 - 1] \|s_k\|_r^3 = \alpha(3 + 3\alpha + \alpha^2) \|s_k\|_r^3,$$

and thus (5.9) becomes

$$\Delta m(\alpha) \geq \alpha \left[ \|g_k\|_{r,1} \|s_k\|_r - \frac{1}{2}\alpha \langle H v_k, v_k \rangle - \frac{1}{6}\sigma(3 + 3\alpha + \alpha^2) \|s_k\|_r^3 \right] \stackrel{\text{def}}{=} \alpha q_1(\alpha) \quad (5.10)$$

with

$$q_1(\alpha) = (\|g_k\|_{r,1}\|s_k\|_r - \beta) - \alpha \left( \frac{1}{2} \langle H v_k, v_k \rangle + \beta \right) - \alpha^2 \left( \frac{1}{3} \beta \right),$$

where we have defined  $\beta \stackrel{\text{def}}{=} \frac{1}{2} \sigma \|s_k\|_r^3$ . Note that the constant term is positive because we have assumed that  $\|g_k\|_{r,1} > \frac{1}{2} \sigma \|s_k\|_r^2$ . Moreover,  $q_1(\alpha)$  is strictly concave. As above, this implies that the equation  $q_1(\alpha) = 0$  has a positive real root and we may then apply Lemma 5.1 with  $\nu = \frac{1}{2} \|s_k\|_r^2$  to deduce that

$$q_1(\alpha_1) > \frac{1}{2} (\|g_k\|_{r,1}\|s_k\|_r - \beta),$$

where

$$\alpha_1 \stackrel{\text{def}}{=} \min \left[ \frac{\|g_k\|_{r,1}\|s_k\|_r - \beta}{\frac{1}{2} \|s_k\|_r^2 + 3 \left| \frac{1}{2} \langle H v_k, v_k \rangle + \beta \right|}, \frac{\sqrt{\|g_k\|_{r,1}\|s_k\|_r - \beta}}{3\sqrt{\frac{1}{3}\beta}} \right],$$

and hence, from (5.10),

$$\begin{aligned} \Delta m(\alpha_1) &\geq \alpha_1 q(\alpha_1) \\ &\geq \frac{1}{2} (\|g_k\|_{r,1}\|s_k\|_r - \beta) \min \left[ \frac{\|g_k\|_{r,1}\|s_k\|_r - \beta}{\frac{1}{2} \|s_k\|_r^2 + 3 \left| \frac{1}{2} \langle H v_k, v_k \rangle + \beta \right|}, \frac{\sqrt{\|g_k\|_{r,1}\|s_k\|_r - \beta}}{3\sqrt{\frac{1}{3}\beta}} \right] \\ &= \frac{1}{2} \min \left[ \frac{(\|g_k\|_{r,1}\|s_k\|_r - \beta)^2}{\frac{1}{2} \|s_k\|_r^2 + 3 \left| \frac{1}{2} \langle H v_k, v_k \rangle + \beta \right|}, \frac{(\|g_k\|_{r,1}\|s_k\|_r - \beta)^{\frac{3}{2}}}{3\sqrt{\frac{1}{3}\beta}} \right]. \end{aligned}$$

Using the identity  $\|v_k\|_r = \|s_k\|_r$  and substituting the definition of  $\beta$ , this finally gives that

$$\Delta m(\alpha_1) \geq \frac{1}{2} \min \left[ \frac{(\|g_k\|_{r,1} - \frac{1}{2} \sigma \|s_k\|_r^2)^2}{1 + \frac{3}{2} (\|H\|_{r,2} + \sigma \|s_k\|_r)}, \frac{(\|g_k\|_{r,1} - \frac{1}{2} \sigma \|s_k\|_r^2)^{\frac{3}{2}}}{3\sqrt{\sigma}} \right]. \quad (5.11)$$

Combining now (5.8) and (5.11) gives (5.6).  $\square$

Note that the ‘‘Cauchy step’’  $d_k$  in (5.7) is made in the direction of the steepest descent for the *unregularized* quadratic, that is *ignoring the regularization term*.

We now consider the alternative to (5.5), which, as (5.3) indicates, means that the step  $s_k$  is too large. It therefore makes sense to consider moving back from  $s_k$  towards the origin.

**Lemma 5.3** Let  $s_k \in \mathbb{R}^n$  such that  $m(s_k) \leq m(0)$  and

$$\|g_k\|_{r,1} < \frac{1}{2} \sigma \|s_k\|_r^2. \quad (5.12)$$

Then

$$m(s_k) - m(s_k^R) \geq \frac{1}{2} \min \left[ \frac{\left| \|g_k\|_{r,1} - \frac{1}{2} \sigma \|s_k\|_r^2 \right|^2}{1 + \frac{3}{2} (\|H\|_{r,2} + \sigma \|s_k\|_r)}, \frac{\left| \|g_k\|_{r,1} - \frac{1}{2} \sigma \|s_k\|_r^2 \right|^{\frac{3}{2}}}{3\sqrt{\sigma}} \right], \quad (5.13)$$

where  $g_k = g + H s_k$  and

$$s_k^R = (1 - \alpha_k^R) s_k \quad \text{with} \quad \alpha_k^R = \arg \min_{\alpha > 0} m(s_k - \alpha s_k). \quad (5.14)$$

**Proof.** Note that (5.12) implies that  $s_k \neq 0$ . Then, from (5.4) with  $d_k = -s_k$ ,

$$\Delta m(\alpha) \geq \alpha \langle g_k, s_k \rangle - \frac{1}{2} \alpha^2 \langle H s_k, s_k \rangle - \frac{1}{6} \sigma \|s_k - \alpha s_k\|_r^3 + \frac{1}{6} \sigma \|s_k\|_r^3. \quad (5.15)$$

Now

$$\|s_k - \alpha s_k\|_r^3 - \|s_k\|_r^3 = [(1 - \alpha)^3 - 1] \|s_k\|_r^3 = -\alpha(3 - 3\alpha + \alpha^2) \|s_k\|_r^3,$$

so that, from (5.15),

$$\Delta m(\alpha) \geq \alpha \left[ \langle g_k, s_k \rangle - \frac{1}{2} \alpha \langle H s_k, s_k \rangle + \frac{1}{6} \sigma (3 - 3\alpha + \alpha^2) \|s_k\|_r^3 \right] \stackrel{\text{def}}{=} \alpha q_2(\alpha), \quad (5.16)$$

where

$$q_2(\alpha) = (\langle g_k, s_k \rangle + \beta) + \alpha \left( -\frac{1}{2} \langle H s_k, s_k \rangle - \beta \right) + \alpha^2 \left( \frac{1}{3} \beta \right).$$

