Journal of Convex Analysis
Volume 29 (2022), No. 3, 669–701

# Dennis-Moré Condition for Set-Valued Vector Fields and the Superlinear Convergence of Broyden Updates in Riemannian Manifolds

**Samir Adly**
*Laboratoire XLIM, Université de Limoges, Limoges, France*
*samir.adly@unilim.fr*

**Huynh Van Ngai***, **Van Vu Nguyen**†
*Dept. of Mathematics and Statistics, Quy Nhon University, Quy Nhon, Binh Dinh, Vietnam*
*ngaivn@yahoo.com, nguyenvanvu@qnu.edu.vn*

Received: July 7, 2021
Accepted: August 7, 2021

This paper deals with the quasi-Newton type scheme for solving generalized equations involving set-valued vector fields on Riemannian manifolds. We establish some conditions ensuring the superlinear convergence for the iterative sequence which approximates a solution of the generalized equations. Such conditions can be viewed as an extension of the classical theorem of J. E. Dennis and J. J. Moré [see: *A characterization of superlinear convergence and its application to quasi-Newton methods*, Math. Computation 28/126 (1974) 549–560] as well as the Riemannian Dennis-Moré condition established by K. A. Gallivan, C. Qi and P.-A. Absil [*A Riemannian Dennis-Moré Condition*, in: *High-Performance Scientific Computing: Algorithms and Applications*, M. W. Berry et al. (eds.), Springer, London (2012) 281–293]. Furthermore, we also apply these results to consider the convergence of a Broyden-type update for the problem of solving generalized equations in Riemannian context. Our results are new even for classical equations defined by single-valued vector fields.

*Keywords:* Variational inclusion, point-to-set vector fields, quasi-Newton methods, Riemannian manifold, Dennis-Moré condition, superlinear convergence.

*2010 Mathematics Subject Classification:* 65J99, 65K15, 58C06, 47H04.

## 1. Introduction

Let $\mathbb{M}$ be a complete (finite dimensional) Riemannian manifold. As usual, denote $T_p\mathbb{M}$ the tangent space of $\mathbb{M}$ at $p \in \mathbb{M}$ and the corresponding tangent bundle $\mathbf{TM}$. By a vector field $V$ defined on $\mathbb{M}$, we mean a map (frequently smooth) from $\mathbb{M}$ into $\mathbf{TM}$ such that $V(p) \in T_p\mathbb{M}$. An important example for such a vector field is the gradient $V = \operatorname{grad} f$ associated with smooth function $f \in C^\infty(\mathbb{M})$. Generally, one can consider a set-valued mapping $\varPi : \mathbb{M} \rightrightarrows \mathbf{TM}$, said to be a set-valued (another terminology *point-to-set* appeared in [9]) vector field whose value at a point $p$ is a subset of $T_p\mathbb{M}$. A convention, we identify $\varPi(p)$ with a subset of $T_p\mathbb{M}$.

* The research of this author was supported by Vingroup Innovation Foundation (VINIF), under Grant VINIF.2019.DA09.
† The research of this author was funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 101.01-2018.309.

ISSN 0944-6532 / $ 2.50 

We consider the problem of finding a solution to an inclusion of the form

$$0_p \in \xi(p) + \Pi(p),\ p \in \mathbb{M}, \tag{1.1}$$

where $0_p$ stands for the origin of $T_p\mathbb{M}$, $\xi$ is a vector field on $\mathbb{M}$ and $\Pi$ is a set-valued vector field. Throughout this paper, $\xi$ is smooth and $\Pi$ is closed in the sense that, if $p_k \to p$ in $\mathbb{M}$ and $u_k \to u$ in the tangent bundle $\mathbf{TM}$ with $u_k \in \Pi(p_k)$, then $u \in \Pi(p)$.

The motivation behind studying abstract model (1.1) is active through a lot of recent researches. One of the the earliest particular issue that was investigated widely is the *variational inequality*, which was mentioned first in the work by Németh [29]. After this pioneer paper, several authors has been examined such a problem, for examples, see [26, 27] and the references therein for further discussions. Especially, in the very recent contribution [7], the authors have considered *variational inclusion problems*, which is exactly the same as (1.1) but in the context of Hadamard manifold setting only. Even if more general, the paper [13] incorporated in the problem of finding solutions to the zeros of the sum of two *multi-valued vector fields*, which is obviously more abstract than the one in (1.1). The appearance of set-valued term is sometimes natural by the intrinsic feature of the practice issues, since the single-value models may not be sufficient to interpret those complex phenomena. The involved multi-valued maps makes our models to be useful by using the advanced nonsmooth techniques on manifolds [21, 24].

Back to our problem of consideration, by suitably choosing the multi-valued term $\Pi$, (1.1) covers many important ones being studied in the literature. The simplest case is let $\Pi$ to be zero field, it reduces immediately the problem of finding the singularities of a vector field $\xi$ [17, 41]. If we identify $\mathbb{M}$ with a Euclidean space $\mathbb{R}^n$, (1.1) is nothing but the so-called *generalized equation* (GE) due to the terminology by Robinson [33, 34]. Let $\Pi(p) = \partial f(p)$ to be the set of subgradients to a function $f : \mathbb{M} \longrightarrow \mathbb{R}$ (see [21, 24]), (1.1) describes somewhat subsumes to necessary optimality condition for a local minimum to $f$.

With respect to the classical GE in Euclidean and even Banach setting, there had been abundant works that consider the various of iterative schemes of type Newton or quasi-Newton, see e.g. [4, 5, 8, 19, 23] and the reference therein. Newton-type methods still demonstrate their success in term of local rate of convergence, however, they seems to be difficult to globalize, and complexity of computing derivatives is a big challenge. Those drawbacks leads to the study of quasi-Newton methods that inherit the good convergence rate while avoid the complexity of derivative computational [12]. One of the advantages of quasi-Newton methods is that superlinear convergence can be guaranteed under some mild assumptions. In the seminal paper [11], Dennis and Moré have provided such a sufficient condition for nonlinear equations. Let us recall this important theorem here.

**Theorem 1.1.** ([11, Theorem 2.2]) *Let $F : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be differentiable on the open, convex set $\Omega$ in $\mathbb{R}^n$, and assume that for some $x^* \in \Omega$ its derivative $F'$ is continuous at $x^*$ and $F'(x^*)$ is nonsingular. Let $(B_k)$ be a sequence of $n$-by-$n$ nonsingular matrices and suppose that for some $x_0$ in $\Omega$ the sequence $(x_k)$ where*

$$x_{k+1} = x_k - B_k^{-1}F(x_k) \tag{1.2}$$

*remains in $\Omega$ and converges to $x^*$. Then $(x_k)$ converges Q-superlinearly to $x^*$ and $F(x^*) = 0$ if and only if*

$$\lim_{k\to\infty} \frac{\|[B_k - F'(x^*)](x_{k+1} - x_k)\|}{\|x_{k+1} - x_k\|} = 0. \tag{1.3}$$

This famous result is referred to the Dennis-Moré theorem which serves as a key ingredient to establish the super-linear convergence of the quasi-Newton type algorithms [11, 12, 20, 28, 36]. Due to its importance in both the theoretical and numerical aspects, the Dennis-Moré theorem was developed to generalized equations defined on finite dimensional spaces or more generally on Banach spaces, e.g. [5, 8, 15].

Recently, the study of optimization methods to functions defined on smooth manifolds has attracted many researchers, due to its applications in several areas, for instance, the eigenvalue and singular value problems, matrix approximations, independent component analysis (ICA), pose estimation problem, etc. The readers are referred to [1, 2, 25, 40] for theoretical results as well as their applications. In the aspect of algorithms, several Newton/quasi-Newton type methods have developed for solving nonlinear equations defined on Riemannian manifolds (e.g., see [3, 17, 22, 30, 39]). Especially, the Riemannian BFGS algorithm had been developed through the works [30, 31, 38] and an extension of Dennis-Moré condition onto the framework of Riemannian manifolds had been established in [18]. In [6], the same authors of the present paper have investigated the Newton type method for generalized equations on Riemannian manifolds.

The present paper concentrates on some quasi-Newton type frameworks to approximate the solution of problem (1.1). Inspired from the classical quasi-Newton method, our scheme for solving (1.1) can be described briefly as follows. Let us take $p_0 \in \mathbb{M}$ (usually expected to be nearby a solution) as an initial point. At each step, suppose that the current iterate $p_k$ and a linear map $B_k : T_{p_k}\mathbb{M} \longrightarrow T_{p_k}\mathbb{M}$ was well-defined. Associated with a retraction $R_{p_k} : T_{p_k}\mathbb{M} \longrightarrow \mathbb{M}$ (it will be specified in Section 2 later), we search a tangent vector $u_k \in T_{p_k}\mathbb{M}$ such that

$$0_{p_k} \in \xi(p_k) + B_k(u_k) + \left(\mathcal{T}_{u_k}^{-1} \varPi R_{p_k}\right)(u_k). \tag{1.4}$$

Then, the new iteration is updated by $p_{k+1} = R_{p_k}(u_k)$. As mentioned later, $\mathcal{T}_{u_k} : T_{p_k}\mathbb{M} \longrightarrow T_{R_{p_k}(u_k)}\mathbb{M}$ stands for the linear map inherited from a vector transport $\mathcal{T}$ corresponding to the retraction $R$ (cf. [31, 1]). Naturally, it is possible to take $\mathcal{T}_{u_k}$ as the parallel translation along the curve $\chi_{u_k}(t) = R_{p_k}(tu_k)$. In the subproblem (1.4), the single-valued part $\xi(p_k) + B_k(u_k)$ may be viewed as a first-order approximate model for $\xi$ near $p_k$, where the operator $B_k$ plays a similar role as an approximation of the Jacobian matrix in the classical Euclidean quasi-Newton scheme (1.2). More general, an inexact version of (1.4) is the one given by

$$0_{p_k} \in \xi(p_k) + B_k(u_k) + \left(\mathcal{T}_{u_k}^{-1} \varPi R_{p_k}\right)(u_k) + r_k\mathbb{B}_{p_k}. \tag{1.5}$$

The appearance of inexact term $r_k\mathbb{B}_{p_k}$ involved in (1.5) makes the framework to be more flexible and could allow us to improve the applicability of practical algorithms.

The purpose of the paper is twofold: Firstly, to establish a general Dennis-Moré type condition for the superlinear convergence of the scheme (1.4); secondly, to apply this general condition to convergence analysis of the Broyden type update. We would like to mention that the extension from the single-valued case to the set-valued one is not trivial. The main difficulty in the proof of the convergence of (1.4) is due to the fact that the set-valued parts in scheme (1.4) change and depend on the geometric structure of the reference manifold at each iteration.

The paper is organized as follows. In the next section we recall some basic notations from differential geometry which will be used in the sequel. Section 3 presents some results (Theorem 3.3 and 3.7) related to Riemannian Dennis-Moré conditions for inclusion (1.1). Section 4 deals with a class of Broyden type update in order to solve (1.1). We establish in Theorem 4.10 the existence and the linear convergence of sequence given by our abstract scheme (1.5). At the end, the superlinear rate is attained in Theorem 4.16 with the help of Dennis-Moré type condition in Theorem 3.3.

## 2. Notations and preliminaries

All fundamental objects considered in what follows are in the scope of differential geometry. We follow [14, 37] in the use of the basic concepts and notation concerning Riemannian manifolds.

Let $\mathbb{M}$ be a (geodesically) complete Riemannian manifold of a dimension $m$ endowed with a metric $g$. The space of all smooth real functions defined on $\mathbb{M}$ is denoted by $C^\infty(\mathbb{M})$. For any smooth curve $\gamma : I \longrightarrow \mathbb{M}$ ($I \subset \mathbb{R}$ is always meant to be an interval), its tangent vector $\gamma'(t_0)$ (or also $\dot\gamma(t_0)$, $\frac{d\gamma}{dt}(t_0)$) at instant $t_0 \in I$ is viewed as a derivation by the formula

$$\gamma'(t_0)f := (f \circ \gamma)'(t_0), \quad f \in C^\infty(\mathbb{M}).$$

If $p \in \mathbb{M}$, the tangent space $T_p\mathbb{M}$ to $\mathbb{M}$ at $p$ is the collection of all tangent vectors $\gamma'(t_0)$ such that $\gamma(t_0) = p$. Formally, the tangent bundle $\mathbf{TM}$ of $\mathbb{M}$ is identified with the disjoint union of all tangent spaces $\mathbf{TM} = \big\{(p, u) : p \in \mathbb{M}, u \in T_p\mathbb{M}\big\}$.

