RAL-TR-2019-001

Convergence and evaluation-complexity analysis of a regularized tensor-Newton method for solving nonlinear least-squares problems subject to convex constraints[1,2]

Nicholas I. M. Gould[3], Tyrone Rees[3] and Jennifer A. Scott[3,4]

ABSTRACT

Given a twice-continuously differentiable vector-valued function $r(x)$, a local minimizer of $\|r(x)\|_2$ within a convex set is sought. We propose and analyse tensor-Newton methods, in which $r(x)$ is replaced locally by its second-order Taylor approximation. Convergence is controlled by regularization of various orders. We establish global convergence to a constrained first-order critical point of $\|r(x)\|_2$, and provide function evaluation bounds that agree with the best-known bounds for methods using second derivatives. Numerical experiments comparing tensor-Newton methods with regularized Gauss-Newton and Newton methods demonstrate the practical performance of the newly proposed method in the unconstrained case.

---

[1] This is a highly revised and extended version of RAL-P-2017-009.

[2] This work was supported by the EPSRC grant EP/M025179/1.

[3] Scientific Computing Department, Rutherford Appleton Laboratory,
Chilton, Oxfordshire, OX11 0QX, England, EU.
Email: {nick.gould,tyrone.rees,jennifer.scott}@stfc.ac.uk .
Current reports available from "http://www.numerical.rl.ac.uk/reports/reports.shtml".

[4] Department of Mathematics and Statistics, University of Reading,
Reading, Berkshire, RG6 6AX, England, EU.
Email: jennifer.scott@reading.ac.uk .

Scientific Computing Department
STFC-Rutherford Appleton Laboratory
Oxfordshire OX11 0QX

January 8th, 2019

1 Introduction

Consider a given, smooth, vector-valued residual function $r : \mathcal{C} \longrightarrow \mathbb{R}^m$ defined on a closed, convex, non-empty set $\mathcal{C} \subseteq \mathbb{R}^n$, and let $\|\cdot\|$ be the Euclidean norm. Our goal is to design effective methods for finding values of $x \in \mathcal{C}$ for which $\|r(x)\|$ is (locally) as small as possible. Since $\|r(x)\|$ is generally non smooth, it is common to consider the equivalent problem of minimizing

$$\Phi(x) := \tfrac{1}{2}\|r(x)\|^2 \tag{1.1}$$

over $\mathcal{C}$, and to tackle the resulting problem using a generic method for convexly-constrained optimization, or one that exploits the special structure of Φ.

To put our proposal into context, arguably the most widely used method for solving nonlinear least-squares problems in $\mathbb{R}^n$ is the Gauss-Newton method and its variants. These iterative methods all build locally-linear (Taylor) approximations to $r(x_k+s)$ about x_k, and then minimize the approximation as a function of s in the least-squares sense to derive the next iterate $x_{k+1} = x_k + s_k$ [23,24,26]. The iteration is usually stabilized either by imposing a trust-region constraint on the permitted s, or by including a quadratic regularization term [3,25]. While these methods are undoubtedly popular in practice, they often suffer when the optimal value of the norm of the residual is large. To counter this, regularized Newton methods for minimizing (1.1) have also been proposed [7,18,19]. Although this usually provides a cure for the slow convergence of Gauss-Newton-like methods on non-zero-residual problems, the global behaviour is sometimes less attractive; we attribute this to the Newton model not fully reflecting the sum-of-squares nature of the original problem.

With this in mind, we consider instead the obvious nonlinear generalization of Gauss-Newton in which a locally quadratic (Taylor) "tensor-Newton" approximation to the residuals is used instead of a locally linear one. Of course, the resulting least-squares model is now quartic rather than quadratic (and thus in principle is harder to solve), but our experiments [21] have indicated that this results in more robust global behaviour than Newton-type methods and an improved performance on non-zero-residual problems than seen for Gauss-Newton variants. Our intention here is to explore the convergence behaviour of a tensor-Newton approach.

We mention in passing that we are not the first authors to consider higher-order models for least-squares problems. The earliest approach we are aware of [4,5] uses a quadratic model of $r(x_k + s)$ in which the Hessian of each residual is approximated by a low-rank matrix that is intended to compensate for any small singular values of the Jacobian. Another approach, known as geodesic acceleration [31,32], aims to modify Gauss-Newton-like steps with a correction that allows for higher-order derivatives. More recently, derivative-free methods that aim to build quadratic models of $r(x_k + s)$ by interpolation/regression of past residual values have been proposed [33,34], although these ultimately more resemble Gauss-Newton variants. While each of these methods has been shown to improve performance relative to Gauss-Newton-like approaches, none makes full use of the residual Hessians. Our intention is thus to investigate the convergence properties of methods based on the tensor-Newton model.

There has been a long-standing interest in establishing the global convergence of general smooth unconstrained optimization methods, that is, in ensuring that a method for minimizing a function $f(x)$ starting from an arbitrary initial guess ultimately delivers an iterate for which a measure of optimality is small. A more recent concern has focused on *how many* evaluations of $f(x)$ and its derivatives are necessary to reduce the optimality measure below a specified

(small) $\epsilon > 0$ from the initial guess. If the measure is $\|g(x)\|$, where $g(x) := \nabla_x f(x)$, it is known that some well-known schemes (including steepest descent and generic second-order trust-region methods) may require $\Theta(\epsilon^{-2})$ evaluations under standard assumptions [6], while this may be improved to $\Theta(\epsilon^{-3/2})$ evaluations for second-order methods with cubic regularization or using specialised trust-region tools [8, 17, 28]. Here and hereafter $O(\cdot)$ indicates a term that is of at worst a multiple of its argument, while $\Theta(\cdot)$ indicates additionally there are instances for which the bound holds.

For the problem we consider here, an obvious approach in the unconstrained case is to apply any of the aforementioned algorithms to minimize (1.1), and to terminate as soon as

$$\|\nabla_x \Phi(x)\| \leq \epsilon, \quad \text{where } \nabla_x \Phi(x) = J^T(x) r(x) \text{ and } J(x) := \nabla_x r(x). \tag{1.2}$$

However, it has been argued [9] that this ignores the possibility that it may suffice to stop instead when $r(x)$ is small, and that a more sensible criterion is to terminate when

$$\|r(x)\| \leq \epsilon_p \quad \text{or} \quad \|g_r(x)\| \leq \epsilon_d, \tag{1.3}$$

where $\epsilon_p > 0$ and $\epsilon_d > 0$ are required accuracy tolerances and $g_r(x)$ is the scaled gradient given by

$$g_r(x) := \begin{cases} \dfrac{J^T(x) r(x)}{\|r(x)\|}, & \text{whenever } r(x) \neq 0; \\ 0, & \text{otherwise.} \end{cases} \tag{1.4}$$

We note that $g_r(x)$ in (1.4) is precisely the gradient of $\|r(x)\|$ whenever $r(x) \neq 0$, while if $r(x) = 0$, we are at the global minimum of r and so $g_r(x) = 0 \in \partial(\|r(x)\|)$, the sub-differential of $r(x)$. Furthermore $\|g_r(x)\|$ is less sensitive to scaling than $\|J^T(x) r(x)\|$. It has been shown that a second-order method based on cubic regularization will satisfy (1.3) after $O\left(\max(\epsilon_d^{-3/2}, \epsilon_p^{-1/2})\right)$ evaluations [9, Theorem 3.2]. One of our aims here is to show similar bounds for the tensor-Newton method we are advocating.

Since it is easy to do so, we consider the more general problem in which the variables are constrained to lie in $\mathcal{C}$. To this effect, let $P_{\mathcal{C}}[x]$ be the orthogonal projection of x onto $\mathcal{C}$, $\langle \cdot, \cdot \rangle$ denote the Euclidean inner product, and $\| \cdot \|_\chi$ be an arbitrary norm, We then consider the (continuous) first-order criticality measures [14, 15, Thm.12.1.6]

$$\pi_h(x) := \|P_{\mathcal{C}}[x - \nabla_x h(x)] - x\|, \tag{1.5}$$

and

$$\chi_h(x) := \left| \min_{x+d \in \mathcal{C}, \|d\|_\chi \leq 1} \langle \nabla_x h(x), d \rangle \right|, \tag{1.6}$$

of a given $x \in \mathcal{C}$ and an arbitrary function h defined over $\mathcal{C}$. Trivially $\pi_h(x) = \chi_h(x) = \|\nabla_x h(x)\|$ in the unconstrained case $\mathcal{C} = \mathbb{R}^n$ when $\| \cdot \|_\chi = \| \cdot \|$. Based on this, we generalize (1.2)–(1.4) using (1.5) so that, for $x \in \mathcal{C}$, either

$$\pi_\Phi(x) \leq \epsilon \tag{1.7}$$

or preferably

$$\|r(x)\| \leq \epsilon_p \quad \text{or} \quad \psi_r(x) \leq \epsilon_d, \tag{1.8}$$

using the scaled projected gradient

$$\psi_r(x) := \begin{cases} \dfrac{\pi_\Phi(x)}{\|r(x)\|}, & \text{whenever } r(x) \neq 0; \\ 0, & \text{otherwise.} \end{cases} \tag{1.9}$$

The obvious alternatives using the measure (1.6) are to find $x \in \mathcal{C}$ so that either

$$\chi_\Phi(x) \leq \epsilon \tag{1.10}$$

or preferably

$$\|r(x)\| \leq \epsilon_p \quad \text{or} \quad \zeta_r(x) \leq \epsilon_d, \tag{1.11}$$

using the scaled conditional gradient

$$\zeta_r(x) := \begin{cases} \dfrac{\chi_\Phi(x)}{\|r(x)\|}, & \text{whenever } r(x) \neq 0, \text{ or} \\ 0, & \text{otherwise.} \end{cases} \tag{1.12}$$

We propose a regularized tensor-Newton method in §2, and analyse both its global convergence and its evaluation complexity in §3; we focus on (1.7)–(1.9) in the latter, but consider the alternative (1.10)–(1.12) in §3.1. The theory in §3 only holds when the regularization order lies between 2 and 3, and so in §4 we introduce a modified algorithm for which $r > 3$ is possible. In §5 we illustrate the performance of our tensor-Newton approach when applied to unconstrained test examples. We make further comments and draw general conclusions in §6.

2 The tensor-Newton method

Suppose that $r(x) \in C^2$ has components $r_i(x)$ for $i = 1, \ldots, m$. Let $t(x, s)$ be the vector whose components are

$$t_i(x, s) := r_i(x) + s^T \nabla_x r_i(x) + \tfrac{1}{2} s^T \nabla_{xx} r_i(x) s \tag{2.1}$$

for $i = 1, \ldots, m$. We build the *tensor-Newton* approximation

$$m(x, s) := \tfrac{1}{2} \|t(x, s)\|^2 \tag{2.2}$$

of $\Phi(x + s)$, and define the regularized model

$$m^{\mathrm{R}}(x, s, \sigma) := m(x, s) + \frac{1}{r} \sigma \|s\|^r, \tag{2.3}$$

as well as the model criticality measure

$$\pi_m^{\mathrm{R}}(x, s, \sigma) := \|P_{\mathcal{C}}[x + s - \nabla_s m^{\mathrm{R}}(x, s, \sigma)] - x - s\|, \tag{2.4}$$

where $r \geq 2$ is given, and

$$\nabla_s m^{\mathrm{R}}(x, s, \sigma) = \nabla_s m(x, s) + \sigma \|s\|^{r-2} s. \tag{2.5}$$

We consider the following algorithm (Algorithm 2.1 on the following page) to find a critical point of $\Phi(x)$ within $\mathcal{C}$ using the measure (1.5).

Algorithm 2.1: Adaptive Tensor-Newton Regularization.

A starting point $x_0 \in \mathcal{C}$, an initial and a minimal regularization parameter $\sigma_0 \geq \sigma_{\min} > 0$ and algorithmic parameters $\theta > 0$, $\gamma_3 \geq \gamma_2 > 1 > \gamma_1 > 0$ and $1 > \eta_2 \geq \eta_1 > 0$, are given. Evaluate $\Phi(x_0)$. For $k = 0, 1, \ldots$, until **termination**, do:

1. If the termination test has not been satisfied, compute derivatives of $r(x)$ at x_k.

2. Compute a step s_k by approximately minimizing $m^{\rm R}(x_k, s, \sigma_k)$ within $\mathcal{C}$ so that

$$x_k + s_k \in \mathcal{C}, \tag{2.6}$$

$$m^{\rm R}(x_k, s_k, \sigma_k) < m^{\rm R}(x_k, 0, \sigma_k) \tag{2.7}$$

and

$$\pi^{\rm R}_m(x_k, s_k, \sigma_k) \leq \theta \|s_k\|^{r-1}. \tag{2.8}$$

3. Compute $\Phi(x_k + s_k)$ and

$$\rho_k = \frac{\Phi(x_k) - \Phi(x_k + s_k)}{m(x_k, 0) - m(x_k, s_k)}. \tag{2.9}$$

If $\rho_k \geq \eta_1$, set $x_{k+1} = x_k + s_k$. Otherwise set $x_{k+1} = x_k$.

4. Set

$$\sigma_{k+1} \in \begin{cases} [\max(\sigma_{\min}, \gamma_1\sigma_k), \sigma_k] & \text{if } \rho_k \geq \eta_2 & \text{[very successful iteration]} \\ [\sigma_k, \gamma_2\sigma_k] & \text{if } \eta_1 \leq \rho_k < \eta_2 & \text{[successful iteration]} \\ [\gamma_2\sigma_k, \gamma_3\sigma_k] & \text{otherwise} & \text{[unsuccessful iteration]}, \end{cases} \tag{2.10}$$

and go to Step 2 if $\rho_k < \eta_1$.

At the very least, we insist that (trivial) termination should occur in Step 1 of Algorithm 2.1 if $\psi_\Phi(x_k) = 0$, but in practice a rule such as (1.7) or (1.8) at $x = x_k$ will be preferred. We note that by construction all iterates are feasible, i.e., $x_k \in \mathcal{C}$ for all $k \geq 0$.

At the heart of Algorithm 2.1 is the need (Step 2) to find a feasible vector s_k that both reduces $m^{\rm R}(x_k, s, \sigma_k)$ and satisfies $\|\nabla_s m^{\rm R}(x_k, s_k, \sigma_k)\| \leq \theta \|s_k\|^{r-1}$ (see, e.g., [1]). Since $m^{\rm R}(x_k, s, \sigma_k)$ is bounded from below (and grows as s approaches infinity), we may apply any descent-based local optimization method that is designed to find a constrained critical point of $m^{\rm R}(x_k, s, \sigma_k)$, starting from $s = 0$, as this will generate an s_k that is guaranteed to satisfy both Step 2 stopping requirements. Crucially, such a minimization is on the model $m^{\rm R}(x_k, s, \sigma_k)$, not the true objective, and thus involves no true objective evaluations. We do not claim that this calculation is trivial, but it might, for example, be achieved by applying a safeguarded projection-based Gauss-Newton method to the least-squares problem involving the extended residuals $(t(x_k, s), \sqrt{\sigma_k \|s\|^{r-2}} s)$.

We define the index set of successful iterations, in the sense of (2.10), up to iteration k to be $\mathcal{S}_k := \{0 \leq l \leq k \mid \rho_l \geq \eta_1\}$ and let $\mathcal{S} := \{k \geq 0 \mid \rho_k \geq \eta_1\}$ be the set of all successful iterations.

3 Convergence analysis

We make the following blanket assumption:

A1 each component $r_i(x)$ and its first two derivatives are Lipschitz continuous on an open set containing the intervals $[x_k, x_k + s_k]$ generated by Algorithm 2.1 (or its successor).

It has been shown [10, Lemma 3.1] that AS.1 implies that $\Phi(x)$ and its first two derivatives are Lipschitz on $[x_k, x_k + s_k]$.

We define

$$H(x, y) := \sum_{i=1}^{m} y_i \nabla_{xx} r_i(x)$$

and let $q(x, s)$ be the vector whose ith component is

$$q_i(x, s) := s^T \nabla_{xx} r_i(x) s$$

for $i = 1, \ldots, m$. In this case

$$t(x, s) = r(x) + J(x)s + \tfrac{1}{2} q(x, s).$$

Since $m(x_k, s)$ is a second-order accurate model of $\Phi(x_k + s)$, we expect bounds of the form

$$|\Phi(x_k + s_k) - m(x_k, s_k)| \leq L_{\mathrm{f}} \|s_k\|^3 \tag{3.1}$$

and

$$|\nabla_x \Phi(x_k + s_k) - \nabla_s m(x_k, s_k)| \leq L_{\mathrm{g}} \|s_k\|^2 \tag{3.2}$$

for some $L_{\mathrm{f}} > 0$ and (without loss of generality) $L_{\mathrm{g}} \geq 1$ and all $k \geq 0$ for which $\|s_k\| \leq 1$ (see Appendix A).

