Local convergence analysis of Inexact Newton method with relative residual error tolerance under majorant condition in Riemannian Manifolds

T. Bittencourt * O. P. Ferreira†

March 14, 2014

Abstract

A local convergence analysis of Inexact Newton's method with relative residual error tolerance for finding a singularity of a differentiable vector field defined on a complete Riemannian manifold, based on majorant principle, is presented in this paper. We prove that under local assumptions, the inexact Newton method with a fixed relative residual error tolerance converges Q -linearly to a singularity of the vector field under consideration. Using this result we show that the inexact Newton method to find a zero of an analytic vector field can be implemented with a fixed relative residual error tolerance. In the absence of errors, our analysis retrieve the classical local theorem on the Newton method in Riemannian context.

Keywords: Inexact Newton's method, majorant principle, local convergence analysis, Riemannian manifold.

1 Introduction

Newton's method and its variations, including the inexact Newton methods, are the most efficient methods known for solving nonlinear equations in Banach spaces. Besides its practical applications, Newton's method is also a powerful theoretical tool with a wide range of applications in pure and applied mathematics, see [6, 11, 16, 24, 25, 27, 34, 35]. In particular, Newton's method has been instrumental in the modern complexity analysis of the solution of polynomial or analytical equations [6, 26], linear and quadratic programming problems and linear semi-definite programming problems

*IME/UFG, CP-131, CEP 74001-970 - Goiânia, GO, Brazil (Email: tiberio.b@gmail.com). This author was supported by CAPES.

†IME/UFG, CP-131, CEP 74001-970 - Goiânia, GO, Brazil (Email: orizon@ufg.br). This author was supported by CNPq Grants 302024/2008-5, 480101/2008-6 and 473756/2009-9, PRONEX–Optimization(FAPERJ/CNPq) and FUNAPE/UFG.

[15, 16, 24, 25]. In all these applications, homotopy methods are combined with Newton's method, which helps the algorithm to keep track of the solution of a parametrized perturbed version of the original problem.

In classic Newton's method, a linear equation system is solved in each iteration which can be expensive and unnecessary when the problem size is large. Inexact Newton's method comes up to overcome such drawback and can effectively cut down the computational cost by solving the linear equations approximately, see [9, 12, 22]. It would be most desirable to have an *a priori* prescribed residual error tolerance in the iterative solutions of linear system for computing the Inexact Newton steps, in order to avoid under-solving or over-solving the linear system in question. The advantage of working with an error tolerance on the residual rests in the fact that the exact Newton step need not to be know for evaluating this error, which makes this criterion attractive for practical applications, see [15, 16].

Newton's method has been extended to Riemannian manifolds with many different purposes. In particular, in the last few years, a couple of papers have dealt with the issue of convergence analysis of Newton's method for finding a singularity of a differentiable vector field defined on a complete Riemannian manifold, see [1, 2, 3, 4, 5, 8, 13, 14, 19, 20, 21, 27, 29, 30, 31, 32, 33]. Extensions to Riemannian manifolds of analyses of Newton's method under the γ-condition was given in [8, 19, 20, 21]. Although the local convergence analysis of Inexact Newton's method in Banach space with relative errors tolerance in the residue [7, 9, 22] are well understood, as far as we know, the convergence analysis of the method in Riemannian manifolds context under general local assumptions, assuming *only* bounded relative residual errors, is a new contribution of this paper. It is worth to point out that, for null error tolerance, the analysis presented merge in the usual local convergence analysis on Newton's method in Riemannian manifold under a majorant condition, see [13]. In our analysis, the classical Lipschitz condition is relaxed using a majorant function which provides a clear relationship between the majorant function and the vector field under consideration. Moreover, several unrelated previous results pertaining to Newton's method are unified (see [8, 19, 20]), now in the Riemannian context.

The organization of the paper is as follows. In Section 2, the notations and basic results used in the paper are presented. In Section 3 the main result is stated and in Section 4 some properties of the majorant function are established and the main relationships between the majorant function and the vector field used in the paper are presented. In Section 5 the main result is proved and two applications of this result are given in Section 6. Some final remarks are made in Section 7.

2 Notation and auxiliary results

In this section we recall some notations, definitions and basic properties of Riemannian manifolds used throughout the paper, they can be found, for example in [10] and [18].

Throughout the paper, $\mathcal{M}$ is a smooth manifold and $C^1(\mathcal{M})$ is the class of all continuously differentiable functions on $\mathcal{M}$. The space of vector fields on $\mathcal{M}$ is denoted by $\mathcal{X}(\mathcal{M})$, by $T_p\mathcal{M}$ we denote the tangent space of $\mathcal{M}$ at p and by $T\mathcal{M} = \bigcup_{x\in\mathcal{M}} T_x\mathcal{M}$ the *tangent bundle* of $\mathcal{M}$. Let $\mathcal{M}$ be endowed with a Riemannian metric $\langle\cdot,\cdot\rangle$, with corresponding norm denoted by $\|\cdot\|$, so that $\mathcal{M}$ is now a *Riemannian manifold.* Let us recall that the metric can be used to define the length of a piecewise C^1 curve $\zeta : [a,b] \to \mathcal{M}$ joining p to q, i.e., such that $\zeta(a) = p$ and $\zeta(b) = q$, by $l(\zeta) = \int_a^b \|\zeta'(t)\|dt$. Minimizing this length functional over the set of all such curves we obtain a distance $d(p,q)$, which induces the original topology on M. The open and closed balls of radius $r > 0$ centered at p are defined, respectively, as

$$B_r(p) := \{q \in M : d(p,q) < r\}, \qquad \overline{B}_r(p) := \{q \in M : d(p,q) \le r\}.$$

Also the metric induces a map $f \in C^1(\mathcal{M}) \mapsto \text{grad} f \in \mathcal{X}(\mathcal{M})$, which associates to each f its *gradient* via the rule $\langle \text{grad} f, X\rangle = df(X)$, for all $X \in \mathcal{X}(\mathcal{M})$. The chain rule generalizes to this setting in the usual way: $(f\circ\zeta)'(t) = \langle \text{grad} f(\zeta(t)), \zeta'(t)\rangle$, for all curves $\zeta \in C^1$. Let ζ be a curve joining the points p and q in $\mathcal{M}$ and let ∇ be a Levi-Civita connection associated to $(\mathcal{M}, \langle,\rangle)$. For each $t \in [a,b]$, ∇ induces an isometry, relative to $\langle\cdot,\cdot\rangle$,

$$\begin{aligned} P_{\zeta,a,t}\colon T_{\zeta(a)}\mathcal{M} &\longrightarrow T_{\zeta(t)}\mathcal{M} \\ v &\longmapsto P_{\zeta,a,t}\, v = V(t), \end{aligned} \tag{1}$$

where V is the unique vector field on ζ such that $\nabla_{\zeta'(t)}V(t) = 0$ and $V(a) = v$, the so-called *parallel translation* along ζ from $\zeta(a)$ to $\zeta(t)$. Note also that

$$P_{\zeta,b_1,b_2} \circ P_{\zeta,a,b_1} = P_{\zeta,a,b_2}, \qquad P_{\zeta,b,a} = P_{\zeta,a,b}{}^{-1}.$$

A vector field V along ζ is said to be *parallel* if $\nabla_{\zeta'}V = 0$. If ζ' itself is parallel, then we say that ζ is a *geodesic.* The geodesic equation $\nabla_{\zeta'}\zeta' = 0$ is a second order nonlinear ordinary differential equation, so the geodesic ζ is determined by its position p and velocity v at p. It is easy to check that $\|\zeta'\|$ is constant. We say that ζ is *normalized* if $\|\zeta'\| = 1$. A geodesic $\zeta : [a,b] \to \mathcal{M}$ is said to be *minimal* if its length is equal the distance of its end points, i.e. $l(\zeta) = d(\zeta(a), \zeta(b))$.