Observe now that, because  $s_k \neq 0$  and, since we have assumed that  $m(s_k) \leq m(0)$ , we have that  $q_2(1) = \Delta m(1) \leq 0$ . Moreover, the Cauchy-Schwarz inequality yields that

$$|\langle g_k, s_k \rangle| \leq \|g_k\|_{r,1} \|s_k\|_r < \frac{1}{2} \sigma \|s_k\|_r^3 = \beta$$

and hence

$$q_2(0) = \langle g_k, s_k \rangle + \beta > 0.$$

This in turn implies the existence of a real root of  $q_2(\alpha)$  in  $(0, 1]$ , and we may then again apply Lemma 5.1 with  $\nu = \frac{1}{2} \|s_k\|_r^2$  to deduce that

$$q_2(\alpha_2) > \frac{1}{2} (\langle g_k, s_k \rangle + \beta), \quad (5.17)$$

where

$$\alpha_2 = \min \left[ \frac{\langle g_k, s_k \rangle + \beta}{\frac{1}{2} \|s_k\|_r^2 + 3 \left| -\frac{1}{2} \langle H s_k, s_k \rangle + \beta \right|}, \frac{1}{3} \sqrt{\frac{\langle g_k, s_k \rangle + \beta}{\frac{1}{3} \beta}} \right]. \quad (5.18)$$

Moreover

$$\langle g_k, s \rangle + \beta \geq -\|g_k\|_{r,1} \|s\|_r + \frac{1}{2} \sigma \|s\|_r^3 > 0.$$

Combining this bound with (5.16), (5.18) and (5.17), we obtain that

$$\Delta m(\alpha_2) \geq \frac{1}{2} \min \left[ \frac{(\frac{1}{2} \sigma \|s_k\|_r^2 - \|g_k\|_{r,1})^2}{1 + \frac{3}{2} (\|H\|_{r,2} + \sigma \|s_k\|_r)}, \frac{(\frac{1}{2} \sigma \|s_k\|_r^2 - \|g_k\|_{r,1})^{\frac{3}{2}}}{3\sqrt{\sigma}} \right], \quad (5.19)$$

which yields (5.13).  $\square$

Remarkably, (5.6) and (5.13) give identical lower bounds for the model decrease. Lemmas 5.2 and 5.3 generalize [9, Lemma 2.1] to the case where  $s_k \neq 0$  and general norms are allowed.

We may now complete the analysis of what can happen at iterate  $s_k$  (of the still unspecified RQMIN algorithm) if the second-order necessary condition of Theorem (4.2) fails. We first state an easy lemma giving lower and upper bounds on the step  $s_k$ , dependent on the ‘‘most negative curvature’’ of the quadratic given by (4.1).

**Lemma 5.4** Suppose that, for some  $s_k$  and some  $\beta \geq 0$ ,

$$m(s_0) - m(s_k) \geq \beta. \quad (5.20)$$

Then,

$$\|s_k\|_r \leq \frac{\frac{1}{2}\|H\|_{r,2} + \sqrt{\|H\|_{r,2}^2 + \frac{2}{3}\sigma\|g\|_{r,1}}}{\frac{1}{3}\sigma} \stackrel{\text{def}}{=} \kappa_{s,\text{upp}}, \quad (5.21)$$

and, if  $\beta > 0$ ,

$$\|s_k\|_r \geq \begin{cases} \frac{\sqrt{\|g\|_{r,1}^2 + 2\beta|\lambda_r[H]|} - \|g\|_{r,1}}{|\lambda_r[H]|} & \text{if } \lambda_r[H] < 0 \\ \frac{\beta}{\|g\|_{r,1}} & \text{otherwise.} \end{cases} \quad (5.22)$$

**Proof.** Since  $s_0 = 0$  and  $m(s_k) \geq m(0) + \langle g, s_k \rangle + \frac{1}{2}\langle Hs_k, s_k \rangle$ , (5.20) implies that

$$-\|g\|_{r,1}\|s_k\|_r + \frac{1}{2} \min[0, \lambda_r[H]] \|s_k\|_r^2 \leq \langle g, s_k \rangle + \frac{1}{2}\langle Hs_k, s_k \rangle \leq m(s_k) - m(0) \leq -\beta,$$

which gives (5.22). Observe now that (5.20) implies that

$$\frac{1}{6}\sigma\|s_k\|_r^3 \leq |\langle g, s_k \rangle| + \frac{1}{2}|\langle Hs_k, s_k \rangle| \leq \|g\|_{r,1}\|s_k\|_r + \frac{1}{2}\|H\|_{r,2}\|s_k\|_r^2,$$

which yields (5.21).  $\square$

Armed with this result, we now derive the model decrease when negative curvature is present.

**Lemma 5.5** Suppose that  $\lambda_a < 0$  where  $\lambda_a$  and  $u_a$  satisfy (4.2) and that the sequence  $\{m(s_k)\}_{k \geq 0}$  is non-increasing. For  $k \geq 0$ , define

$$s_k^E = s_k + \alpha_k^E u_k \quad \text{where } u_k = -\text{sign}(\langle g_k, u_a \rangle) u_a \quad \text{and } \alpha_k^E = \arg \min_{\alpha > 0} m(s_k + \alpha u_k). \quad (5.23)$$

Then

$$m(s_0) - m(s_1) \geq m(s_0) - m(s_0^E) \geq \frac{9|\lambda_a|^3}{16\sigma^2} \quad (5.24)$$

and there exists a constant  $\kappa_s$  such that, for  $k \geq 1$ ,

$$\|s_k\|_r \geq \kappa_s. \quad (5.25)$$

Moreover, if

$$\lambda_a + \sigma\omega(s_k)\|s_k\|_r < 0 \quad (5.26)$$

at iteration  $k \geq 1$ , then one has that

$$m(s_k) - m(s_k^E) \geq \frac{9\|s_k\|_r^2}{16\sigma^2} \left| \lambda_a + \sigma\omega(s_k)\|s_k\|_r \right|^3. \quad (5.27)$$