Also, since $\|r(x)\|$ decreases monotonically,

$$\|J^T(x_k) r(x_k)\| \leq \|J^T(x_k)\| \|r(x_k)\| \leq L_{\mathrm{J}} \|r(x_0)\| \tag{3.3}$$

and

$$\|H(x_k, r(x_k))\| \leq L_{\mathrm{H}} \|r(x_k)\| \leq L_{\mathrm{H}} \|r(x_0)\| \tag{3.4}$$

for some $L_{\mathrm{J}}, L_{\mathrm{H}} > 0$ and all $k \geq 0$ (again, see Appendix A).

Our first result derives simple conclusions from the basic requirement that the step s_k in our algorithm is chosen to reduce the regularized model.

Lemma 3.1. Algorithm 2.1 ensures that

$$m(x_k,0) - m(x_k,s_k) > \tfrac{1}{r}\sigma_k\|s_k\|^r \tag{3.5}$$

In addition, if $r = 2$, at least one of

$$\sigma_k < 2\|H(x_k, r(x_k))\| \tag{3.6}$$

or

$$\sigma_k\|s_k\| < 4\|J^T(x_k)r(x_k)\| \tag{3.7}$$

holds, while if $r > 2$,

$$\|s_k\| < \max\left(\left(\frac{r\|H(x_k,r(x_k))\|}{\sigma_k}\right)^{1/(r-1)}, \left(\frac{2r\|J^T(x_k)r(x_k)\|}{\sigma_k}\right)^{1/(r-2)}\right). \tag{3.8}$$

Proof. It follows from (2.7), (2.3) and (2.2) that

$$\begin{aligned}
0 &> 2\left(m(x_k,s_k) + \tfrac{1}{r}\sigma_k\|s_k\|^r - m(x_k,0)\right)\\
&= \|r(x_k) + J(x_k)s_k + \tfrac{1}{2}q(x_k,s_k)\|^2 + \tfrac{2}{r}\sigma_k\|s_k\|^r - \|r(x_k)\|^2\\
&= \|J(x_k)s_k + \tfrac{1}{2}q(x_k,s_k)\|^2 + 2r^T(x_k)\left(J(x_k)s_k + \tfrac{1}{2}q(x_k,s_k)\right) + \tfrac{2}{r}\sigma_k\|s_k\|^r\\
&= \|J(x_k)s_k + \tfrac{1}{2}q(x_k,s_k)\|^2 + 2s_k^TJ^T(x_k)r(x_k) + s_k^TH(x_k,r(x_k))s_k + \tfrac{2}{r}\sigma_k\|s_k\|^r\\
&\geq \|J(x_k)s_k + \tfrac{1}{2}q(x_k,s_k)\|^2 - 2\|J^T(x_k)r(x_k)\|\|s_k\| - \|H(x_k,r(x_k))\|\|s_k\|^2 + \tfrac{2}{r}\sigma_k\|s_k\|^r.
\end{aligned} \tag{3.9}$$

Inequality (3.5) follows immediately from the first inequality in (3.9). When $r = 2$, inequality (3.9) becomes

$$\begin{aligned}
0 > \|J(x_k)s_k + \tfrac{1}{2}q(x_k,s_k)\|^2 +\\
\left(\tfrac{1}{2}\sigma_k\|s_k\| - 2\|J^T(x_k)r(x_k)\|\right)\|s_k\| + \left(\tfrac{1}{2}\sigma_k - \|H(x_k,r(x_k))\|\right)\|s_k\|^2.
\end{aligned}$$

In order for this to be true, it must be that at least one of the last two terms is negative, and this provides the alternatives (3.6) and (3.7). By contrast, when $r > 2$, inequality (3.9) becomes

$$\begin{aligned}
0 > \|J(x_k)s_k + \tfrac{1}{2}q(x_k,s_k)\|^2 +\\
\left(\tfrac{1}{r}\sigma_k\|s_k\|^{r-1} - 2\|J^T(x_k)r(x_k)\|\right)\|s_k\| + \left(\tfrac{1}{r}\sigma_k\|s_k\|^{r-2} - \|H(x_k,r(x_k))\|\right)\|s_k\|^2,
\end{aligned}$$

and this implies that

$$\tfrac{1}{r}\sigma_k\|s_k\|^{r-1} < 2\|J^T(x_k)r(x_k)\| \quad\text{or}\quad \tfrac{1}{r}\sigma_k\|s_k\|^{r-2} < \|H(x_k,r(x_k))\|$$

(or both), which gives (3.8). □

Our next task is to show that σ_k is bounded from above. Let

$$\mathcal{B}_\gamma := \left\{j \geq 0 \mid \sigma_j \geq \gamma r \max\left(\|H\left(x_j, r(x_j)\right)\|, 2\|J^T(x_j)r(x_j)\|\right)\right\}$$

and

$$\mathcal{B} := \mathcal{B}_1,$$

and note that Lemma 3.1 implies that

$$\|s_k\| \leq 1 \text{ if } k \in \mathcal{B}_\gamma \text{ when } \gamma \geq 1,$$

and in particular

$$\|s_k\| \leq 1 \text{ for all } k \in \mathcal{B}. \tag{3.10}$$

We consider first the special case for which $r = 2$.

Lemma 3.2. Suppose that AS.1 holds, $r = 2$, $k \in \mathcal{B}$ and

$$\sigma_k \geq \sqrt{\frac{8L_{\mathrm{f}}L_{\mathrm{J}}\|r(x_0)\|}{1-\eta_2}}. \tag{3.11}$$

Then iteration k of Algorithm 2.1 is very successful.

Proof. Since $k \in \mathcal{B}$, Lemma 3.1 implies that (3.7) and (3.10) hold. Then (2.9), (3.1) and (3.5) give that

$$|\rho_k - 1| = \frac{|\Phi(x_k + s_k) - m(x_k, s_k)|}{m(x_k, 0) - m(x_k, s_k)} \leq \frac{2L_{\mathrm{f}}\|s_k\|}{\sigma_k}$$

and hence

$$|\rho_k - 1| \leq \frac{8L_{\mathrm{f}}\|J^T(x_k)r(x_k)\|}{\sigma_k^2} \leq \frac{8L_{\mathrm{f}}L_{\mathrm{J}}\|r(x_0)\|}{\sigma_k^2} \leq 1 - \eta_2$$

from (3.3), (3.7) and (3.11). Thus it follows from (2.10) that the iteration is very successful. □

Lemma 3.3. Suppose that AS.1 holds and $r = 2$. Then Algorithm 2.1 ensures that

$$\sigma_k \leq \sigma_{\max} := \gamma_3 \max\left(\sqrt{\frac{8L_{\mathrm{f}}L_{\mathrm{J}}\|r(x_0)\|}{1-\eta_2}}, \sigma_0, 2\max(L_{\mathrm{H}}, 2L_{\mathrm{J}})\|r(x_0)\|\right) \tag{3.12}$$

for all $k \geq 0$.

Proof. Let

$$\sigma_{\max}^{\mathrm{B}} = \gamma_3 \max\left(\sqrt{\frac{8L_{\mathrm{f}}L_{\mathrm{J}}\|r(x_0)\|}{1-\eta_2}}, \sigma_0\right).$$

Suppose that $k+1 \in \mathcal{B}_{\gamma_3}$ is the first iteration for which $\sigma_{k+1} \geq \sigma_{\max}^{\mathrm{B}}$. Then, since $\sigma_k < \sigma_{k+1}$, iteration k must have been unsuccessful, $x_k = x_{k+1}$ and (2.10) gives that $\sigma_{k+1} \leq \gamma_3\sigma_k$. Thus

$$\begin{aligned}\gamma_3\sigma_k &\geq \sigma_{k+1} \geq 2\gamma_3 \max(\|H(x_{k+1}, r(x_{k+1}))\, 2\|J^T(x_{k+1})r(x_{k+1})\|) \\ &= 2\gamma_3 \max(\|H(x_k, r(x_k)), 2\|J^T(x_k)r(x_k)\|)\end{aligned}$$

since $k+1 \in \mathcal{B}_{\gamma_3}$, which implies that $k \in \mathcal{B}$. Furthermore,

$$\gamma_3 \sigma_k \geq \sigma_{k+1} \geq \sigma_{\max}^{\mathrm{B}} \geq \gamma_3 \sqrt{\frac{8 L_{\mathrm{f}} L_{\mathrm{J}} \|r(x_0)\|}{1-\eta_2}},$$

which implies that (3.11) holds. But then Lemma 3.2 implies that iteration k must be very successful. This contradiction ensures that

$$\sigma_k < \sigma_{\max}^{\mathrm{B}} \tag{3.13}$$

for all $k \in \mathcal{B}_{\gamma_3}$. For all other iterations, we have that $k \notin \mathcal{B}_{\gamma_3}$, and for these the definition of $\mathcal{B}_{\gamma_3}$, and the bounds (3.3) and (3.4) give

$$\sigma_k < 2\gamma_3 \max(\|H(x_k, r(x_k))\|, 2\|J^T(x_k) r(x_k))\|) \leq 2\gamma_3 \max(L_{\mathrm{H}}, 2L_{\mathrm{J}}) \|r(x_0)\|. \tag{3.14}$$

Combining (3.13) and (3.14) gives (3.12). □

We now turn to the general case for which $2 < r \leq 3$.

Lemma 3.4. Suppose that AS.1 holds, $2 < r \leq 3$, $k \in \mathcal{B}$ and

$$\sigma_k \geq \max\left(\left(\frac{r L_{\mathrm{f}} (r L_{\mathrm{H}} \|r(x_0)\|)^{\frac{3-r}{r-1}}}{1-\eta_2} \right)^{\frac{r-1}{2}}, \left(\frac{r L_{\mathrm{f}} (2r L_{\mathrm{J}} \|r(x_0)\|)^{\frac{3-r}{r-2}}}{1-\eta_2} \right)^{r-2} \right) \tag{3.15}$$

Then iteration k of Algorithm 2.1 is very successful.

Proof. Since $k \in \mathcal{B}$, it follows from (2.9), (3.10), (3.1), (3.5), (3.8), (3.3), (3.4) and (3.15) that

$$\begin{aligned} |\rho_k - 1| = & \frac{|\Phi(x_k + s_k) - m(x_k, s_k)|}{m(x_k, 0) - m(x_k, s_k)} \leq \frac{r L_{\mathrm{f}} \|s_k\|^{3-r}}{\sigma_k} \\ < & r L_{\mathrm{f}} \max\Big((r\|H(x_k, r(x_k))\|)^{(3-r)/(r-1)} \sigma_k^{-2/(r-1)}, \\ & \qquad \big(2r\|J^T(x_k) r(x_k)\|\big)^{(3-r)/(r-2)} \sigma_k^{-1/(r-2)} \Big) \\ \leq & r L_{\mathrm{f}} \max\Big((r L_{\mathrm{H}} \|r(x_0)\|)^{(3-r)/(r-1)} \sigma_k^{-2/(r-1)}, \\ & \qquad (2r L_{\mathrm{J}} \|r(x_0)\|)^{(3-r)/(r-2)} \sigma_k^{-1/(r-2)} \Big) \\ \leq & 1 - \eta_2. \end{aligned} \tag{3.16}$$

As before, (2.10) then ensures that the iteration is very successful. □

Lemma 3.5. Suppose that AS.1 holds and $2 < r \leq 3$. Then Algorithm 2.1 ensures that

$$\sigma_k \leq \sigma_{\max} := \gamma_3 \max\left(\left(\frac{rL_{\mathrm{f}}\left(pL_{\mathrm{H}}\|r(x_0)\|\right)^{\frac{3-r}{r-1}}}{1-\eta_2}\right)^{\frac{r-1}{2}}, \left(\frac{rL_{\mathrm{f}}\left(2rL_{\mathrm{J}}\|r(x_0)\|\right)^{\frac{3-r}{r-2}}}{1-\eta_2}\right)^{r-2}, \sigma_0, r\max(L_{\mathrm{H}}, 2L_{\mathrm{J}})\|r(x_0)\|\right) \tag{3.17}$$

for all $k \geq 0$.

Proof. The proof mimics that of Lemma 3.3. First, suppose that $k \in \mathcal{B}_{\gamma_3}$ and that iteration $k+1$ is the first for which

$$\sigma_{k+1} \geq \sigma_{\max}^{\mathrm{B}} := \gamma_3 \max\left(\left(\frac{rL_{\mathrm{f}}\left(rL_{\mathrm{H}}\|r(x_0)\|\right)^{\frac{3-r}{r-1}}}{1-\eta_2}\right)^{\frac{r-1}{2}}, \left(\frac{rL_{\mathrm{f}}\left(2rL_{\mathrm{J}}\|r(x_0)\|\right)^{\frac{3-r}{r-2}}}{1-\eta_2}\right)^{r-2}, \sigma_0\right).$$

Then, since $\sigma_k < \sigma_{k+1}$, iteration k must have been unsuccessful and (2.10) gives that

$$\gamma_3\sigma_k \geq \sigma_{k+1} \geq \sigma_{\max}^{\mathrm{B}},$$

which implies that $k \in \mathcal{B}$ and (3.15) holds. But then Lemma 3.4 implies that iteration k must be very successful. This contradiction provides the first three terms in the bound (3.17), while the others arise as for the proof of Lemma 3.3 when $k \notin \mathcal{B}_{\gamma_3}$. □

Next, we bound the number of iterations in terms of the number of successful ones.

Lemma 3.6. [8, Theorem 2.1]. The adjustment (2.10) in Algorithm 2.1 ensures that

$$k \leq \kappa_u|\mathcal{S}_k| + \kappa_s, \quad \text{where } \kappa_u := \left(1 - \frac{\log\gamma_1}{\log\gamma_2}\right), \kappa_s := \frac{1}{\log\gamma_2}\log\left(\frac{\sigma_{\max}}{\sigma_0}\right), \tag{3.18}$$

and $\sigma_{\max}$ is any known upper bound on σ_k.

Our final ingredient is to find a useful bound on the smallest model decrease as the algorithm proceeds. Let $\mathcal{L} := \{k \mid \|s_k\| \leq 1\}$, and let $\mathcal{G} := \{k \mid \|s_k\| > 1\}$ be its compliment. We then have the following crucial bounds.

Lemma 3.7. Suppose that AS.1 holds and $2 \leq r \leq 3$. Then Algorithm 2.1 ensures that

$$m(x_k, 0) - m(x_k, s_k) \geq \begin{cases} \frac{1}{r}\sigma_{\min}\left(\frac{\pi_\Phi(x_k+s_k)}{L_{\rm g}+\theta+\sigma_{\max}}\right)^{\frac{r}{r-1}} & \text{if } k \in \mathcal{L} \\ \frac{1}{r}\sigma_{\min} & \text{if } k \in \mathcal{G}, \end{cases} \tag{3.19}$$

where $\sigma_{\max}$ is given by (3.12) when $r = 2$ and by (3.17) for $2 < r \leq 3$.

Proof. Using, in order, the definition (1.7), the Cauchy-Schwarz inequality, the contractive property of projection operators in Hilbert spaces, the definition (2.4) and identity (2.5), and the Cauchy-Schwarz inequality again, we find that

$$\begin{aligned} \pi_\Phi(x_k+s_k) &= \|P_{\mathcal{C}}[x_k+s_k-\nabla_x\Phi(x_k+s_k)] - x_k - s_k\| \\ &= \|P_{\mathcal{C}}[x_k+s_k-\nabla_x\Phi(x_k+s_k)] - P_{\mathcal{C}}[x_k+s_k-\nabla_s m^{\rm R}(x_k,s_k,\sigma_k)]+ \\ &\quad P_{\mathcal{C}}[x_k+s_k-\nabla_s m^{\rm R}(x_k,s_k,\sigma_k)] - x_k - s_k\| \\ &\leq \|P_{\mathcal{C}}[x_k+s_k-\nabla_x\Phi(x_k+s_k)] - P_{\mathcal{C}}[x_k+s_k-\nabla_s m^{\rm R}(x_k,s_k,\sigma_k)]\|+ \\ &\quad \|P_{\mathcal{C}}[x_k+s_k-\nabla_s m^{\rm R}(x_k,s_k,\sigma_k)] - x_k - s_k\| \\ &\leq \|[x_k+s_k-\nabla_x\Phi(x_k+s_k)] - [x_k+s_k-\nabla_s m^{\rm R}(x_k,s_k,\sigma_k)]\|+ \\ &\quad \pi^{\rm R}_m(x_k,s_k,\sigma_k) \\ &= \|\nabla_x\Phi(x_k+s_k) - \nabla_s m^{\rm R}(x_k,s_k,\sigma_k)\| + \pi^{\rm R}_m(x_k,s_k,\sigma_k) \\ &= \|\nabla_x\Phi(x_k+s_k) - \nabla_s m(x_k,s_k) - \sigma_k\|s_k\|^{r-2}s\| + \pi^{\rm R}_m(x_k,s_k,\sigma_k) \\ &\leq \|\nabla_x\Phi(x_k+s_k) - \nabla_s m(x_k,s_k)\| + \sigma_k\|s_k\|^{r-1} + \pi^{\rm R}_m(x_k,s_k,\sigma_k). \end{aligned} \tag{3.20}$$

Consider $k \in \mathcal{L}$. Combining (3.20) with (3.2), (2.8), (3.12), (3.17) and $\|s_k\| \leq 1$ we have

$$\pi_\Phi(x_k+s_k) \leq L_{\rm g}\|s_k\|^2 + \theta\|s_k\|^{r-1} + \sigma_{\max}\|s_k\|^{r-1} \leq (L_{\rm g}+\theta+\sigma_{\max})\|s_k\|^{r-1}$$

and thus that

$$\|s_k\| \geq \left(\frac{\pi_\Phi(x_k+s_k)}{L_{\rm g}+\theta+\sigma_{\max}}\right)^{\frac{1}{r-1}}.$$

But then, combining this with (3.5), the lower bound

$$\sigma_k \geq \sigma_{\min} \tag{3.21}$$

imposed by Algorithm 2.1 and (3.5) provides the first possibility in (3.19).