A Riemannian manifold is *complete* if its geodesics are defined for any values of t. The Hopf-Rinow's theorem asserts that if this is the case then any pair of points, say p and q, in $\mathcal{M}$ can be joined by a (not necessarily unique) minimal geodesic segment. Moreover, $(\mathcal{M}, d)$ is a complete metric space and bounded and closed subsets are compact.

The *exponential map* at p, $\exp_p : T_p\mathcal{M} \to \mathcal{M}$ is defined by $\exp_p v = \zeta_v(1)$, where ζ_v is the geodesic defined by its position p and velocity v at p and $\zeta_v(t) = \exp_p tv$ for any value of t. For $p \in \mathcal{M}$, let

$$r_p := \sup\left\{ r > 0 : \exp_{p|_{B_r(o_p)}} \text{ is a diffeomorphism} \right\},$$

where o_p denotes the origin of $T_p\mathcal{M}$ and $B_r(o_p) := \{v \in T_p\mathcal{M} : \|v - o_p\| < r\}$. Note that if $0 < \delta < r_p$ then $\exp_p B_\delta(o_p) = B_\delta(p)$. The number r_p is called the *injectivity radius* of $\mathcal{M}$ at p.

Definition 1. *Let $p \in \mathcal{M}$ and r_p the radius of injectivity at p. Define the quantity*

$$K_p := \sup \left\{ \frac{d(\exp_q u, \exp_q v)}{\|u - v\|} \ : \ q \in B_{r_p}(p),\ u, v \in T_q\mathcal{M},\ u \neq v,\ \|v\| \leq r_p,\ \|u - v\| \leq r_p \right\}.$$

Remark 1. *The quantity K_p measures how fast the geodesics spread apart in $\mathcal{M}$. In particular, when $u = 0$ or more generally when u and v are on the same line through o_q,*

$$d(\exp_q u, \exp_q v) = \|u - v\|.$$

So $K_p \geq 1$ for all $p \in \mathcal{M}$. When $\mathcal{M}$ has non-negative sectional curvature, the geodesics spread apart less than the rays ([10], Chap. 5) so that

$$d(\exp_q u, \exp_q v) \leq \|u - v\|.$$

As a consequence $K_p = 1$ for all $p \in \mathcal{M}$. Finally it is worth mentioning that radii less than r_p could be used as well (although this would require added notation such as $K_p(\rho)$ for r_p). In this case, the measure by which geodesics spread apart might decrease, thereby providing slightly stronger results so long as the radius was not too much less than r_p.

Let X be a C^1 vector field on $\mathcal{M}$. The covariant derivative of X determined by the Levi-Civita connection ∇ defines at each $p \in \mathcal{M}$ a linear map $\nabla X(p) : T_p\mathcal{M} \to T_p\mathcal{M}$ given by

$$\nabla X(p)v := \nabla_Y X(p), \tag{2}$$

where Y is a vector field such that $Y(p) = v$.

Definition 2. *Let $\mathcal{M}$ be a complete Riemannian manifold and $Y_1, \ldots, Y_n$ be vector fields on $\mathcal{M}$. Then, the n-th covariant derivative of X with respect to $Y_1, \ldots, Y_n$ is defined inductively by*

$$\nabla^2_{\{Y_1, Y_2\}} X := \nabla_{Y_2} \nabla_{Y_1} X, \qquad \nabla^n_{\{Y_i\}_{i=1}^n} X := \nabla_{Y_n}(\nabla_{Y_{n-1}} \cdots \nabla_{Y_1} X).$$

Definition 3. *Let $\mathcal{M}$ be a complete Riemannian manifold, and $p \in \mathcal{M}$. Then, the n-th covariant derivative of X at p is the n-th multilinear map $\nabla^n X(p) : T_p\mathcal{M} \times \ldots \times T_p\mathcal{M} \to T_p\mathcal{M}$ defined by*

$$\nabla^n X(p)(v_1, \ldots, v_n) := \nabla^n_{\{Y_i\}_{i=1}^n} X(p),$$

where $Y_1, \ldots, Y_n$ are vector fields on $\mathcal{M}$ such that $Y_1(p) = v_1, \ldots, Y_n(p) = v_n$.

We remark that Definition 3 only depends on the n-tuple of vectors $(v_1, \ldots, v_n)$ since the covariant derivative is tensorial in each vector field Y_i.

Definition 4. *Let $\mathcal{M}$ be a complete Riemannian manifold and $p \in \mathcal{M}$. The norm of an n-th multilinear map $A : T_p\mathcal{M} \times \ldots \times T_p\mathcal{M} \to T_p\mathcal{M}$ is defined by*

$$\|A\| = \sup\left\{\|A(v_1, \ldots, v_n)\| \ : \ v_1, \ldots, v_n \in T_p\mathcal{M},\ \|v_i\| = 1,\ i = 1, \ldots, n\right\}.$$

In particular the norm of the n-th covariant derivative of X at p is given by

$$\|\nabla^n X(p)\| = \sup\left\{\|\nabla^n X(p)(v_1, \ldots, v_n)\| \ : \ v_1, \ldots, v_n \in T_p\mathcal{M},\ \|v_i\| = 1,\ i = 1, \ldots, n\right\}.$$

Lemma 1. *Let Ω be an open subset of $\mathcal{M}$, X a C^1 vector field defined on Ω and $\zeta : [a, b] \to \Omega$ a C^∞ curve. Then*

$$P_{\zeta,t,a} X(\zeta(t)) = X(\zeta(a)) + \int_a^t P_{\zeta,s,a} \nabla X(\zeta(s))\, \zeta'(s)\, ds, \qquad t \in [a, b].$$

Proof. See [14]. □

Lemma 2. *Let Ω be an open subset of $\mathcal{M}$, X a C^2 vector field defined on Ω and $\zeta : [a, b] \to \Omega$ a C^∞ curve. Then for all $Y \in \mathcal{X}(\mathcal{M})$ we have that*

$$P_{\zeta,t,a} \nabla X(\zeta(t))\, Y(\zeta(t)) = \nabla X(\zeta(a)) Y(\zeta(a)) + \int_a^t P_{\zeta,s,a} \nabla^2 X(\zeta(s))(Y(\zeta(s)), \zeta'(s))\, ds, \qquad t \in [a, b].$$

Proof. See [19]. □

Lemma 3 (Banach's Lemma). *Let B be a linear operator and let I_p be the identity operator in T_pM. If $\|B - I_p\| < 1$ then B is invertible and $\|B^{-1}\| \leq 1/\left(1 - \|B - I_p\|\right)$.*

Proof. Under the hypothesis, it is easily shown that $B^{-1} = \sum_{i=0}^{\infty}(B - I_p)^i$ and hence $\|B^{-1}\| \leqslant \sum_{i=0}^{\infty} \|B - I_p\|^i = 1/(1 - \|(B - I_p)\|)$. □

3 Local analysis for Inexact Newton method

Our goal is to prove in Riemannian manifold context the following version of Inexact Newton method with relative residual error tolerance under majorant condition.