**Proof.** The first inequality in (5.24) results from (5.31) and the second is a direct consequence of the proof of Theorem 4.2 (see (4.8)). The existence of  $\kappa_s$  such that (5.25) holds for  $k \geq 1$  then follows from Lemma 5.4 with  $\beta = \frac{9}{16}|\lambda_a|^3/\sigma^2$  and our assumption that  $\{m(s_k)\}$  is non-increasing. We now prove (5.27). From (5.26), we have that

$$\lambda_a + \sigma\omega(s_k)\|s_k\|_r = -\mu\sigma\|s_k\|_r \quad (5.28)$$

for some  $\mu > 0$ . But (4.7) implies that  $0 > \lambda_a + \sigma\|s_k\|_r = -|\lambda_a + \sigma\|s_k\|_r|$ , and thus, from (4.7),

$$|\lambda_a + \sigma\|s_k\|_r| = \frac{2\sqrt{\psi(s_k)}}{\sqrt{3}}\sigma\|s_k\|_r + \mu\sigma\|s_k\|_r, \quad (5.29)$$

from which we obtain that

$$(\lambda_a + \sigma\|s_k\|_r)^2 = \frac{4}{3}\psi(s_k)\sigma^2\|s_k\|_r^2 + \sigma^2\|s_k\|_r^2 \left( \mu^2 + \frac{4\sqrt{\psi(s_k)}}{\sqrt{3}}\mu \right).$$

Substituting this inequality in (4.3), then gives that there exists an  $\alpha > 0$  such that

$$\begin{aligned} m(s_k) - m(s_k + \alpha\|s_k\|_r u_a) &\geq \frac{3(\lambda_a + \sigma\|s_k\|_r)}{4\sigma^2} \left[ -\frac{3}{4}\sigma^2\|s_k\|_r^2 \left( \mu^2 + \frac{4\sqrt{\psi(s_k)}}{\sqrt{3}}\mu \right) \right] \\ &= \frac{9}{16} |\lambda_a + \sigma\|s_k\|_r| \|s_k\|_r^2 \left( \mu^2 + \frac{4\sqrt{\psi(s_k)}}{\sqrt{3}}\mu \right). \end{aligned}$$

But (5.29) implies that  $|\lambda_a + \sigma\|s_k\|_r| \geq \mu\sigma\|s_k\|_r$ , and thus

$$m(s_k) - m(s_k + \alpha\|s_k\|_r u_a) \geq \frac{9\sigma\|s_k\|_r^3}{16} \left( \mu^3 + \frac{4\sqrt{\psi(s_k)}}{\sqrt{3}}\mu^2 \right) \geq \frac{9\sigma\|s_k\|_r^3}{16} \mu^3.$$

The inequality (5.27) then follows from (5.28) and (5.23).  $\square$

We now have all ingredients to describe the RQMIN algorithm. It hinges on (5.2), (5.3) and (5.5) and proceeds by successive one-dimensional minimizations of  $m$  along the directions  $s_k^C$  or  $s_k^R$  (depending on the sign of  $\|g_k\|_{r,1} - \frac{1}{2}\sigma\|s\|_r^2$ ) and, if needed,  $s_k^E$ . It is formally stated on the following page.

Note that the mechanism of the algorithm, which proceeds by successive unidimensional minimizations, guarantees that the sequence  $\{m(s_k)\}$  is monotonically decreasing, as announced.

Having established, in Lemma 5.2, 5.3 and 5.5, lower bounds on the decrease in  $m$  for all steps produced by the RQMIN algorithm, we are now ready to state its iteration complexity<sup>(2)</sup>.

<sup>(2)</sup>At variance with its evaluation complexity, which would be irrelevant here since evaluating  $m(s)$  as many times as necessary does not require evaluating  $f_0$ ,  $g$  and  $H$  more than once (when the algorithm is called).

**Algorithm 5.1: An algorithm for minimization of a regularized quadratic (RQMIN)**

The value  $f_0$ , gradient  $g$  and Hessian  $H$  of the quadratic at  $s = 0$  are given, as well as a regularization parameter  $\sigma$  and accuracy requests  $\epsilon_1 > 0$  and  $\epsilon_2 > 0$ .

**Step 0: Initialization** If unavailable, compute  $\lambda_a$  and  $u_a$  according to (4.2). Set  $k = 0$ ,  $s_0 = 0$  and  $g_0 = g$ .

**Step 1: Check for termination.** Terminate if

$$\|g_k\|_{r,1} - \frac{1}{2}\sigma\|s_k\|_r^2 \leq \epsilon_1 \quad \text{and} \quad \lambda_a + \theta_2\omega(s_k)\sigma\|s_k\|_r \geq 0. \quad (5.30)$$

**Step 2: Negative gradient step.** If  $\|g_k\|_{r,1} > \frac{1}{2}\sigma\|s_k\|_r^2$ , compute  $s_k^C$  according to (5.7), set  $m_{k,1} = m(s_k^C)$  and go to Step 4.

**Step 3: Retraction step.** If  $\|g_k\|_{r,1} < \frac{1}{2}\sigma\|s_k\|_r^2$ , compute  $s_k^C$  according to (5.14) and set  $m_{k,1} = m(s_k^R)$ .

**Step 4: Eigenvalue step.** If  $\lambda_a + \omega(s_k)\sigma\|s_k\|_r < -\epsilon_2\sigma\|s_k\|_r$ , compute  $s_k^E$  according to (5.23) and set  $m_{k,2} = m(s_k^E)$ . Else, set  $m_{k,2} = m(s_k)$ .

**Step 5: Move.** Set

$$s_{k+1} = \begin{cases} s_k^C & \text{if } m_{k,1} \leq m_{k,2}, \\ s_k^E & \text{otherwise,} \end{cases} \quad \text{and} \quad g_{k+1} = g_k + H(s_{k+1} - s_k). \quad (5.31)$$

Increment  $k$  by one and go to Step 1.