As usual, denote by $\|\cdot\|_p$ the norm induced by $g$ onto the linear space $T_p\mathbb{M}$. If $\gamma : I \longrightarrow \mathbb{M}$ is a piecewise smooth curve, i.e., there are a finite number of points $a = t_0 < t_1 < \ldots < t_k = b$ for which $\gamma|_{(t_i,t_{i+1})}$ is smooth, then the length of $\gamma_0 = \gamma|_{[a,b]}$ is given by the following quantity [14]

$$\ell_a^b(\gamma_0) := \sum_{i=0}^{k-1} \int_{t_i}^{t_{i+1}} \|\gamma'(t)\|_{\gamma(t)}\, dt. \tag{2.1}$$

The *Riemannian distance* on $\mathbb{M}$ is defined by

$$d(p, q) := \inf \Big\{ \ell_a^b(\gamma) \;\Big|\; \gamma \text{ is piecewise smooth, } \gamma(a) = p,\ \gamma(b) = q \Big\}. \tag{2.2}$$

The fundamental theorem of Riemannian geometry asserts that there exists a unique symmetric affine connection $\nabla$ on $\mathbb{M}$ which is compatible with the metric $g$ (named as Riemannian connection or Levi-Civita connection). Let us fix a smooth curve $\gamma : I \longrightarrow \mathbb{M}$ and $t_0, t_1 \in I$. The *parallel translation* $P_\gamma^{t_0,t_1} : T_{\gamma(t_0)}\mathbb{M} \longrightarrow T_{\gamma(t_1)}\mathbb{M}$ along

$\gamma$ is a map determined by $P_\gamma^{t_0,t_1}(u) = V(u,t_0;t_1)$, in which $V(u,t_0;\cdot)$ is the unique solution of an ordinary differential system

$$V(u,t_0;t_0) = u, \quad \nabla_{\gamma'} V(u,t_0;\cdot) \equiv 0. \tag{2.3}$$

All parallel translations are linearly isometric, and $P_\gamma^{t_0,t_1} = \left(P_\gamma^{t_1,t_0}\right)^{-1}$. Moreover, the following relation does hold

$$\nabla_{\gamma'(t_0)} V = \frac{d}{dt}\left\{P_\gamma^{t,t_0} V(\gamma(t))\right\}\Big|_{t=t_0}, \tag{2.4}$$

where $\frac{d}{dt}$ is meant to be the usual derivative with respect to real variable $t$.

A *geodesic* $\gamma$ in $\mathbb{M}$ is a smooth curve whose acceleration $\nabla_{\gamma'}\gamma'$ is vanishing. It follows from the definition that the speed $\|\gamma'(t)\|$ does not vary with respect to $t$. As mentioned at the beginning of this section, $\mathbb{M}$ is geodesically complete, that is, any geodesic can be extended to a map defining on the whole real line $\mathbb{R}$. Now let $(p,u)$ be in the tangent bundle $\mathbf{TM}$. Then, it has a unique geodesic $\gamma(p,u;\cdot)$ passing through $p$ at $t=0$ with initial velocity $\frac{d}{dt}\gamma(p,u;\cdot)\big|_{t=0} = u$. The exponential map $\mathrm{Exp} : \mathbf{TM} \longrightarrow \mathbb{M}$ is defined by $(p,u) \in \mathbf{TM} \mapsto \mathrm{Exp}(p,u) := q = \gamma(p,u;1) \in \mathbb{M}$. Associated with a point $p \in \mathbb{M}$, the restriction of Exp onto $T_p\mathbb{M}$ defines the exponential $\exp_p$ at $p$. Geodesic and exponential map possess many important properties. More details about Riemannian geometry can be found, for instance, in the textbooks [14, 37].

We will need later the concepts of retraction together with vector transportation. The readers are referred to [1, 31] for further discussions. A retraction is a smooth map $R$ from the tangent bundle $\mathbf{TM}$ onto $\mathbb{M}$ such that, for $R_p : T_p\mathbb{M} \longrightarrow \mathbb{M}$ being the retraction at $p$ (i.e, restriction of $R$ onto $T_p\mathbb{M}$), the following conditions are fulfilled:

(1) $R_p(0_p) = p$;

(2) (rigidity) $DR_p(0_p) = \mathrm{id}_{T_p\mathbb{M}}$ under the identification $T_{0_p}(T_p\mathbb{M}) \simeq T_p\mathbb{M}$, where $DR_p(0_p)$ denotes the derivative of $R_p$ at the origin $0_p$.

Obviously, the exponential map Exp is a standard example for retractions. Inspired from the standard properties of exponential map, we will say a neighborhood $\Omega_p$ of $p$ is *R-normal* if $R_q^{-1}$ is a single-valued and smooth whenever $q$ is in $\Omega_p$.

Corresponding to a given retraction $R$, a vector transport $\mathcal{T}$ defined on the Whitney sum $\mathbf{TM} \oplus \mathbf{TM}$ is a map such that the following three conditions are fulfilled:

(i) $\mathcal{T}$ is associated with $R$ in the sense that $\mathcal{T}_u v := \mathcal{T}(u,v) \in T_{R_p(u)}\mathbb{M}$ for every $u, v \in T_p(\mathbb{M})$;

(ii) (consistency) $\mathcal{T}_{0_p} u = u$ for any $(p,u) \in \mathbf{TM}$;

(iii) (linearity) the map $\mathcal{T}_u : T_p\mathbb{M} \longrightarrow T_{R_p(u)}\mathbb{M}$ is linear.

Intuitively, $\mathcal{T}_u(v) \in T_{R_p(u)}\mathbb{M}$ denotes the tangent vector in transporting $v \in T_p\mathbb{M}$ along a retraction departing from $p$ with velocity $u \in T_p\mathbb{M}$. As a particular case, we can take $\mathcal{T}_u v = P_{\gamma_u}^{0,1} v$ for $\gamma_u(t) := R_p(tu)$. Another choice is inherited from the differential of retraction $R_p$, i.e. $\mathcal{T}_u v = (dR_p)_u(v)$. To avoid the complexity of the multivalued inverse for $\mathcal{T}_u^{-1}$, let us assume that for each $p$ the linear map $\mathcal{T}_u : T_p\mathbb{M} \longrightarrow T_{R_p(u)}\mathbb{M}$ is injective whenever $u$ lies on a neighborhood around the origin $0_p$.

Similarly to the Dennis-Moré theorem in the Euclidean setting, in [18], to establish a Riemannian Dennis-Moré condition for smooth vector fields, the invertibility of the covariant derivative is needed. In the setting of set-valued vector fields considered in the present paper, we will need some metric (strong)(sub-)regularities of set-valued vector fields. Recall that (see, e.g., [16, 35]), a set-valued mapping $F : X \rightrightarrows Y$ between two Banach spaces $X$ and $Y$ is a correspondence assigning to each $x \in X$ a subset $F(x) \subset Y$. $F$ is said to be *strongly metrically subregular* with a modulus $\tau > 0$ around $(\bar{x}, \bar{y})$ for $\bar{y} \in F(\bar{x})$ if there is a neighborhood $U$ of $\bar{x}$ such that

$$\|x - \bar{x}\| \leqslant \tau d\big(\bar{y}, F(x)\big), \quad \text{for all } x \in U. \tag{2.5}$$

$F$ is said to be *metrically regular* around $(\bar{x}, \bar{y}) \in X \times Y$, if there are $\tau > 0$ and a neighborhood $U \times V$ of $(\bar{x}, \bar{y})$ such that

$$d(x, F^{-1}(y)) \leqslant \tau d\big(y, F(x)\big), \quad \text{for all } (x, y) \in U \times V. \tag{2.6}$$

The strong metric subregularity and the metric regularity are extended naturally to set-valued vector fileds on manifolds as follows.

**Definition 2.1.** ($(R, \mathcal{T})$-strong metric subregularity) Let $\mathcal{T}$ be a vector transport associated to retraction $R$ and let $\Upsilon$ be a point-to-set vector field. We say that $\Upsilon$ is $(R, \mathcal{T})$-*strongly metrically subregular* around $\bar{p}$ relative to $\bar{\omega} \in \Upsilon(\bar{p})$ provided that there exist parameters $\kappa > 0$ and $r > 0$, $r' > 0$ such that

$$d(p, \bar{p}) < r \Longrightarrow d(p, \bar{p}) \leqslant \kappa \inf \left\{ \left\|\bar{\omega} - \mathcal{T}_u^{-1}(\omega)\right\| \;\middle|\; \begin{array}{c} u = R_{\bar{p}}^{-1}(p),\ \omega \in \Upsilon(p), \\ \left\|\bar{\omega} - \mathcal{T}_u^{-1}(\omega)\right\| < r' \end{array} \right\}. \tag{2.7}$$

By $\operatorname{subreg} \Upsilon_{R,\mathcal{T}}(\bar{p}, \bar{\omega})$, we mean the infimum of all constants $\kappa > 0$ verifying (2.7). Otherwise, we set $\operatorname{subreg} \Upsilon_{R,\mathcal{T}}(\bar{p}, \bar{\omega}) = +\infty$ when $\Upsilon$ is not $\mathcal{T}$-strongly metrically subregular around $\bar{p}$ relative to $\bar{\omega}$. When the manifold $M$ coincides with a Euclidean space: $\mathbb{M} = \mathbb{R}^m$, by identifying $T_p\mathbb{M} = \{p\} \times \mathbb{R}^m$, $\mathbf{TM} = \mathbb{R}^m \times \mathbb{R}^m$, and by taking $R_p(p, u) = p + u$ and $\mathcal{T}_{(p,u)}(p, v) = (R_p(u), v)$, obviously Definition 2.1 becomes exactly the usual strong metric subregularity.

**Definition 2.2.** ($(R, \mathcal{T})$-metric regularity) Let $\mathcal{T}$ be a vector transport associated with retraction $R$ and let $\Upsilon$ be a given set-valued vector field. One says that $\Upsilon$ is *metrically regular* with respect to the pair of retraction-transport $(R, \mathcal{T})$ at $(\bar{p}, \bar{u})$ for $\bar{u} \in \Upsilon(\bar{p})$ if there exist some positive parameters $r_1$, $r_2$ and $\kappa$ such that

$$\left\{ \begin{array}{l} d(p, \bar{p}) < r_1,\ v \in T_p\mathbb{M}, \\ \left\|v - \mathcal{T}_{R_{\bar{p}}^{-1}(p)}\bar{u}\right\| < r_2 \end{array} \right\} \Longrightarrow d\Big(p, \Big\{q : \Upsilon(q) \ni \mathcal{T}_{R_p^{-1}(q)} v\Big\}\Big) \leqslant \kappa d\big(v, \Upsilon(p)\big). \tag{2.8}$$

Denote by $\operatorname{reg}_{R,\mathcal{T}} \Upsilon(\bar{p}, \bar{u})$ the infimum over all moduli $\kappa$ satisfying (2.8). Otherwise, we set $\operatorname{reg}_{R,\mathcal{T}} \Upsilon(\bar{p}, \bar{u}) = +\infty$. In the case $R = \operatorname{Exp}$ and $\mathcal{T}$ is generated by the parallel transport, we will say shortly that $\Upsilon$ is metrically regular at $(\bar{p}, \bar{u})$, and write $\operatorname{reg} \Upsilon(\bar{p}, \bar{u})$ instead of $\operatorname{reg}_{R,\mathcal{T}} \Upsilon(\bar{p}, \bar{u})$. Also recall that, for the two subsets $A, B$ of a metric space $X$ endowed with a metric $d$, the *excess* of $A$ over $B$ is defined by

$$e(A, B) := \sup_{a \in A} d(a, B).$$

## 3. Riemannian Dennis-Moré condition for inclusions involving set-valued vector fields

Consider the inclusion (1.1) associated to a vector field $\xi$ and a set-valued vector field $\Pi$ defined on a Riemannian manifold $\mathbb{M}$. For a given pair $(R, \mathcal{T})$, the (exact) scheme of quasi-Newton type for solving (1.1) is described as

$$\begin{cases} 0_p \in \xi(p_k) + B_k(u_k) + (\mathcal{T}_{u_k}^{-1} \Pi R_{p_k})(u_k), \\ p_{k+1} = R_{p_k}(u_k); \end{cases} \tag{3.1}$$

and an inexact version of (3.1) is given by

$$\begin{cases} 0_p \in \xi(p_k) + B_k(u_k) + (\mathcal{T}_{u_k}^{-1} \Pi R_{p_k})(u_k) + r_k \mathbb{B}_{p_k}, \\ p_{k+1} = R_{p_k}(u_k), \end{cases} \tag{3.2}$$

where, $\mathbb{B}_{p_k}$ stands the closed unit ball of the tangent space $T_{p_k}\mathbb{M}$. Suppose that the sequence $(p_k)$ defined through either inclusion (3.1) or the one in (3.2) and that $p_k \to \bar{p}$. Let $\Omega_{\bar{p}}$ be a $R$-normal neighborhood at $\bar{p}$, and assume, upon skipping several initial terms, $p_k \in \Omega_{\bar{p}}$. Denote $v_k = R_{\bar{p}}^{-1}(p_k)$, $T_k = \mathcal{T}_{u_k}(\cdot)$ and $S_k = \mathcal{T}_{v_k}(\cdot)$ (recall that the inverse of the retraction $R_{\bar{p}}^{-1}$ is well-defined around $\bar{p}$ see [1] for more details). We will need some preparatory lemmas.