Theorem 4. *Let $\mathcal{M}$ be a Riemannian manifold, $\Omega \subseteq \mathcal{M}$ an open set and $X : \Omega \to T\mathcal{M}$ a continuously differentiable vector field. Let $p_* \in \Omega$, $R > 0$ and $\kappa := \sup\{t \in [0, R) : B_t(p_*) \subset \Omega\}$. Suppose that $X(p_*) = 0$, $\nabla X(p_*)$ is invertible and there exists an $f : [0,\ R) \to \mathbb{R}$ continuously differentiable such that*

$$\left\|\nabla X(p_*)^{-1}[P_{\zeta,1,0}\, \nabla X(p) - \ P_{\zeta,\tau,0} \nabla X(\zeta(\tau)) P_{\zeta,1,\tau}]\right\| \leq f'\left(d(p_*, p)\right) - f'\left(\tau d(p_*, p)\right), \qquad (3)$$

for all $\tau \in [0, 1]$, $p \in B_\kappa(p_*)$, where $\zeta : [0, 1] \to \mathcal{M}$ is a minimizing geodesic from p_* to p and

h1) *$f(0) = 0$ and $f'(0) = -1$;*

h2) *f' is strictly increasing.*

Let $0 \le \vartheta < 1/K_{p_*}$, $\nu := \sup\{t \in [0, R) : f'(t) < 0\}$, $\rho := \sup\{\delta \in (0, \nu) : [(1+\vartheta)|t - f(t)/f'(t)|/t + \vartheta] < 1/K_{p_*},\ t \in (0, \delta)\}$ and

$$r := \min\{\kappa,\ \rho,\ r_{p_*}\}.$$

Then the sequence generated by the Inexact Newton method for solving $X(p) = 0$ with starting point $p_0 \in B_r(p_*) \setminus \{p_*\}$ and residual relative error tolerance θ,

$$p_{k+1} = \exp_{p_k}(S_k), \qquad \|X(p_k) + \nabla X(p_k) S_k\| \le \theta \|X(p_k)\|, \qquad k = 0, 1, \ldots, \tag{4}$$

$$0 \le cond(\nabla X(p_*))\theta \le \vartheta / \left[2/|f'(d(p_*, p_0))| - 1\right], \tag{5}$$

is well defined (for any particular choice of each $S_k \in T_{p_k}M$), the sequence $\{p_k\}$ is contained in $B_r(p_*)$ and converges to the point p_* which is the unique zero of X in $B_\sigma(p_*)$, where $\sigma := \sup\{t \in (0, \kappa) : f(t) < 0\}$, and we have that:

$$d(p_*,\, p_{k+1}) \le K_{p_*} \left[(1+\vartheta) \frac{\left| d(p_*, p_k) - \dfrac{f(d(p_*, p_k))}{f'(d(p_*, p_k))} \right|}{d(p_*, p_k)} + \vartheta \right] d(p_*,\, p_k), \qquad k = 0, 1, \ldots, \tag{6}$$

and $\{p_k\}$ converges linearly to p_*. If, in additional, the function f satisfies the following condition

h3) *f' is convex,*

then there holds

$$d(p_*,\, p_{k+1}) \le K_{p_*} \left[(1+\vartheta) \frac{\left| d(p_*, p_0) - \dfrac{f(d(p_*, p_0))}{f'(d(p_*, p_0))} \right|}{d^2(p_*, p_0)} d(p_*, p_k) + \vartheta \right] d(p_*,\, p_k), \quad k = 0, 1, \ldots. \tag{7}$$

as a consequence, the sequence $\{p_k\}$ converges to p_* with linear rate as follows

$$d(p_*,\, p_{k+1}) \le K_{p_*} \left[(1+\vartheta) \frac{\left| d(p_*, p_0) - \dfrac{f(d(p_*, p_0))}{f'(d(p_*, p_0))} \right|}{d(p_*, p_0)} + \vartheta \right] d(p_*,\, p_k), \quad k = 0, 1, \ldots. \tag{8}$$

Remark 2. *First note that from simple algebraic manipulation we have the following equality*

$$\frac{\left| d(p_*, p_k) - \dfrac{f(d(p_*, p_k))}{f'(d(p_*, p_k))} \right|}{d(p_*, p_k)} = \left| 1 - \frac{1}{f'(d(p_*, p_k))} \frac{f(d(p_*, p_k)) - f(0)}{d(p_*, p_k) - 0} \right|.$$

Since the sequence $\{p_k\}$ *is contained in* $B_r(p_*)$ *and converges to the point* p_* *then it is easy to see that right hand side of last equality goes to zero as* k *goes to infinity. Therefore in Theorem 4 if taking* $\vartheta = \vartheta_k$ *in each iteration and letting* ϑ_k *goes to zero (in this case,* $\theta = \theta_k$ *also goes to zero) as* k *goes to infinity, then* (6) *implies that* $\{p_k\}$ *converges to* p_* *with asymptotic superlinear rate.*

Note that letting $\vartheta = 0$ *in Theorem 4 which implies from* (5) *that* $\theta = 0$*, the linear equation in* (4) *is solved exactly. Therefore* (7) *implies that* $\{p_k\}$ *converges to* p_* *with quadratic rate.*

From now on, we assume that the hypotheses of Theorem 4 hold with the exception of **h3**, which will be considered to hold only when explicitly stated.

4 Preliminary results

The scalar function f in Theorem 4 is called a *majorant function* for vector field X at a point p_*. In this section we analyze some basic properties of f and the main relationships between f and X.

4.1 The majorant function

We begin by proving that the constants κ, ν and σ are positives.

Proposition 5. *The constants* κ, ν *and* σ *are positives and* $t - f(t)/f'(t) < 0$ *for all* $t \in (0, \nu)$.

Proof. Since Ω is open and $p_* \in \Omega$, we conclude that $\kappa > 0$. As f' is continuous in 0 with $f'(0) = -1$, there exists $\delta > 0$ such that $f'(t) < 0$ for all $t \in (0, \delta)$, so $\nu > 0$. Because $f(0) = 0$ and f' is continuous in 0 with $f'(0) = -1$, there exists $\delta > 0$ such that $f(t) < 0$ for all $t \in (0, \delta)$, hence $\sigma > 0$.

Assumption **h2** implies that f is strictly convex, so using the strict convexity of f and the first equality in assumption **h1** we have $f(t) - tf'(t) < f(0) = 0$ for all $t \in (0, R)$. If $t \in (0, \nu)$ then $f'(t) < 0$, which combined with the last inequality yields the desired inequality. □

According to **h2** and definition of ν, we have $f'(t) < 0$ for all $t \in [0, \nu)$. Therefore Newton iteration map for f is well defined in $[0, \nu)$. Let us call it n_f,

$$\begin{aligned} n_f : [0, \nu) &\to (-\infty, 0], \\ t &\mapsto t - f(t)/f'(t). \end{aligned} \tag{9}$$

Because $f'(t) \neq 0$ for all $t \in [0, \nu)$ the Newton iteration map n_f is a continuous function.

Proposition 6. $\lim_{t\to 0} |n_f(t)|/t = 0$. *As a consequence* $\rho > 0$ *and* $(1+\vartheta)|n_f(t)|/t + \vartheta < 1/K_{p_*}$ *for all* $t \in (0,\, \rho)$.

Proof. Using definition in (9), Proposition 5, $f(0) = 0$ and definition of ν, a simple algebraic manipulation gives

$$\frac{|n_f(t)|}{t} = \frac{f(t)/f'(t) - t}{t} = \frac{1}{f'(t)} \frac{f(t) - f(0)}{t - 0} - 1, \qquad t \in (0,\, \nu). \tag{10}$$

Because $f'(0) \neq 0$ the first statement follows by taking the limit in (10) as t goes to 0.