**Theorem 5.6** Given  $\epsilon_1 > 0$  and  $\theta_2 > 1$ , there exist a constant  $\kappa_{\text{RQMIN}} > 0$  independent of  $k$  such that the RQMIN algorithm requires at most

$$\kappa_{\text{RQMIN}} \max \left[ \epsilon_1^{-2}, \epsilon_1^{-\frac{3}{2}}, \epsilon_2^{-3} \right] \quad (5.32)$$

iterations to produce an iterate  $s_k$  such that

$$\|g_k\|_{r,1} - \frac{1}{2}\sigma\|s_k\|_r^2 \leq \epsilon_1 \quad \text{and} \quad \lambda_a + \omega(s_k)\sigma\|s_k\|_r \geq -\epsilon_2\sigma\|s_k\|_r. \quad (5.33)$$

**Proof.** If the RQMIN algorithm terminates at  $k = 0$ , then the bound (5.32) is trivially satisfied. Assume therefore that termination does not occur at  $s_0$ . We then have that, for  $k \geq 1$  before termination,

$$\text{either } \|g_k\|_{r,1} - \frac{1}{2}\theta_1\sigma\|s_k\|_r^2 > \epsilon_1 \quad \text{or} \quad \lambda_a + \theta_2\omega(s_k)\sigma\|s_k\|_r < 0.$$

Let us define  $\mathcal{N} \stackrel{\text{def}}{=} \{k \geq 0 \mid s_k = s_k^E\}$  and note that, by construction, this set is non-empty only if  $\lambda_a < 0$ . We then obtain from  $g_0 = g$ , (5.31), (5.6) and (5.24) that

$$m(s_0) - m(s_1) \geq \begin{cases} \frac{1}{2} \max \left\{ \min \left[ \frac{\|g\|_{r,1}^2}{1 + \frac{3}{2}\|H\|_{r,2}}, \frac{\|g\|_{r,1}^{\frac{3}{2}}}{3\sqrt{\sigma}} \right], \frac{9|\lambda_a|^3}{16\sigma^2} \right\} & \text{if } 0 \in \mathcal{N}, \\ \frac{1}{2} \min \left[ \frac{\|g\|_{r,1}^2}{1 + \frac{3}{2}\|H\|_{r,2}}, \frac{\|g\|_{r,1}^{\frac{3}{2}}}{3\sqrt{\sigma}} \right] & \text{otherwise.} \end{cases} \quad (5.34)$$

Observe now that the second part of (5.33) cannot hold as long as

$$-(\lambda_a + \omega(s_k)\sigma\|s_k\|_r) > \epsilon_2\omega(s_k)\sigma\|s_k\|_r.$$

Hence (5.27) give that, for  $k \geq 1$  and  $k \in \mathcal{N}$  before termination,

$$m(s_k) - m(s_k^E) \geq \frac{9\|s_k\|_r}{16\sigma^2} [(\theta_2 - 1)\omega(s_k)]^3. \quad (5.35)$$

Because Lemma 5.4 with  $\beta$  chosen as the relevant right-hand side in (5.34) guarantees the existence of  $\kappa_{s,\text{low}} > 0$  such that  $\|s_k\|_r \geq \kappa_{s,\text{low}}$  for all  $k \geq 1$ , and, because  $\omega(s_k) \geq 1$ , (5.35) ensures that, before termination and for  $1 \leq k \in \mathcal{N}$ ,

$$m(s_k) - m(s_k^E) \geq \frac{9\kappa_{s,\text{low}}^2}{16\sigma^2} \epsilon_2^3.$$

Using this together with (5.31) and (5.6) gives that for  $k \geq 1$  before termination,

$$m(s_k) - m(s_{k+1}) \geq \begin{cases} \frac{1}{2} \max \left\{ \min \left[ \frac{\epsilon_1^2}{1 + \frac{3}{2}(\|H\|_{r,2} + \sigma\|s_k\|_r)}, \frac{\epsilon_1^{\frac{3}{2}}}{3\sqrt{\sigma}} \right], \frac{9\epsilon_2^3\kappa_{s,\text{low}}^2}{16\sigma^2} \right\} & \text{if } k \in \mathcal{N}, \\ \min \left[ \frac{\epsilon_1^2}{1 + \frac{3}{2}(\|H\|_{r,2} + \sigma\|s_k\|_r)}, \frac{\epsilon_1^{\frac{3}{2}}}{3\sqrt{\sigma}} \right] & \text{otherwise,} \end{cases}$$

and therefore, using (5.21),

$$m(s_k) - m(s_{k+1}) \geq \frac{1}{2} \min \left[ \frac{1}{1 + \frac{3}{2} (\|H\|_{r,2} + \sigma \kappa_{s,\text{upp}})}, \frac{1}{3\sqrt{\sigma}}, \frac{9\kappa_{s,\text{low}}^2}{16\sigma^2} \right] \min \left[ \epsilon_1^2, \epsilon_1^{\frac{3}{2}}, \epsilon_2^3 \right] \\ \stackrel{\text{def}}{=} \kappa_* \min \left[ \epsilon_1^2, \epsilon_1^{\frac{3}{2}}, \epsilon_2^3 \right]. \quad (5.36)$$

We now observe that the definition of  $m(s)$  and (5.21) together imply that

$$m(s) \geq m(0) - \|g\|_{r,1} \kappa_{s,\text{upp}} - \frac{1}{2} \|H\|_{r,2} \kappa_{s,\text{upp}}^2 \stackrel{\text{def}}{=} m_{\text{low}}.$$

Therefore (5.36) implies that the number of iterations required by the RQMIN algorithm to produce an iterate such that (5.33) holds cannot exceed

$$\frac{m(0) - \beta - m_{\text{low}}}{\kappa_* \min \left[ \epsilon_1^2, \epsilon_1^{\frac{3}{2}}, \epsilon_2^3 \right]}$$

which is (5.32) with  $\kappa_{\text{RQMIN}} = (m(0) - \beta - m_{\text{low}})/\kappa_*$ .  $\square$