By contrast, if $k \in \mathcal{G}$, (3.5), $\|s_k\| > 1$ and (3.21) ensure the second possibility in (3.19). □

Corollary 3.8. Suppose that AS.1 holds and $2 \leq r \leq 3$. Then Algorithm 2.1 ensures that

$$\Phi(x_k) - \Phi(x_{k+1}) \geq \begin{cases} \frac{1}{r}\eta_1\sigma_{\min}\left(\frac{\pi_\Phi(x_{k+1})}{L_{\rm g}+\theta+\sigma_{\max}}\right)^{\frac{r}{r-1}} & \text{if } k \in \mathcal{L}\cap\mathcal{S} \\ \frac{1}{r}\eta_1\sigma_{\min} & \text{if } k \in \mathcal{G}\cap\mathcal{S}, \end{cases} \tag{3.22}$$

where $\sigma_{\max}$ is given by (3.12) when $r = 2$ and by (3.17) for $2 < r \leq 3$.

Proof. The result follows directly from and (2.9) and (3.19). □

We now provide our three main convergence results. Firstly, we establish the global convergence[1] of our algorithm to first-order critical points of $\Phi(x)$.

Theorem 3.9. Suppose that AS.1 holds and $2 \leq r \leq 3$. Then the iterates $\{x_k\}$ generated by Algorithm 2.1 satisfy

$$\lim_{k \to \infty} \pi_\Phi(x_k) = 0 \tag{3.23}$$

if no non-trivial termination test is provided.

Proof. Suppose that $\epsilon > 0$, and consider any successful iteration for which

$$\pi_\Phi(x_k) \geq \epsilon > 0. \tag{3.24}$$

Then it follows from (3.22) that

$$\Phi(x_k) - \Phi(x_{k+1}) \geq \delta := \frac{\eta_1 \sigma_{\min}}{r} \min\left(\left(\frac{\epsilon}{L_{\mathrm{g}} + \theta + \sigma_{\max}} \right)^{\frac{r}{r-1}}, 1 \right) > 0. \tag{3.25}$$

Consider the set $\mathcal{U}_\epsilon = \{k \in \mathcal{S} \mid \|\nabla_x \Phi(x_k)\| \geq \epsilon\}$, suppose that $\mathcal{U}_\epsilon$ is infinite, and let k_i be the i-th entry of $\mathcal{U}_\epsilon$. Now consider

$$i_\epsilon = \lceil \tfrac{1}{2} \|r(x_0)\|^2 / \delta \rceil + 1.$$

Thus summing (3.25) over successful iterations, recalling that $\Phi(x_0) = \frac{1}{2}\|r(x_0)\|^2$, $\Phi(x_k) \geq 0$, and that Φ decreases monotonically and using (3.25), we have that

$$\tfrac{1}{2}\|r(x_0)\|^2 \geq \Phi(x_0) - \Phi(x_{k_{i_\epsilon}+1}) \geq \sum_{k \in \mathcal{U}_\epsilon, k \leq k_{i_\epsilon}} \Phi(x_k) - \Phi(x_{k+1}) \geq i_\epsilon \delta > \tfrac{1}{2}\|r(x_0)\|^2. \tag{3.26}$$

This contradiction shows that $\mathcal{U}_\epsilon$ is finite for any $\epsilon > 0$, and therefore (3.23) holds. □

Secondly, we provide an evaluation complexity result based on the stopping criterion (1.7).

Theorem 3.10. Suppose that AS.1 holds and $2 \leq r \leq 3$. Then Algorithm 2.1 requires at most $\kappa_u \#_{\mathrm{S}} + \kappa_s + 1$ evaluations of $r(x)$ and $\#_{\mathrm{S}} + 1$ evaluations of its derivatives, involving

$$\#_{\mathrm{S}} := \left\lceil \frac{\|r(x_0)\|^2 r \left(L_{\mathrm{g}} + \theta + \sigma_{\max}\right)^{\frac{r}{r-1}}}{2\eta_1 \sigma_{\min}} \epsilon^{-\frac{r}{r-1}}. \right\rceil$$

successful iterations, to find an iterate x_k for which the termination test

$$\pi_\Phi(x_k) \leq \epsilon$$

is satisfied for given $0 < \epsilon \leq 1$, where κ_u and κ_s are defined in (3.18) and $\sigma_{\max}$ is given by (3.12) when $r = 2$ and by (3.17) for $2 < r \leq 3$.

[1] Our proof avoids the traditional route via a lim inf result, and is indebted to [16].

Proof. If the algorithm has not terminated, (3.24) holds, so summing (3.25) as before

$$\tfrac{1}{2}\|r(x_0)\|^2 \geq \Phi(x_0) - \Phi(x_{k+1}) \geq |\mathcal{S}_k|\delta = |\mathcal{S}_k|\frac{\eta_1\sigma_{\min}}{r}\left(\frac{\epsilon}{L_{\rm g}+\theta+\sigma_{\max}}\right)^{\frac{r}{r-1}} \tag{3.27}$$

since $\epsilon \leq 1$ and $L_{\rm g} \geq 1$, and thus $|\mathcal{S}_k| \leq \#_{\rm S}$. Combining this with (3.18) and remembering that we need to evaluate the function and gradient at the final x_{k+1} yields the required evaluation bounds. □

Notice how the evaluation complexity improves from $O(\epsilon^{-2})$ evaluations with quadratic ($r = 2$) regularization to $O(\epsilon^{-3/2})$ evaluations with cubic ($r = 3$) regularization. It is not clear if these bounds are sharp.

Finally, we refine this analysis to provide an alternative complexity result based on the stopping rule (1.8). The proof of this follows similar arguments in [9, §3.2; 11, §3] and crucially depends upon the following elementary result.

Lemma 3.11. Suppose that $a > b \geq 0$. Then

$$a^2 - b^2 \geq c \;\text{ implies that }\; a^{1/2^i} - b^{1/2^i} \geq \frac{c}{2^{i+1}a^{\frac{2^{i+1}-1}{2^i}}}$$

for all integers $i \geq -1$.

Proof. The result follows directly by induction using the identity $A^2 - B^2 = (A-B)(A+B)$ with $A = a^{1/2^j} > B = b^{1/2^j}$ for increasing $j \leq i$. □

Theorem 3.12. Suppose that AS.1 holds, $2 < r \leq 3$ and that the integer

$$i \geq i_0 := \left\lceil \log_2\left(\frac{r-1}{r-2}\right)\right\rceil \tag{3.28}$$

is given. Then Algorithm 2.1 requires at most $\kappa_u\#_{\rm S} + \kappa_s + 1$ evaluations of $r(x)$ and $\#_{\rm S} + 1$ evaluations of its derivatives, involving

$$\#_{\rm S} := \left\lceil \max\left(\kappa_c^{-1}, \kappa_g^{-1}\epsilon_d^{-r/(r-1)}, \kappa_r^{-1}\epsilon_p^{-1/2^i}\right)\right\rceil \tag{3.29}$$

successful iterations, to find an iterate x_k for which the termination test

$$\|r(x_k)\| \leq \epsilon_p \quad \text{or} \quad \psi_r(x_k) \leq \epsilon_d, \tag{3.30}$$

is satisfied for given $\epsilon_p > 0$ and $\epsilon_d > 0$, where κ_u and κ_s are defined in (3.18), κ_c, κ_g and κ_r

are given by

$$\begin{aligned} \kappa_c &:= \tfrac{1}{2}^{i+1} \frac{\eta_1 \sigma_{\min}}{r} \|r(x_0)\|^{-(2^{i+1}-1)/2^i}, \\ \kappa_g &:= \frac{\tfrac{1}{2}^i \eta_1 \sigma_{\min} \beta^{r/(r-1)}}{r(L+\theta+\sigma_{\max})^{r/(r-1)}} \|r(x_0)\|^{\left(r/(r-1)-(2^{i+1}-1)/2^i\right)} \\ \text{and } \kappa_r &:= \frac{1-\beta^{1/2^i}}{\beta^{1/2^i}}, \end{aligned} \tag{3.31}$$

$\sigma_{\max}$ is given by (3.17), and $\beta \in (0,1)$ is a fixed problem-independent constant.

Proof. Consider

$$\mathcal{S}_\beta := \{l \in \mathcal{S} \mid \|r(x_{l+1})\| > \beta \|r(x_l)\|\}, \tag{3.32}$$

and let i be the smallest integer for which

$$\frac{2^{i+1}-1}{2^i} \geq \frac{r}{r-1}, \tag{3.33}$$

that is i satisfies (3.28).

First, consider $l \in \mathcal{G} \cap \mathcal{S}$. Then (3.22) gives that

$$\|r(x_l)\|^2 - \|r(x_{l+1})\|^2 \geq \frac{\eta_1 \sigma_{\min}}{r} \tag{3.34}$$

and, since

$$\|r(x_{l+1})\| < \|r(x_l)\| \leq \|r(x_0)\| \tag{3.35}$$

for all $l \in \mathcal{S}$, Lemma 3.11 implies that

$$\begin{aligned} \|r(x_l)\|^{1/2^i} - \|r(x_{l+1})\|^{1/2^i} &\geq \tfrac{1}{2}^{i+1} \frac{\eta_1 \sigma_{\min}}{r} \|r(x_l)\|^{-(2^{i+1}-1)/2^i} \\ &\geq \tfrac{1}{2}^{i+1} \frac{\eta_1 \sigma_{\min}}{r} \|r(x_0)\|^{-(2^{i+1}-1)/2^i}. \end{aligned} \tag{3.36}$$

By contrast, for $l \in \mathcal{L} \cap \mathcal{S}$, (3.22) gives that

$$\|r(x_l)\|^2 - \|r(x_{l+1})\|^2 \geq \kappa (\pi_\Phi(x_{l+1}))^{r/(r-1)}, \text{ where } \kappa = \frac{2\eta_1 \sigma_{\min}}{r(L+\theta+\sigma_{\max})^{r/(r-1)}}. \tag{3.37}$$

If additionally $l \in \mathcal{S}_\beta$, (3.37) may be refined as

$$\begin{aligned} \|r(x_l)\|^2 - \|r(x_{l+1})\|^2 &\geq \kappa \left(\frac{\pi_\Phi(x_{l+1})}{\|r(x_{l+1})\|} \right)^{r/(r-1)} \|r(x_{l+1})\|^{r/(r-1)} \\ &\geq \kappa \beta^{r/(r-1)} \left[\psi_r(x_{l+1})\right]^{r/(r-1)} \|r(x_l)\|^{r/(r-1)} \end{aligned} \tag{3.38}$$

from (1.9) and the requirement that $\|r(x_{l+1})\| > \beta \|r(x_l)\|$. Using (3.38), (3.35), Lemma 3.11 and (3.33), we then obtain the bound

$$\begin{aligned} \|r(x_l)\|^{1/2^i} - \|r(x_{l+1})\|^{1/2^i} &\geq \tfrac{1}{2}^{i+1} \kappa \beta^{r/(r-1)} [\psi_r(x_{l+1})]^{r/(r-1)} \|r(x_l)\|^{\left(r/(r-1)-(2^{i+1}-1)/2^i\right)} \\ &\geq \tfrac{1}{2}^{i+1} \kappa \beta^{r/(r-1)} \|r(x_0)\|^{\left(r/(r-1)-(2^{i+1}-1)/2^i\right)} [\psi_r(x_{l+1})]^{r/(r-1)} \end{aligned} \tag{3.39}$$

for all $l \in \mathcal{L} \cap \mathcal{S}_\beta$. Finally, consider $l \in \mathcal{S} \setminus \mathcal{S}_\beta$, for which $\|r(x_{l+1})\| \leq \beta \|r(x_l)\|$ and hence $\|r(x_{l+1})\|^{1/2^i} \leq \beta^{1/2^i} \|r(x_l)\|^{1/2^i}$. Thus we have that

$$\begin{aligned} \|r(x_l)\|^{1/2^i} - \|r(x_{l+1})\|^{1/2^i} &\geq (1 - \beta^{1/2^i}) \|r(x_l)\|^{1/2^i} \\ &\geq \tfrac{1-\beta^{1/2^i}}{\beta^{1/2^i}} \|r(x_{l+1})\|^{1/2^i} \end{aligned} \tag{3.40}$$

for all $l \in \mathcal{L} \cap (\mathcal{S} \setminus \mathcal{S}_\beta)$. Thus, combining (3.36), (3.39) and (3.40), we have that

$$\|r(x_l)\|^{1/2^i} - \|r(x_{l+1})\|^{1/2^i} \geq \min\left(\kappa_c, \kappa_g [\psi_r(x_{l+1})]^{r/(r-1)}, \kappa_r \|r(x_{l+1})\|^{1/2^i}\right), \tag{3.41}$$

for κ_c, κ_g and κ_r given by (3.31), for all $l \in \mathcal{S}$.

Now suppose that the stopping rule (3.30) has not been satisfied up until the start of iteration $k+1$, and thus that

$$\|r(x_{l+1})\| > \epsilon_p \quad \text{and} \quad \psi_r(x_{l+1}) > \epsilon_d \tag{3.42}$$

for all $l \in \mathcal{S}_k$. Combining this with (3.41), we have that

$$\|r(x_l)\|^{1/2^i} - \|r(x_{l+1})\|^{1/2^i} \geq \min\left(\kappa_c, \kappa_g \epsilon_d^{r/(r-1)}, \kappa_r \epsilon_p^{1/2^i}\right),$$

and thus, summing over $l \in \mathcal{S}_k$ and using (3.35),

$$\|r(x_0)\|^{1/2^i} \geq \|r(x_0)\|^{1/2^i} - \|r(x_{k+1})\|^{1/2^i} \geq |\mathcal{S}_k| \min\left(\kappa_c, \kappa_g \epsilon_d^{r/(r-1)}, \kappa_r \epsilon_p^{1/2^i}\right). \tag{3.43}$$

As before, combining this with (3.18) and remembering that we need to evaluate the function and gradient at the final x_{k+1} yields the claimed evaluation bounds. □

If $i < i_0$, a weaker bound that includes $r = 2$ is possible. The key is to note that the purpose of (3.33) is to guarantee the second inequality in (3.39). Without this, we have instead

$$\|r(x_l)\|^{1/2^i} - \|r(x_{l+1})\|^{1/2^i} \geq \tfrac{1}{2}^{i+1} \kappa \beta^{r/(r-1)} [\psi_r(x_{l+1})]^{r/(r-1)} \|r(x_{l+1})\|^{\left(r/(r-1) - (2^{i+1}-1)/2^i\right)} \tag{3.44}$$

for all $l \in \mathcal{L} \cap \mathcal{S}_\beta$, and this leads to

$$\|r(x_l)\|^{1/2^i} - \|r(x_{l+1})\|^{1/2^i} \geq \min\left(\kappa_c, \kappa_g \epsilon_d^{r/(r-1)} \epsilon_p^{\left(r/(r-1) - (2^{i+1}-1)/2^i\right)}, \kappa_r \epsilon_p^{1/2^i}\right),$$

where

$$\kappa_g \prime := \frac{\tfrac{1}{2}^i \eta_1 \sigma_{\min} \beta^{r/(r-1)}}{r(L_{\rm g} + \theta + \sigma_{\max})^{r/(r-1)}}.$$

if (3.42) holds. This results in a bound of $O\left(\max(1, \epsilon_d^{r/(r-1)} \cdot \epsilon_p^{\left(r/(r-1) - (2^{i+1}-1)/2^i\right)}, \epsilon_p^{1/2^i})\right)$ function evaluations, which approaches that in Theorem 3.12 as i increases to infinity when $r = 2$.