Since $\lim_{t\to 0} |n_f(t)|/t = 0$ and $\vartheta < 1/K_{p_*}$ the first equality in (10) implies that there exists $\delta > 0$ such that

$$(1+\vartheta)[f(t)/f'(t) - t]/t + \vartheta < 1/K_{p_*}, \qquad t \in (0, \delta).$$

Therefore from definition of ρ and (9) the last result of the proposition follows. □

Proposition 7. *If* f *satisfies* **h3** *then the function* $(0,\, \nu) \ni t \mapsto n_f(t)|/t^2$ *is increasing.*

Proof. Using definition of n_f in (9), Proposition 5 and **h1** we obtain, after simples algebraic manipulation, that

$$\frac{|n_f(t)|}{t^2} = \frac{1}{|f'(t)|} \int_0^1 \frac{f'(t) - f'(\tau t)}{t}\, d\tau, \qquad \forall\, t \in (0,\, \nu). \tag{11}$$

On the other hand as f' is strictly increasing the map $[0,\, \nu) \ni t \mapsto [f'(t) - f'(\tau t)]/t$ is positive for all $\tau \in (0,1)$. From **h3** f' is convex, so we conclude that the last map is increasing. Hence the second term in the right hand side of (11) is positive and increasing. Assumption **h2** and definition of ν imply that the first term in the right hand side of (11) is also positive and strictly increasing. Therefore we conclude that the left hand side of (11) is increasing and the statement of the proposition follows. □

4.2 Relationship between the majorant function and the vector field

We present the main relationships between the majorant function f and the vector field X.

Lemma 8. *Let* $p \in \Omega \subseteq \mathcal{M}$. *If* $d(p_*, p) < \min\{\kappa, \nu\}$ *then* $\nabla X(p)$ *is invertible and*

$$\|\nabla X(p)^{-1} P_{\zeta,0,1} \nabla X(p_*)\| \leq 1/|f'(d(p_*, p))|$$

where $\zeta : [0,1] \to \mathcal{M}$ *is a minimizing geodesic from* p_* *to* p. *In particular* $\nabla X(p)$ *is invertible for all* $p \in B_r(p_*)$ *where* r *is as defined in Theorem 4.*

Proof. See Lemma 4.4 of [13]. □

Lemma 9. *Let $p \in \Omega \subseteq \mathcal{M}$. If $d(p_*, p) \leq d(p_*, p_0) < \min\{\kappa, \nu\}$, then there holds*

$$cond(\nabla X(p)) \leq cond(\nabla X(p_*)) \left[2/|f'(d(p_*, p_0))| - 1\right].$$

As a consequence, $\theta cond(\nabla X(p)) \leq \vartheta$.

Proof. Let $I_{p_*} : T_{p_*}\mathcal{M} \to T_{p_*}\mathcal{M}$ the identity operator, $p \in B_\kappa(p_*)$ and $\zeta : [0, 1] \to \mathcal{M}$ a minimizing geodesic from p_* to p. Since $P_{\zeta,0,0} = I_{p_*}$ and $P_{\zeta,0,1}$ is an isometry we obtain

$$\left\|\nabla X(p_*)^{-1} P_{\zeta,1,0} \nabla X(p) P_{\zeta,0,1} - I_{p_*}\right\| = \left\|\nabla X(p_*)^{-1}[P_{\zeta,1,0} \nabla X(p) - P_{\zeta,0,0} \nabla X(p_*) P_{\zeta,1,0}]\right\|.$$

As $d(p_*, p) < \nu$ we have $f'(d(p_*, p)) < 0$. Using the last equation, (3) and **h1** we conclude that

$$\|\nabla X(p_*)^{-1} P_{\zeta,1,0} \nabla X(p) P_{\zeta,0,1} - I_{p_*}\| \leq f'(d(p_*, p)) + 1.$$

Since $P_{\zeta,0,1}$ is an isometry and $\|\nabla X(p)\| \leq \|\nabla X(p_*)\| \|\nabla X(p_*)^{-1} P_{\zeta,1,0} \nabla X(p) P_{\zeta,0,1}\|$, triangular inequality together with above inequality imply

$$\|\nabla X(p)\| \leq \|\nabla X(p_*)\| \left[f'(d(p_*, p)) + 2\right].$$

On the other hand, it is easy to see from Lemma 8 that $\|\nabla X(p)^{-1}\| \leq \|\nabla X(p_*)^{-1}\|/|f'(d(p_*, p))|$. Therefore, combining two last inequalities and definition of condition number we obtain

$$\mathrm{cond}(\nabla X(p)) \leq \mathrm{cond}(\nabla X(p_*)) \left[2/|f'(d(p_*, p))| - 1\right].$$

Since f' is strictly increasing, $f' < 0$ in $[0, \nu)$ and $d(p_*, p) \leq d(p_*, p_0) < \min\{\kappa, \nu\}$, the first inequality of the lemma follows from last inequality.

The last inequality of the lemma follows from (5) and first inequality. □

The linearization error of X at a point in $B_\kappa(p_*)$ is defined by:

$$E_X(p_*, p) := X(p_*) - P_{\alpha,0,1}\left[X(p) + \nabla X(p)\alpha'(0)\right], \qquad p \in B_\kappa(p_*), \tag{12}$$

where $\alpha : [0, 1] \to \mathcal{M}$ is a minimizing geodesic from p to $p*$. We will bound this error by the error in the linearization on the majorant function f,

$$e_f(t, u) := f(u) - \left[f(t) + f'(t)(u - t)\right], \qquad t,\, u \in [0, R). \tag{13}$$

Lemma 10. *Let $p \in \Omega \subseteq \mathcal{M}$. If $d(p_*, p) \leq \kappa$ then $\|\nabla X(p_*)^{-1} E_X(p_*, p)\| \leq e_f(d(p_*, p), 0)$.*

Proof. See Lemma 4.5 of [13]. □

Lemma 11. *Let $p \in \Omega \subseteq \mathcal{M}$. If $d(p_*, p) < r$ then*

$$\left\|\nabla X(p)^{-1} X(p)\right\| \leq \frac{f(d(p_*,\, p))}{f'(d(p_*,\, p))}, \qquad p \in B_r(p_*).$$

Proof. Since $X(p_*) = 0$, the inequality is trivial for $p = p_*$. Now assume that $0 < d(p_*, p) < r$. Lemma 8 implies that $\nabla X(p)$ is invertible. Let $\alpha : [0,1] \to \mathcal{M}$ be a minimizing geodesic from p to p_*. Because $X(p_*) = 0$, the definition of $E_X(p_*, p)$ in (12) and direct manipulation yields

$$-\nabla X(p)^{-1} P_{\alpha,1,0} E_X(p, p_*) = \nabla X(p)^{-1} X(p) + \alpha'(0).$$

Using the above equation, Lemma 8 and Lemma 10, it is easy to conclude that

$$\begin{aligned}\|\nabla X(p)^{-1} X(p) + \alpha'(0)\| &\leq \| - \nabla X(p)^{-1} P_{\alpha,1,0} \nabla X(p_*)\| \|\nabla X(p_*)^{-1} E_F(p, p_*)\| \\ &\leq e_f(d(p_*, p), 0)/|f'(d(p_*, p))|.\end{aligned}$$

As $f(0) = 0$, definition of e_f gives $e_f(d(p_*, p), 0)/|f'(d(p_*, p))| = -d(p_*, p) + f(d(p_*, p))/f'(d(p_*, p))$, which combined with last inequality yields

$$\|\nabla X(p)^{-1} X(p) + \alpha'(0)\| \leq -d(p_*, p) + f(d(p_*, p))/f'(d(p_*, p)).$$

Since $\|\alpha'(0)\| = d(p_*, p)$, after simples algebraic manipulation we conclude

$$\|\nabla X(p)^{-1} X(p)\| \leq \|\nabla X(p)^{-1} X(p) + \alpha'(0)\| + d(p_*, p),$$

which combined with last inequality yields the desired result. □

The outcome of an Inexact Newton iteration is any point satisfying some error tolerance. Hence, instead of a mapping for Newton iteration, we shall deal with a *family* of mappings describing all possible inexact iterations.