We now consider applying the RQMIN algorithm to find a step  $s_k$  in Step 2 of the AR2GN method<sup>(3)</sup>. This latter methods requires the conditions (2.5), (2.6) and (4.11) to hold. We immediately note that (2.5) automatically holds because of the monotonically decreasing nature of the values of  $m$  in the RQMIN algorithm. Moreover, (4.11) and the second part of (5.30) are identical. However, the first part of (5.30) is too strong, because it imposes a two-sided inequality on  $\|g_k\|_{r,1} - \frac{1}{2}\sigma_k\|s_k\|_r$  while (2.6) only requests

$$\|g_k\|_{r,1} - \frac{1}{2}\sigma_k\|s_k\|_r^2 \leq \epsilon_{1s} \stackrel{\text{def}}{=} \frac{1}{2}(\theta_1 - 1)\sigma_k\|s_k\|_r^2 \quad (5.37)$$

but allows for  $\|g_k\|_{r,1} - \frac{1}{2}\sigma_k\|s_k\|_r^2$  to be negative. In Figure 4.1, this amounts to removing the outer blue curve, thus considerably enlarging the admissible regions containing the minimizers. A modified variant of the RQMIN algorithm is therefore suitable if our only objective is to satisfy (2.5), (2.6) and (4.11). This variant, which we call the RQMIN1 algorithm, differs from RQMIN in that

1. the first part of (5.30) is replaced by requiring that (5.37) holds,
2. Step 3 of RQMIN is skipped (as there is no need to correct for negative  $\|g_k\|_{r,1} - \frac{1}{2}\sigma_k\|s_k\|_r^2$ ).

In addition, because (5.37) is weaker than the first part of (5.30), termination of the RQMIN1 algorithm cannot happen later than that what would happen if applying the RQMIN algorithm with  $\epsilon_1 = \epsilon_{1s}$ . This allows us to derive the following upper bound on the number of iterations of the RQMIN1 algorithm that are necessary to compute a step  $s_k$  in Step 2 of AR2GN.

**Corollary 5.7** Given  $\theta_1 > 1$  and  $\theta_2 > 1$ , there exist a constant  $\kappa_{\text{RQMIN1}} > 0$  independent of  $k$  such that the RQMIN1 algorithm requires at most

$$\kappa_{\text{RQMIN1}} \max \left[ (\theta_1 - 1)^{-2}, (\theta_1 - 1)^{-\frac{3}{2}}, (\theta_2 - 1)^{-3} \right] \quad (5.38)$$

iterations to produce an iterate  $s_k$  such that (2.5), (2.6) and (4.11) hold.

<sup>(3)</sup>With  $f(x_k) = f_0$ ,  $\nabla_x^1 f(x_k) = g$ ,  $\nabla_x^2 f(x_k) = H$ ,  $\sigma_k = \sigma$  and  $x_k = 0$ .



**Proof.** The desired result immediately follows by noting that, by virtue of (2.8), (5.37) and the definition of  $\kappa_{s,\text{low}}$  in the proof of Theorem 5.6,

$$\epsilon_{1s} \geq \frac{1}{2}(\theta_1 - 1)\sigma\kappa_{s,\text{low}}^2 \quad \text{and} \quad \epsilon_2 = \theta_2 - 1.$$

The bound (5.38) then follows with

$$\kappa_{\text{RQMIN1}} \stackrel{\text{def}}{=} \kappa_{\text{RQMIN}} \min \left[ \left( \frac{1}{2}\sigma\kappa_{s,\text{low}}^2 \right)^2, \left( \frac{1}{2}\sigma\kappa_{s,\text{low}}^2 \right)^{\frac{3}{2}} \right].$$

□

The reader may have wondered why we did consider the RQMIN method and its two-sided condition at all, since its one-sided version RQMIN1 is sufficient for the purpose of computing a step in AR2GN. Our motivation for RQMIN is that it is likely to achieve a larger model decrease, hopefully reducing the number of iterations needed by AR2GN to terminate. Whether this motivation translates in practice remains to be explored.

But the story does not finish here. As we have alluded to in Section 2, an even simpler variant of the RQMIN algorithm can be used to compute  $s_k$  in Step 2 of the AR1 $p$ GN algorithm when  $p = 2$ . Since the only requirements on  $s_k$  are then (2.5) and (2.6), we may define the RQMIN2 algorithm as a variant of RQMIN where

1. the whole of (5.30) is replaced by requiring that (5.37) holds,
2. Step 3 and Step 4 of RQMIN are skipped and the first part of (5.31) replaced by  $s_{k+1} = s_k^C$ .

Removing all bounds related to the second-order condition in Theorem 5.6, we then obtain the following iteration bound for the RQMIN2 algorithm (as needed in Step 2 of AR1 $p$ GN with  $p = 2$ ).

**Corollary 5.8** Given  $\theta_1 > 1$ , there exist a constant  $\kappa_{\text{RQMIN2}} > 0$  independent of  $k$  such that the RQMIN2 algorithm requires at most

$$\kappa_{\text{RQMIN2}} \max \left[ (\theta_1 - 1)^{-2}, (\theta_1 - 1)^{-\frac{3}{2}} \right] \tag{5.39}$$

iterations to produce an iterate  $s_k$  such that (2.5), and (2.6) hold.

Note that the RQMIN2 algorithm reduces to a standard first-order method (in the  $\|\cdot\|_r$  norm), but applied to the quadratic alone, instead of to the complete regularized model.

## 6 Measuring curvature approximately

As we have observed at the end of Section 4, a poor capacity of approximating  $\lambda_r[H]$  in the sense of (4.2), resulting in a large value of  $\tau^{-1}$ , can very quickly lead to extremely large bounds (see the definition of  $\kappa_{\text{AR2GN}}$  in Theorem 4.5). For example, consider the following

naive strategy using  $\lambda_{\min}[H]$ , the smallest standard (Euclidean) eigenvalue of  $H$  and  $u_2$  one of its associated eigenvector. Observe that, for  $\lambda_r[H] < 0$ ,

$$\lambda_r[H] \leq \lambda_{\min}[H] \max_{v \neq 0} \frac{\|v\|_2^2}{\|v\|_r^2} \quad \text{and} \quad \lambda_{\min}[H] \leq \lambda_r[H] \max_{v \neq 0} \frac{\|v\|_r^2}{\|v\|_2^2}.$$