**Lemma 3.1.** *Let $\mathcal{A}_k : T_{\bar{p}}\mathbb{M} \rightrightarrows T_{\bar{p}}\mathbb{M}$ be a sequence of mappings. Define*

$$\begin{cases} a_k = d\left(B_k(u_k), S_k \mathcal{A}_k S_k^{-1} u_k\right), \\ b_k = \sup_{w \in S_k \mathcal{A}_k S_k^{-1} u_k} \left\| T_k^{-1} \xi(p_{k+1}) - \xi(p_k) - w \right\|, \end{cases} \tag{3.3}$$

*where $(p_k)$ is generated by the scheme (3.2). Then one has*

$$d\left(0_{p_k}, T_k^{-1} \xi(p_{k+1}) + T_k^{-1} \Pi(p_{k+1})\right) \leqslant a_k + b_k + r_k. \tag{3.4}$$

**Proof.** Based on (3.2), we can select $\omega_{k+1} \in \Pi(p_{k+1})$ and $w_k \in \mathbb{B}_{p_k}$ satisying $-\xi(p_k) - T_k^{-1}\omega_{k+1} - r_k w_k = B_k(u_k)$. Thus we have

$$d\left(0_{p_{k+1}}, \xi(p_{k+1}) + \Pi(p_{k+1})\right) \leqslant \|\xi(p_{k+1}) + \omega_{k+1}\|.$$

Let $w$ be in the set $S_k \mathcal{A}_k S_k^{-1} u_k$. By the representation

$$T_k^{-1}\xi(p_{k+1}) + T_k^{-1}\omega_{k+1} = \left[T_k^{-1}\xi(p_{k+1}) - \xi(p_k) - w\right] - \left[-\xi(p_k) - T_k^{-1}\omega_{k+1} - w\right],$$

we deduce

$$\begin{aligned} & d\left(0_{p_k}, T_k^{-1}\xi(p_{k+1}) + T_k^{-1}\Pi(p_{k+1})\right) \\ & \qquad \leqslant \sup_{w \in S_k \mathcal{A}_k S_k^{-1} u_k} \left\|T_k^{-1}\xi(p_{k+1}) - \xi(p_k) - w\right\| + \left\|-\xi(p_k) - T_k^{-1}\omega_{k+1} - w\right\| \\ & \qquad = b_k + \left\|-\xi(p_k) - T_k^{-1}\omega_{k+1} - w\right\|. \end{aligned}$$

Fixing $k$ and taking the infimum over $w \in S_k\mathcal{A}_kS_k^{-1}u_k$, we find

$$d\left(0_{p_k}, T_k^{-1}\xi(p_{k+1})+T_k^{-1}\Pi(p_{k+1})\right) \leqslant b_k+d\left(-\xi(p_k)-T_k^{-1}\omega_{k+1}, S_k\mathcal{A}_kS_k^{-1}u_k\right)$$
$$\leqslant b_k+d\left(B_k(u_k), S_k\mathcal{A}_kS_k^{-1}u_k\right)+r_k = b_k+a_k+r_k,$$

where the last inequality follows by the inclusion $-\xi(p_k)-T_k^{-1}\omega_{k+1}-r_kw_k = B_k(u_k)$. Hence, (3.4) is done. □

**Lemma 3.2.** *Consider the sequences defined by*

$$\begin{cases} c_k = e\left(S_k\mathcal{A}_kS_k^{-1}u_k, T_k^{-1}\xi(p_{k+1}) + B_k(u_k) + T_k^{-1}\Pi(p_{k+1})\right), \\ d_k = e\left(S_k\mathcal{A}_kS_k^{-1}u_k, T_k^{-1}S_{k+1}\xi(\bar{p}) + B_k(u_k) + T_k^{-1}\Pi(p_{k+1})\right). \end{cases} \tag{3.5}$$

*Suppose that* $(p_k)$ *satisfies* (3.2) *(or, particular* (3.1)*), we have*

$$\begin{cases} c_k \leqslant b_k + r_k, \\ d_k \leqslant b_k + \left\|T_k^{-1}S_{k+1}\right\| \left\|S_{k+1}^{-1}\xi(p_{k+1}) - \xi(\bar{p})\right\| + r_k. \end{cases} \tag{3.6}$$

**Proof.** Choose $\omega_{k+1} \in \Pi(p_{k+1})$, $w_k \in B_k(u_k)$ and $w_k' \in \mathbb{B}_{p_k}$ such that

$$\xi(p_k) + w_k + T_k^{-1}\omega_{k+1} + r_kw_k' = 0.$$

Let $w \in S_k\mathcal{A}_kS_k^{-1}u_k$, we have

$$d\big(w, T_k^{-1}\xi(p_{k+1}) + B_k(u_k) + T_k^{-1}\Pi(p_{k+1})\big) \leqslant \left\|w - T_k^{-1}\xi(p_{k+1}) - w_k - T_k^{-1}\omega_{k+1}\right\|$$
$$= \left\|w - T_k^{-1}\xi(p_{k+1}) + \xi(p_k) + r_kw_k'\right\| \leqslant \left\|w - T_k^{-1}\xi(p_{k+1}) + \xi(p_k)\right\| + r_k.$$

Taking the supremum with respect to $w$, we obtain the first inequality in (3.6). For the other one, if $w \in S_k\mathcal{A}_kS_k^{-1}u_k$ then

$$d\left(w, T_k^{-1}S_{k+1}\xi(\bar{p}) + B_k(u_k) + T_k^{-1}\Pi(p_{k+1})\right)$$
$$\leqslant \left\|w - T_k^{-1}S_{k+1}\xi(\bar{p}) - w_k - T_k^{-1}\omega_{k+1}\right\| = \left\|w - T_k^{-1}S_{k+1}\xi(\bar{p}) + \xi(p_k) + r_kw_k'\right\|$$

Next, making use of the following identical expression

$$w - T_k^{-1}S_{k+1}\xi(\bar{p}) + \xi(p_k)$$
$$= \left[-T_k^{-1}\xi(p_{k+1}) + w + \xi(p_k)\right] + T_k^{-1}S_{k+1}\left[S_{k+1}^{-1}\xi(p_{k+1}) - \xi(\bar{p})\right] + r_kw_k',$$

we get

$$d\left(w, \xi(\bar{p}) + S_{k+1}^{-1}T_kB_k(u_k) + S_{k+1}^{-1}\Pi(p_{k+1})\right)$$
$$\leqslant \left\|T_k^{-1}\xi(p_{k+1}) - w - \xi(p_k)\right\| + \left\|T_k^{-1}S_{k+1}\right\| \left\|S_{k+1}^{-1}\xi(p_{k+1}) - \xi(\bar{p})\right\| + r_k.$$

Hence, the second estimation in (3.6) is involved. □

We are now going to present the main results of this section, which provides an abstract Dennis-Moré condition for set-valued vector fields.

**Theorem 3.3.** *Given $\xi$ and $\Pi$ where $\Pi$ is closed in the sense mentioned in the introductory section. Let $R$ be given a bundle retraction and $\mathcal{T}$ be a continuous vector transport associated with $R$. Suppose the algorithm* (1.5) *generates a sequence $p_k$ converging to $\bar{p}$ in $\mathbb{M}$ with $u_k \neq 0_{p_k}$. Let $\mathcal{A}_k : T_{\bar{p}}\mathbb{M} \rightrightarrows T_{\bar{p}}\mathbb{M}$ be a sequence of mappings such that*

$$\text{(C1)} \qquad \lim_{k\to\infty} \frac{\sup_{w\in S_k\mathcal{A}_k S_k^{-1}u_k} \left\|T_k^{-1}\xi(p_{k+1}) - \xi(p_k) - w\right\|}{\|u_k\|} = 0.$$

*Here, we use again the notations $S_k = \mathcal{T}_{v_k}(\cdot)$, $T_k = \mathcal{T}_{u_k}(\cdot)$ and $v_k = R_{\bar{p}}^{-1}(p_k)$. In addition, assume that the following conditions are satisfied*

$$\text{(C2)} \qquad \limsup_{k\to\infty} \frac{r_k}{d(p_k,\bar{p})} = 0,$$

$$\text{(C3)} \qquad 0 < \liminf_{k\to\infty} \frac{d(p_k,p_{k+1})}{\|u_k\|} \leqslant \limsup_{k\to\infty} \frac{d(p_k,p_{k+1})}{\|u_k\|} < +\infty,$$

*and*

$$\text{(C4)} \qquad \limsup_{k\to\infty} \left\{ \max\left(\|T_k\|, \left\|T_k^{-1}\right\|, \|S_k\|, \left\|S_k^{-1}\right\|\right) \right\} < +\infty.$$

*Then, the following statements hold.*

(i) *When $p_k$ converges to $\bar{p}$ superlinearly, one has*

$$\lim_{k\to\infty} \frac{e\left(S_k\mathcal{A}_k S_k^{-1}u_k, T_k^{-1}\xi(p_{k+1}) + B_k(u_k) + T_k^{-1}\Pi(p_{k+1})\right)}{\|u_k\|} = 0. \tag{3.7}$$

*Moreover, if*

$$\limsup_{k\to\infty} \frac{\|S_k^{-1}\xi(p_k) - \xi(\bar{p})\|}{\|u_k\|} < +\infty, \tag{3.8}$$

*then*

$$\lim_{k\to\infty} \frac{e\left(S_k\mathcal{A}_k S_k^{-1}u_k, T_k^{-1}S_{k+1}\xi(\bar{p}) + B_k(u_k) + T_k^{-1}\Pi(p_{k+1})\right)}{\|u_k\|} = 0. \tag{3.9}$$

(ii) *Suppose now*

$$\lim_{k\to\infty} \frac{d\left(B_k(u_k), S_k\mathcal{A}_k S_k^{-1}u_k\right)}{\|u_k\|} = 0. \tag{3.10}$$

*If the map $h\colon u \in T_{\bar{p}}\mathbb{M} \longmapsto \mathcal{T}_u^{-1}\xi\left(R_{\bar{p}}(u)\right)$ is continuous, then $\bar{p}$ is a solution of problem* (1.1), *i.e.* $0_{\bar{p}} \in \xi(\bar{p}) + \Pi(\bar{p})$. *Further, if $\xi + \Pi$ is $\mathcal{T}$-strongly subregular around $\bar{p}$ relative to $0_{\bar{p}}$, then the sequence $(p_k)$ converges to $\bar{p}$ superlinearly.*

**Proof.** To simplify notation let us define $\alpha_k = d(p_k,\bar{p}) > 0$ and $\beta_k = d(p_k, p_{k+1}) > 0$. Moreover, we also adopt the parameters $a_k$, $b_k$, $c_k$ and $d_k$ defined as in Propositions 3.1 and 3.2.

**Proof of (i).** Suppose $p_k \to \bar{p}$ superlinearly, then $\alpha_{k+1} = t_k\alpha_k$ for some positive sequence $t_k \to 0$. Thus, for $k$ large enough one has

$$\beta_k = d(p_k,p_{k+1}) \geqslant |d(p_k,\bar{p}) - d(p_{k+1},\bar{p})| = |\alpha_k - \alpha_{k+1}| = (1-t_k)\,\alpha_k.$$

With the same notations, Lemma 3.2 gives

$$\begin{cases} c_k \leqslant b_k + r_k, \\ d_k \leqslant b_k + \left\|T_k^{-1}S_{k+1}\right\| \left\|S_{k+1}^{-1}\xi(p_{k+1}) - \xi(\bar{p})\right\| + r_k. \end{cases} \tag{3.11}$$

Recall that $b_k = \sup_{w\in S_k\mathcal{A}_k S_k^{-1}u_k} \left\|T_k^{-1}\xi(p_{k+1}) - \xi(p_k) - w\right\|$.

We are going to verify the following relations

$$\lim_{k\to\infty} \frac{r_k}{\|u_k\|} = \lim_{k\to\infty} \left\{ \frac{1}{\|u_k\|} \left\|T_k^{-1}S_{k+1}\right\| \left\|S_{k+1}^{-1}\xi(p_{k+1}) - \xi(\bar{p})\right\| \right\} = 0. \tag{3.12}$$

Indeed, notice that $t_k \to 0$ and since

$$\frac{r_k}{\|u_k\|} = \left(\frac{r_k}{\alpha_k}\right)\left(\frac{\alpha_k}{\beta_k}\right)\left(\frac{\beta_k}{\|u_k\|}\right) \leqslant \left(\frac{r_k}{\alpha_k}\right)\left(\frac{1}{1-t_k}\right)\left(\frac{\beta_k}{\|u_k\|}\right),$$

assumptions (C2), (C3) imply that $\limsup\limits_{k\to\infty} \dfrac{r_k}{\|u_k\|} = 0$.

On the other hand, we recall that

$$\limsup_{k\to\infty} \left\{ \|u_{k+1}\|^{-1} \left\|S_{k+1}^{-1}\xi(p_{k+1}) - \xi(\bar{p})\right\| \right\} < +\infty. \tag{3.13}$$

By the triangle inequality, we have

$$\beta_{k+1} = d(p_{k+1}, p_{k+2}) \leqslant d(\bar{p}, p_{k+1}) + d(\bar{p}, p_{k+2}) = \alpha_{k+1} + \alpha_{k+2},$$

so $\beta_{k+1} = o(\alpha_k) = o(\beta_k)$. Thanks to (C3), (C4) and (3.13), we obtain

$$\begin{aligned}
&\limsup_{k\to\infty} \left\{ \frac{1}{\|u_k\|} \left\|T_k^{-1}S_{k+1}\right\| \left\|S_{k+1}^{-1}\xi(p_{k+1}) - \xi(\bar{p})\right\| \right\} \\
&= \limsup_{k\to\infty} \left\{ \left(\frac{\beta_{k+1}}{\beta_k}\right)\left(\frac{\beta_k}{\|u_k\|}\right)\left(\left\|T_k^{-1}S_{k+1}\right\|\right)\left[\|u_{k+1}\|^{-1}\left\|S_{k+1}^{-1}\xi(p_{k+1}) - \xi(\bar{p})\right\|\right] \right\} = 0.
\end{aligned}$$

Thus, (3.12) is valid. Finally, combining (C1), (3.11) and (3.12) we get (3.7) and (3.8).