Definition 5. *For $0 \leq \theta$, $\mathcal{N}_\theta$ is the family of maps $N_\theta : B_r(p_*) \to X$ such that*

$$\left\|X(p) + \nabla X(p) \exp_p^{-1} N_\theta(p)\right\| \leq \theta \|X(p)\|, \qquad p \in B_r(p_*). \tag{14}$$

If $p \in B_r(p_*)$ then $\nabla X(p)$ is non-singular. Therefore for $\theta = 0$ the family $\mathcal{N}_0$ has a single element, namely, the exact Newton iteration map

$$\begin{aligned} N_0 : B_r(p_*) &\to \mathcal{M} \\ p &\mapsto \exp_p\left(-\nabla X(p)^{-1} X(p)\right). \end{aligned} \tag{15}$$

Trivially, if $0 \leq \theta \leq \theta'$ then $\mathcal{N}_0 \subset \mathcal{N}_\theta \subset \mathcal{N}_{\theta'}$. Hence $\mathcal{N}_\theta$ is non-empty for all $\theta \geq 0$.

Remark 3. *For any $\theta \in (0,1)$ and $N_\theta \in \mathcal{N}_\theta$*

$$N_\theta(p) = p \iff X(p) = 0, \qquad p \in B_r(p_*).$$

This means that the fixed points of the Inexact Newton iteration N_θ are the same fixed points of the exact *Newton iteration, namely, the zeros of X.*

Lemma 12. *Let θ be such that $0 \leq \theta cond(\nabla X(p_*)) \leq \vartheta / [1 + 2/|f'(d(p_*, p_0))|]$ and $p \in \Omega \subseteq \mathcal{M}$. If $d(p_*, p) \leq d(p_*, p_0) < r$ and $N_\theta \in \mathcal{N}_\theta$ then*

$$d(p_*, N_\theta(p)) \leq K_{p_*} \left[(1+\vartheta) \frac{|n_f(d(p_*, p))|}{d(p_*, p)} + \vartheta \right] d(p_*, p), \qquad p \in B_r(p_*).$$

As a consequence, $N_\theta(B_r(p_*)) \subset B_r(p_*)$.

Proof. Since $X(p_*) = 0$, the inequality is trivial for $p = p_*$. Now, assume that $0 < d(p_*, p) \leq r$. Let $\alpha : [0, 1] \to \mathcal{M}$ be a minimizing geodesic from p to p_*. After simple algebraic manipulations, triangular inequality and definition of the linearization error we obtain

$$\| \exp_p^{-1} N_\theta(p) - \alpha'(0) \| \leq \left\| \nabla X(p)^{-1} \left[\nabla X(p) \exp_p^{-1} N_\theta(p) + X(p) \right] \right\| + \left\| \nabla X(p)^{-1} E_X(p_*, p) \right\|. \quad (16)$$

Using Definition (5) the first term in the right hand side of the above inequality is bounded by

$$\left\| \nabla X(p)^{-1} \left[\nabla X(p) \exp_p^{-1} N_\theta(p) + X(p) \right] \right\| \leq \left\| \nabla X(p)^{-1} \right\| \theta \|X(p)\|.$$

Now, since $\|X(p)\| \leq \|\nabla X(p)\| \left\| \nabla X(p)^{-1} X(p) \right\|$ we obtain from Lemma (11) that

$$|X(p)\| \leq \|\nabla X(p)\| \frac{f(d(p_*, p))}{f'(d(p_*, p))}.$$

Definition of condition number and two above inequalities imply

$$\left\| \nabla X(p)^{-1} \left[\nabla X(p) \exp_p^{-1} N_\theta(p) + X(p) \right] \right\| \leq \theta \mathrm{cond}(\nabla X(p)) \frac{f(d(p_*, p))}{f'(d(p_*, p))}. \quad (17)$$

Now, combining Lemma (10) and Lemma (8) the second term in (16) is bounded by

$$\left\| \nabla X(p)^{-1} E_X(p_*, p), \right\| \leq \frac{1}{|f'(d(p_*, p))|} e_f(d(p_*, p), 0).$$

Therefore, (16), (17) and last inequality give us

$$\| \exp_p^{-1} N_\theta(p) - \alpha'(0) \| \leq \theta \mathrm{cond}(\nabla X(p)) \frac{f(d(p_*, p))}{f'(d(p_*, p))} + \frac{1}{|f'(d(p_*, p))|} e_f(d(p_*, p), 0).$$

Since Lemma (9) implies $\theta \mathrm{cond}(\nabla X(p)) \leq \vartheta$, after simple algebraic manipulation and taking in account definitions of e_f and n_f the above inequaliy becomes

$$\| \exp_p^{-1} N_\theta(p) - \alpha'(0) \| \leq \left[(1+\vartheta) \frac{|n_f(d(p_*, p))|}{d(p_*, p)} + \vartheta \right] d(p_*, p).$$

Note that, as $d(p_*, p) \leq r < \rho$, second part of Proposition (6) implies that the term in brackets of last inequality is less than $1/K_{p_*} \leq 1$. So left hand side of last inequality is less than $r \leq r_{p_*}$. Therefore letting $p = p_*$, $q = p$, $v = \alpha'(0)$, $u = \exp_p^{-1} N_\theta(p)$ in Definition 1 we conclude that

$$d(p_*, N_\theta(p)) \leq K_{p_*} \| \exp_p^{-1} N_\theta(p) - \alpha'(0) \|.$$

Finally combining two above inequalities the inequality of the lemma follows.

Take $p \in B_r(p_*)$. Since $d(p_*, p) < r$ and $r \leq \rho$, the first part of the lemma and the second part of Proposition 6 imply that $d(p_*, N_X(p)) < d(p_*, p)$ and the result follows. □

5 The Newton sequence

In this section we prove Theorem 4. Let $0 \leq \theta$ satisfying (5) and $N_\theta \in \mathcal{N}_\theta$, where $\mathcal{N}_\theta$ is defined in Definition 5. Therefore (4) together with Definition 5 implies that the sequence $\{p_k\}$ satisfies

$$p_{k+1} = N_\theta(p_k), \qquad k = 0, 1, \ldots, \tag{18}$$

which is indeed an equivalent definition of this sequence.

Proof of Theorem 4: Since $p_0 \in B_r(p_*)$, $r \leq \nu$ and $0 < \theta \mathrm{cond}(\nabla X(p_*)) \leq \vartheta / \left[2/|f'(d(p_*, p_0))| - 1\right]$, combining (18), the inclusion $N_\theta(B_r(p_*)) \subset B_r(p_*)$ in Lemma 12 and Lemma 8, it is easy to conclude that by an induction argument the sequence $\{p_k\}$ is well defined and remains in $B_r(p_*)$.

Now we are going to prove that $\{p_k\}$ converges towards p_*. Since $d(p_*, p_k) < r$, for $k = 0, 1, \ldots$, we obtain from (18) and Lemma 12 that

$$d(p_*, p_{k+1}) \leq K_{p_*} \left[(1 + \vartheta) \frac{|n_f(d(p_*, p_k))|}{d(p_*, p_k)} + \vartheta \right] d(p_*, p_k). \tag{19}$$

As $d(p_*, p_k) < r \leq \rho$, for $k = 0, 1, \ldots$, using second statement in Proposition 6 and last inequality we conclude that $0 \leq d(p_*, p_{k+1}) < d(p_*, p_k)$, for $k = 0, 1, \ldots$. So $\{d(p_*, p_k)\}$ is strictly decreasing and bounded below which implies that it converges. Let $\ell_* := \lim_{k\to\infty} d(p_*, p_k)$. Because $\{d(p_*, p_k)\}$ rests in $(0, \rho)$ and is strictly decreasing we have $0 \leq \ell_* < \rho$. We are going to show that $\ell_* = 0$. If $0 < \ell_*$ then letting k goes to infinity in (19), the continuity of n_f in $[0, \rho)$ and Proposition 6 imply that

$$\ell_* \leq K_{p_*} \left[(1 + \vartheta) \frac{|n_f(\ell_*)|}{\ell_*} + \vartheta \right] \ell_* < \ell_*, \tag{20}$$

which is an absurd. Hence we must have $\ell_* = 0$. Therefore the convergence of $\{p_k\}$ to p_* is proved. The uniqueness of p_* in $B_\sigma(p_*)$ was proved in Lemma 5.1 of [13].