If we now define

$$\kappa_1 \stackrel{\text{def}}{=} \min_{v \neq 0} \frac{\|v\|_r}{\|v\|_2} \leq \max_{v \neq 0} \frac{\|v\|_r}{\|v\|_2} \stackrel{\text{def}}{=} \kappa_2,$$

then

$$\lambda_r[H] \leq \frac{\lambda_{\min}[H]}{\kappa_2^2} \leq \left(\frac{\kappa_1}{\kappa_2}\right)^2 \lambda_r[H],$$

which is (4.2) with  $\lambda_a = \lambda_{\min}[H]/\kappa_2^2$  and  $\tau = (\kappa_1/\kappa_2)^2 \in (0, 1] \dots$  and  $\kappa_{\text{AR2GN}}$  would involve the twelfth power of the norm equivalence constant  $(\kappa_1/\kappa_2)^{-1}$ , a truly fearsome prospect. Unfortunately, for fully general norms, the question of finding a practical approximation scheme remains open. For norms whose unit balls are well approximated by an ellipsoid, a generalized eigenvalue calculation is an option leading to an acceptably small  $\kappa_1/\kappa_2$ .

We conclude by examining two other options. The first is, unsurprisingly, to resort to a heuristic method for computing the approximations  $\lambda_a$  and  $u_a$ . For instance, if the unit ball has a manageable geometry, direct attempts to minimize  $\langle Hv, v \rangle / \|v\|_r^2$  can be considered, or we could call on a stochastic sampling method to approximate  $\lambda_r[H]$  directly. Heuristics approaches are however unlikely to give guarantees of (approximate) global minimization, the AR2GN using the resulting approximation may miss significant negative curvature at iteration  $k$ , and thus terminate early. As a consequence, the evaluation complexity upper bound obtained for  $\tau = 1$  still holds with the proviso that the second part of (4.14) can only be guaranteed within the ‘‘power’’ of the chosen heuristic. In other words, the only ‘‘second-order’’ property holding at  $x_\epsilon$  is that the used heuristic is unable to detect significant negative curvature (in the  $\|\cdot\|_r$  norm) locally. For example, if a stochastic sampling method is used, it might be possible to ensure (4.2) with a given probability, and the second part of (4.14) can then be guaranteed with the same probability. This type of degraded second-order property might well be acceptable in practice. The second option is to give up imposing curvature conditions in the  $\|\cdot\|_r$  norm and to limit one’s ambition to ensuring that  $\lambda_{\min}[\nabla_x^2 f(x_\epsilon)] \geq -\epsilon_2$  at termination (instead of the second part of (4.14)) This has the disadvantage of being theoretically less coherent, but the advantage of being very easily computable. As it turns out, *the theory we have developed in Section 4 remains valid with a modified (4.14)<sup>(4)</sup>* if  $\lambda_r[H]$  is replaced by  $\lambda_{\min}[H]$ ,  $u_a$  is chosen as a corresponding eigenvector with  $\|u_a\|_2 = 1$  and  $\tau$  is set to one if (4.2), and if the Lipschitz constant  $L_{f,2}$  in the  $\|\cdot\|_r$  norm used in AS.1, the proof of Lemma (4.4) and its consequences is replaced by the Lipschitz constant in the Euclidean norm. This is of course meaningful if these Lipschitz constants do not differ too much in magnitude. Whether this modification of the algorithm’s purpose is acceptable is of course strongly application-dependent.

<sup>(4)</sup>The interested reader will find a version of the present paper using this option in [26].

## 7 Discussion

We have presented simple modifications of the already simplified<sup>(5)</sup> first- and second-order evaluation complexity theory for adaptive regularization methods proposed by [6], [20] and [14]. These modifications allow a direct use of general possibly non-smooth regularization norms. The new algorithms (AR1 $p$ GN and AR2GN) differ from their predecessors in that the conditions for approximate model minimization now use the norm of the gradient and the smallest eigenvalue of the Hessian *of the Taylor expansion of  $f$* , instead of those of the regularized model. The resulting first- and second-order complexity theory is even more compact than that presented in [6] and [14]. This theory does not depend on the norm-equivalence constants as far as finding approximate first-order minimizers is concerned, but it does depend on the factor  $\tau$ , the relative accuracy of the approximation of  $\lambda_R[H]$ , in the second-order case. Because this factor may itself depend strongly on norm-equivalence constants, our objective of avoiding such constants is only partially successful if approximate second-order minimizers (in the  $\|\cdot\|_r$  norm) are wanted at a reasonable cost. Practical alternatives were discussed, leading to weaker second-order optimality guarantees. We have also presented and analyzed a new method (RQMIN) for computing (potentially very) approximate second-order minimizers of quadratic polynomials regularized with a general norm. Variants of this method were finally discussed which provide implementable ways of solving the subproblems arising in the AR1 $p$ GN (with  $p = 2$ ) and AR2GN algorithms.

Because of its general nature and strong links with existing theory for regularization in smooth norms, we anticipate that the approach discussed here can be extended to other contexts for which analysis is available for the smooth-norm case. In particular, we think of the minimization of composite functions (along the lines of [16, 25] or [1]) and of the case where function's and derivatives' values are allowed to be inexact but controllable either deterministically (using ideas of [2, 25]) or probabilistically (along the lines of [7] for instance). Another useful extension which seems likely is adding convex constraints to the problem (see [15]).

While it might well be possible to extend our subproblem termination rule to higher order by putting requests on  $T_{f,j}(x_k, s_k)$ , it not yet clear that a step satisfying the resulting condition may be shown to exist. This topic and its potential use in our context are the subject of ongoing research.

Experimenting numerically with the ideas proposed here is also of interest. A first research question is to measure the impact on performance of the more flexible termination rules for the step computation within the standard adaptive cubic regularization algorithm. But we anticipate that further developments are possible, for instance to provide possibly more efficient alternatives to the RQMIN algorithm or to improve the strategies for measuring curvature in the  $\|\cdot\|_r$  norm.

### Acknowledgements

The authors are indebted to Sadok Jerad for his careful reading of the manuscript.

---

<sup>(5)</sup>Compared with [34] or [10].