**Proof of (ii).** Let $a_k$ and $b_k$ be the quantities defined as in Lemma 3.1. From (3.10), we are able to write $a_k = s_k \|u_k\|$, $b_k = s'_k \|u_k\|$, for some sequences $(s_k)$ and $(s'_k)$ converging to 0. In view of Lemma 3.1, it holds that

$$d\left(0_{p_k}, T_k^{-1}\xi(p_{k+1}) + T_k^{-1}\Pi(p_{k+1})\right) \leqslant a_k + b_k + r_k.$$

The closedness of $\Pi$ ensures that $\Pi(p_{k+1})$ is closed. As a result, there is an element $\omega_{k+1} \in \Pi(p_{k+1})$ such that

$$\left\|T_k^{-1}\xi(p_{k+1}) + T_k^{-1}\omega_{k+1}\right\| = d\left(0_{p_k}, T_k^{-1}\xi(p_{k+1}) + T_k^{-1}\Pi(p_{k+1})\right) \leqslant a_k + b_k + r_k.$$

Thus,

$$\begin{aligned}
\left\|-S_{k+1}^{-1}\omega_{k+1} + \xi(\bar{p})\right\| &\leqslant \left\|S_{k+1}^{-1}\omega_{k+1} + S_{k+1}^{-1}\xi(p_{k+1})\right\| + \left\|S_{k+1}^{-1}\xi(p_{k+1}) - \xi(\bar{p})\right\| \\
&\leqslant \left\|S_{k+1}^{-1}S_kT_k\right\| \left\|T_k^{-1}\omega_{k+1} + T_k^{-1}\xi(p_{k+1})\right\| + \left\|S_{k+1}^{-1}\xi(p_{k+1}) - \xi(\bar{p})\right\| \\
&\leqslant \left\|S_{k+1}^{-1}S_kT_k\right\| (a_k + b_k + r_k) + \|h(v_{k+1}) - h(0_{\bar{p}})\|.
\end{aligned}$$

Taking into account the assumptions (C1), (C2), (C3) and (3.10), the continuity of $h$ gives us $\lim_{k\to\infty} \left\|-S_{k+1}^{-1}\omega_{k+1} + \xi(\bar{p})\right\| = 0$, and in consequence the inclusion $0_{\bar{p}} \in \xi(\bar{p}) + \Pi(\bar{p})$ is shown.

We include now the hypothesis $\operatorname{subreg}(\xi + \Pi)_{\mathcal{T}}(\bar{p}, 0_{\bar{p}}) < +\infty$ which is in the sense of Definition 2.1. Choosing some parameters $\kappa > 0$, $r > 0$ and $r' > 0$ such that

$$d(p, \bar{p}) < r \Longrightarrow d(p, \bar{p}) \leqslant \kappa \inf \left\{ \left\| \mathcal{T}_u^{-1} \zeta \right\| \;\middle|\; \begin{array}{l} R_{\bar{p}}^{-1}(p) = u,\ \zeta \in \xi(p) + \Pi(p), \\ \left\| \mathcal{T}_u^{-1} \zeta \right\| < r' \end{array} \right\}. \tag{3.14}$$

Recall that $\left\| T_k^{-1} \xi(p_{k+1}) + T_k^{-1} \omega_{k+1} \right\| = d\big(0_{p_k}, T_k^{-1} \xi(p_{k+1}) + T_k^{-1} \Pi(p_{k+1})\big)$. Define $\zeta_{k+1} = \xi(p_{k+1}) + \omega_{k+1}$, then this new element is in $\xi(p_{k+1}) + \Pi(p_{k+1})$. From Lemma 3.1, we infer

$$\left\| S_{k+1}^{-1} \zeta_{k+1} \right\| \leqslant \left\| S_{k+1}^{-1} T_k \right\| \left\| T_k^{-1} \zeta_{k+1} \right\| \leqslant \left\| S_{k+1}^{-1} T_k \right\| (a_k + b_k + r_k),$$

which yields $\left\| S_{k+1}^{-1} \zeta_{k+1} \right\| \to 0$ by invoking (C1), (C2), (C4) and (3.10). Consequently, $d(p_{k+1}, \bar{p}) < r$ and $\left\| S_{k+1}^{-1} \zeta_{k+1} \right\| < r'$ when $k$ is large. Applying (3.14) for $p = p_{k+1}$ and $\zeta = \zeta_{k+1}$, we find

$$d(p_{k+1}, \bar{p}) \leqslant \kappa \left\| S_{k+1}^{-1} \zeta_{k+1} \right\| \leqslant \kappa \left\| S_{k+1}^{-1} T_k \right\| (a_k + b_k + r_k) = \nu_k \left\| u_k \right\| + \nu_k' \alpha_k, \tag{3.15}$$

where $\nu_k = \kappa \left\| S_{k+1}^{-1} T_k \right\| (s_k + s_k')$ and $\nu_k' = \kappa \left\| S_{k+1}^{-1} T_k \right\| \alpha_k^{-1} r_k$. It follows from the triangle inequality in $\mathbb{M}$ that

$$\beta_k = d(p_k, p_{k+1}) \leqslant d(\bar{p}, p_k) + d(\bar{p}, p_{k+1}) = (1 + \nu_k')\alpha_k + \nu_k \left\| u_k \right\|. \tag{3.16}$$

Define $\mu_k = d(p_{k+1}, p_k) / \left\| u_k \right\|$. Then we deduce from (3.16)

$$\mu_k \beta_k \leqslant \mu_k (1 + \nu_k') \alpha_k + \nu_k \beta_k. \tag{3.17}$$

Taking into account (C3), one has

$$0 < \liminf_{k \to \infty} \mu_k \leqslant \limsup_{k \to \infty} \mu_k < +\infty. \tag{3.18}$$

Further, using (C4), the sequence $\nu_k = \kappa \left\| S_{k+1}^{-1} T_k \right\| (s_k + s_k')$ converges to zero, thanks to the fact that of $s_k \to 0$ and $s_k' \to 0$. Thus, $\lim_{k \to \infty} \frac{\nu_k}{\mu_k} = 0$, which permits us to conclude after skipping finite many indices $k$ that

$$\beta_k \leqslant \frac{\mu_k (1 + \nu_k')}{\mu_k - \nu_k} \alpha_k. \tag{3.19}$$

Next, invoking both (3.15) and (3.19), we obtain

$$\alpha_{k+1} \leqslant \nu_k \mu_k^{-1} \beta_k + \nu_k' \alpha_k \leqslant \nu_k \frac{1 + \nu_k'}{\mu_k - \nu_k} \alpha_k + \nu_k' \alpha_k = \frac{\nu_k}{\mu_k} \frac{1 + \nu_k'}{1 - \nu_k / \mu_k} \alpha_k + \nu_k' \alpha_k. \tag{3.20}$$

Finally, by observing that $\nu_k' \to 0$ and $\nu_k / \mu_k \to 0$, we deduce

$$\limsup_{k \to \infty} \frac{d(p_{k+1}, \bar{p})}{d(p_k, \bar{p})} = \limsup_{k \to \infty} \frac{\alpha_{k+1}}{\alpha_k} \leqslant \limsup_{k \to \infty} \left( \frac{\nu_k}{\mu_k} \frac{1 + \nu_k'}{1 - \nu_k / \mu_k} + \nu_k' \right) = 0, \tag{3.21}$$

and the proof is complete. □

**Remark 3.4.** It is obvious that the conditions (C1), (C2), (C3) and (C4) play essential roles in the proof of Theorem 3.3. (C4) is of course trivial when considering the vector transport $\mathcal{T}$ induced by parallel translation due to the isometric property. The condition (C3) is automatically valid if $R = \text{Exp}$ [37]. For an abstract retraction, the paper [32] provides a class of retractions guaranteeing not only (C3) but something also stronger. In the case $\xi$ and $\mathcal{T}$ are smooth enough, e.g. they are of class $C^2$, an immediate candidate for choosing $\mathcal{A}_k$ which ensures (C1) is the composition $\mathcal{A}_k = S_k^{-1} Dh_k(0_{p_k}) S_k$. Here, $h_k$ stands for the map sending $v \in T_{p_k}\mathbb{M}$ into $\mathcal{T}_v^{-1}\xi\left(R_{p_k}(v)\right)$, and $Dh_k(0_{p_k})$ indicates the derivative of $h_k$ at the origin $0_{p_k}$. The validity of (C1) with $\mathcal{A}_k = S_k^{-1} Dh_k(0_{p_k}) S_k$ was mentioned in the work [18]. Further, as proved in [18, Lemma 14.6], in the case $\bar{p}$ is a nondegenerate zero to $\xi$, the condition (3.8) holds automatically. □

By subsuming (3.1) as a particular case of (3.2), we obtain a similar result for the convergence of algorithm (1.4) which is an extension of Riemannian Dennis-Moré investigated in [18]. The next result is in this sense.

**Corollary 3.5.** *With $\xi$, $\varPi$, $R$ and $\mathcal{T}$ as in Theorem* 3.3*, assume the convergent sequence* $(p_k)$ *is now induced by the exact quasi-Newton scheme* (3.1)*. Let us suppose now three assumptions* (C1), (C3) *and* (C4) *are fulfilled. Then both statements* (i) *and* (ii) *of Theorem* 3.3 *are still valid as well.*

**Proof.** The proof can be obtained by taking $r_k = 0$ in Theorem 3.3. □

**Remark 3.6.** If we let $\varPi(p) = \{0_p\}$, then Corollary 3.5 (more general, Theorem 3.3) covers the Riemannian Dennis-Moré condition obtained in the work [18]. Indeed, assuming $\xi$ and $\mathcal{T}$ are both $C^2$ and the map $v \in T_{\bar{p}}\mathbb{M} \longmapsto \nabla_v \xi$ is nonsingular. Let $\mathcal{A}_k = S_k^{-1} Dh(0_{\bar{p}}) S_k$ for the map $h$ mentioned in the statement of Theorem 3.3. Due to [18], the conditions (C1) and (3.8) are satisfied (see in [18, Lemma 14.6] and also the proof of [18, Theorem 14.1]). Thus, Theorem 3.3 recovers the one in [18, Theorem 14.1]. Note that for such a situation, the two relations (3.9) and (3.10) become the same. □

In Theorem 3.3 and Corollary 3.5, the sufficient conditions ensuring the superlinear convergence of frameworks (3.1) and (3.2) depend on the behavior of $\xi + \varPi$ around an expected solution $\bar{p}$ to the reference problem (1.1). Nonetheless, it is not easy to localize in practice the information about $\bar{p}$. Those aforementioned results would be more useful if their assumptions do not depend on the solution $\bar{p}$. The next theorem is stated in this sense.

**Theorem 3.7.** *Let $\xi$ and $\varPi$ be as similar as in Theorem* 3.3*. Suppose that the sequence* $(p_k)$ *in* $\mathbb{M}$ *is generated by either* (3.1) *or* (3.2) *converging to* $\bar{p}$*. Let* $\mathcal{H}_k : T_{p_k}\mathbb{M} \rightrightarrows T_{p_k}\mathbb{M}$ *be a sequence of set-valued mappings. Assume that the hypotheses* (C2), (C3) *and* (C4) *in Theorem* 3.3 *are fulfilled, while* (C1) *is replaced by the following*

$$\lim_{k\to\infty} \left\{ \frac{1}{\|u_k\|} \sup_{w \in \mathcal{H}_k(u_k)} \left\| T_k^{-1}\xi(p_{k+1}) - \xi(p_k) - w \right\| \right\} = 0. \qquad \text{(C1')}$$

*Then the same conclusions as in the statements* (i) *and* (ii) *of Theorem* 3.3 *are still valid by substituting* $S_k \mathcal{A}_k S_k^{-1}$ *with* $\mathcal{H}_k$*.*

**Proof.** The proof is very similar to the one of Theorem 3.3. Indeed, it suffices to interchange the role of $S_k \mathcal{A}_k S_k^{-1}$ in the proof of Theorem 3.3 by $\mathcal{H}_k$ and invoke (C1') instead of (C1). $\square$

## 4. Local convergence of the Riemannian Broyden-type update

This section is devoted to study a class of Broyden type updates for solving the model (1.1). Such a strategy was inspired from the classical quasi-Newton method applied to nonlinear equations in vectorial spaces and also to the framework of manifolds. We refer to [12, 28] for a survey to quasi-Newton method and [5, 8] for the corresponding extensions applied to generalized equation in Banach spaces setting. Further discussions about the well-known Broyden method can be found in [20, 36], and the paper [22] deals with a Riemannian Broyden update for Riemannian optimization problems.

We shall focus only on the case $R = \text{Exp}$ whereas $\mathcal{T}$ induced from the parallel transport as comments in the beginning of the paper. In other words, this section will deal with the following scheme

$$\begin{cases} 0_{p_k} \in \xi(p_k) + B_k(u_k) + \mathcal{T}_{u_k}^{-1} \varPi\big(\exp_{p_k}(u_k)\big), \\ p_{k+1} = \exp_{p_k}(u_k), \end{cases} \tag{4.1}$$

for $k = 0, 1, 2, \ldots$, where $B_{k+1} : T_{p_{k+1}}\mathbb{M} \longrightarrow T_{p_{k+1}}\mathbb{M}$ is obtained through the following Broyden-type updating formula

$$B_{k+1}(\cdot) = \widetilde{B}_k(\cdot) + \frac{\langle \mathcal{T}_{u_k} u_k, \cdot\rangle}{\|u_k\|^2}\big[\xi(p_{k+1}) - \mathcal{T}_{u_k}\xi(p_k) - \mathcal{T}_{u_k} B_k(u_k)\big], \tag{4.2}$$

for $\widetilde{B}_k = \mathcal{T}_{u_k} B_k \mathcal{T}_{u_k}^{-1}$. With the goal of establishing local convergence results, we shall keep throughout this section that $\bar{p}$ is a solution of the problem $0 \in \xi(p) + \varPi(p)$. The following standing assumptions will be crucial in our analyses.

**Assumption 4.1.** Let $i_{\mathbb{M}}(p)$ denote the *injectivity radius* of $\mathbb{M}$ at the reference point $p \in \mathbb{M}$ (see [37]). We suppose that there exist two positive constants $\sigma_1$, $\sigma_2$ satisfying

$$\sigma_1 \|u - v\| \leqslant d\big(\exp_p(u), \exp_p(v)\big) \leqslant \sigma_2 \|u - v\| \tag{4.3}$$

if $d(p, \bar{p}) < i_{\mathbb{M}}(\bar{p})$ and $u, v$ in $T_p\mathbb{M}$ with $\|u\| < i_{\mathbb{M}}(p)$, $\|v\| < i_{\mathbb{M}}(p)$.