An improved complexity result may be obtained if $\psi_r(x_k) \geq \psi_{\min}$ for some constant $\psi_{\min} > 0$ at all generated x_k; for unconstrained problems, this would be the case if the smallest singular value of $J(x_k)$ is uniformly bounded away from zero. To see why this is so, we show that such a restriction implies that the sequence $\{\|r(x_k)\|\}_{k \geq k_0}$ decreases linearly on successful iterations once $\|r(x_{k_0})\|$ is sufficiently small; the result is a variation of [11, Thm.3.6].

Theorem 3.13. Suppose that AS.1, $2 < r \leq 3$ and (3.28) hold, that Algorithm 2.1 is applied with the termination test (3.30), and that

$$\psi_r(x_k) \geq \psi_{\min} > 0 \text{ for all } k \text{ until termination.} \tag{3.45}$$

Then Algorithm 2.1 requires at most $\kappa_u \#_{\mathrm{S}} + \kappa_s + 1$ evaluations of $r(x)$ and $\#_{\mathrm{S}} + 1$ evaluations of its derivatives, involving

$$\#_{\mathrm{S}} := \begin{cases} \left\lceil \max\left(\kappa_c^{-1}, \kappa_g^{-1} \max(\epsilon_d, \psi_{\min})^{-r/(r-1)}, \kappa_r^{-1} \rho^{-1/2^i}\right)\right\rceil + \left\lceil |\log_\beta(\epsilon_p/\rho)|\right\rceil & \text{if } \epsilon_p < \rho \\ \left\lceil \max\left(\kappa_c^{-1}, \kappa_g^{-1} \max(\epsilon_d, \psi_{\min})^{-r/(r-1)}, \kappa_r^{-1} \epsilon_p^{-1/2^i}\right)\right\rceil & \text{otherwise,} \end{cases} \tag{3.46}$$

successful iterations, to find an iterate x_k for which (3.30) holds, where κ_u and κ_s are defined in (3.18) using $\sigma_{\max}$ from (3.17), κ_c, κ_g and κ_r are defined in (3.31),

$$\rho := \min\left((0.99\kappa)^{\frac{r-1}{r-2}} (\beta\psi_{\min})^{\frac{r}{r-2}}, \sqrt{\eta_1 \sigma_{\min}\left(\frac{\beta^2}{1-\beta^2}\right)} \right), \tag{3.47}$$

and $\beta \in (0,1)$ is a fixed problem-independent constant.

Proof. First, observe that since Theorem 3.12 shows that Algorithm 2.1 ensures (3.30) for any given $\epsilon_p, \epsilon_d > 0$, and as (3.45) forces $\psi_r(x_k) > \epsilon_d$ whenever $\epsilon_d < \psi_{\min}$, it must be that the algorithm terminates with $\|r(x)\| \leq \epsilon_p$ in this case. As this is true for arbitrary ϵ_p, we conclude that $\|r(x)\|$ may be made arbitrarily small within $\mathcal{C}$ by picking ϵ_p appropriately small.

Let k_0 be the smallest k for which $\|r(x_k)\| \leq \rho$, in which case (3.47) implies

$$\|r(x_k)\|^{\frac{r-2}{r-1}} \leq 0.99\kappa(\beta\psi_{\min})^{\frac{r}{r-1}} < \kappa(\beta\psi_{\min})^{\frac{r}{r-1}} \tag{3.48}$$

and, since (3.35) holds,

$$\eta_1 \sigma_{\min} \geq \left(\frac{1-\beta^2}{\beta^2}\right) \|r(x_k)\|^2 \geq \left(\frac{1-\beta^2}{\beta^2}\right) \|r(x_{l+1})\|^2. \tag{3.49}$$

Now consider the set $\mathcal{S}_\beta$ just as in (3.32) in the proof of Theorem 3.12 Then if $k \in \mathcal{L} \cap \mathcal{S}_\beta$, (3.38) holds, and combining this with (3.45) and (3.48) we find that

$$\|r(x_k)\|^2 - \|r(x_{k+1})\|^2 \geq \kappa(\beta\psi_{\min})^{r/(r-1)} \|r(x_k)\|^{r/(r-1)} > \|r(x_k)\|^2. \tag{3.50}$$

Since this then implies $\|r(x_{k+1})\|^2 < 0$, which is impossible, we must have that if $k \in \mathcal{L}$, $k \in \mathcal{S} \setminus \mathcal{S}_\beta$ for all successful $k \geq k_0$, and thus

$$\|r(x_{k+1})\| \leq \beta \|r(x_k)\| \text{ for all } k \geq k_0 \in \mathcal{L} \cap \mathcal{S}. \tag{3.51}$$

It also follows from (3.34) and (3.49) that if $k \geq k_0 \in \mathcal{G} \cap \mathcal{S}$

$$\|r(x_k)\|^2 - \|r(x_{k+1})\|^2 \geq \eta_1 \sigma_{\min} \geq \left(\frac{1}{\beta^2} - 1\right) \|r(x_{k+1})\|^2$$

and thus that

$$\|r(x_{k+1})\| \leq \beta \|r(x_k)\| \text{ for all } k \geq k_0 \in \mathcal{G} \cap \mathcal{S}. \tag{3.52}$$

Thus, combining (3.51) and (3.52), it follows that

$$\|r(x_{k+1})\| \leq \beta \|r(x_k)\| \text{ for all } k \geq k_0 \in \mathcal{S}. \tag{3.53}$$

We may then invoke Theorem 3.12 to deduce that Algorithm 2.1 will find the first k, $k = k_f$ say, for which

$$\|r(x_k))\| \leq \max(\epsilon_p, \rho) \quad \text{or} \quad \psi_r(x_k)) \leq \max(\epsilon_d, \psi_{\min}) \tag{3.54}$$

after at most

$$\#_{\rm R} := \left\lceil \max\left(\kappa_c^{-1}, \kappa_g^{-1} \max(\epsilon_d, \psi_{\min})^{-r/(r-1)}, \kappa_r^{-1} \max(\epsilon_p, \rho)^{-1/2^i}\right)\right\rceil$$

successful iterations.

There are four possibilities:

1. $\rho \leq \epsilon_p$ and $\psi_{\min} \leq \epsilon_d$. In this case, x_{k_f} satisfies (3.30).
2. $\rho \leq \epsilon_p$ and $\psi_{\min} > \epsilon_d$. Then since $\psi_r(x_{k_f}) \geq \psi_{\min} > \epsilon_d$, it must be that $\|r(x_{k_f})\| \leq \epsilon_p$, and again x_{k_f} satisfies (3.30).
3. $\rho > \epsilon_p$ and $\psi_{\min} > \epsilon_d$. As in the previous case, $\psi_r(x_{k_f}) > \epsilon_d$, and thus $\|r(x_{k_f})\| \leq \rho$. Since $\|r(x_{k_f-1})\| > \rho$ (as otherwise Algorithm 2.1 would have terminated if, or before, iteration k_f-1), we have that $k_f = k_0$. Furthermore, if this happens, additional iterations will be required to guarantee that $\|r(x_k)\| \leq \epsilon_p$.
4. $\rho > \epsilon_p$ and $\psi_{\min} \leq \epsilon_d$. In this case, either $\psi_r(x_{k_f}) \leq \epsilon_d$ or $\|r(x_{k_f})\| \leq \rho$ (or both). In the former case, again x_{k_f} satisfies (3.30), while in the latter, as in possibility 3. we have $k_f = k_0$ and require further iterations.

In summary, if $\rho \leq \epsilon_p$, x_{k_f} satisfies (3.30), while if $\rho > \epsilon_p$ further iterations may be required, and if they are, we have $k_f = k_0$. Since (3.53) holds, we require at most $\#_{\rm L} := \lceil |\log_\beta(\epsilon_p/\rho)| \rceil$ additional iterations to achieve $\|r(x_k)\| \leq \epsilon_p$. Combining the successful iteration counts $\#_{\rm R}$ and $\#_{\rm L}$ with Lemma 3.6, and remembering that we need to evaluate the function and gradient at the final x_{k+1} yields desired evaluation bounds. □

There might be concerns that (3.46) indicates an evaluation bound dependence on ϵ_d and ϵ_p because of the terms $\epsilon_p^{-1/2^i}$—when $\epsilon_p \geq \rho$—and $\max(\epsilon_d, \psi_{\min})^{-r/(r-1)}$. However, the (weaker) bounds

$$\epsilon_p^{-1/2^i} \leq \rho - 1/2^i \text{ and } \max(\epsilon_d, \psi_{\min})^{-r/(r-1)} \leq \psi_{\min}^{-r/(r-1)}$$

in this case should allay any such fears. The only true dependence on either tolerance is via the very mild term $\lceil |\log_\beta(\epsilon_p/\rho)| \rceil$ when $\epsilon_p < \rho$.

3.1 Using the alternative χ criticality measure

We now return to the alternative criticality measure (1.6). Our aim is briefly to indicate that the results obtained in §3 for the measure (1.5) may be mirrored so long as we adapt Algorithm 2.1 in an appropriate way. To do so, let

$$\chi_m^{\mathrm{R}}(x,s,\sigma) := \left| \min_{\substack{x+s+d\in\mathcal{C},\\ \|d\|_\chi\leq 1}} \langle \nabla_s m^{\mathrm{R}}(x,s,\sigma), d\rangle \right|, \tag{3.55}$$

where $\nabla_s m^{\mathrm{R}}(x,s,\sigma)$ is defined by (2.5). We then consider the following algorithm (Algorithm 3.1) to find a critical point of $\Phi(x)$ within $\mathcal{C}$ using the measure (1.6).

Algorithm 3.1: Adaptive Tensor-Newton Regularization based on χ.

Use Algorithm 2.1, but replace (2.8) by

$$\chi_m^{\mathrm{R}}(x_k,s_k,\sigma_k) \leq \theta\|s_k\|^{r-1}. \tag{3.56}$$

In what follows, let $\kappa_\chi > 0$ be the norm equivalence constant for which

$$\|v\| \leq \kappa_\chi \|v\|_\chi \ \text{ for all } \ v \in \mathbb{R}^n. \tag{3.57}$$

Careful examination of the results in §3 reveal that the only use of (2.8) is in the proof of Lemma 3.7 and its corollary. We may replace these by the following result.

Corollary 3.14. Suppose that AS.1 holds and $2 \leq r \leq 3$. Then Algorithm 3.1 ensures that

$$\Phi(x_k) - \Phi(x_{k+1}) \geq \begin{cases} \frac{1}{r}\eta_1\sigma_{\min}\left(\dfrac{\chi_\Phi(x_{k+1})}{2\kappa_\chi(L_{\mathrm{g}}+\theta+\sigma_{\max})}\right)^{\frac{r}{r-1}} & \text{if } k\in\mathcal{L}\cap\mathcal{S} \\ \frac{1}{r}\eta_1\sigma_{\min} & \text{if } k\in\mathcal{G}\cap\mathcal{S}, \end{cases} \tag{3.58}$$

where $\sigma_{\max}$ is given by (3.12) when $r=2$ and by (3.17) for $2<r\leq 3$.

Proof. Suppose that $k \in \mathcal{L}\cup\mathcal{S}$. Since $k\in\mathcal{S}$, $x_{k+1} = x_k + s_k$, and it follows from (3.2), (2.5), (3.12), (3.17) and the bound $\|s_k\| \leq 1$, as $k\in\mathcal{L}$, that

$$\|\nabla\Phi(x_{k+1}) - \nabla_s m^{\mathrm{R}}(x_k,s_k,\sigma_k)\| \leq L_{\mathrm{g}}\|s_k\|^2 + \sigma_k\|s_k\|^{r-1} \leq (L_{\mathrm{g}}+\sigma_{\max})\|s_k\|^{r-1}. \tag{3.59}$$

Moreover

$$\begin{aligned} \chi_\Phi(x_{k+1}) &:= |\langle \nabla_x\Phi(x_{k+1}), d_{k+1}\rangle| \\ &\leq |\langle \nabla_x\Phi(x_{k+1}) - \nabla_s m^{\mathrm{R}}(x_k,s_k,\sigma_k), d_{k+1}\rangle| + |\langle \nabla_s m^{\mathrm{R}}(x_k,s_k,\sigma_k), d_{k+1}\rangle|, \end{aligned} \tag{3.60}$$

where the first equality defines the vector d_{k+1} with

$$\|d_{k+1}\|_\chi \leq 1. \tag{3.61}$$

Assume now, for the purpose of deriving a contradiction, that

$$\|s_k\|^{r-1} < \frac{\chi_\Phi(x_{k+1})}{2\kappa_\chi(L_{\rm g} + \theta + \sigma_{\max})}. \tag{3.62}$$

Using the Cauchy-Schwarz inequality, (3.57), (3.61), (3.59), the assumption (3.62), and (3.60) in turn, we find that

$$\begin{aligned}
\langle \nabla_s m^{\rm R}(x_k, s_k, \sigma_k), d_{k+1}\rangle - \langle \nabla_x \Phi(x_{k+1}), d_{k+1}\rangle & \\
&\leq |\langle \nabla_x \Phi(x_{k+1}), d_{k+1}\rangle - \langle \nabla_s m^{\rm R}(x_k, s_k, \sigma_k), d_{k+1}\rangle| \\
&\leq \|\nabla_x \Phi(x_{k+1}) - \nabla_s m^{\rm R}(x_k, s_k, \sigma_k)\| \, \|d_{k+1}\| \\
&\leq \kappa_\chi(L_{\rm g} + \sigma_{\max})\|s_k\|^{r-1} \\
&\leq \kappa_\chi(L_{\rm g} + \theta + \sigma_{\max})\|s_k\|^{r-1} \\
&\leq \tfrac{1}{2}\chi_\Phi(x_{k+1}) \\
&= -\tfrac{1}{2}\langle \nabla_x \Phi(x_{k+1}), d_{k+1}\rangle,
\end{aligned}$$

which ensures that

$$\langle \nabla_s m^{\rm R}(x_k, s_k, \sigma_k), d_{k+1}\rangle \leq \tfrac{1}{2}\langle \nabla_x \Phi(x_{k+1}), d_{k+1}\rangle < 0.$$

Moreover, $x_{k+1} + d_{k+1} \in \mathcal{C}$ by definition of $\chi_\Phi(x_{k+1})$, and hence, using (3.61),

$$|\langle \nabla_s m^{\rm R}(x_k, s_k, \sigma_k), d_{k+1}\rangle| \leq \chi^{\rm R}_m(x_{k+1}).$$

Substituting this into (3.60) and using the Cauchy-Schwarz inequality, (3.57) and (3.61) again to deduce that

$$\chi_\Phi(x_{k+1}) \leq \|\nabla_x \Phi(x_{k+1}) - \nabla_s m^{\rm R}(x_k, s_k, \sigma_k)\| + \chi^{\rm R}_m(x_{k+1}) \leq \kappa_\chi(L_{\rm g} + \theta + \sigma_{\max})\|s_k\|^{r-1}$$

where the last inequality results from (3.59) and (3.56). But this contradicts the assumption that (3.62) holds, and hence

$$\|s_k\| \geq \left[\frac{\chi_\Phi(x_{k+1})}{2\kappa_\chi(L_{\rm g} + \theta + \sigma_{\max})}\right]^{\frac{1}{r-1}} \quad \text{for all} \quad k \in \mathcal{L} \cup \mathcal{S}. \tag{3.63}$$

Since

$$\|s_k\| \geq 1 \quad \text{for all} \quad k \in \mathcal{G} \cup \mathcal{S} \tag{3.64}$$

by definition of $\mathcal{G}$, we may combine the fact that $\rho_k \geq \eta_1$ in (2.9) when $k \in \mathcal{S}$, with (3.5), (3.63) and (3.64) to deduce (3.58). □

Armed with Corollary (3.14), the remaining main results in §3 (Theorems 3.9, 3.10, 3.12 and 3.13) remain true if we replace the stopping tests (1.7) and (1.8) by (1.10) and (1.11), and account for the extra factor $2\kappa_\chi$ that occurs in (3.58) compared with (3.22).

4 A modified algorithm for cubic-and-higher regularization

For the case where $r > 3$, the proof of Lemma 3.4 breaks down as there is no obvious bound on the quantity $\|s_k\|^{3-r}/\sigma_k$. One way around this defect is to modify Algorithm 2.1 so that such a bound automatically occurs. We consider the following variant; our development follows very closely that in [12], itself inspired by [22]. For completeness, we allow $r = 3$ in this new framework since it is trivial to do so.

Algorithm 4.1: Adaptive Tensor-Newton Regularization when $r \geq 3$.