For proving the equality in (6) it is sufficient to use equation (19) and definition of n_f in (9). As $d(p_*, p_k) < r \leq \rho$, for $k = 0, 1, \ldots$, $\lim_{k\to\infty} d(p_*, p_k) = 0$ and by hypothesis $\vartheta < 1/K_{p_*}$ thus using definition of n_f and first statement in Proposition 6 we conclude

$$\lim_{k\to\infty} K_{p_*}\left[(1+\vartheta)\frac{\left|d(p_*,p_k)-\dfrac{f(d(p_*,p_k))}{f'(d(p_*,p_k))}\right|}{d(p_*,p_k))}+\vartheta\right]=K_{p_*}\vartheta<1.$$

which implies the linear convergence of $\{p_k\}$ to p_* in (6).

Now we are going to prove the inequality in (7): If f satisfies **h3** then using definition of n_f and Proposition 7 we conclude

$$(1+\vartheta)\frac{\left|d(p_*,p_k)-\dfrac{f(d(p_*,p_k))}{f'(d(p_*,p_k))}\right|}{d^2(p_*,p_k))}d(p_*,p_k))+\vartheta\le(1+\vartheta)\frac{\left|d(p_*,p_0)-\dfrac{f(d(p_*,p_0))}{f'(d(p_*,p_0))}\right|}{d^2(p_*,p_0))}d(p_*,p_k))+\vartheta.$$

As the quantity of the left hand side of the last inequality is equal to quantity in the brackets of (6), the inequality in (7) follows from (6) and last inequality.

Since $\{d(p_*,p_k)\}$ is strictly decreasing, the inequality in (8) follows from (7) and we conclude the proof of the theorem. □

6 Special Cases

The affine majorant condition is crucial for our analysis. It is worth pointing out that to construct a majorizing function for a given nonlinear function is a very difficult problem and this is not our aim in this moment. On the other hand, there exist some classes of well known functions which a majorant function is available, below we will present two examples, namely, the classes of functions satisfying the a affine invariant Hölder-like and Smale's conditions, respectively. In this sense, the results obtained in Theorem 4 unify the convergence analysis for the classes of inclusion problems involving these functions.

6.1 Convergence result under Hölder-like condition

For null error tolerance, the next theorem on Inexact Newton's method under a Hölder-like condition merges in Theorem 7.1 of [13].

Theorem 13. *Let $\mathcal{M}$ be a Riemannian manifold, $\Omega \subseteq \mathcal{M}$ an open set and $X:\Omega\to T\mathcal{M}$ a continuously differentiable vector field. Take $p_*\in\Omega$, $R>0$ and let $\kappa:=\sup\{t\in[0,R): B_t(p_*)\subset \Omega\}$. Suppose that $X(p_*)=0$, $\nabla X(p_*)$ is invertible and there exist constants $L>0$ and $0\le\mu<1$ such that*

$$\left\|\nabla X(p_*)^{-1}[P_{\zeta,1,0}\,\nabla X(p)-\,P_{\zeta,\tau,0}\nabla X(\zeta(\tau))P_{\zeta,1,\tau}]\right\|\le L(1-\tau^{\mu})d(p_*,p)^{\mu}, \tag{21}$$

for all $\tau \in [0,1]$ and $p \in B_\kappa(p_*)$, where $\zeta : [0,1] \to \mathcal{M}$ is a minimizing geodesic from p_* to p. Let r_{p_*} be the injectivity radius of $\mathcal{M}$ in p_*, K_{p_*} as in Definition 1, $0 \le \vartheta < 1/K_{p_*}$ and

$$r := \min\left\{\kappa,\ \left[(\mu+1)/\left(L\left(\frac{1+K_{p_*}}{1-K_{p_*}\vartheta}\mu+1\right)\right)\right]^{1/\mu},\ r_{p_*}\right\}.$$

Then the sequence generated by the Inexact Newton method for solving $X(p)=0$ with starting point $p_0 \in B_r(p_*) \setminus \{p_*\}$ and residual relative error tolerance θ,

$$p_{k+1} = \exp_{p_k}(S_k), \qquad \|X(p_k) + \nabla X(p_k) S_k\| \le \theta \|X(p_k)\|, \qquad k = 0,1,\ldots, \tag{22}$$

$$0 \le cond(\nabla X(p_*))\theta \le \vartheta\, \frac{1+Ld(p_*,p_0)^\mu}{1-Ld(p_*,p_0)^\mu}, \tag{23}$$

is well defined (for any particular choice of each $S_k \in T_{p_k}M$), the sequence $\{p_k\}$ is contained in $B_r(p_*)$ and converges to the point p_* which is the unique zero of X in $B_{[(\mu+1)/L]^{1/\mu}}(p_*)$ and we have that:

$$d(p_*,p_{k+1}) \le K_{p_*}\left[(1+\vartheta)\frac{\mu L d(p_*,p_k)^\mu}{(\mu+1)\left[1-Ld(p_*,p_k)^\mu\right]} + \vartheta\right] d(p_*,p_k), \qquad k=0,1,\ldots,$$

and $\{p_k\}$ converges linearly to p_*. If, in additional, $\mu = 1$ then there holds

$$d(p_*,p_{k+1}) \le K_{p_*}\left[(1+\vartheta)\frac{L}{2\left[1-Ld(p_*,p_0)\right]} d(p_*,p_k) + \vartheta\right] d(p_*,p_k) \qquad k=0,1,\ldots \tag{24}$$

as a consequence, the sequence $\{p_k\}$ converges to p_* with linear rate as follows

$$d(p_*,p_{k+1}) \le K_{p_*}\left[(1+\vartheta)\frac{Ld(p_*,p_0)}{2\left[1-Ld(p_*,p_0)\right]} + \vartheta\right] d(p_*,p_k) \qquad k=0,1,\ldots$$

Proof. We can prove that X, p_* and $f : [0,+\infty) \to \mathbb{R}$, defined by $f(t) = Lt^{\mu+1}/(\mu+1) - t$, satisfy the inequality (3) and the conditions **h1** and **h2** in Theorem 4. Moreover, if $\mu = 1$ then f satisfies condition **h3**. It is easy to see that ρ, ν and σ, as defined in Theorem 4, satisfy

$$\rho = \left[\frac{(\mu+1)}{L\left(\dfrac{1+K_{p_*}}{1-K_{p_*}\vartheta}\mu+1\right)}\right]^{1/\mu} \le\ \nu = \frac{1}{L^{1/\mu}}, \qquad \sigma = [(\mu+1)/L]^{1/\mu}.$$

Therefore, the result follows by invoking Theorem 4. □

Remark 4. *Note that if vector field X is Lipschitz with constant L then it satisfies the condition (21) with $\mu = 1$.*

We remark that letting $\vartheta = 0$ in Theorem 13 which implies from (23) that $\theta = 0$, the linear equation in (22) is solved exactly. Therefore (24) implies that if $\mu = 1$ then $\{p_k\}$ converges to p_* with quadratic rate.

6.2 Convergence result under Smale's condition

For null error tolerance, the next theorem on Inexact Newton's method under Smale's condition merges in Theorem 7.2 of [13]. We note that Theorem 7.2 of [13] extends to the Riemannian context Theorem 1.1 of [8] (see also Theorem 3.1 of [31]) which generalizes to the Riemannian context Corollary of Proposition 3 on p. 195 of [26], see also Proposition 1 p. 157 and Remark 1 p. 158 of [6].