**Remark 4.2.** By considering $u = 0_p$ in (4.3), we get $\sigma_1 \leqslant 1 \leqslant \sigma_2$ since we have $\|v\| = d(p, \exp_p(v))$. As remarked in [10], if $\mathbb{M}$ has a non-negative sectional curvature (e.g., the unit sphere $\mathbb{S}^m$) then $\sigma_2 = 1$. Conversely, in the case $\mathbb{M}$ inherits a non-positive sectional curvature (e.g., Hadamard spaces), then one can choose $\sigma_1 = 1$ (see [14]). Of course, it is trivial to see that $i_{\mathbb{M}}(p) = +\infty$ as well as $\sigma_1 = \sigma_2 = 1$ for any Euclidean space. $\square$

In order to study the convergence behavior of the proposed scheme (4.2), the notion of Riemannian curvature plays an important role in our analysis. Let $\mathcal{R}$ denote the Riemannian curvature of the manifold $\mathbb{M}$ (see [37])

$$\mathcal{R}(X, Y)Z := \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z, \tag{4.4}$$

where $[\cdot,\cdot]$ stands for the Lie bracket. The *sectional curvature* $K_\Sigma = K_p(x,y)$ of a plane $\Sigma \subset T_p\mathbb{M}$ spanned by two linearly independent vectors $x, y \in T_p\mathbb{M}$ is the quantity

$$K_\Sigma = K_p(x,y) := \frac{\langle \mathcal{R}(X,Y)Y, X\rangle}{\|X\|^2\|Y\|^2 - \langle X, Y\rangle^2}\Big|_p; \quad X_p = x, \quad Y_p = y. \tag{4.5}$$

Note that the expression in (4.5) does not depend on the choice of vector fields $X$ and $Y$. The next assumption imposed on the curvature is needed.

**Assumption 4.3.** Let $\Omega_{\bar p}$ be a totally normal neighborhood at $\bar p \in \mathbb{M}$, i.e., the one such that for any pair of its points $(p,q)$ there always exists a minimizing geodesic segment lying inside $\Omega_{\bar p}$ and joining $p$ to $q$. There exists some positive constant $C_{\mathcal{R}}$ such that for all $p \in \Omega_{\bar p}$ and for $v, w$ belonging to the unit sphere of $T_p\mathbb{M}$, one has

$$\|\mathcal{R}(v,w)\| \leqslant C_{\mathcal{R}}. \tag{4.6}$$

Here the expression $\|\mathcal{R}(v,w)\|$ stands for the norm of the linear transformation $\mathcal{R}(v,w)(\cdot) : T_p\mathbb{M} \longrightarrow T_p\mathbb{M}$ taken in the corresponding inner product of $T_p\mathbb{M}$.

**Remark 4.4.** Assumption 4.3 is a quantitative property depending on the behavior of curvature tensor $\mathcal{R}$. It is satisfied for manifolds satisfying some suitable bound imposed on the sectional curvature. As an illustration, we consider a simple case where $M$ has a constant sectional curvature $K$. In this case, according to [14, Chapter 4], $\mathcal{R}$ verifies the following rule

$$\langle \mathcal{R}(X,Y)Z, W\rangle = K\big(\langle X,Z\rangle\langle Y,W\rangle - \langle Y,Z\rangle\langle X,W\rangle\big). \tag{4.7}$$

As a result, (4.6) is straightforward with $C_{\mathcal{R}} \geqslant |K|$. □

**Remark 4.5.** Suppose that Assumption 4.3 is fulfilled. Since $\mathcal{R}$ is trilinear, one has

$$\big\|\mathcal{R}(X,Y)Z\big\| \leqslant C_{\mathcal{R}}\|X\|\|Y\|\|Z\|. \tag{4.8}$$

By using (4.8), we may infer some useful upper bound results as in the next two propositions. Further, since $\mathcal{R}$ is fixed by the manifold, we shall ignore the subscript and write briefly $C$ instead of $C_{\mathcal{R}}$.

**Proposition 4.6.** (Bound for the Jacobi field) *Suppose that $J$ is a vector field along geodesic $\gamma$ on $\mathbb{M}$ satisfying the differential equation $\nabla_{\gamma'}\nabla_{\gamma'}J = \mathcal{R}(\gamma', J)\gamma'$ with the initial data $J(0) = J_0$ and $\nabla_{\gamma'}J\big|_{t=0} = \dot J_0$. If the image of $\gamma$ is contained inside the neighborhood $\Omega_{\bar p}$ as in Assumption 4.3, then one has*

$$\begin{aligned} \|J(t)\| \leqslant \|\dot J_0\| &+ \big(\|J_0\| - \|\dot J_0\|\big)\cosh\big(\sqrt{C}\|\gamma'\|t\big) \\ &+ \big(\|\dot J_0\| - C\|\gamma'\|^2\|J_0\|\big)\frac{\sinh\big(\sqrt{C}\|\gamma'\|t\big)}{\sqrt{C}\|\gamma'\|}. \end{aligned} \tag{4.9}$$

**Proof.** Denoting $V = \gamma'$, by Taylor's expansion for $J$, we have

$$\|J\| = \left\|J_0 + \int_0^t P_\gamma^{s,0}\nabla_V J\, ds\right\| \leqslant \|J_0\| + \int_0^t \|\nabla_V J\|\, ds. \tag{4.10}$$

Similarly we have for $T := \nabla_V J$

$$\begin{aligned}\|T\| &\leqslant \|\dot{J}_0\| + \int_0^t \|\nabla_V T\|\, ds \leqslant \|\dot{J}_0\| + \int_0^t \|R(V,J)V\| ds \\ &\leqslant \|\dot{J}_0\| + C\|V\|^2 \int_0^t \|J\|\, ds. \end{aligned} \tag{4.11}$$

Setting $\Lambda_1 = \|J_0\| + \int_0^t \|J\|\, ds$, $\Lambda_2 = \|\dot{J}_0\| + \int_0^t \|T\|\, ds$ and $\Lambda = \begin{bmatrix}\Lambda_1 & \Lambda_2\end{bmatrix}^T$ we obtain the following system of differential equations

$$\frac{d}{dt}\Lambda = M\Lambda + b, \tag{4.12}$$

where, the matrix $M = \begin{bmatrix} 0 & 1 \\ \rho^2 & 0 \end{bmatrix}$ does not depend on the variable $t$. The vector $b = [b_1 \quad b_2]^T$ is given by

$$b_1 = \|J\| - \|\dot{J}_0\| - \int_{[0,t]} \|T\| \quad \text{and} \quad b_2 = \|T\| - \rho^2 \left( \|J_0\| + \int_{[0,t]} \|J\| \right)$$

for $\rho = \sqrt{C}\|V\|$. We are now able to integrating w.r.t. $t$ from (4.12) and obtain

$$\Lambda(t) = \exp(tM) \times \int_0^t \exp(-sM) b\, ds + \Lambda(0).$$

Here, the notation $\exp(\cdot)$ stands for the usual exponential matrix. After a simple computation we get

$$\exp(sM) = \frac{1}{2\rho} \begin{bmatrix} 1 & 1 \\ \rho & -\rho \end{bmatrix} \begin{bmatrix} e^{s\rho} & 0 \\ 0 & e^{-s\rho} \end{bmatrix} \begin{bmatrix} \rho & 1 \\ \rho & -1 \end{bmatrix} = \frac{1}{\rho} \begin{bmatrix} \rho\cosh(s\rho) & \sinh(s\rho) \\ \rho^2 \sinh(s\rho) & \rho\cosh(s\rho) \end{bmatrix},$$

and hence,

$$\Lambda = \frac{1}{\rho} \int_0^t \left( \begin{bmatrix} \rho\cosh((t-s)\rho) & \sinh((t-s)\rho) \\ \rho^2\sinh((t-s)\rho) & \rho\cosh((t-s)\rho) \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \right) ds + \begin{bmatrix} \|J_0\| \\ \|\dot{J}_0\| \end{bmatrix}. \tag{4.13}$$

Taking into account (4.10) and (4.11), it holds that $b_1 \leqslant \|J_0\| - \|\dot{J}_0\|$ as well as $b_2 \leqslant \|\dot{J}_0\| - \rho^2\|J_0\|$. Thus, $\Lambda_2$ does not exceed in the quantity

$$\frac{1}{\rho}\int_0^t \Big\{ \rho^2 \big(\|J_0\| - \|\dot{J}_0\|\big) \sinh((t-s)\rho) + \rho\big(\|\dot{J}_0\| - \rho^2\|J_0\|\big) \cosh((t-s)\rho) \Big\}\, ds + \|\dot{J}_0\|.$$

Combining this with the inequality (4.10), we obtain

$$\|J\| \leqslant \big(\|J_0\| - \|\dot{J}_0\|\big)\cosh(t\rho) + \big(\|\dot{J}_0\| - \rho^2\|J_0\|\big)\frac{\sinh(t\rho)}{\rho} + \|\dot{J}_0\|. \tag{4.14}$$

The proof is thereby complete. $\square$

**Proposition 4.7.** (Upper bound for the change of parallelism) *Assume that Assumption 4.3 is fulfilled. Let $p \in \Omega_{\bar{p}}$ be given and let $\gamma : [0,1] \longrightarrow \mathbb{M}$ be a geodesic segment whose image lies into $\Omega_{\bar{p}}$. Then the following estimation is valid*

$$\left\| P_\gamma^{1,0} \tau_{p,\gamma(1)} - \tau_{p,\gamma(0)} \right\| \leqslant \sqrt{C} \ell_0^1(\gamma) \int_0^1 \varphi\big(\rho_{p,\gamma}(t)\big)\, dt \tag{4.15}$$

*for $\rho_{p,\gamma}(t) = \sqrt{C} d\big(p, \gamma(t)\big)$ and $\varphi(\rho) = \sigma_1^{-1}(\rho \cosh \rho - \sinh \rho)$.*

Here and in what follows, by $\tau_{p,q}$ we mean the parallel transportation along the minimizing geodesic joining $p$ to $q$.

**Proof.** Let us fix a unit tangent vector $v_p \in T_p\mathbb{M}$ and define a local vector field $X : q \in \Omega_{\bar{p}} \longmapsto \tau_{p,q} v_p$. In order to find a differential equation for $X$, we define a smooth map $\chi(s,t) = \exp_p\big(s \log_p(\gamma(t))\big)$ where $\log_p$ denotes the inverse of $\exp_p$. Let $V = \frac{\partial \chi}{\partial s} = d\chi\left(\frac{\partial}{\partial s}\right)$ and $T = \frac{\partial \chi}{\partial t} = d\chi\left(\frac{\partial}{\partial t}\right)$ be vector fields along $\chi$. $T$ and $V$ satisfy $[T,V] = -[V,T] = 0$, and hence

$$\nabla_V \nabla_T X = \nabla_T \nabla_V X + \mathcal{R}(V,T)X. \tag{4.16}$$

Nevertheless, we have

$$\begin{aligned} \nabla_V X\big|_{s=s_0} &= \frac{d}{ds} P^{s,s_0}_{\chi(\cdot,t)} X\big(\chi(s,t)\big)\Big|_{s=s_0} = \frac{d}{ds} P^{s,s_0}_{\chi(\cdot,t)} \tau_{p,\chi(s,t)} v_p\Big|_{s=s_0} \\ &= \frac{d}{ds} P^{s,s_0}_{\chi(\cdot,t)} P^{0,s}_{\chi(\cdot,t)} v_p\Big|_{s=s_0} = \frac{d}{ds} P^{0,s_0}_{\chi(\cdot,t)} v_p\Big|_{s=s_0} = 0, \end{aligned}$$

which gives us

$$\nabla_V \nabla_T X = \mathcal{R}(V,T)X. \tag{4.17}$$

Thus, by Taylor's expansion, one has

$$P^{s,0}_{\chi(\cdot,t)} \nabla_T X = \nabla_T X\big|_{s=0} + \int_0^s P^{s',0}_{\chi(\cdot,t)} \nabla_V \nabla_T X\, ds' = \int_0^s P^{s',0}_{\chi(\cdot,t)} \mathcal{R}(V,T)X\, ds'. \tag{4.18}$$

The second equality is due to the fact $T\big|_{s=0} = \frac{d}{dt}\chi(0,t) = 0$. Similarly, we get

$$\begin{aligned} \left(P_\gamma^{1,0}\tau_{p,\gamma(1)} - \tau_{p,\gamma(0)}\right) v_p &= P_\gamma^{1,0} X(\gamma(1)) - X(\gamma(0)) = \int_0^1 \left( P_\gamma^{t,0} \nabla_T X\big|_{s=1}\right) dt \\ &= \int_0^1 P_\gamma^{t,0} \left( \tau_{p,\chi(1,t)} \int_0^1 \tau^{-1}_{p,\chi(s,t)} \mathcal{R}(V,T)X\, ds \right) dt \end{aligned}$$

where the relation $\tau_{p,\chi(s,t)} = P^{0,s}_{\chi(\cdot,t)}$ was used. As a result, we derive

$$\begin{aligned} \left\| \left(P_\gamma^{1,0}\tau_{p,\gamma(1)} - \tau_{p,\gamma(0)}\right) v_p \right\| &\leqslant \int_0^1 \left\| \int_0^1 \tau^{-1}_{p,\chi(r,t)} \mathcal{R}(V,T)X\, ds \right\| dt \\ \leqslant \int_0^1 \left( \int_0^1 \left\| \mathcal{R}(V,T)X \right\| ds \right) dt &\leqslant \int_0^1 \left( \int_0^1 C\|V\|\|T\|\|X\|\, ds \right) dt. \end{aligned}$$