A starting point $x_0 \in \mathcal{C}$, an initial regularization parameter $\sigma_0 > 0$ and algorithmic parameters $\theta > 0$, $\alpha \in (0, \frac{1}{3}]$, $\gamma_3 \geq \gamma_2 > 1 > \gamma_1 > 0$ and $1 > \eta_2 \geq \eta_1 > 0$, are given. Evaluate $\Phi(x_0)$, and test for termination at x_0.
For $k = 0, 1, \ldots$, until **termination**, do:

1. Compute derivatives of $r(x)$ at x_k.

2. Compute a step s_k by approximately minimizing $m^{\mathrm{R}}(x_k, s, \sigma_k)$ within $\mathcal{C}$ so that

$$x_k + s_k \in \mathcal{C},$$

$$m^{\mathrm{R}}(x_k, s_k, \sigma_k) < m^{\mathrm{R}}(x_k, 0, \sigma_k)$$

and

$$\pi_m^{\mathrm{R}}(x_k, s_k, \sigma_k) \leq \theta \|s_k\|^2 \tag{4.1}$$

hold.

3. Test for termination at $x_k + s_k$.

4. Compute $\Phi(x_k + s_k)$ and

$$\rho_k = \frac{\Phi(x_k) - \Phi(x_k + s_k)}{m(x_k, 0) - m(x_k, s_k)}.$$

If $\rho_k \geq \eta_1$ and

$$\sigma_k \|s_k\|^{r-1} \geq \alpha \pi_\Phi(x_k + s_k), \tag{4.2}$$

set $x_{k+1} = x_k + s_k$. Otherwise set $x_{k+1} = x_k$.

5. Set

$$\sigma_{k+1} \in \begin{cases} [\gamma_1 \sigma_k, \sigma_k] & \text{if } \rho_k \geq \eta_2 \text{ and (4.2) holds} \\ [\sigma_k, \gamma_2 \sigma_k] & \text{if } \eta_1 \leq \rho_k < \eta_2 \text{ and (4.2) holds} \\ [\gamma_2 \sigma_k, \gamma_3 \sigma_k] & \text{if } \rho_k < \eta_1 \text{ or (4.2) fails,} \end{cases} \tag{4.3}$$

and go to Step 2 if $\rho_k < \eta_1$ or (4.2) fails.

It is important that termination is tested at Step 3 as deductions from computations in subsequent steps rely on this. We modify our definition of a successful step accordingly so that

now $\mathcal{S}_k = \{0 \leq l \leq k \mid \rho_l \geq \eta_1 \text{ and (4.2) holds}\}$ and $\mathcal{S} = \{k \geq 0 \mid \rho_k \geq \eta_1 \text{ and (4.2) holds}\}$, and note in particular that Lemma 3.6 continues to hold in this case since it only depends on the adjustments in (4.3). Likewise, a very successful iteration is now one for which $\rho_k \geq \eta_2$ and (4.2) holds. Note that (4.3), unlike (2.10) in Algorithm 2.1, does not impose a nonzero lower bound on the generated regularization weight; this will be reflected in our derived complexity bound (cf Theorems 3.12 and 4.7).

As is now standard, our first task is to establish an upper bound on σ_k.

Lemma 4.1. Suppose that AS.1 holds, $r \geq 3$ and

$$\sigma_k \|s_k\|^{r-3} \geq \kappa_2, \text{ where } \kappa_2 := \frac{rL}{1-\eta_2} \text{ and } L = \max(L_{\mathrm{f}}, L_{\mathrm{g}}, \theta). \tag{4.4}$$

Then iteration k of Algorithm 4.1 is very successful.

Proof. It follows immediately from (2.9), (3.1), (3.5) and (4.4) that

$$|\rho_k - 1| = \frac{|\Phi(x_k + s_k) - m(x_k, s_k)|}{m(x_k, 0) - m(x_k, s_k)} \leq \frac{rL_{\mathrm{f}}\|s_k\|^{3-r}}{\sigma_k} \leq \frac{rL\|s_k\|^{3-r}}{\sigma_k} \leq 1 - \eta_2,$$

and thus $\rho_k \geq \eta_2$. Observe that

$$\kappa_2 \geq L \tag{4.5}$$

since $1 - \eta_2 \leq 1$ and $r \geq 1$. We also have from (3.20), (3.2) and (4.1) that

$$\pi_\Phi(x_k + s_k) \leq L_{\mathrm{g}}\|s_k\|^2 + \theta\|s_k\|^2 + \sigma_k\|s_k\|^{r-1} = (L_{\mathrm{g}} + \theta + \sigma_k\|s_k\|^{r-3})\|s_k\|^2 \tag{4.6}$$

and thus from (4.4), (4.5) and the algorithmic restriction $3 \leq 1/\alpha$ that

$$\pi_\Phi(x_k + s_k) \leq (2L + \sigma_k\|s_k\|^{r-3})\|s_k\|^2 \leq (3\sigma_k\|s_k\|^{r-3})\|s_k\|^2 = 3\sigma_k\|s_k\|^{r-1} \leq \frac{\sigma_k}{\alpha}\|s_k\|^{r-1}.$$

Thus (4.2) is also satisfied, and hence iteration k is very successful. □

Lemma 4.2. Suppose that AS.1 holds, $r \geq 3$ and

$$\sigma_k \geq \kappa_1[\pi_\Phi(x_k + s_k)]^{(3-r)/2}, \text{ where } \kappa_1 := \kappa_2(3\kappa_2)^{(r-3)/2} \tag{4.7}$$

and κ_2 is defined in the statement of Lemma 4.1. Then iteration k of Algorithm 4.1 is very successful.

Proof. It follows from Lemma 4.1 that it suffices to show that (4.7) implies (4.4). The result is immediate when $r = 3$ since then (4.4) is (4.7). Suppose therefore that $r > 3$ and that (4.4) is not true, that is

$$\sigma_k\|s_k\|^{r-3} < \kappa_2. \tag{4.8}$$

Then (4.6), (4.8) and (4.5) imply that

$$\pi_\Phi(x_k + s_k) \leq (2L + \kappa_2)\|s_k\|^2 < 3\kappa_2\|s_k\|^2 < 3\kappa_2 \left(\frac{\kappa_2}{\sigma_k}\right)^{2/(r-3)}$$

which contradicts (4.7). Thus (4.4) holds. □

Unlike in our previous analysis of Algorithm 2.1 when $r \leq 3$, we are unable to deduce an upper bound on σ_k without further consideration. With this in mind, we now suppose that all the iterates $x_k + s_k$ generated by Algorithm 4.1 satisfy

$$\pi_\Phi(x_k + s_k) \geq \epsilon \tag{4.9}$$

for some $\epsilon > 0$ and all $0 \leq k \leq l$, and thus, from (4.2), that

$$\sigma_k\|s_k\|^{r-1} \geq \alpha\epsilon \tag{4.10}$$

for $k \in \mathcal{S}_l$. In this case, we can show that σ_k is bounded from above.

Lemma 4.3. Suppose that AS.1 holds and $r \geq 3$. Then provided that (4.9) holds for all $0 \leq k \leq l$, Algorithm 4.1 ensures that

$$\sigma_k \leq \sigma_{\max} := \gamma_3 \max\left(\kappa_1\epsilon^{(3-r)/2}, \sigma_0\right) \tag{4.11}$$

and κ_1 is defined in the statement of Lemma 4.2.

Proof. The proof is similar to the first part of that of Lemma 3.5. Suppose that iteration $k+1$ (with $k \leq l$) is the first for which $\sigma_{k+1} \geq \sigma_{\max}$. Then, since $\sigma_k < \sigma_{k+1}$, iteration k must have been unsuccessful and (4.3) gives that

$$\gamma_3\sigma_k \geq \sigma_{k+1} \geq \sigma_{\max},$$

i.e., that

$$\sigma_k \geq \max\left(\kappa_1\epsilon^{(3-r)/2}, \sigma_0\right) \geq \kappa_1\epsilon^{(3-r)/2} \geq \kappa_1[\pi_\Phi(x_k + s_k)]^{(3-r)/2}$$

because of (4.9). But then Lemma 4.2 implies that iteration k must be very successful. This contradiction establishes (4.11). □

We may also show that a successful step ensures a non-trivial reduction in $\Phi(x)$.

Lemma 4.4. Suppose that AS.1 holds and $r \geq 3$. Suppose further that (4.9) holds for all $0 \leq k \leq l$. Then provided that (4.9) holds for all $0 \leq k \leq l$ and some $0 < \epsilon \leq 1$, Algorithm 4.1 guarantees that

$$\Phi(x_k) - \Phi(x_{k+1}) \geq \kappa_4 \epsilon^{3/2} > 0 \tag{4.12}$$

for all $k \in \mathcal{S}$, where

$$\kappa_4 := \frac{\eta \alpha^{r/(r-1)}}{r \kappa_3^{1/(r-1)}}, \quad \kappa_3 := \gamma_3 \max(\kappa_1, \sigma_0), \tag{4.13}$$

and κ_1 is defined in the statement of Lemma 4.2.

Proof. Since $0 < \epsilon \leq 1$, (4.11) ensures that $\sigma_{\max} \leq \kappa_3 \epsilon^{(3-r)/2}$ and thus if $k \in \mathcal{S}$, it follows from (3.5) and (4.10) that

$$\begin{aligned}
\Phi(x_k) - \Phi(x_{k+1}) &\geq \eta_1 (m(x_k, 0) - m(x_k, s_k)) > \frac{\eta_1}{r} \sigma_k \|s_k\|^r \\
&= \frac{\eta_1}{r} (\sigma_k \|s_k\|^{r-1}) \|s_k\| \geq \frac{\eta_1}{r} \alpha \epsilon \frac{(\alpha \epsilon)^{1/(r-1)}}{\sigma_k^{1/(r-1)}} \geq \frac{\eta (\alpha \epsilon)^{r/(r-1)}}{r \sigma_{\max}^{1/(r-1)}} \\
&\geq \frac{\eta \alpha^{r/(r-1)}}{r \kappa_3^{1/(r-1)}} \frac{\epsilon^{r/(r-1)}}{\left(\epsilon^{(3-r)/2}\right)^{1/(r-1)}} = \kappa_4 \epsilon^{3/2} > 0,
\end{aligned}$$

as required. □

These introductory lemmas now lead to our main convergence results. First we establish global convergence to a critical point of $\Phi(x)$.

Theorem 4.5. Suppose that AS.1 holds and $r \geq 3$. Then the iterates $\{x_k\}$ generated by Algorithm 4.1 satisfy

$$\lim_{k \to \infty} \inf \pi_\Phi(x_k) = 0 \tag{4.14}$$

if no non-trivial termination test is provided.

Proof. Suppose that (4.14) does not hold, in which case (4.9) holds for some $0 < \epsilon \leq 1$ and all $k \geq 0$. We then deduce by summing the reduction in $\Phi(x)$ guaranteed by Lemma 4.4 over successful iterations that

$$\tfrac{1}{2} \|r(x_0)\|^2 \geq \Phi(x_0) - \Phi(x_{k+1}) \geq |\mathcal{S}_k| \kappa_4 \epsilon^{3/2}.$$

Just as in the proof of Theorem 3.10, this ensures that there are only a finite number of successful iterations. If iteration k is the last of these, all subsequent iterations are unsuccessful, and thus σ_k grows without bound. But as this contradicts Lemma 4.3, (4.9) cannot be true, and thus (4.14) holds. □

Next, we give an evaluation complexity result based on the stopping criterion (1.7).

Theorem 4.6. Suppose that AS.1 holds and $r \geq 3$. Then Algorithm 4.1 requires at most

$$\begin{cases} \left\lceil \kappa_u \dfrac{\|r(x_0)\|^2}{2\kappa_4} \epsilon^{-3/2} + \kappa_b \right\rceil + 1 & \text{if } r = 3, \\ \left\lceil \kappa_u \dfrac{\|r(x_0)\|^2}{2\kappa_4} \epsilon^{-3/2} + \kappa_i + \kappa_e \log \epsilon^{-1} \right\rceil + 1 & \text{if } r > 3 \text{ and } \epsilon < \left(\dfrac{\kappa_1}{\sigma_0}\right)^{2/(r-3)}, \\ \left\lceil \kappa_u \dfrac{\|r(x_0)\|^2}{2\kappa_4} \epsilon^{-3/2} + \kappa_a \right\rceil + 1 & \text{otherwise} \end{cases} \tag{4.15}$$

evaluations of $r(x)$ and its derivatives to find an iterate x_k for which the termination test

$$\pi_\Phi(x_k) \leq \epsilon$$

is satisfied for given $0 < \epsilon < 1$, where

$$\kappa_b := \frac{\log(\kappa_3/\sigma_0)}{\log \gamma_2}, \quad \kappa_i := \frac{\log(\gamma_3 \kappa_1/\sigma_0)}{\log \gamma_2}, \quad \kappa_e := \frac{r-3}{2 \log \gamma_2} \quad \text{and} \quad \kappa_a := \frac{\log \gamma_3}{\log \gamma_2}, \tag{4.16}$$

κ_u is defined in (3.18), κ_1 in (4.7) and κ_3 in (4.13).

Proof. If the algorithm has not terminated on or before iteration k, (4.9) holds, and so summing (4.12) over successful iterations and recalling that $\Phi(x_0) = \frac{1}{2}\|r(x_0)\|^2$ and $\Phi(x_k) \geq 0$, we have that

$$\tfrac{1}{2}\|r(x_0)\|^2 \geq \Phi(x_0) - \Phi(x_{k+1}) \geq |\mathcal{S}_k| \kappa_4 \epsilon^{3/2}.$$

Thus there at most

$$|\mathcal{S}_k| \leq \frac{\|r(x_0)\|^2}{2\kappa_4} \epsilon^{-3/2}$$

successful iterations. Combining this with Lemma 3.6, accounting for the max in (4.11) and remembering that we need to evaluate the function and gradient at the final x_{k+1} yields the bound (4.15). □

We note in passing that in order to derive Theorem 4.6, we could have replaced the test (4.2) in Algorithm 4.1 by the normally significantly-weaker requirement (4.10).

Our final result examines the evaluation complexity under the stopping rule (3.30).

Theorem 4.7. Suppose that AS.1 holds, $r \geq 3$ and an $i \geq 1$ is given. Then Algorithm 4.1 requires at most

$$\begin{array}{l} \left\lceil \kappa_u \|r(x_0)\|^{1/2^i} \max\left(\kappa_g^{-1}\epsilon_d^{-3/2}, \kappa_r^{-1}\epsilon_p^{-1/2^i}\right) + \kappa_b \right\rceil + 1 \\ \text{when } \ r = 3 \\ \left\lceil \kappa_u \|r(x_0)\|^{1/2^i} \max\left(\kappa_g^{-1}\epsilon_d^{-3/2}, \kappa_r^{-1}\epsilon_p^{-1/2^i}\right) + \kappa_i + \kappa_e(\log \epsilon_d^{-1} + \log \epsilon_p^{-1}) \right\rceil + 1 \\ \text{when } \ r > 3 \ \text{ and } \ \epsilon_p\epsilon_d < \left(\dfrac{\kappa_1}{\sigma_0}\right)^{2/(r-3)}, \quad \text{or otherwise} \\ \left\lceil \kappa_u \|r(x_0)\|^{1/2^i} \max\left(\kappa_g^{-1}\epsilon_d^{-3/2}, \kappa_r^{-1}\epsilon_p^{-1/2^i}\right) + \kappa_a \right\rceil + 1, \end{array} \tag{4.17}$$

evaluations of $r(x)$ and its derivatives to find an iterate x_k for which the termination test

$$\|r(x_k)\| \leq \epsilon_p \quad \text{or} \quad \psi_r(x_k) \leq \epsilon_d,$$

is satisfied for given $0 < \epsilon_p, \epsilon_d \leq 1$, where κ_u is defined in (3.18),

$$\kappa_g := \frac{\eta_1 \alpha^{r/(r-1)}}{2^i r \gamma_3^{1/(r-1)}} \min\left(\frac{1}{\kappa_1}, \frac{1}{\sigma_0}\right)^{1/(r-1)} \|r(x_0)\|^{(3/2-(2^{i+1}-1)/2^i)}, \ \ \kappa_r := (1-\beta^{1/2^i}), \tag{4.18}$$

κ_1 is defined in (4.7), κ_b, κ_i, κ_e and κ_a in (4.16), and $\beta \in (0,1)$ is a fixed problem-independent constant.