Theorem 14. *Let $\mathcal{M}$ be an analytic Riemannian manifold, $\Omega \subseteq \mathcal{M}$ an open set and $X : \Omega \to T\mathcal{M}$ an analytic vector field. Take $p_* \in \Omega$, $R > 0$ and let $\kappa := \sup\{t \in [0, R) : B_t(p_*) \subset \Omega\}$. Suppose that $X(p_*) = 0$, $\nabla X(p_*)$ is invertible and*

$$\gamma := \sup_{n>1} \left\| \frac{\nabla X(p_*)^{-1} \nabla^n X(p_*)}{n!} \right\|^{1/(n-1)} < +\infty. \tag{25}$$

Let r_{p_*} be the injectivity radius of $\mathcal{M}$ in p_*, K_{p_*} as in Definition 1, $0 \le \vartheta < 1/K_{p_*}$ and

$$r := \min \left\{ \kappa, \ \frac{K_{p*}(1-3\vartheta) + 4 - \sqrt{K_{p*}^2(1-6\vartheta+\vartheta^2) + 8K_{p*}(1-\vartheta) + 8}}{4\gamma(1-K_{p_*}\vartheta)}, \ r_{p_*} \right\}.$$

Then the sequence generated by the Inexact Newton method for solving $X(p) = 0$ with starting point $p_0 \in B_r(p_*) \setminus \{p_*\}$ and residual relative error tolerance θ,

$$p_{k+1} = \exp_{p_k}(S_k), \qquad \|X(p_k) + \nabla X(p_k) S_k\| \le \theta \|X(p_k)\|, \qquad k = 0, 1, \ldots, \tag{26}$$

$$0 \le cond(\nabla X(p_*))\theta \le \vartheta \left[2[1 - \gamma d(p_*, p_0)]^2 - 1 \right], \tag{27}$$

is well defined (for any particular choice of each $S_k \in T_{p_k}M$), the sequence $\{p_k\}$ is contained in $B_r(p_*)$ and converges to the point p_* which is the unique zero of X in $B_{1/(2\gamma)}(p_*)$ and we have that:

$$d(p_*, p_{k+1}) \le K_{p_*} \left[(1+\vartheta) \frac{\gamma}{2\left[1 - \gamma d(p_*, p_0)\right]^2 - 1} d(p_*, p_k) + \vartheta \right] d(p_*, p_k), \quad k = 0, 1, \ldots. \tag{28}$$

as a consequence, the sequence $\{p_k\}$ converges to p_* with linear rate as follows

$$d(p_*, p_{k+1}) \le K_{p_*} \left[(1+\vartheta) \frac{\gamma d(p_*, p_0)}{2\left[1 - \gamma d(p_*, p_0)\right]^2 - 1} + \vartheta \right] d(p_*, p_k), \quad k = 0, 1, \ldots. \tag{29}$$

We need the following results to prove the above theorem.

Lemma 15. *Let $\mathcal{M}$ be an analytic Riemannian manifold, $\Omega \subseteq \mathcal{M}$ an open set and $X : \Omega \to T\mathcal{M}$ an analytic vector field. Suppose that $p_* \in \Omega$, $\nabla X(p_*)$ is invertible, $\gamma < +\infty$ and that $B_{1/\gamma}(p_*) \subset \Omega$, where γ is defined in* (25). *Then, for all $p \in B_{1/\gamma}(p_*)$,*

$$\|\nabla X(p_*)^{-1} P_{\zeta,1,0} \nabla^2 X(p))\| \leq (2\gamma)/(1 - \gamma d(p_*, p))^3,$$

where $\zeta : [0, 1] \to \mathcal{M}$ is a minimizing geodesic from p_* to p.

Proof. The proof follows the pattern of Lemma 5.3 of [1]. □

The next result is the Lemma 7.4 of [13], it gives an alternative condition for checking condition (3), whenever the vector field under consideration is twice continuously differentiable.

Lemma 16. *Let $\mathcal{M}$ be an analytic Riemannian manifold, $\Omega \subseteq \mathcal{M}$ an open set and $X : \Omega \to T\mathcal{M}$ an analytic vector field. Suppose that $p_* \in \Omega$ and $\nabla X(p_*)$ is invertible. If there exists an $f : [0, R) \to \mathbb{R}$ twice continuously differentiable such that*

$$\|\nabla X(p_*)^{-1} P_{\alpha,1,0} \nabla^2 X(q))\| \leqslant f''(d(p_*, q)), \qquad \forall\, q \in B_\kappa(p_*), \tag{30}$$

where $\alpha : [0, 1] \to \mathcal{M}$ is a minimizing geodesic from p_* to q, then X and f satisfy (3).

Corollary 17. *Let $\mathcal{M}$ be an analytic Riemannian manifold, $\Omega \subseteq \mathcal{M}$ an open set and $X : \Omega \to T\mathcal{M}$ an analytic vector field. Take $p_* \in \Omega$ and let $\kappa := \sup\{t \in [0, R) : B_t(p_*) \subset \Omega\}$ and $\gamma < +\infty$ be as defined in* (25). *Suppose that $\nabla X(p_*)$ is invertible. Then*

$$\left\|\nabla X(p_*)^{-1}[P_{\zeta,1,0}\, \nabla X(p) - \, P_{\zeta,\tau,0} \nabla X(\zeta(\tau)) P_{\zeta,1,\tau}]\right\| \leq \frac{1}{(1 - \gamma d(p_*, p))^2} - \frac{1}{(1 - \tau\gamma d(p_*, p))^2}$$

for all $\tau \in [0, 1]$, $p \in B_{1/\gamma}(p_*)$, where $\zeta : [0, 1] \to \mathcal{M}$ a minimizing geodesic from p_* to p

Proof. The proof follows by a combination of Lemma 16 with Lemma 15. □

[Proof of Theorem 14]. Assume that all hypotheses of Theorem 14 hold. Consider the real analytical function $f : [0, 1/\gamma) \to \mathbb{R}$ defined by

$$f(t) = \frac{t}{1 - \gamma t} - 2t.$$

It is straightforward to show that f is analytic and that

$$f(0) = 0, \quad f'(t) = 1/(1-\gamma t)^2 - 2, \quad f'(0) = -1, \quad f''(t) = (2\gamma)/(1-\gamma t)^3, \quad f'''(t) = 6\gamma^2/(1-\gamma t)^4.$$

It follows from the last equalities that f satisfies **h1**, **h2** and **h3**. Now, since $f'(t) = 1/(1 - \gamma t)^2 - 2$ we conclude from Corollary 17 that X and f satisfy (3) with $R = 1/\gamma$. In this case, it is easy to see that the constants ν, ρ and σ, as defined in Theorem 4, satisfy

$$\rho = \frac{K_{p*}(1 - 3\vartheta) + 4 - \sqrt{K_{p*}^2(\vartheta^2 - 6\vartheta + 1) + 8K_{p*}(1 - \vartheta) + 8}}{4\gamma(1 - K_{p_*}\vartheta)} \leq \nu = \frac{\sqrt{2} - 1}{\gamma\sqrt{2}},$$

$\sigma = 1/(2\gamma)$ and $f(0) = f(1/(2\gamma)) = 0$ and $f(t) < 0$ for all $t \in (0,\ 1/(2\gamma))$. Therefore, the result follows by invoking Theorem 4. □

Remark 5. *We remark that letting $\vartheta = 0$ in Theorem 14 which implies from* (27) *that $\theta = 0$, the linear equation in* (26) *is solved exactly. Therefore* (28) *implies that $\{p_k\}$ converges to p_* with quadratic rate.*

7 Final remarks

The results in Theorem 4 are dependent on the injective radius of the exponential map. It would be interesting to establish the convergence radius independent of the injective radius of the exponential map.