## References

- [1] A. Aravkin, R. Baraldi, and D. Orban. A proximal quasi-Newton trust-region method for nonsmooth regularized optimization. arXiv:2103.15993, 2021.
- [2] S. Bellavia, G. Gurioli, B. Morini, and Ph. L. Toint. Adaptive regularization algorithms with inexact evaluations for nonconvex optimization. *SIAM Journal on Optimization*, 29(4):2881–2915, 2019.
- [3] E. Bergou, Y. Diouane, and S. Gratton. On the use of the energy norm in trust-region and adaptive cubic regularization subproblems. Optimization Online, April 2017.
- [4] T. Bianconcini, G. Liuzzi, B. Morini, and M. Sciandrone. On the use of iterative methods in cubic regularization for unconstrained optimization. *Computational Optimization and Applications*, 60(1):35–57, 2015.
- [5] T. Bianconcini and M. Sciandrone. A cubic regularization algorithm for unconstrained optimization using line search and nonmonotone techniques. *Optimization Methods and Software*, 31(5):1008–1035, 2016.
- [6] E. G. Birgin, J. L. Gardenghi, J. M. Martínez, S. A. Santos, and Ph. L. Toint. Worst-case evaluation complexity for unconstrained nonlinear optimization using high-order regularized models. *Mathematical Programming, Series A*, 163(1):359–368, 2017.
- [7] J. Blanchet, C. Cartis, M. Menickelly, and K. Scheinberg. Convergence rate analysis of a stochastic trust region method via supermartingales. *INFORMS Journal on Optimization*, 1(2):92–119, 2019.
- [8] C. Cartis, N. I. M. Gould, and Ph. L. Toint. Trust-region and other regularization of linear least-squares problems. *BIT*, 49(1):21–53, 2009.
- [9] C. Cartis, N. I. M. Gould, and Ph. L. Toint. Adaptive cubic overestimation methods for unconstrained optimization. Part I: motivation, convergence and numerical results. *Mathematical Programming, Series A*, 127(2):245–295, 2011.
- [10] C. Cartis, N. I. M. Gould, and Ph. L. Toint. Adaptive cubic overestimation methods for unconstrained optimization. Part II: worst-case function-evaluation complexity. *Mathematical Programming, Series A*, 130(2):295–319, 2011.
- [11] C. Cartis, N. I. M. Gould, and Ph. L. Toint. On the evaluation complexity of composite function minimization with applications to nonconvex nonlinear programming. *SIAM Journal on Optimization*, 21(4):1721–1739, 2011.
- [12] C. Cartis, N. I. M. Gould, and Ph. L. Toint. On the oracle complexity of first-order and derivative-free algorithms for smooth nonconvex minimization. *SIAM Journal on Optimization*, 22(1):66–86, 2012.
- [13] C. Cartis, N. I. M. Gould, and Ph. L. Toint. Worst-case evaluation complexity and optimality of second-order methods for nonconvex smooth optimization. In B. Sirakov, P. de Souza, and M. Viana, editors, *Invited Lectures, Proceedings of the 2018 International Conference of Mathematicians (ICM 2018)*, vol. 4, Rio de Janeiro, pages 3729–3768. World Scientific Publishing Co Pte Ltd, 2018.
- [14] C. Cartis, N. I. M. Gould, and Ph. L. Toint. A concise second-order evaluation complexity for unconstrained nonlinear optimization using high-order regularized models. *Optimization Methods and Software*, 35(2):243–256, 2020.
- [15] C. Cartis, N. I. M. Gould, and Ph. L. Toint. Sharp worst-case evaluation complexity bounds for arbitrary-order nonconvex optimization with inexpensive constraints. *SIAM Journal on Optimization*, 30(1):513–541, 2020.
- [16] C. Cartis, N. I. M. Gould, and Ph. L. Toint. Strong evaluation complexity bounds for arbitrary-order optimization of nonconvex nonsmooth composite functions. arXiv:2001.10802, 2020.
- [17] X. Chen and Ph. L. Toint. High-order evaluation complexity for convexly-constrained optimization with non-Lipschitzian group sparsity terms. *Mathematical Programming, Series A*, 187(1-2):47–78, 2020.
- [18] X. Chen, Ph. L. Toint, and H. Wang. Partially separable convexly-constrained optimization with non-Lipschitzian singularities and its complexity. *SIAM Journal on Optimization*, 29:874–903, 2019.
- [19] F. Clarke, Yu. Ledyaev, R. Stern, and P. Wolenski. *Nonsmooth Analysis and Control Theory*. Springer Verlag, New York, 1998.
- [20] J. P. Dussault. Simple unified convergence proofs for the trust-region and a new ARC variant. Technical report, University of Sherbrooke, Sherbrooke, Canada, 2015.

- [21] R. Fletcher. A model algorithm for composite nondifferentiable optimization problems. *Mathematical Programming Studies*, 17:67–76, 1982.
- [22] G. N. Grapiglia, J. Yuan, and Y. Yuan. On the convergence and worst-case complexity of trust-region and regularization methods for unconstrained optimization. *Mathematical Programming, Series A*, 152:491–520, 2015.
- [23] S. Gratton, C. W. Royer, and L. N. Vicente. A decoupled first/second-order steps technique for nonconvex nonlinear unconstrained optimization with improved complexity bounds. Technical Report TR 17-21, Department of Mathematics, University of Coimbra, Coimbra, Portugal, 2017.
- [24] S. Gratton, A. Sartenaer, and Ph. L. Toint. Recursive trust-region methods for multiscale nonlinear optimization. *SIAM Journal on Optimization*, 19(1):414–444, 2008.
- [25] S. Gratton, E. Simon, and Ph. L. Toint. An algorithm for the minimization of nonsmooth nonconvex functions using inexact evaluations and its worst-case complexity. *Mathematical Programming, Series A*, 187(1-2):1–24, 2020.
- [26] S. Gratton and Ph. L. Toint. Adaptive regularization minimization algorithms with non-smooth norms and Euclidean curvature. arXiv:2105.07765, 2021.
- [27] A. Griewank. The modification of Newton’s method for unconstrained optimization by bounding cubic terms. Technical Report NA/12, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, United Kingdom, 1981.
- [28] J. B. Hiriart-Urruty and C. Lemaréchal. *Convex analysis and minimization algorithms*. Springer Verlag, Heidelberg, Berlin, New York, 1996.
- [29] M. Knossalla. Concepts on generalized  $\epsilon$ -subdifferentials for minimizing locally Lipschitz continuous functions. *J. Nonlinear Var. Anal.*, 1(2):265–279, 2017.
- [30] A. S. Lewis and S. J. Wright. A proximal method for composite minimization. *Mathematical Programming, Series A*, 158:501–546, 2016.
- [31] J. M. Martínez. On high-order model regularization for constrained optimization. *SIAM Journal on Optimization*, 27:2447–2458, 2017.
- [32] K. G. Murty and S. N. Kabadi. Some NP-complete problems in quadratic and nonlinear programming. *Mathematical Programming*, 39(2):117–129, 1987.
- [33] Yu. Nesterov. Gradient methods for minimizing composite objective functions. *Mathematical Programming, Series A*, 140:125–161, 2013.
- [34] Yu. Nesterov and B. T. Polyak. Cubic regularization of Newton method and its global performance. *Mathematical Programming, Series A*, 108(1):177–205, 2006.
- [35] M. Nouièhed, J. Lee, and M. Razaviyayn. Convergence to second-order stationarity for constrained non-convex optimization. arXiv:1810.02024, 2018.