Proof. As in the proof of Theorem 3.12, let $\mathcal{S}_\beta := \{l \in \mathcal{S} \mid \|r(x_{l+1})\| > \beta \|r(x_l)\|\}$ for a given $\beta \in (0,1)$. We suppose that Algorithm 4.1 has not terminated prior to iteration $l+1$, and thus that

$$\|r(x_k)\| > \epsilon_p \quad \text{and} \quad \psi_r(x_k) > \epsilon_d \tag{4.19}$$

for all $k \leq l+1$. If $l \in \mathcal{S}_\beta$, it follows from (3.5), (4.2) and the definition (1.9) that

$$\begin{aligned} \|r(x_l)\|^2 - \|r(x_{l+1})\|^2 &\geq 2\eta_1(m(x_l,0) - m(x_l,s_l)) > \frac{2\eta_1}{r}\sigma_l\|s_l\|^r \\ &= \frac{2\eta_1}{r}(\sigma_l\|s_l\|^{r-1})\|s_l\| \geq \frac{2\eta_1}{r}\alpha^{r/(r-1)}\sigma_l^{-1/(r-1)}\|\nabla_x\Phi(x_{l+1})\|^{r/(r-1)} \\ &\geq \frac{2\eta_1}{r}\alpha^{r/(r-1)}\sigma_l^{-1/(r-1)}[\psi_r(x_{l+1})]^{r/(r-1)}\|r(x_{l+1})\|^{r/(r-1)} \\ &\geq \frac{2\eta_1}{r}\alpha^{r/(r-1)}\sigma_l^{-1/(r-1)}[\psi_r(x_{l+1})]^{r/(r-1)}\|r(x_l)\|^{r/(r-1)}\beta^{r/(r-1)} \end{aligned}$$

and thus applying Lemma 3.11 with $i \geq 1$,

$$\begin{aligned} &\|r(x_l)\|^{1/2^i} - \|r(x_{l+1})\|^{1/2^i} \\ &\quad \geq \ \frac{\eta_1\alpha^{r/(r-1)}}{2^i r}\beta^{r/(r-1)}\sigma_l^{-1/(r-1)}[\psi_r(x_{l+1})]^{r/(r-1)}\|r(x_l)\|^{(r/(r-1)-(2^{i+1}-1)/2^i)} \\ &\quad = \ \frac{\eta_1\alpha^{r/(r-1)}}{2^i r}\beta^{r/(r-1)}\sigma_l^{-1/(r-1)}[\psi_r(x_{l+1})]^{r/(r-1)}\|r(x_l)\|^{(r/(r-1)-3/2)}\|r(x_l)\|^{(3/2-(2^{i+1}-1)/2^i)} \\ &\quad \geq \ \kappa_d\sigma_l^{-1/(r-1)}[\psi_r(x_{l+1})]^{r/(r-1)}\|r(x_l)\|^{(r/(r-1)-3/2)}, \end{aligned} \tag{4.20}$$

where $\kappa_d := \dfrac{\eta_1\alpha^{r/(r-1)}}{2^i r}\beta^{r/(r-1)}\|r(x_0)\|^{(3/2-(2^{i+1}-1)/2^i)}$, as $3/2 \leq (2^{i+1}-1)/2^i$ and (3.35) holds.

In particular (4.20) becomes

$$\|r(x_l)\|^{1/2^i} - \|r(x_{l+1})\|^{1/2^i} \geq \kappa_d \sigma_l^{-1/(r-1)} \epsilon_p^{(3-r)/2(r-1)} \epsilon_d^{r/(r-1)} \tag{4.21}$$

and (4.10) holds with $\epsilon = \epsilon_p \epsilon_d$, and so

$$\sigma_l \leq \sigma_{\max} := \gamma_3 \max\left(\kappa_1 \epsilon_p^{(3-r)/2} \epsilon_d^{(3-r)/2}, \sigma_0\right) \tag{4.22}$$

from Lemma 4.3. Consider the possibility

$$\kappa_1 \epsilon_p^{(3-r)/2} \epsilon_d^{(3-r)/2} \geq \sigma_0. \tag{4.23}$$

In this case, (4.22) implies that

$$\sigma_l^{-1/(r-1)} \geq \frac{1}{(\gamma_3 \kappa_1)^{1/(r-1)}} \epsilon_p^{(r-3)/2(r-1)} \epsilon_d^{(r-3)/2(r-1)}$$

and hence combining with (4.21), we find that

$$\|r(x_l)\|^{1/2^i} - \|r(x_{l+1})\|^{1/2^i} \geq \frac{\kappa_d}{(\gamma_3 \kappa_1)^{1/(r-1)}} \epsilon_d^{3/2} \tag{4.24}$$

If (4.23) does not hold,

$$\sigma_l^{-1/(r-1)} \geq \frac{1}{(\gamma_3 \sigma_0)^{1/(r-1)}}$$

and thus (4.21) implies that

$$\|r(x_l)\|^{1/2^i} - \|r(x_{l+1})\|^{1/2^i} \geq \frac{\kappa_d}{(\gamma_3 \sigma_0)^{1/(r-1)}} \epsilon_p^{(3-r)/2(r-1)} \epsilon_d^{r/(r-1)} \geq \frac{\kappa_d}{(\gamma_3 \sigma_0)^{1/(r-1)}} \epsilon_d^{3/2} \tag{4.25}$$

since ϵ_p and $\epsilon_d \leq 1$ and $r \geq 3$. Hence (4.24) and (4.25) hold when $l \in \mathcal{S}_\beta$,
For $l \in \mathcal{S} \setminus \mathcal{S}_\beta$, for which $\|r(x_{l+1})\| \leq \beta \|r(x_l)\|$ and hence $\|r(x_{l+1})\|^{1/2^i} \leq \beta^{1/2^i} \|r(x_l)\|^{1/2^i}$. Thus in view of (4.19), we have that

$$\|r(x_l)\|^{1/2^i} - \|r(x_{l+1})\|^{1/2^i} \geq (1 - \beta^{1/2^i}) \|r(x_l)\|^{1/2^i} \geq (1 - \beta^{1/2^i}) \epsilon_p^{1/2^i} \tag{4.26}$$

for all $l \in \mathcal{S} \setminus \mathcal{S}_\beta$. Thus, combining 01(4.24),(4.25) and (4.26), we have that

$$\|r(x_l)\|^{1/2^i} - \|r(x_{l+1})\|^{1/2^i} \geq \min\left(\kappa_g \epsilon_d^{3/2}, \kappa_r \epsilon_p^{1/2^i}\right)$$

for all $l \in \mathcal{S}$, where κ_g and κ_r are given by (4.18). Summing over $l \in \mathcal{S}_k$ and using (3.35),

$$\|r(x_0)\|^{1/2^i} \geq \|r(x_0)\|^{1/2^i} - \|r(x_{k+1})\|^{1/2^i} \geq |\mathcal{S}_k| \min\left(\kappa_g \epsilon_d^{3/2}, \kappa_r \epsilon_p^{1/2^i}\right)$$

and thus that there are at most

$$|\mathcal{S}_k| \leq \|r(x_0)\|^{1/2^i} \max\left(\kappa_g^{-1} \epsilon_d^{-3/2}, \kappa_r^{-1} \epsilon_p^{-1/2^i}\right).$$

successful iterations. As before, combining this with Lemma 3.6 for $\epsilon = \epsilon_p \epsilon_d$, accounting for the max in (4.11) and remembering that we need to evaluate the function and gradient at the final x_{k+1} yields the bound (4.17). □

Comparing the bounds in Theorems 3.12 and 4.7, there seems little theoretical advantage (aside from constants) in using regularization of order more than three. We note, however, that the constants in the complexity bounds in Section 3 depend (inversely) on $\sigma_{\min}$, while those in Section 4 do not; whether this is important in practice for small chosen $\sigma_{\min}$ depends on quite how tight our bounds actually are when $r = 3$.

5 Numerical Experiments

We compare the performance of the newly proposed algorithm in the unconstrained (i.e., $\mathcal{C} = \mathbb{R}^n$) case with a Gauss-Newton method, with regularization of order two, and a Newton method, with regularization of order three. We use implementations of these algorithms found in our `RALFit` software [30], which is an open-source Fortran package for solving nonlinear least-squares problems. We apply tensor-Newton methods with regularization powers $r = 2$ and 3, and we solve the subproblem (Step 2 of Algorithm 2.1) by calling the `RALFit` code recursively; see [21] for details.

Table 5.1 reports the number of iterations, function evaluations, and Jacobian evaluations needed to solve the 26 problems in the NIST nonlinear regression test set [29]. We also include the median numbers over all tests.

Table 5.1 reports that, for most problems in the test set, the tensor-Newton methods required fewer iterations, function evaluations, and Jacobian evaluations. We can learn more about the performance of individual problems by looking at convergence curves that plot the gradient, $\|J^T r\|$, at each iteration; we give these for a number of the problems, chosen to represent different behaviours, in Figure 5.1. As should be expected, the asymptotic convergence rate of the Newton approximation is better than that of Gauss-Newton. We also see that, despite the inferior asymptotic convergence rate of Gauss-Newton, it often converges in fewer iterations that Newton due to the fact that it takes longer for Newton to enter this asymptotic regime (see, e.g., [13]). This is the case in Figures 5.1a, 5.1b, and 5.1c (see also Table 5.1). Our newly proposed tensor-Newton algorithm seems to converge at the same asymptotic rate as Newton, but with this regime being entered into much earlier, as is typical of Gauss-Newton. We credit this behaviour to the fact that, unlike Newton, the Gauss-Newton and tensor-Newton models are themselves sums-of-squares. We note that, although we observe something close to quadratic convergence in practice, whether this is always the asymptotic convergence rate is an open question (but see Appendix B).

Figure 5.1d shows convergence curves for one of the few tests where the performance of tensor-Newton is worse than that of the alternatives. All four methods struggle with this problem initially, but Gauss-Newton and Newton fall into the asymptotic regime first. Figure 5.1c, by contrast, shows an example where both variants of tensor-Newton perform much better than Gauss-Newton/Newton, which both suffer from a long period of stagnation.

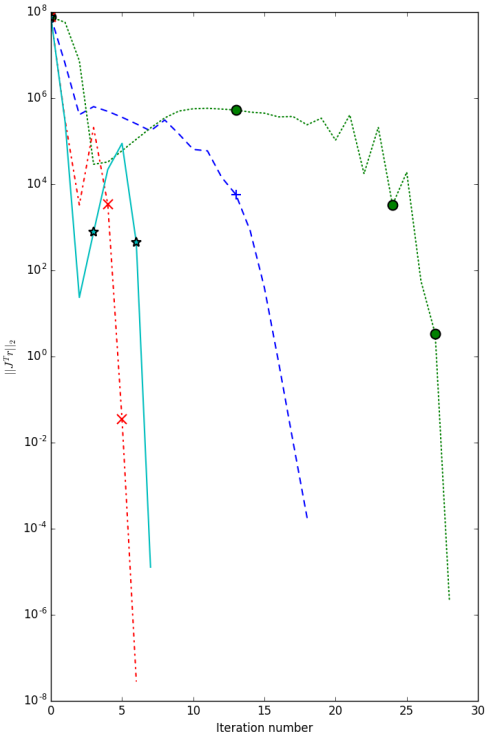
The NIST examples are generally too small to make useful time comparisons. In Table 5.2 we report timings for those where at least one of the solvers took over 0.5s. These computations were performed on a desktop machine running Linux Mint 18.2, with an Intel Core i7-7700 and 16GB RAM, and we used the `gfortran` compiler.

We see that the cost of carrying out an iteration of the tensor-Newton method is significantly higher than that of Gauss-Newton/Newton, but there are examples (e.g., BENNETT5, MGH17) where it is the fastest.

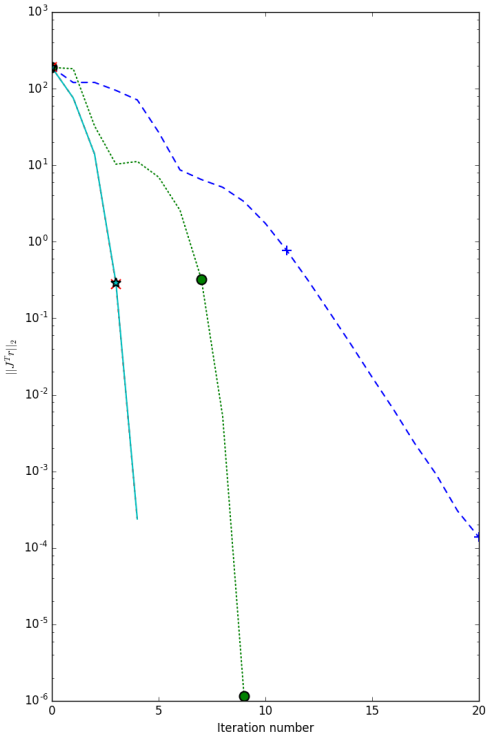
In order to demonstrate the behaviour of the algorithms with an expensive function evaluation, we performed an experiment where we read in the data at each function/derivative evaluation from a directory stored on a remote computer. We performed this test for the example closest to the median behaviour in Table 5.1: MISRA1B. Here, Gauss-Newton took 0.108 seconds, Newton 0.148 seconds, and tensor-Newton 0.004 seconds. This highlights that, while more work needs to be done per iteration in the tensor-Newton method, once the function has been evaluated and

Problem	Gauss-Newton			Newton			tensor Newton ($r = 2$)			tensor Newton ($r = 3$)		
	it	fe	je	it	fe	je	it	fe	je	it	fe	je
BENNETT5	429	436	430	597	880	598	**4**	**5**	**5**	**4**	**5**	**5**
BOXBOD	36	64	37	6	8	7	**3**	**4**	**4**	4	5	5
CHWIRUT1	14	20	15	13	16	14	**4**	**5**	**5**	**4**	**5**	**5**
CHWIRUT2	13	19	14	11	14	12	**4**	**5**	**5**	**4**	**5**	**5**
DANWOOD	7	8	8	10	11	11	**4**	**5**	**5**	**4**	**5**	**5**
ECKERLE4	21	40	22	**1**	**2**	**2**	3	4	4	3	4	4
ENSO	20	26	21	9	12	10	**4**	**5**	**5**	**4**	**5**	**5**
GAUSS1	5	6	6	7	8	8	**3**	**4**	**4**	**3**	**4**	**4**
GAUSS2	6	7	7	7	8	8	**3**	**4**	**4**	**3**	**4**	**4**
GAUSS3	7	8	8	9	10	10	**3**	**4**	**4**	**3**	**4**	**4**
HAHN1	19	**20**	20	50	86	51	17	29	18	**16**	26	**17**
LANCZOS1	67	68	68	35	48	36	38	63	39	**28**	**45**	**29**
LANCZOS2	68	69	69	35	49	36	38	63	39	**28**	**45**	**29**
LANCZOS3	121	122	122	36	51	37	41	66	42	**30**	**47**	**31**
MGH09	141	156	142	-5000	-7296	-5001	54	101	55	**32**	**50**	**33**
MGH10	-5000	-5016	-5001	481	840	482	86	168	87	**55**	**96**	**56**
MGH17	37	67	38	3113	3340	3114	**3**	**4**	**4**	7	9	8
MISRA1A	22	24	23	34	49	35	**6**	**7**	**7**	8	9	9
MISRA1B	18	20	19	28	40	29	**6**	**7**	**7**	7	8	8
MISRA1C	10	11	11	29	40	30	**6**	**7**	**7**	7	8	8
MISRA1D	13	14	14	34	47	35	**6**	**7**	**7**	7	8	8
NELSON	**71**	**81**	**72**	81	124	82	167	310	168	341	462	342
RAT42	9	10	10	31	50	32	**4**	**5**	**5**	**4**	**5**	**5**
RAT43	18	19	19	25	32	26	7	11	8	**5**	**6**	**6**
ROSZMAN1	21	30	**22**	**17**	142	18	24	**25**	25	146	147	147
THURBER	33	34	34	26	27	27	**5**	**6**	**6**	9	11	10
median	20.5	25.0	21.5	28.5	43.5	29.5	**5.5**	**6.5**	**6.5**	7.0	8.0	8.0

Table 5.1: Results for the NIST test set. A negative value indicates that the method did not converge within 5000 iterations. `it`: iterations, `fe`: function evaluations, `je`: Jacobian evaluations. The best performer in each category is boldface.