Now we are going to explain the difference between the affine majorant condition used here and the similar affine generalized Lipschitz condition due to X-H. Wang [34] which is extensively used for studying Newton's method. For sake of simplicity we assume that $\mathcal{M} = \mathbb{R}^n$ and suppose that there exists a positive integrable function $L : [0,\ R) \to \mathbb{R}$ such that

$$\left\|F'(x_*)^{-1}\left[F'(x) - F'(x_* + \tau(x - x_*))\right]\right\| \leq \int_{\tau\|x-x_*\|}^{\|x-x_*\|} L(u)\mathrm{d}u, \tag{31}$$

for all $\tau \in [0, 1]$, $x \in B(x_*, \kappa)$. Observe that if the positive integrable function $L : [0,\ R) \to \mathbb{R}$ is nondecreasing then the strictly increasing function $f' : [0,\ R) \to \mathbb{R}$, defined by

$$f'(t) = \int_0^t L(u)\mathrm{d}u - 1,$$

is convex. In this case, is not hard to prove that the inequalities (3) and (31) are equivalents. On the other hand, if f' is strictly increasing and non necessary convex then the inequalities (3) and (31) are not equivalents, because there exists functions strictly increasing, continuous, with derivative zero almost everywhere, see [28] (see also [23]). Note that these functions are not absolutely continuous, so they can not be represented by an integral.

References

[1] F. Alvarez, J. Bolte, and J. Munier. A unifying local convergence result for Newton's method in Riemannian manifolds. *Found. Comput. Math.*, 8(2):197–226, 2008.

[2] S. Amat, I. K. Argyros, S. Busquier, R. Castro, S. Hilout, and S. Plaza. Traub-type high order iterative procedures on Riemannian manifolds. *S$\vec{e}$MA Journal. Boletin de la Sociedad Espanñola de Matemática Aplicada*, 63:27–52, 2014.

[3] S. Amat, I. K. Argyros, S. Busquier, R. Castro, S. Hilout, and S. Plaza. Newton-type methods on Riemannian manifolds under Kantorovich-type conditions. *Appl. Math. Comput.*, 227:762–787, 2014.

[4] S. Amat, S. Busquier, R. Castro, and S. Plaza. Third-order methods on Riemannian manifolds under Kantorovich conditions. *J. Comput. Appl. Math.*, 255:106–121, 2014.

[5] S. Amat, I. K. Argyros, S. Busquier, R. Castro, S. Hilout, and S. Plaza. On a bilinear operator free third order method on Riemannian manifolds. *Appl. Math. Comput.*, 219(14):7429–7444, 2013.

[6] L. Blum, F. Cucker, M. Shub, and S. Smale. *Complexity and real computation*. Springer-Verlag, New York, 1998. With a foreword by Richard M. Karp.

[7] J. Chen and W. Li. Convergence behaviour of inexact Newton methods under weak Lipschitz condition. *J. Comput. Appl. Math.*, 191(1):143–164, 2006.

[8] J.-P. Dedieu, P. Priouret, and G. Malajovich. Newton's method on Riemannian manifolds: convariant alpha theory. *IMA J. Numer. Anal.*, 23(3):395–419, 2003.

[9] R. S. Dembo, S. C. Eisenstat, and T. Steihaug. Inexact Newton methods. *SIAM J. Numer. Anal.*, 19(2):400–408, 1982.

[10] M. P. Do Carmo. *Riemannian Geometry*. Birkhauser, 1992.

[11] A. L. Dontchev and R. T. Rockafellar. *Implicit functions and solution mappings*. Springer Monographs in Mathematics. Springer, Dordrecht, 2009. A view from variational analysis.

[12] A. L. Dontchev and R. T. Rockafellar. Convergence of inexact Newton methods for generalized equations. *Math. Program.*, 139(1-2):115–137, 2013.

[13] O. P. Ferreira and R. C. M. Silva. Local convergence of Newton's method under a majorant condition in Riemannian manifolds. *IMA J. Numer. Anal.*, 32(4):1696–1713, 2012.

[14] O. P. Ferreira and B. F. Svaiter. Kantorovich's theorem on Newton's method in Riemannian manifolds. *J. Complexity*, 18(1):304–329, 2002.

[15] J. Gondzio. Interior point method 25 years later. *European J. Operational Research*, 218:587–601, 2012.

[16] J. Gondzio. Convergence analysis of an inexact feasible interior point method for convex quadratic programming. *SIAM J. Optim.*, 23(3):1810–1527, 2013.

[17] Z. Huang. The convergence ball of newton's method and the uniqueness ball of equations under hölder-type continuos derivatives. *Comput. Math. Appl.*, 47:247–251, 2004.

[18] S. Lang. *Differential and Riemannian Manifolds.* Springer-Verlag, 1995.

[19] C. Li and J. Wang. Newton's method on Riemannian manifolds: Smale's point estimate theory under the γ-condition. *IMA J. Numer. Anal.*, 26(2):228–251, 2006.

[20] C. Li and J. Wang. Newton's method for sections on Riemannian manifolds: generalized covariant α-theory. *J. Complexity*, 24(3):423–451, 2008.

[21] C. Li, J.-H. Wang, and J.-P. Dedieu. Smale's point estimate theory for Newton's method on Lie groups. *J. Complexity*, 25(2):128–151, 2009.

[22] B. Morini. Convergence behaviour of inexact Newton methods. *Math. Comp.*, 68(228):1605–1613, 1999.

[23] H. Okamoto and M. Wunsch, A geometric construction of continuous, strictly increasing singular functions. *Proc. Japan Acad. Ser. A Math. Sci.* 83 (7):114–118, 2007.

[24] Y. Nesterov and A. Nemirovskii. *Interior-point polynomial algorithms in convex programming*, volume 13 of *SIAM Studies in Applied Mathematics.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994.

[25] F. A. Potra. The Kantorovich Theorem and interior point methods. *Math. Program.*, 102(1, Ser. A):47–70, 2005.

[26] S. Smale. Newton's method estimates from data at one point. In *The merging of disciplines: new directions in pure, applied, and computational mathematics (Laramie, Wyo., 1985)*, pages 185–196. Springer, New York, 1986.

[27] S. T. Smith. Optimization techniques on Riemannian manifolds. In *Hamiltonian and gradient flows, algorithms and control*, volume 3 of *Fields Inst. Commun.*, pages 113–136. Amer. Math. Soc., Providence, RI, 1994.

[28] L. Takács, An increasing continuous singular function. *Amer. Math. Monthly*, 85(1):35–37, 1978.

[29] J. H. Wang. Convergence of Newton's method for sections on Riemannian manifolds. *J. Optim. Theory Appl.*, 148(1):125–145, 2011.

[30] J.-H. Wang, S. Huang, and C. Li. Extended Newton's method for mappings on Riemannian manifolds with values in a cone. *Taiwanese J. Math.*, 13(2B):633–656, 2009.

[31] J.-h. Wang and C. Li. Uniqueness of the singular points of vector fields on Riemannian manifolds under the γ-condition. *J. Complexity*, 22(4):533–548, 2006.

[32] J.-H. Wang and C. Li. Kantorovich's theorem for newton's method on lie groups. *Journal of Zhejiang University: Science A*, 8(6):978–986, 2007. cited By (since 1996) 0.

[33] J.-H. Wang, J.-C. Yao, and C. Li. Gauss-Newton method for convex composite optimizations on Riemannian manifolds. *J. Global Optim.*, 53(1):5–28, 2012.

[34] X. Wang. Convergence of Newton's method and inverse function theorem in Banach space. *Math. Comp.*, 68(225):169–186, 1999.

[35] C. E. Wayne. An introduction to KAM theory. In *Dynamical systems and probabilistic methods in partial differential equations (Berkeley, CA, 1994)*, volume 31 of *Lectures in Appl. Math.*, pages 3–29. Amer. Math. Soc., Providence, RI, 1996.