From the definition of $V$ we get $\|V\| = \|V\|\big|_{s=0} = \|\log_p(\gamma(t))\| = \frac{1}{\sqrt{C}}\rho_{p,\gamma}(t)$.
Moreover, by replacing $V$ with $X$ in (4.16) and taking into account $\nabla_V V = 0$ and $\nabla_T V = \nabla_V T + [T,V] = \nabla_V T$, we reach the following equation:

$$\nabla_V \nabla_V T = \mathcal{R}(V,T)V.$$

Thus, Proposition 4.6 yields

$$\begin{aligned}\|T\| \leqslant \|\nabla_V T\|\big|_{s=0} &+ \big(\|T\|\big|_{s=0} - \|\nabla_V T\|\big|_{s=0}\big)\cosh\big(s\rho_{p,\gamma}\big)\\ &+ \big(\|T\|\big|_{s=0} - \rho_{p,\gamma}^2\|\nabla_V T\|\big|_{s=0}\big)\frac{\sinh\big(s\rho_{p,\gamma}\big)}{\rho_{p,\gamma}}.\end{aligned}$$

Observe that $T\big|_{s=0} = \frac{d}{dt}\chi(0,t) = 0$, while for any $t_0 \in [0,1]$,

$$\nabla_V T\big|_{s=0,t=t_0} = \nabla_T V\big|_{s=0,t=t_0} = \lim_{t\to t_0}\frac{P^{t,t_0}_{\chi(0,\cdot)}\frac{\partial\chi}{\partial s}(0,t) - \frac{\partial\chi}{\partial s}(0,t_0)}{t-t_0} = \frac{d}{dt}\log_p(\gamma(t))\Big|_{t=t_0}.$$

Using Assumption 4.1 we deduce for $t,t' \in [0,1]$,

$$\sigma_1\|\log_p(\gamma(t)) - \log_p(\gamma(t'))\| \leqslant d\big(\gamma(t),\gamma(t')\big) \leqslant \Big|\int_t^{t'}\|\gamma'(s)\|\,ds\Big| = |t-t'|\ell_0^1(\gamma),$$

which implies $\big\|\frac{d}{dt}\log_p\big(\gamma(t)\big)\big\| \leqslant \sigma_1^{-1}\ell_0^1(\gamma)$, for all $t \in [0,1]$. In summary, we get

$$\|T\| \leqslant f\big(s,\rho_{p,\gamma}(t)\big)\ell;\quad \ell = \ell_0^1(\gamma),\quad f(s,\rho) = \sigma_1^{-1}\rho\big[1-\cosh(s\rho) + \rho\sinh(s\rho)\big]. \tag{4.19}$$

Finally, since $\|X(q)\| = \|\tau_{p,q}v_p\| = \|v_p\|$, we derive immediately the desired conclusion, so the proof is complete. □

The next theorem provides a sufficient condition for the existence of a solution to the subproblem (4.1) which is crucial for our convergence results thereafter. Since in the present section, we consider only the case of the exponential map and the parallel transport, the short notation $\operatorname{reg}\Upsilon(\bar p, 0_{\bar p})$ should be used instead of $\operatorname{reg}_{R,\mathcal{T}}\Upsilon(\bar p, 0_{\bar p})$.

**Theorem 4.8.** *Suppose that $\Upsilon = \xi + \Pi$ satisfies $\operatorname{reg}\Upsilon(\bar p, 0_{\bar p}) < +\infty$. Assume further that $r_1 \leqslant \iota = \inf\{i_{\mathbb{M}}(p) : p \in \Omega_{\bar p}\}$, $r_2$ and $\kappa$ are positive constants for which the estimation described in* (2.8) *holds (recall here $R = \operatorname{Exp}$ and $\mathcal{T}$ being induced by parallel transport). Let Assumption* 4.1 *be valid with respect to $\Omega_{\bar p} = \mathbb{B}_{\mathbb{M}}(\bar p, r_1)$. Pick some positive parameters $\delta$ and $L$ such that*

$$L\kappa < 1,\quad \frac{\kappa}{1-\kappa L}\,\delta < r_1,$$

*and*

$$\frac{\kappa}{1-\kappa L}\,\delta < r_2. \tag{4.20}$$

*Let $p$ in $\Omega_{\bar p}$ and let $A : T_p\mathbb{M} \longrightarrow T_p\mathbb{M}$ be a linear map. We define locally a vector field $\zeta_A : q \in \Omega_{\bar p} \longmapsto \xi(q) - \tau_{p,q}A\big(\exp_p^{-1}(q)\big) - \tau_{p,q}\xi(p)$. Suppose further:*

(i) $\zeta_A$ *satisfies the following estimation along any geodesic* $\gamma : [0,1] \longrightarrow \mathbb{M}$ *whose image* $\gamma([0,1])$ *lies inside* $\Omega_{\bar{p}}$

$$\left\|P_{\gamma}^{1,0}\zeta_A(\gamma(1)) - \zeta_A(\gamma(0))\right\| \leqslant L\ell_0^1(\gamma); \tag{4.21}$$

(ii) $\left\|\xi(p) + A(\bar{v}) - \mathcal{T}_{\bar{v}}^{-1}\xi(\bar{p})\right\| \leqslant \delta$, *in which* $\bar{v} = \exp_p^{-1}(\bar{p})$.

*Then the following inclusion*

$$0_p \in \xi(p) + A(v) + \mathcal{T}_v^{-1}\Pi\left(\exp_p(v)\right), v \in T_p\mathbb{M} \tag{4.22}$$

*admits a solution* $v^*$ *such that*

$$\sigma_1\|v^* - \bar{v}\| \leqslant \frac{\kappa}{1-\kappa L}\left\|\xi(p) + A(\bar{v}) - \mathcal{T}_{\bar{v}}^{-1}\xi(\bar{p})\right\|. \tag{4.23}$$

According to [37], the function $p \longmapsto i_{\mathbb{M}}(p)$ is continuous, hence, the existence of such a radius $r_1$ in the statement of Theorem 4.8 is evident. We shall need an auxiliary result below.

**Lemma 4.9.** *Let* $q_1, q_2$ *be two given points in* $\Omega_{\bar{p}}$. *Define the points* $u_1 = \exp_p^{-1}(q_1)$, $u_2 = \exp_p^{-1}(q_2)$ *and* $u_{q_1} = \exp_{q_1}^{-1}(q_2)$. *Then one has*

$$\begin{aligned}\left\|\left[\xi(q_2) - \mathcal{T}_{u_2}\left(A\left(u_2\right)\right) - \mathcal{T}_{u_2}\xi(p)\right] - \mathcal{T}_{u_{q_1}}\left[\xi(q_1) - \mathcal{T}_{u_1}\left(A\left(u_1\right)\right) - \mathcal{T}_{u_1}\xi(p)\right]\right\| \\ \leqslant Ld(q_2, q_1).\end{aligned} \tag{4.24}$$

**Proof of Lemma 4.9.** Define the smooth curves

$$\gamma_1(t) = \exp_p(tu_1),\ \gamma_{q_1}(t) = \exp_{q_0}(tu_{q_1}) \text{ and } \gamma_2(s) = \exp_p(su_2).$$

Then $\mathcal{T}_{u_1}$, $\mathcal{T}_{u_2}$ and $\mathcal{T}_{u_{q_1}}$ coincide in turn with the parallel transports $P_{\gamma_1}^{0,1}$, $P_{\gamma_2}^{0,1}$ and $P_{\gamma_{q_1}}^{0,1}$, respectively. By the choice of $\Omega_{\bar{p}}$, both $\gamma_1$ and $\gamma_2$ are minimizing geodesics. Thus, the left-hand side of (4.24) clearly expresses the quantity

$$\left\|\zeta_A\left(\gamma_{q_1}(1)\right) - P_{\gamma_{q_1}}^{0,1}\zeta_A\left(\gamma_{q_1}(0)\right)\right\|.$$

We obtain the necessary conclusion in (4.24) by combining this with (4.21). □

**Proof of Theorem 4.8.** The strategy is to build a suitable sequence $q_k \in \mathbb{M}$ converging to some $q^* \in \mathbb{M}$ with $\exp_p^{-1}(q^*) := v^*$ being a solution of (4.22). At the starting, we set $q_0 = \bar{p}$, $u_0 = \exp_{\bar{p}}^{-1}(q_0)$, $v_0 = \exp_p^{-1}(q_0) \in T_p\mathbb{M}$ and

$$w_{q_0} = \xi(q_0) - \mathcal{T}_{v_0}\left[\xi(p) + A(v_0)\right] \in T_{q_0}\mathbb{M}.$$

As $q_0 = \bar{p}$, we have $w_{q_0} = \xi(\bar{p}) - \mathcal{T}_{\bar{v}}\left[\xi(p) + A(\bar{v})\right]$, which yields

$$\left\|w_{q_0}\right\| = \left\|\xi(p) + A(\bar{v}) - \mathcal{T}_{\bar{v}}^{-1}\xi(\bar{p})\right\| \leqslant \delta < r_2.$$

Therefore, $$d\Big(q_0, \Big\{q : \Upsilon(q) \ni \mathcal{T}_{\exp_{q_0}^{-1}(q)}w_{q_0}\Big\}\Big) \leqslant \kappa d\big(w_{q_0}, \Upsilon(q_0)\big). \tag{4.25}$$

Thus, there exists a point $q_1 \in \mathbb{M}$ at which the distance in the left-hand side of (4.25) attains. In other words,

$$d(q_0, q_1) \leqslant \kappa d\big(w_{q_0}, \Upsilon(q_0)\big) \leqslant \kappa \|w_{q_0}\| = \kappa \left\|\xi(p) + A(\bar{v}) - \mathcal{T}_{\bar{v}}^{-1}\xi(\bar{p})\right\| \leqslant \kappa\delta. \quad (4.26)$$

We notice that $\kappa\delta < r_1$, so $q_1$ still lies inside $\Omega_{\bar{p}}$. To continue the current process, let us define some tangent vectors $v_1 = \exp_p^{-1}(q_1)$, $u_1 = \exp_{q_0}^{-1}(q_1)$. Put

$$w_{q_1} = \xi(q_1) - \mathcal{T}_{v_1}\left[\xi(p) + A(v_1)\right] \in T_{q_1}\mathbb{M},$$

we claim that $\|w_{q_1}\| < r_2$. Indeed, the quantity $\|w_{q_1}\|$ can be dominated as follows

$$\|w_{q_1}\| \leqslant \left\| (w_{q_1} - \mathcal{T}_{u_1} w_{q_0}) + \mathcal{T}_{u_1} w_{q_0}\right\| \leqslant \left\|w_{q_1} - \mathcal{T}_{u_1} w_{q_0}\right\| + \left\|w_{q_0}\right\|.$$

But it is easy to see that

$$\begin{aligned} w_{q_1} - \mathcal{T}_{u_1} w_{q_0} &= \left[\xi(q_1) - \mathcal{T}_{v_1}\left(A(v_1)\right) - \mathcal{T}_{v_1}\xi(p)\right] \\ &\quad - \mathcal{T}_{u_1}\left[\xi(q_0) - \mathcal{T}_{v_0}\left(A(v_0)\right) - \mathcal{T}_{v_0}\xi(p)\right]. \end{aligned}$$

Therefore, we infer from Lemma 4.9 the estimation

$$\|w_{q_1} - \mathcal{T}_{u_1} w_{q_0}\| \leqslant L d(q_1, q_0).$$

Taking into account $\|w_{q_0}\| \leqslant \delta$, (4.20) shows that $\|w_{q_1}\| < r_2$. By applying Definition 2.2 into the pair of $(q_1, w_{q_1})$, we get

$$d\Big(q_1, \Big\{q : \Upsilon(q) \ni \mathcal{T}_{\exp_{q_1}^{-1}(q)} w_{q_1}\Big\}\Big) \leqslant \kappa d\big(w_{q_1}, \Upsilon(q_1)\big). \quad (4.27)$$

Similar to the case $k = 0$ it is possible to select a point $q_2$ such that we have $\Upsilon(q_2) \ni \mathcal{T}_{\exp_{q_1}^{-1}(q_2)} w_{q_1}$ and

$$d(q_1, q_2) \leqslant \kappa d\big(w_{q_1}, \Upsilon(q_1)\big) \leqslant \kappa \big\|w_{q_1} - \mathcal{T}_{u_1} w_{q_0}\big\| \leqslant \kappa L d(q_1, q_0).$$

The second inequality here is due to the choice of $q_1$, that guarantees $\Upsilon(q_1) \ni \mathcal{T}_{u_1} w_{q_0}$.
Passing to the induction step, suppose that $q_0, q_1, \ldots, q_k$ are given points for some $k \geqslant 2$. As suggested from the aforementioned arguments, we involve the following relations:

- for each $j \leqslant k-1$ one has $\mathcal{T}_{u_j} w_{q_j} \in \Upsilon(q_{j+1})$, where $u_j = \exp_{q_j}^{-1}(q_{j+1})$ and
$$w_{q_j} = \xi(q_j) - \mathcal{T}_{v_j}\left[\xi(p) + A(v_j)\right], v_j = \exp_p^{-1}(q_j);$$
- $d(q_j, q_{j+1}) \leqslant \kappa \big\|w_{q_j} - \mathcal{T}_{u_{j-1}}(w_{q_{j-1}})\big\| \leqslant (\kappa L)^j d(q_1, q_0)$, $j = 1, \ldots, k-1$.