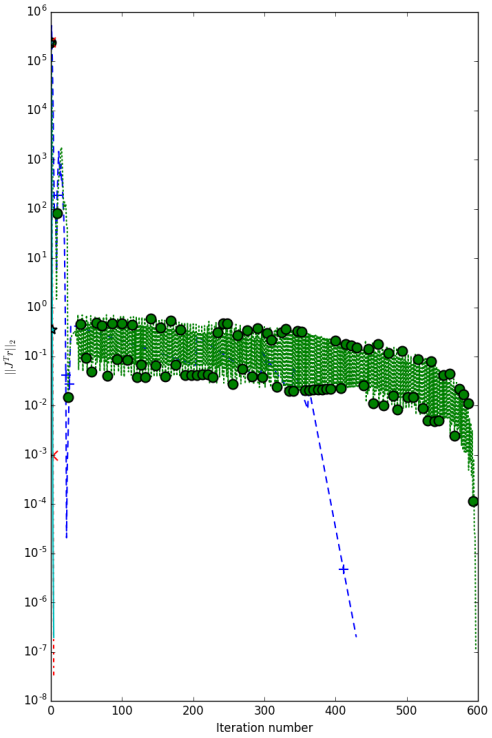


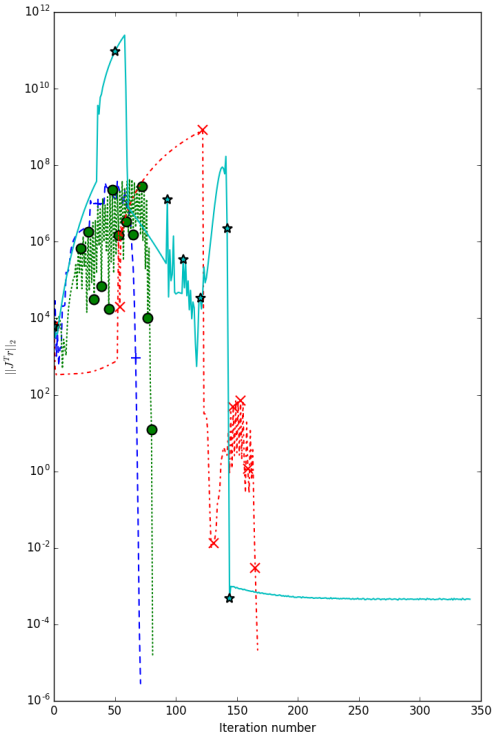
(a) MISRA1B: $\frac{1}{2}\|r(x)\| = 0.0377$



(b) ENSO: $\frac{1}{2}\|r(x)\| = 394.3$

Gauss-Newton Newton tensor-Newton (p=2) tensor-Newton (p=3)





Problem	GaussNewton	Newton	tensor Newton ($r = 2$)	tensor Newton ($r = 3$)
BENNETT5	0.40 (429)	0.59 (597)	0.03 (4)	0.08 (4)
HAHN1	0.01 (19)	0.04 (50)	0.56 (17)	0.58 (16)
LANCZOS1	0.01 (67)	0.02 (35)	0.55 (38)	0.28 (28)
LANCZOS2	$<$ 0.01 (68)	0.02 (35)	0.14 (38)	0.52 (28)
LANCZOS3	0.02 (121)	0.03 (36)	0.56 (41)	0.18 (30)
MGH09	0.01 (141)	0.62 (-5000)	0.23 (54)	0.03 (32)
MGH10	0.31 (-5000)	0.36 (481)	0.18 (86)	0.59 (55)
MGH17	0.01 (37)	0.74 (3113)	$<$ 0.01 (3)	$<$ 0.01 (7)
NELSON	0.01 (71)	0.05 (81)	0.51 (167)	0.90 (341)
ROSZMAN1	$<$ 0.01 (21)	$<$ 0.01 (17)	0.01 (24)	0.55 (146)

Table 5.2: Wallclock timings (seconds), with the number of iterations in brackets, for NIST problems where at least one solver took over 0.5s. A negative number of iterations means the method did not converge.

the derivatives calculated, it makes greater use of the information, which can lead to a faster solution time.

6 Conclusions

We have proposed and analysed a related pair of tensor-Newton algorithms for solving nonlinear least-squares problems. Under reasonable assumptions, the algorithms have been shown to converge globally to a first-order critical point. Moreover, their function-evaluation complexity is as good as the best-known algorithms for such problems. In particular, convergence to an ϵ-first-order critical point of the sum-of-squares objective (1.1) requires at most $O\left(\epsilon^{-\min(r/(r-1),3/2)}\right)$ function evaluations with r-th-order regularization with $r \geq 2$. Moreover, convergence to a point that satisfies the more natural convergence criteria (1.8) takes at most $O\left(\max(\epsilon_d^{-\min(r/(r-1),3/2)}, \epsilon_p^{-1/2^i}\right)$ evaluations for any chosen $i \geq \lceil \log_2 ((r-1)/(r-2)) \rceil$. Whether such bounds may be achieved is an open question.

Although quadratic ($r = 2$) regularization produces the poorest theoretical worst-case bound in the above, in practice it often performs well. Moreover, although quadratic regularization is rarely mentioned for general optimization in the literature (but see [2] for a recent example), it is perhaps more natural in the least-squares setting since the Gauss- and tensor-Newton approximations (2.2) are naturally bounded from below and thus it might be argued that regularization need not be so severe. The rather weak dependence of the second bound above on ϵ_p is worth noting. Indeed, increasing i reduces the influence, but of course the constant hidden by the $O(\cdot)$ notation grows with i. A similar improvement on the related bound in [9, Theorem 3.2] is possible using the same arguments.

It is also possible to imagine generalizations of the methods here in which the quadratic tensor-Newton model in (2.1) is replaced by a $p-$th-order Taylor approximation ($p > 2$). One might then anticipate evaluation-complexity bounds in which the exponents $\min(r/(r-1), 3/2)$ mentioned above are replaced by $\min(r/(r-1), (p+1)/p)$, along the lines considered elsewhere [11,12]. The

limiting applicability will likely be the cost of computing higher-order derivative tensors.

An open question relates to the asymptotic rates of convergence of our methods. It is well known that Gauss-Newton methods converge quadratically for rank-deficient problems under reasonable assumptions, but that a Newton-like method is needed to achieve this rate when the optimal residuals are nonzero. It is not clear what the rate is for our tensor-Newton method. The main obstacle to a convincing analysis is that, unlike its quadratic counterpart, a quartic model such as used by the tensor-Newton may have multiple minimizers. Our inner-iteration stopping criteria make no attempt to distinguish, indeed to do so would require global optimality conditions. In practice, however, we generally observe at least quadratic convergence, sometimes even faster when the optimal residuals are zero. In Appendix B, we indicate that a reasonable choice of the step s_k in Algorithm 2.1 does indeed converge with an asymptotic Q rate of $r-1$ for $2 < r < 3$ under standard assumptions. Extending this to Algorithm 4.1 is less obvious as it is unclear that the additional required acceptance test (4.2) might not deny an otherwise rapidly-converging natural choice of the step.

Our interest in these algorithms has been prompted by observed good behaviour when applied to practical unconstrained problems [21]. The resulting software is available as part of the `RALFit` [30] and GALAHAD [20] software libraries.

Acknowledgement

The authors are grateful to two referees and the editor for their very helpful comments on an earlier version of this paper.

References

[1] 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*, 163(1):359–368, 2017.

[2] E. G. Birgin and J. M. Martinez. Quadratic regularization with cubic descent for unconstrained optimization. Technical Report MCDO271016, State University of Campinas, Brazil, 2016.

[3] Å. Björck. *Numerical Methods for Least Squares Problems*. SIAM, Philadelphia, USA, 1996.

[4] A. Bouaricha and R. B. Schnabel. Algorithm 768: TENSOLVE: a software package for solving systems of nonlinear equations and nonlinear least-squares problems. *ACM Transactions on Mathematical Software*, 23(2):174–195, 1997.

[5] A. Bouaricha and R. B. Schnabel. Tensor methods for large, sparse nonlinear least squares problems. *SIAM Journal on Scientific and Statistical Computing*, 21(4):1199–1221, 1999.

[6] C. Cartis, N. I. M. Gould, and Ph. L. Toint. On the complexity of steepest descent, Newton's method and regularized Newton's methods for nonconvex unconstrained optimization problems. *SIAM Journal on Optimization*, 20(6):2833–2852, 2010.

[7] C. Cartis, N. I. M. Gould, and Ph. L. Toint. Adaptive cubic regularisation methods for unconstrained optimization. Part I: motivation, convergence and numerical results. *Mathematical Programming, Series A*, 127(2):245–295, 2011.

[8] C. Cartis, N. I. M. Gould, and Ph. L. Toint. Adaptive cubic regularisation methods for unconstrained optimization. Part II: worst-case function and derivative-evaluation complexity. *Mathematical Programming, Series A*, 130(2):295–319, 2011.

[9] C. Cartis, N. I. M. Gould, and Ph. L. Toint. On the evaluation complexity of cubic regularization methods for potentially rank-deficient nonlinear least-squares problems and its relevance to constrained nonlinear optimization. *SIAM Journal on Optimization*, 23(3):1553–1574, 2013.

[10] C. Cartis, N. I. M. Gould, and Ph. L. Toint. Evaluation complexity bounds for smooth constrained nonlinear optimization using scaled KKT conditions and high-order models. Report naXys-11-2015(R1), University of Namur, Belgium, 2015.

[11] C. Cartis, N. I. M. Gould, and Ph. L. Toint. Improved worst-case evaluation complexity for potentially rank-deficient nonlinear least-Euclidean-norm problems using higher-order regularized models. Technical Report RAL-TR-2015-011, Rutherford Appleton Laboratory, Chilton, Oxfordshire, England, 2015.

[12] C. Cartis, N. I. M. Gould, and Ph. L. Toint. Universal regularization methods-varying the power, the smoothness and the accuracy. Preprint RAL-P-2016-010, Rutherford Appleton Laboratory, Chilton, Oxfordshire, England, 2016.

[13] P. Chen. Hessian matrix vs. Gauss–Newton Hessian matrix. *SIAM Journal on Numerical Analysis*, 49(4):1417–1435, 2011.

[14] A. R. Conn, N. I. M. Gould, and Ph. L. Toint. Global convergence of a class of trust region algorithms for optimization with simple bound. *SIAM Journal on Numerical Analysis*, 25(2):433–460, 1988. See also same journal 26(3):764–767, 1989.

[15] A. R. Conn, N. I. M. Gould, and Ph. L. Toint. *Trust-Region Methods*. SIAM, Philadelphia, 2000.

[16] F. E. Curtis, Z. Lubberts and D. P. Robinson. Concise complexity analyses for trust-region methods. Technical Report 01-2018, Johns Hopkins University, Baltimore, MD, USA, 2018.

[17] F. E. Curtis, D. P. Robinson and M. Samadi. A trust region algorithm with a worst-case iteration complexity of $O(\epsilon^{-3/2})$ for nonconvex optimization. *Mathematical Programming*, 162(1-2):1–32, 2017.

[18] J. E. Dennis, D. M. Gay, and R. E. Welsh. An adaptive nonlinear least squares algorithm. *ACM Transactions on Mathematical Software*, 7(3):348–368, 1981.

[19] P. E. Gill and W. Murray. Algorithms for the solution of the nonlinear least squares problem. *SIAM Journal on Numerical Analysis*, 15(5):977–992, 1978.

[20] N. I. M. Gould, D. Orban, and Ph. L. Toint. GALAHAD—a library of thread-safe Fortran 90 packages for large-scale nonlinear optimization. *ACM Transactions on Mathematical Software*, 29(4):353–372, 2003.

[21] N. I. M. Gould, T. Rees, and J. A. Scott. A higher order method for solving nonlinear least-squares problems. Technical Report RAL-P-2017-010, STFC Rutherford Appleton Laboratory, 2017.

[22] G. N. Grapiglia and Y. Nesterov. Regularized Newton methods for minimizing functions with Hölder continuous Hessians. *SIAM Journal on Optimization*, 27(1):478:506, 2017.

[23] K. Levenberg. A method for the solution of certain problems in least squares. *Quarterly of Applied Mathematics*, 2(2):164–168, 1944.

[24] D. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *SIAM Journal on Applied Mathematics*, 11(2):431–441, 1963.

[25] J. J. Moré. The Levenberg-Marquardt algorithm: implementation and theory. In G. A. Watson, editor, *Numerical Analysis, Dundee 1977*, number 630 in Lecture Notes in Mathematics, pages 105–116, Heidelberg, Berlin, New York, 1978. Springer Verlag.

[26] D. D. Morrison. Methods for nonlinear least squares problems and convergence proofs. In J. Lorell and F. Yagi, editors, *Proceedings of the Seminar on Tracking Programs and Orbit Determination*, pages 1–9, Pasadena, USA, 1960. Jet Propulsion Laboratory.

[27] Yu. Nesterov. *Introductory lectures on convex optimization*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 2004.

[28] Yu. Nesterov and B. T. Polyak. Cubic regularization of Newton method and its global performance. *Mathematical Programming*, 108(1):177–205, 2006.

[29] NIST Nonlinear Regression Datasets. `http://www.itl.nist.gov/div898/strd/nls/nls_main.shtml`. Accessed June 2018.

[30] `RALFit`. `https://github.com/ralna/RALFit`. Accessed: 2018-07-20.

[31] M. K. Transtrum, B. B. Machta, and J. P. Sethna. Why are nonlinear fits to data so challenging? *Physical Review Letters*, 104(6):060201, 2010.

[32] M. K. Transtrum and J. P. Sethna. Geodesic acceleration and the small-curvature approximation for nonlinear least squares. arXiV.1207.4999, 2012.

[33] H. Zhang and A. R. Conn. On the local convergence of a derivative-free algorithm for least-squares minimization. *Computational Optimization and Applications*, 51(2):481–507, 2012.

[34] H. Zhang, A. R. Conn, and K. Scheinberg. A derivative-free algorithm for least-squares minimization. *SIAM Journal on Optimization*, 20(6):3555–3576, 2010.

Appendix A: Proofs of function bounds (3.1)–(3.4)

We assume that $r_i(x)$, $i = 1, \ldots, m$ are twice-continuously differentiable, and that they and their first two derivatives are Lipschitz on the intervals $\mathcal{F}_k = \{x : x = x_k + \alpha s_k \text{ for some } \alpha \in [0,1]\}$. Therefore

$$\|r(x) - r(y)\| \leq L_{\mathrm{r}}\|x-y\|, \;\; \|J(x) - J(y)\| \leq L_{\mathrm{j}}\|x-y\| \;\text{ and }\; \|\nabla_{xx} r_i(x) - \nabla_{xx} r_i(y)\| \leq L_{\mathrm{h}}\|x-y\| \tag{A.1}$$

for x, $y \in \mathcal{F}_k$. Moreover, these Lipschitz bounds imply that

$$\|\nabla_x r_i(x)\| \leq L_{\mathrm{r}}, \;\; \|J(x)\| \leq L_{\mathrm{r}} \;\text{ and }\; \|\nabla_{xx} r_i(x)\| \leq L_{\mathrm{j}} \tag{A.2}$$

for $x \in \mathcal{F}_k$ [27, Lemma 1.2.2]. It follows from Taylor's theorem and (A.1) that

$$|r_i(x_k + s_k) - t_i(x_k, s_k)| \leq \tfrac{1}{6} L_{\mathrm{h}} \|s_k\|^3, \tag{A.3}$$

and from the definition (2.1) of $t_i(x,s)$, the Cauchy-Schwarz inequality, (A.2) and the monotonicity bound

$$|r_i(x_k)| \leq \|r(x_k)\| \leq \|r(x_0)\| \tag{A.4}$$

that

$$\begin{aligned} |t_i(x_k, s_k)| &\leq |r_i(x_k)| + \|\nabla_x r_i(x_k))\|\|s_k\| + \tfrac{1}{2}\|\nabla_{xx} r_i(x_k)\|\|s_k\|^2 \\ &\leq \|r(x_0)\| + L_{\mathrm{r}}\|s_k\| + \tfrac{1}{2} L_{\mathrm{j}} \|s_k\|^2. \end{aligned} \tag{A.5}$$

But, using (A.3)–(A.5),

$$\begin{aligned} |r_i^2(x_k + s_k) - t_i^2(x_k, s_k)| &= |r_i(x_k + s_k) - t_i(x_k, s_k)||r_i(x_k + s_k) + t_i(x_k, s_k)| \\ &\leq \tfrac{1}{6} L_{\mathrm{h}} \|s_k\|^3 (|2t_i(x_k, s_k)| + L_{\mathrm{h}}\|s_k\|^3) \\ &\leq \tfrac{1}{6} L_{\mathrm{h}} \|s_k\|^3 (2\|r(x_0)\| + 2L_{\mathrm{r}}\|s_k\| + L_{\mathrm{j}}\|s_k\|^2 + L_{\mathrm{h}}\|s_k\|^3). \end{aligned}$$

Thus if $\|s_k\| \leq 1$, it follows from the triangle inequality that

$$|\tfrac{1}{2}\|r(x_k + s_k)\|^2 - \tfrac{1}{2}\|t(x_k, s_k)\|^2| \leq \tfrac{1}{12} m L_{\mathrm{h}} (2\|r(x_0)\| + 2L_{\mathrm{r}} + L_{\mathrm{j}} + L_{\mathrm{h}})$$

which provides the bound (3.1) with $L_{\mathrm{f}} := \frac{1}{12} m L_{\mathrm{h}} (2\|r(x_0)\| + 2L_{\mathrm{r}} + L_{\mathrm{j}} + L_{\mathrm{h}})$.