## Proof of Lemma 3.1

We follow the reasoning of [15, Lemma 2.1] and first recall that

$$\int_0^1 \xi(1 - \xi)^{k-1} d\xi = \frac{(k-1)!}{(k+1)!}. \quad (\text{A.1})$$

Consider now the Taylor identity

$$\psi(1) - \tau_k(1) = \frac{1}{(k-1)!} \int_0^1 (1 - \xi)^{k-1} [\psi^{(k)}(\xi) - \psi^{(k)}(0)] d\xi \quad (\text{A.2})$$

involving a given univariate  $C^k$  function  $\psi(t)$  and its  $k$ -th order Taylor expansion

$$\tau_k(t) = \sum_{i=0}^k \psi^{(i)}(0) \frac{t^i}{i!}$$

expressed in terms of the value  $\psi^{(0)} = \psi$  and  $i$ -th derivatives  $\psi^{(i)}$ ,  $i \in \{1, \dots, k\}$ . Then, picking  $\psi(t) = f(x + ts)$ , for given  $x, s \in \mathbb{R}^n$ , and  $k = p$ , the identity (A.2), and the relationships  $\psi^{(p)}(t) = \nabla_x^p f(x + ts)[s]^p$  and  $\tau_p(1) = T_p(x, s)$  give that

$$f(x + s) - T_p(x, s) = \frac{1}{(p-1)!} \int_0^1 (1-\xi)^{k-1} (\nabla_x^p f(x + \xi s) - \nabla_x^p f(x)) [s]^p d\xi,$$

and thus, using (2.2), AS.A and (A.1) for  $k = p$ , that

$$\begin{aligned} f(x + s) - T_p(x, s) &\leq \frac{1}{(p-1)!} \int_0^1 (1-\xi)^{k-1} \left| (\nabla_x^p f(x + \xi s) - \nabla_x^p f(x)) \left[ \frac{s}{\|s\|_r} \right]^p \right| \|s\|_r^p d\xi \\ &\leq \frac{1}{(p-1)!} \int_0^1 (1-\xi)^{k-1} \max_{\|v\|_r=1} |(\nabla_x^p f(x + \xi s) - \nabla_x^p f(x)) [v]^p| \|s\|_r^p d\xi \\ &= \frac{1}{(p-1)!} \int_0^1 (1-\xi)^{k-1} \|\nabla_x^p f(x + \xi s) - \nabla_x^p f(x)\|_{r,p} d\xi \cdot \|s\|_r^p \\ &\leq \frac{1}{(p-1)!} \int_0^1 \xi^\beta (1-\xi)^{p-1} d\xi \cdot L_{f,p} \|s\|_r^{p+1} = \frac{L_{f,p}}{(p+1)!} \|s\|_r^{p+1} \end{aligned}$$

for all  $x, s \in \mathbb{R}^n$ , which gives (3.1).

Likewise, for an arbitrary unit vector  $v$ , choosing  $\psi(t) = \nabla_x^j f(x + ts)[v]^j$  and  $k = p - j$ , it follows from (A.2), the relationships  $\psi^{(p-j)}(t) = \nabla_x^p f(x + ts)[v]^j [s]^{p-j}$  and  $\tau_{p-j}(1) = \nabla_s^j T_p(x, s)$  that

$$\begin{aligned} &(\nabla_x^j f(x + s) - \nabla_s^j T_p(x, s)) [v]^j \\ &= \frac{1}{(p-j-1)!} \int_0^1 (1-\xi)^{p-j-1} (\nabla_x^p f(x + \xi s) - \nabla_x^p f(x)) [v]^j [s]^{p-j} d\xi. \end{aligned} \tag{A.3}$$

Then, again using (2.2), AS.1 and (A.1) when  $k = p - j$ , we obtain that

$$\begin{aligned} &\|\nabla_x^j f(x + s) - \nabla_s^j T_p(x, s)\|_{r,j} \\ &\leq \frac{1}{(p-j-1)!} \int_0^1 (1-\xi)^{p-j-1} \left| (\nabla_x^p f(x + \xi s) - \nabla_x^p f(x)) [v]^j \left[ \frac{s}{\|s\|_r} \right]^{p-j} \right| \|s\|_r^{p-j} d\xi \\ &\leq \frac{1}{(p-j-1)!} \int_0^1 (1-\xi)^{p-j-1} \max_{\|v\|_r=1} |(\nabla_x^p f(x + \xi s) - \nabla_x^p f(x)) [v]^p| \|s\|_r^{p-j} d\xi \\ &= \frac{1}{(p-j-1)!} \int_0^1 (1-\xi)^{p-j-1} \|\nabla_x^p f(x + \xi s) - \nabla_x^p f(x)\|_{r,p} d\xi \cdot \|s\|_r^{p-j} \\ &\leq \frac{1}{(p-j-1)!} \int_0^1 \xi (1-\xi)^{p-j-1} d\xi \cdot L_{f,p} \|s\|_r^{p-j+1} = \frac{L}{(p-j+1)!} \|s\|_r^{p-j+1} \end{aligned}$$

for all  $x, s \in \mathbb{R}^n$ , which gives (3.2). □