Our goal now is to generate $q_{k+1}$. Using the triangle inequality many times, one has

$$d(q_0, q_j) \leqslant \sum_{i=0}^{j-1} d(q_i, q_{i+1}) \leqslant \sum_{i=0}^{j-1} (\kappa L)^i d(q_1, q_0) \leqslant \frac{1}{1-\kappa L}\kappa\delta < r_1.$$

In particular, all iterations $q_j$ are in $\Omega_{\bar{p}}$. Let's set $v_k = \exp_p^{-1}(q_k)$ and

$$w_{q_k} = \xi(q_k) - \mathcal{T}_{v_k}\left[\xi(p) + A(v_k)\right].$$

Following a mimic strategy as in the case $k = 1$ above we get

$$\left\|w_{q_j} - \mathcal{T}_{u_j} w_{q_{j-1}}\right\| \leqslant Ld(q_j, q_{j-1}), \quad 1 \leqslant j \leqslant k,$$

and this implies

$$\|w_{q_k}\| \leqslant \sum_{j=0}^{k-1} L\,(\kappa L)^j\, d(q_1, q_0) + \|w_{q_0}\| \leqslant \frac{\kappa L}{1-\kappa L}\,\|w_{q_0}\| + \|w_{q_0}\| = \frac{1}{1-\kappa L}\,\|w_{q_0}\|\,.$$

Recalling $\|w_{q_0}\| \leqslant \kappa\delta$, condition (4.20) gives us $\|w_{q_k}\| < r_2$. Invoking once more the regularity property for $\Upsilon$, we deduce

$$d\Big(q_k, \Big\{q : \Upsilon(q) \ni \mathcal{T}_{\exp_{q_k}^{-1}(q)} w_{q_k}\Big\}\Big) \leqslant \kappa d\big(w_{q_k}, \Upsilon(q_k)\big). \tag{4.28}$$

Hence, it enables to define $q_{k+1}$ as a point in $\mathbb{M}$ at which the distance expressed in the left-hand side of (4.28) attains. Denote $u_k = \exp_{q_k}^{-1}(q_{k+1}) \in T_{q_k}\mathbb{M}$, it follows simultaneously that $\Upsilon(q_{k+1}) \ni \mathcal{T}_{u_k} w_{q_k}$ and

$$d(q_k, q_{k+1}) \leqslant \kappa d\big(w_{q_k}, \Upsilon(q_k)\big) \leqslant \kappa\left\|w_{q_k} - \mathcal{T}_{u_{k-1}}(w_{q_{k-1}})\right\| \leqslant \kappa L d(q_k, q_{k-1}).$$

The construction is thereby done.

From the constructive process above, the sequence $(q_k)$ is Cauchy, so it converges in $\mathbb{M}$. Let $q^* \in \mathbb{M}$ be the corresponding limit, and let $v^* = \exp_p^{-1}(q^*)$. To finish the proof, we are going to show that $v^*$ solves (4.22) and establish the inequality (4.23). Indeed, we infer from the construction

$$\mathcal{T}_{\exp_{q_k}^{-1}(q_{k+1})}(w_{q_k}) \in \Upsilon(q_{k+1}). \tag{4.29}$$

By virtue of the continuity for $\mathcal{T}$, and by passing to the limits in (4.29), the closedness of $\Upsilon$ gives us

$$\xi(q^*) - \mathcal{T}_{v^*}\left[\xi(p) + A(v^*)\right] \in \Upsilon(q^*).$$

In other words, $v^*$ solves (4.22). To claim (4.23), let's note that

$$\sigma_1\|\bar{v}_p - v^*\| \leqslant d\left(\exp_p(\bar{v}), \exp_p(v^*)\right) = d(\bar{p}, q^*) = d(q_0, q^*) \leqslant \sum_{k\geqslant 0} d\left(q_k, q_{k+1}\right).$$

Therefore, by including $d(q_k, q_{k+1}) \leqslant (\kappa L)^k\, d(q_1, q_0)$ and $d(q_1, q_0) \leqslant \kappa\delta$, (4.23) is derived as well. $\square$

We are now in the position of stating a local convergence theorem for scheme (4.2) with Broyden-type update (4.1). If the reference manifold $\mathbb{M}$ becomes a Euclidean space, then it subsumes somewhat studied in [8]. We recall first a Lipschitz concept for covariant derivative used in [17]. Let $\mathbb{D}\xi(p) : T_p\mathbb{M} \longrightarrow T_p\mathbb{M}$ be the linear map defined by $\mathbb{D}\xi(p)(v_p) = \nabla_{v_p}\xi$ for $v_p \in T_p\mathbb{M}$. One says that $\mathbb{D}\xi$ is Lipschitz with constant $\lambda > 0$ on the set $\Omega$, written as $\mathbb{D}\xi \in \mathrm{Lip}_\lambda(\Omega)$, providing for each geodesic $\gamma : [a, b] \longrightarrow \mathbb{M}$ with $\gamma([a, b]) \subset \Omega$ it holds that

$$\left\|\mathbb{D}\xi(\gamma(b)) - P_\gamma^{a,b}\mathbb{D}\xi(\gamma(a))P_\gamma^{b,a}\right\| \leqslant \lambda\ell_a^b(\gamma). \tag{4.30}$$

Since the parallel transport is linear isometry, the above inequality implies

$$\sup_{t\in[a,b]} \|\mathbb{D}\xi(\gamma(t))\| \leqslant \|\mathbb{D}\xi(\gamma(a))\| + \lambda\ell_a^b(\gamma). \tag{4.31}$$

**Theorem 4.10.** (Linear convergence for Broyden-type update) *Consider the inclusion $0 \in \xi(p) + \Pi(p)$ with a smooth vector field $\xi$. Suppose that both Assumptions* 4.1, 4.3 *holds in some normal neighborhood $\Omega$ of $\bar{p}$. Let $\mathbb{D}\xi \in \mathrm{Lip}_\lambda(\Omega)$ and let $\operatorname{reg} \Upsilon(\bar{p}, 0_{\bar{p}})$ be finite for $\Upsilon := \xi + \Pi$. Then, there exists $\alpha > 0$ such that the following property holds: for any starting point $p_0 \in \mathbb{B}_{\mathbb{M}}(\bar{p}, \alpha)$ and any linear map $B_0 : T_{p_0}\mathbb{M} \longrightarrow T_{p_0}\mathbb{M}$ satisfying*

$$\kappa \|\mathbb{D}\xi(p_0) - B_0\| < \frac{\sigma_1}{1 + \sigma_2}, \tag{4.32}$$

*where $\sigma_1$ and $\sigma_2$ are defined in Assumption* 4.1*, the scheme* (4.2) *with Broyden-type update* (4.1) *generates a sequence $(p_k)$ converging to $\bar{p}$ linearly.*

The proof of the theorem is based on the following lemmas.

**Lemma 4.11.** *Let $q \in \Omega$ be given, and let $u \in T_q\mathbb{M}$ with $\|u\| \le i_{\mathbb{M}}(q)$ be such that the geodesic segment $t \in [0,1] \longmapsto \exp_q(tu)$ lies inside $\Omega$. If $\mathbb{D}\xi \in \mathrm{Lip}_\lambda(\Omega)$, then one has*

$$\left\| \mathcal{T}_u^{-1} \xi(\exp_q(u)) - [A(u) + \xi(q)] \right\| \leqslant \frac{1}{2}\lambda \|u\|^2 + \|A - \mathbb{D}\xi(q)\| \, \|u\| \tag{4.33}$$

*for any linear map $A : T_q\mathbb{M} \longrightarrow T_q\mathbb{M}$. In particular,*

$$\sup_{d(q',q) \leqslant r} \|\xi(q')\| \leqslant \frac{1}{2}\lambda r^2 + \|\mathbb{D}\xi(q)\| r + \|\xi(q)\|. \tag{4.34}$$

**Proof.** According to [17, Lemma 2.3], it holds that

$$\left\| \mathcal{T}_u^{-1} \xi(\exp_q(u)) - [\mathbb{D}\xi(q)(u) + \xi(q)] \right\| \leqslant \frac{1}{2}\lambda \|u\|^2. \tag{4.35}$$

Thus, we obtain (4.33) as a direct consequence of (4.35). Let $A = \mathbb{D}\xi(q)$ and $u = \exp_q^{-1}(q')$ in (4.33), we deduce (4.34). □

**Lemma 4.12.** *Keep in mind all Assumptions as in Theorem* 4.10*. Consider a geodesic $\gamma$ whose image $\gamma([0,1])$ lies inside the ball $\mathbb{B}(p, \bar{\alpha}) \subset \Omega$ with center $p$ and radius $\bar{\alpha}$. Let $B : T_p\mathbb{M} \longrightarrow T_p\mathbb{M}$ be a linear map. Define the vector field $\zeta_B(q) = \xi(q) - \tau_{p,q} B\left(\exp_p^{-1}(q)\right) - \tau_{p,q}\xi(p)$, then one has*

$$\begin{aligned} \frac{1}{\ell_0^1(\gamma)} \left\| P_\gamma^{1,0} \zeta_B(\gamma(1)) - \zeta_B(\gamma(0)) \right\| &\leqslant \sigma_1^{-1} \|B - \mathbb{D}\xi(p)\| \\ &+ \sigma_1^{-1} [M' + \bar{\rho}\varphi(\bar{\rho}) M][\bar{\rho}\cosh(\bar{\rho}) - \sinh(\bar{\rho})] \\ &+ C\bar{\alpha}\sigma_1^{-2} \ell \big\{ M'[\bar{\rho}\cosh(\bar{\rho}) - \sinh(\bar{\rho})] + \bar{\rho}\varphi(\bar{\rho}) M \big\} [\bar{\rho}\cosh(\bar{\rho}) - \sinh(\bar{\rho})], \end{aligned} \tag{4.36}$$

*where $\bar{\rho} := \bar{\alpha}\sqrt{C}$, $M = \sup_{d(q,\bar{p}) \leqslant \bar{\alpha}} \|\mathbb{D}(q)\|$, $M' = \sup_{d(q,\bar{p}) \leqslant \bar{\alpha}, \|X_q\| \leqslant \bar{\alpha}, \|Y_q\| \leqslant 1} \left\| \nabla_Y \nabla_X \xi \right\|_q$, and $\varphi$ is defined as in Proposition* 4.7.

**Proof.** Throughout this proof we will write briefly $\ell = \ell_0^1(\gamma)$. Define the smooth maps $\phi_B : q \in \Omega \longmapsto \tau_{p,q}^{-1}\xi(q) - B \log_p(q) - \xi(p)$ and $\chi(s,t) = \exp_p(s \log_p(\gamma(t)))$.

Then, it holds that $\zeta_B = \tau_{p,q}\phi_B(q)$. Our goal is to derive an upper bound for $\nabla_{\gamma'}\zeta_B$. By the definition, one has $\zeta_B(\gamma(t)) = X_t(1)$, where the vector field $X(s,t) = X_t(s)$ is a solution of the problem

$$\nabla_{\frac{\partial\chi}{\partial s}} X_t = 0, X_t(0) = \phi_B(\gamma(t)). \tag{4.37}$$

For simplicity, let's denote $T = \frac{\partial\chi}{\partial t}$, $V = \frac{\partial\chi}{\partial s}$. According to [14, Ch. 4], we get

$$\nabla_V \nabla_T X = \mathcal{R}(V,T)X + \nabla_T \nabla_V X = \mathcal{R}(V,T)X. \tag{4.38}$$

Invoking Taylor's expansion for $\nabla_T X$ (see, e.g. [18]), we find

$$P^{1,0}_{\chi(\cdot,t)} \nabla_T X\big|_{s=1} = \nabla_T X\big|_{s=0} + \int_0^1 P^{s,0}_{\chi(\cdot,t)} \nabla_V \nabla_T X \, ds.$$

Note that $\chi(1,t) = \gamma(t)$, so $P^{1,0}_{\chi(\cdot,t)} \nabla_T X\big|_{s=1} = \tau^{-1}_{p,\gamma(t)} \nabla_{\gamma'(t)}\zeta_B$. Further, denoting $\dot{X}_0 = \nabla_T X\big|_{s=0}$, one has

$$\dot{X}_0 = \frac{d}{dt} P^{t,0}_{\chi(0,\cdot)} X_t(0)\Big|_{t=0} = \big(\phi_B\gamma\big)'(0).$$

Remind (4.38), we obtain

$$\nabla_{\gamma'(t)}\zeta_B = \tau_{p,\gamma(t)}\dot{X}_0 + \tau_{p,\gamma(t)} \int_0^1 \tau_{p,\chi(s,t)} \mathcal{R}(V,T)X \, ds,$$

which implies

$$\begin{aligned} \left\|P^{1,0}_{\gamma}\zeta_B(\gamma(1)) - \zeta_B(\gamma(0))\right\| &= \left\|\int_0^1 P^{t,0}_{\gamma} \nabla_{\gamma'(t)}\zeta_B \, dt\right\| \\ &\leqslant \int_{[0,1]} \left(\|\dot{X}_0\| + \int_{[0,1]} \|\mathcal{R}(V,T)X\| \, ds\right) dt \\ &\leqslant \int_{[0,1]} \|\dot{X}_0\| \, dt + \int_{[0,1]} \left(\int_{[0,1]} C\|V\|\|T\|\|X\| \, ds\right) dt. \end{aligned} \tag{4.39}$$

Both quantities $\|T\|$ and $\|V\|$ are estimated as similar as in the proof of Proposition 4.7. Next, we shall establish the estimations for $\|X\|$ and $\|\dot{X}_0\|$. Indeed, by the definition of $X$, it holds that

$$\|X(s,t)\| = \|X(0,t)\| = \|(\phi_B\gamma)(t)\|. \tag{4.40}$$

Restricting the vector field $\xi$ along the geodesic $\chi(\cdot,t)$, we obtain

$$\tau^{-1}_{p,\gamma(t)}\xi(\gamma(t)) = \xi(p) + \nabla_{\frac{\partial\chi}{\partial s}(0,t)}\xi + \int_0^1 \left[P^{s,0}_{\chi(\cdot,t)} \nabla_{\frac{\partial\chi}{\partial s}}\xi - \nabla_{\frac{\partial\chi}{\partial s}(0,t)}\xi\right] ds.$$