Taylor's theorem once again gives that

$$\|\nabla_x r_i(x_k + s_k) - \nabla_s t_i(x_k, s_k)\| \leq \tfrac{1}{2} L_{\mathrm{j}} \|s\|^2. \tag{A.6}$$

But then the triangle inequality together with (A.3), (A.5) and (A.6) give

$$\begin{aligned} &\|r_i(x_k + s_k)\nabla_x r_i(x_k + s_k) - t_i(x_k, s_k)\nabla_s t_i(x_k, s_k)\| \\ &= \|(r_i(x_k + s_k) - t_i(x_k, s_k))\nabla_x r_i(x_k + s_k) + t_i(x_k, s_k)(\nabla_x r_i(x_k + s_k) - \nabla_s t_i(x_k, s_k))\| \\ &\leq |r_i(x_k + s_k) - t_i(x_k, s_k)|\|\nabla_x r_i(x_k + s_k)\| + |t_i(x_k, s_k)|\|\nabla_x r_i(x_k + s_k) - \nabla_s t_i(x_k, s_k)\| \\ &\leq \tfrac{1}{6} L_{\mathrm{h}} L_{\mathrm{j}} \|s_k\|^3 + \tfrac{1}{2} L_{\mathrm{j}} (\|r(x_0)\| + L_{\mathrm{r}}\|s_k\| + \tfrac{1}{2} L_{\mathrm{j}} \|s_k\|^2) \|s_k\|^2. \end{aligned}$$

Hence, if $\|s_k\| \leq 1$, we have that

$$|\Phi(x_k + s_k) - m(x_k, s_k)| \leq m\left(\tfrac{1}{6} L_{\mathrm{h}} L_{\mathrm{j}} + \tfrac{1}{2} L_{\mathrm{j}} (\|r(x_0)\| + L_{\mathrm{r}} + \tfrac{1}{2} L_{\mathrm{j}})\right),$$

which is (3.2) with $L_{\mathrm{g}} := m(\frac{1}{6} L_{\mathrm{h}} L_{\mathrm{j}} + \frac{1}{2} L_{\mathrm{j}} (\|r(x_0)\| + L_{\mathrm{r}} + \frac{1}{2} L_{\mathrm{j}})$.

The bound (3.3) follows immediately from Cauchy-Schwarz and (A.2) with $L_{\rm J} := L_{\rm r}$. Finally (A.2), (A.4) and the well-known relationship $\|\cdot\|_1 \leq \sqrt{m}\|\cdot\|$ between the ℓ_1 and Euclidean norms give

$$\|H(x_k, r(x_k))\| = \left\| \sum_{i=1}^{m} r_i(x_k) \nabla_{xx} r_i(x_k) \right\| \leq \sum_{i=1}^{m} |r_i(x_k)| \|\nabla_{xx} r_i(x_k)\| \leq \|r(x_k)\|_1 L_{\rm j} \leq \sqrt{m} L_{\rm j} \|r(x_0)\|,$$

which is (3.4) with $L_{\rm H} := \sqrt{m} L_{\rm j}$.

Appendix B: Superlinear convergence

We focus on Algorithm 2.1 and[2] the case $2 < r < 3$. Denote the leftmost eigenvalue of a generic real symmetric matrix H by $\lambda_{\min}[H]$. Consider the gradient $\nabla_s m^{\rm R}(x, s, \sigma)$ of the regularized model given by (2.5). It follows from (2.1) and (2.2) that

$$\begin{aligned} &\nabla_s m^{\rm R}(x_k, s, \sigma_k) \\ &= \sum_{i=1}^{m} \left(r_i(x_k) + s^T \nabla_x r_i(x_k) + \tfrac{1}{2} s^T \nabla_{xx} r_i(x_k) s \right) \left(\nabla_x r_i(x_k) + \nabla_{xx} r_i(x_k) s \right) + \sigma_k \|s\|^{r-2} s \\ &= g_k + (H_k + \sigma_k \|s\|^{r-2} I) s + \sum_{i=1}^{m} \left(s^T \nabla_x r_i(x_k) \right) \nabla_{xx} r_i(x_k) s \\ &\quad + \tfrac{1}{2} \sum_{i=1}^{m} \left(s^T \nabla_{xx} r_i(x_k) s \right) \nabla_x r_i(x_k) + \tfrac{1}{2} \sum_{i=1}^{m} \left(s^T \nabla_{xx} r_i(x_k) s \right) \nabla_{xx} r_i(x_k) s, \end{aligned} \tag{B.1}$$

where for brevity we have written

$$g_k := \nabla_x \Phi(x_k) \equiv J^T(x_k) r(x_k) \ \text{ and } \ H_k := \nabla_{xx} \Phi(x_k) \equiv H(x_k, r(x_k)) + J^T(x_k) J(x_k).$$

Ideally one might hope to choose s in (B.1) to make $\nabla_s m^{\rm R}(x_k, s, \sigma_k) = 0$, but this is generally unrealistic as $\nabla_s m^{\rm R}(x, s, \sigma)$ is a combination of a cubic function and the derivative of the regularization term. A tractable compromise is to pick $s = s_k^{\rm N}$, so that

$$(H_k + \lambda_k I) s_k^{\rm N} = -g_k, \tag{B.2}$$

where

$$\lambda_k := \sigma_k \|s_k^{\rm N}\|^{r-2} \geq 0. \tag{B.3}$$

since this provides a zero of the lower-order terms in (B.1).

We will try $s_k = s_k^{\rm N}$ if H_k is positive definite, with leftmost eigenvalue $\lambda_{\min,k} := \lambda_{\min}[H_k] > 0$, and three essential properties hold, namely that

$$m^{\rm R}(x_k, s_k^{\rm N}, \sigma_k) < m^{\rm R}(x_k, 0, \sigma_k), \tag{B.4}$$

$$\|\nabla_s m^{\rm R}(x_k, s_k^{\rm N}, \sigma_k)\| \leq \theta \|s_k^{\rm N}\|^{r-1} \ \text{ and} \tag{B.5}$$

$$\frac{\Phi(x_k) - \Phi(x_k + s_k^{\rm N})}{m(x_k, 0) - m(x_k, s_k^{\rm N})} \geq \eta_1 \tag{B.6}$$

[2] It is unclear what happens when $r = 2$ or 3.

If so, s_k^{N} provides a successful step in Algorithm 2.1, since (B.4)–(B.6) are then that (2.7)–(2.8) and $\rho_k \geq \eta_1$ hold, We are not specific about how s_k is chosen when H_k is not positive definite, nor how s_k might be chosen if s_k^{N} does not provide a successful step.

Consider the sub-sequence of iterates $\{x_k\}$, $k \in \mathcal{K}$, whose limit is x_* (and thus for which $g_* := \nabla_x \Phi(x_*) = 0$ because of Theorem 3.9), suppose that $\nabla_x \Phi(x)$ is Lipschitz continuous in an open neighbourhood of x_* and that $\lambda_{\min,*} := \lambda_{\min}[\nabla_{xx}\Phi(x_*)] > 0$. Then, for all $k \in \mathcal{K}$ sufficiently large, $\lambda_{\min,k} \geq \frac{1}{2}\lambda_{\min,*}$. This ensures that

$$\|(H_k + \lambda_k I)^{-1}\| \leq \frac{1}{\lambda_{\min,k} + \lambda_k} \leq \frac{1}{\lambda_{\min,k}} \leq \frac{2}{\lambda_{\min,*}}, \tag{B.7}$$

and hence (B.2) and (B.7) provides the bound

$$\|s_k^{\mathrm{N}}\| \leq \|(H_k + \lambda_k I)^{-1}\| \|g_k\| \leq \frac{2\|g_k\|}{\lambda_{\min,*}}. \tag{B.8}$$

But Lipschitz continuity and Taylor's theorem applied to $\nabla_x \Phi(x)$ yields

$$\|g_k\| = \|g_* - g_k\| \leq L_1 \|x_* - x_k\|$$

and

$$\|g_* - g_k - H_k (x_* - x_k)\|_2 \leq L_2 \|x_* - x_k\|_2^2 \tag{B.9}$$

for some constants $L_1, L_2 > 0$, and thus

$$\|s_k^{\mathrm{N}}\| \leq \frac{2L_1}{\lambda_{\min,*}} \|x_* - x_k\| \tag{B.10}$$

because of (B.8).

Define

$$\kappa_s := \frac{2L_2}{\lambda_{\min,*}} \left(L_2 + \sigma_{\max} \left(\frac{2L_1}{\lambda_{\min,*}} \right)^{r-2} \right), \tag{B.11}$$

where $\sigma_{\max}$ is given by (3.17), and suppose that $x_k \in \mathcal{X}$, where

$$\mathcal{X} = \left\{ x \;\middle|\; \begin{array}{l} x \in \mathcal{B}_\delta \;\text{ and }\; \|x - x_*\| \leq \min\left(\left(\dfrac{1}{2\kappa_s} \right)^{1/(r-2)}, \right. \\ \left. \dfrac{\lambda_{\min,*}}{2L_1} \min\left[1, \min\left(\dfrac{(r-2)\sigma_{\min}}{mrL_{\mathrm{j}}(L_{\mathrm{r}} + L_{\mathrm{j}})}, \dfrac{2\theta}{mL_{\mathrm{j}}(3L_{\mathrm{r}} + L_{\mathrm{j}})}, \dfrac{\sigma_{\min}(1-\eta_2)}{rL_{\mathrm{f}}} \right)^{1/(3-r)} \right] \right) \end{array} \right\}. \tag{B.12}$$

and $\mathcal{B}_\delta = \{x \mid \|x - x_*\| \leq \delta\}$ is any ball around x_* of fixed radius $\delta > 0$ for which $\lambda_{\min}[\nabla_{xx}\Phi(x)] \geq \frac{1}{2}\lambda_{\min,*}$ for all $x \in \mathcal{B}_\delta$. In this case (B.10) guarantees that

$$\|s_k^{\mathrm{N}}\| \leq \min\left[1, \left(\frac{(r-2)\sigma_{\min}}{mrL_{\mathrm{j}}(L_{\mathrm{r}} + L_{\mathrm{j}})} \right)^{1/(3-r)}, \left(\frac{2\theta}{mL_{\mathrm{j}}(3L_{\mathrm{r}} + L_{\mathrm{j}})} \right)^{1/(3-r)}, \left(\frac{\sigma_{\min}(1-\eta_2)}{rL_{\mathrm{f}}} \right)^{1/(3-r)} \right], \tag{B.13}$$

and hence, trivially,

$$\|s_k^{\mathrm{N}}\|^3 \leq \|s_k^{\mathrm{N}}\|^2 \leq \|s_k^{\mathrm{N}}\|^{r-1} \leq \|s_k^{\mathrm{N}}\|. \tag{B.14}$$

We now establish the required bounds (B.4)–(B.6). Firstly, expanding the definition (2.2) of $m(x,s)$ gives

$$\begin{aligned} m(x_k,s_k^{\rm N}) - m(x_k,0) &= g_k^T s_k^{\rm N} + \tfrac{1}{2} {s_k^{\rm N}}^T H_k s_k^{\rm N} + e(x_k,s_k^{\rm N}),\\ \text{where } \; e(x_k,s_k^{\rm N}) &:= \tfrac{1}{2}\sum_{i=1}^m {s_k^{\rm N}}^T \nabla_x r_i(x_k) {s_k^{\rm N}}^T \nabla_{xx} r_i(x_k) s_k^{\rm N} + \tfrac{1}{8}\sum_{i=1}^m \left({s_k^{\rm N}}^T \nabla_{xx} r_i(x_k) s_k^{\rm N}\right)^2, \end{aligned} \tag{B.15}$$

and it follows directly from the Cauchy-Schwarz inequality, (A.2) and (B.14) that

$$e(x_k,s_k^{\rm N}) \le \tfrac{1}{2} m \|s_k^{\rm N}\|^3 L_{\rm r} L_{\rm j} + \tfrac{1}{8} m L_{\rm j}^2 \|s_k^{\rm N}\|^4 < \tfrac{1}{2} m L_{\rm j}(L_{\rm r}+L_{\rm j})\|s_k^{\rm N}\|^3. \tag{B.16}$$

Substituting (B.15) into the definition (2.3) of the regularized model $m^{\rm R}(x,s,\sigma)$ gives

$$\begin{aligned} m^{\rm R}(x_k,s_k^{\rm N},\sigma_k) - m^{\rm R}(x_k,0,\sigma_k) &= m(x_k,s_k^{\rm N}) - m(x_k,0) + \tfrac{\sigma_k}{r}\|s_k^{\rm N}\|^r\\ &= g_k^T s_k^{\rm N} + \tfrac{1}{2}{s_k^{\rm N}}^T H_k s_k^{\rm N} + \tfrac{\sigma_k}{r}\|s_k^{\rm N}\|^r + e(x_k,s_k^{\rm N})\\ &= -\tfrac{1}{2}{s_k^{\rm N}}^T (H_k+\lambda_k I) s_k^{\rm N} - \tfrac{r-2}{2r}\sigma_k \|s_k^{\rm N}\|^r + e(x_k,s_k^{\rm N})\\ &< -\tfrac{1}{2}\tfrac{r-2}{r}\sigma_{\min}\|s_k^{\rm N}\|^r + \tfrac{1}{2} m L_{\rm j}(L_{\rm r}+L_{\rm j})\|s_k^{\rm N}\|^3 < 0 \end{aligned}$$

because of the positive semi-definiteness of $H_k+\lambda_k I$, the requirement that $\sigma_k \ge \sigma_{\min} > 0$, and the bounds (A.2) and (B.16) and the second term in (B.13). This provides the required bound (B.4).

It also follows immediately from (B.1) and (B.2) that

$$\|\nabla_s m^{\rm R}(x_k,s_k^{\rm N},\sigma_k)\| \le \tfrac{3}{2} m L_{\rm r} L_{\rm j} \|s_k^{\rm N}\|^2 + \tfrac{1}{2} m L_{\rm j}^2 \|s_k^{\rm N}\|^3 \le \theta \|s_k^{\rm N}\|^{r-1}$$

using the triangle inequality, (A.2) and the third term in (B.13), which establishes (B.5).

Finally, it follows precisely as in (3.16) that

$$|\rho_k - 1| = \frac{|\Phi(x_k+s_k^{\rm N}) - m(x_k,s_k^{\rm N})|}{m(x_k,0)-m(x_k,s_k^{\rm N})} \le \frac{r L_{\rm f}}{\sigma_k}\|s_k^{\rm N}\|^{3-r} \le \frac{r L_{\rm f}}{\sigma_{\min}}\|s_k^{\rm N}\|^{3-r}$$

since $\sigma_k \ge \sigma_{\min} > 0$. Combining this with the fourth term in (B.13) immediately gives that $|\rho_k - 1| \le 1-\eta_2$ and hence that (B.6) holds. Thus we have shown that $s_k^{\rm N}$ is allowed by Step 2 of Algorithm 2.1, and leads to a successful iteration for which $x_{k+1} = x_k + s_k^{\rm N}$.

Our intention is to show that

$$\|x_{k+1} - x_*\| \le \kappa \|x_k - x_*\|^{r-1} \tag{B.17}$$

for some $\kappa > 0$, and hence the resulting iteration ultimately converges at a (Q-order $r-1$) superlinear rate. The iterate $x_{k+1} = x_k + s_k^{\rm N}$ satisfies

$$\begin{aligned} x_{k+1} - x_* &= x_k + s_k^{\rm N} - x_*\\ &= x_k - x_* - (H_k+\lambda_k I)^{-1} g_k\\ &= x_k - x_* - (H_k+\lambda_k I)^{-1}(g_k - g_*)\\ &= (H_k+\lambda_k I)^{-1}\left(g_* - g_k - (H_k+\lambda_k I)(x_* - x_k)\right)\\ &= (H_k+\lambda_k I)^{-1}\left(g_* - g_k - H_k(x_*-x_k) - \lambda_k(x_*-x_k)\right). \end{aligned} \tag{B.18}$$

Taking norms and combining this with (B.9) gives

$$\begin{aligned} \|x_{k+1}-x_*\| &\le L_2 \|(H_k+\lambda_k I)^{-1}\| \left(L_2\|x_*-x_k\|^2 + \lambda_k \|x_*-x_k\|\right)\\ &\le \frac{2L_2}{\lambda_{\min,*}}\left(L_2\|x_*-x_k\|^2 + \sigma_{\max}\left(\frac{2L_1}{\lambda_{\min,*}}\right)^{r-2}\|x_*-x_k\|^{r-1}\right)\\ &\le \kappa_s \|x_*-x_k\|^{r-1} \end{aligned} \tag{B.19}$$

using (B.18), (B.7), (B.3), (B.10) and (B.11) and the appropriate bound $\sigma_k \leq \sigma_{\max}$ from (3.17). Thus (B.17) holds. Moreover, it also follows from (B.19) and the first term in (B.12) that

$$\|x_{k+1} - x_*\| \leq \tfrac{1}{2}\|x_k - x_*\|,$$

in which case $x_{k+1} \in \mathcal{X}$ and thus (B.12) continues to hold at iteration $k+1$. Hence once an iterate enters $\mathcal{X}$, it will remain there, and the remaining sequence will converge superlinearly to x_*.