Since $\frac{\partial\chi}{\partial s}(0,t) = \log_p(\gamma(t))$, the latter is equivalent to

$$(\phi_B\gamma)(t) = \int_0^1 \left[P^{s,0}_{\chi(\cdot,t)} \nabla_{\frac{\partial\chi}{\partial s}(s,t)}\xi - \nabla_{\frac{\partial\chi}{\partial s}(0,t)}\xi\right] ds - [B - \mathbb{D}\xi(p)]\big(\log_p(\gamma(t))\big). \tag{4.41}$$

For any $s,t$ belonging to the interval $[0,1]$, let us write

$$
\begin{aligned}
&P^{s,0}_{\chi(\cdot,t)}\nabla_{\frac{\partial\chi}{\partial s}}\xi - P^{s,0}_{\chi(\cdot,0)}\nabla_{\frac{\partial\chi}{\partial s}(s,0)}\xi \\
&= P^{s,0}_{\chi(\cdot,t)}P^{0,t}_{\chi(s,\cdot)}\left(P^{t,0}_{\chi(s,\cdot)}\nabla_{\frac{\partial\chi}{\partial s}}\xi - \nabla_{\frac{\partial\chi}{\partial s}(s,0)}\xi\right) \\
&\qquad + \left[P^{s,0}_{\chi(\cdot,0)}\left(P^{0,s}_{\chi(\cdot,0)} - P^{t,0}_{\chi(s,\cdot)}P^{0,s}_{\chi(\cdot,t)}\right)P^{s,0}_{\chi(\cdot,t)}P^{0,t}_{\chi(s,\cdot)}\right]\nabla_{\frac{\partial\chi}{\partial s}(s,0)}\xi \\
&= P^{s,0}_{\chi(\cdot,t)}P^{0,t}_{\chi(s,\cdot)}\left(P^{t,0}_{\chi(s,\cdot)}\nabla_{\frac{\partial\chi}{\partial s}}\xi - \nabla_{\frac{\partial\chi}{\partial s}(s,0)}\xi\right) \\
&\qquad + \left[P^{s,0}_{\chi(\cdot,0)}\left(\tau_{p,\chi(s,0)} - P^{t,0}_{\chi(s,\cdot)}\tau_{p,\chi(s,t)}\right)P^{s,0}_{\chi(\cdot,t)}P^{0,t}_{\chi(s,\cdot)}\right]\nabla_{\frac{\partial\chi}{\partial s}(s,0)}\xi.
\end{aligned}
$$

Consequently,

$$
\begin{aligned}
\left\|P^{s,0}_{\chi(\cdot,t)}\nabla_{\frac{\partial\chi}{\partial s}}\xi - P^{s,0}_{\chi(\cdot,0)}\nabla_{\frac{\partial\chi}{\partial s}}\xi\right\| &\leqslant \left\|P^{t,0}_{\chi(s,\cdot)}\nabla_{\frac{\partial\chi}{\partial s}}\xi - \nabla_{\frac{\partial\chi}{\partial s}(s,0)}\xi\right\| \\
&\quad + \left\|\tau_{p,\chi(s,0)} - P^{t,0}_{\chi(s,\cdot)}\tau_{p,\chi(s,t)}\right\|\left\|\nabla_{\frac{\partial\chi}{\partial s}(s,0)}\xi\right\|. \qquad (4.42)
\end{aligned}
$$

By Proposition 4.7, the second term in the right-hand side of (4.43) is dominated by

$$
\begin{aligned}
&\sqrt{C}\Big(\int_0^1 \varphi\big(\sqrt{C}d(p,\chi(s,t))\big)\,dt\Big)\ell_0^t(\chi(s,\cdot))\left\|\nabla_{\frac{\partial\chi}{\partial s}}\xi\right\|_{t=0} \qquad (4.43)\\
&\leqslant \sqrt{C}\Big(\int_0^1 \varphi\big(\sqrt{C}d(p,\chi(s,t))\big)\,dt\Big)\left(\int_0^t \left\|\frac{\partial\chi}{\partial t}\right\|dt\right)\left\|\mathbb{D}\xi(\chi(s,0))\right\|\left\|\frac{\partial\chi}{\partial s}\right\|_{t=0}.
\end{aligned}
$$

Observe that $\chi(\cdot,0)$ is a geodesic, it follows that $\left\|\frac{\partial\chi}{\partial s}\right\|_{t=0} = \left\|\frac{\partial\chi}{\partial s}\right\|_{(0,0)}$. On the other hand, we have

$$
P^{t,0}_{\chi(s,\cdot)}\nabla_{\frac{\partial\chi}{\partial s}}\xi = \nabla_{\frac{\partial\chi}{\partial s}}\xi\Big|_{(s,0)} + \int_0^t P^{\theta,0}_{\chi(s,\cdot)}\nabla_{\frac{\partial\chi}{\partial t}(s,\theta)}\nabla_{\frac{\partial\chi}{\partial s}}\xi\,d\theta. \qquad (4.44)
$$

Taking into account $\left\|\frac{\partial\chi}{\partial s}\right\| = d\big(p,\gamma(t)\big)$, it is possible to write

$$
\left\|\nabla_{\frac{\partial\chi}{\partial t}}\nabla_{\frac{\partial\chi}{\partial s}}\xi\right\| \leqslant M'\left\|\frac{\partial\chi}{\partial t}\right\|. \qquad (4.45)
$$

Thus, we can derive that

$$
\begin{aligned}
&\|X\| = \|(\phi_B\gamma)(t)\| \qquad (4.46)\\
&\quad\leqslant M'\int_0^1\left(\int_0^t\left\|\frac{\partial\chi}{\partial t}\right\|dt\right)ds + \varphi(\bar\alpha\sqrt{C})M\sqrt{C}\left\|\frac{\partial\chi}{\partial s}\right\|_{(0,0)}\left(\int_0^t\left\|\frac{\partial\chi}{\partial t}\right\|dt\right).
\end{aligned}
$$

since $\chi(s,t)\in\mathbb{B}(p,\bar\alpha)$. Furthermore, after differentiating with respect to $t$ in (4.41), one obtains

$$
\begin{aligned}
\dot X_0 &= -[B-\mathbb{D}\xi(p)]\left(\frac{d}{dt}\big(\log_p(\gamma(t))\big)\right)_{t=0} \\
&\qquad + \int_0^1 \frac{d}{dt}\left\{P^{s,0}_{\chi(\cdot,t)}\nabla_{\frac{\partial\chi}{\partial s}(s,t)}\xi - \nabla_{\frac{\partial\chi}{\partial s}(0,t)}\xi\right\}_{t=0} ds. \qquad (4.47)
\end{aligned}
$$

Combining (4.42), (4.43) and (4.44) we find

$$\begin{aligned}&\left\|\frac{d}{dt}\left\{P^{s,0}_{\chi(\cdot,t)}\nabla_{\frac{\partial\chi}{\partial s}(s,t)}\xi\right\}_{t=0}\right\|\\&\leqslant\left\|\nabla_{\frac{\partial\chi}{\partial t}}\nabla_{\frac{\partial\chi}{\partial s}}\xi\right\|_{t=0}+M\varphi\big(\bar\alpha\sqrt{C}\big)\sqrt{C}\left\|\frac{\partial\chi}{\partial s}\right\|_{(0,0)}\left\|\frac{\partial\chi}{\partial t}\right\|_{t=0}.\end{aligned}\tag{4.48}$$

Thanks to Assumption 4.1 we have

$$\big\|\log_p(\gamma(t))-\log_p(\gamma(0))\big\|\leqslant\sigma_1^{-1}d\big(\gamma(t),\gamma(0)\big)=\sigma_1^{-1}\ell_0^t(\gamma)=\sigma_1^{-1}\int_0^t\|\gamma'(t)\|\,dt,$$

and we get $\left\|\frac{d}{dt}\log_p(\gamma(t))\right\|_{t=0}\leqslant\sigma_1^{-1}\|\gamma'(0)\|=\sigma_1^{-1}\ell$. Summing up we reach the following estimation

$$\begin{aligned}\|\dot X_0\|&\leqslant\sigma_1^{-1}\|B-\mathbb{D}\xi(p)\|\ell\\&+\left[M'+M\varphi\big(\bar\alpha\sqrt{C}\big)\sqrt{C}\left\|\frac{\partial\chi}{\partial s}\right\|_{(0,0)}\right]\int_0^1\left(\left\|\frac{\partial\chi}{\partial t}\right\|_{(s,0)}+\left\|\frac{\partial\chi}{\partial t}\right\|_{(0,0)}\right)ds.\end{aligned}\tag{4.49}$$

The fields $V=\frac{\partial\chi}{\partial s}$ and $T=\frac{\partial\chi}{\partial t}$ satisfy the estimation that (see the proof of Proposition 4.7)

$$\|V\|\leqslant d\big(p,\gamma(t)\big),\ \|T\|\leqslant\sigma_1^{-1}\rho\big[1-\cosh(s\rho)+\rho\sinh(s\rho)\big]\ell,\ \rho=\sqrt{C}d\big(p,\gamma(t)\big).\tag{4.50}$$

Invoking (4.46), (4.49), (4.50) and using the fact that $\left\|\frac{\partial\chi}{\partial s}\right\|_{(0,0)}=d\big(p,\gamma(0)\big)\leqslant\bar\alpha$, the proof of Lemma 4.12 is thereby complete. □

**Remark 4.13.** Assume that the iterative procedure (4.2) with updating formula (4.1) generates a sequence $(p_k)$ in some normal ball with center $\bar p$ and radius $\alpha\leqslant\bar\alpha$. Since $\lim_{t\to0}(t\cosh t-\sinh t)/t=0$ when $\bar\alpha$ is sufficiently small, we get from (4.36)

$$\big\|P^{1,0}_\gamma\zeta_{B_k}(\gamma(1))-\zeta_{B_k}(\gamma(0))\big\|\leqslant\ell_0^1(\gamma)\left[\sigma_1^{-1}\beta_k+(\sigma_1^{-1}+\sigma_1^{-2})\alpha\sqrt{C}\right],\tag{4.51}$$

for $\beta_k=\|\mathbb{D}\xi(p_k)-B_k\|$ and the geodesic $\gamma$ such that $d\big(\bar p,\gamma(t)\big)<\alpha$. □

**Lemma 4.14.** *Let $\Omega$ be a normal neighborhood of $\bar p$. Assume that some iteration $p_k\in\Omega$ was known and $u_k$, $p_{k+1}$ are obtained via* (4.2). *If $B_{k+1}$ are updated through the formula* (4.1) *while $p_{k+1}$ still lies into $\Omega$, then*

$$\|B_{k+1}-\mathbb{D}\xi(p_{k+1})\|\leqslant\|B_k-\mathbb{D}\xi(p_k)\|+\frac{3}{2}\lambda\big[d(p_{k+1},\bar p)+d(p_k,\bar p)\big].\tag{4.52}$$

**Proof.** We have

$$\begin{aligned}B_{k+1}-\mathbb{D}\xi(p_{k+1})&=\mathcal{T}_{u_k}\big(B_k-\mathbb{D}\xi(p_k)\big)\mathcal{T}_{u_k}^{-1}-\frac{\langle\mathcal{T}_{u_k}u_k,\cdot\rangle}{\|u_k\|^2}\mathcal{T}_{u_k}\big(B_k-\mathbb{D}\xi(p_k)\big)u_k\\&\quad-\big[\mathbb{D}\xi(p_{k+1})-\mathcal{T}_{u_k}\mathbb{D}\xi(p_k)\mathcal{T}_{u_k}^{-1}\big]\\&\quad+\|u_k\|^{-2}\langle\mathcal{T}_{u_k}u_k,\cdot\rangle\big[\xi(p_{k+1})-\mathcal{T}_{u_k}\xi(p_k)-\mathcal{T}_{u_k}\mathbb{D}\xi(p_k)u_k\big].\end{aligned}\tag{4.53}$$

Define $C_k = B_k - \mathbb{D}\xi(p_k)$. Because $\mathcal{T}_{u_k} : T_{p_k}\mathbb{M} \longrightarrow T_{p_{k+1}}\mathbb{M}$ is a linear isometry, it is possible to write

$$\big\|\mathcal{T}_{u_k} C_k \mathcal{T}_{u_k}^{-1} - \|u_k\|^{-2}\langle \mathcal{T}_{u_k} u_k, \cdot\rangle \mathcal{T}_{u_k} C_k u_k\big\| = \big\|C_k - \|u_k\|^{-2}\langle u_k, \cdot\rangle C_k u_k\big\|.$$

Thus we derive due to [8, Lemma 4.1]

$$\big\|\mathcal{T}_{u_k} C_k \mathcal{T}_{u_k}^{-1} - \|u_k\|^{-2}\langle \mathcal{T}_{u_k} u_k, \cdot\rangle \mathcal{T}_{u_k} C_k u_k\big\| \leqslant \|C_k\|. \tag{4.54}$$

Recall that $\mathbb{D}\xi$ is Lipschitz, the norm of operator $\mathbb{D}\xi(p_{k+1}) - \mathcal{T}_{u_k}\mathbb{D}\xi(p_k)\mathcal{T}_{u_k}^{-1}$ does not exceed in $\lambda d(p_k, p_{k+1})$. Further, by invoking Lemma 4.11, we obtain

$$\|\xi(p_{k+1}) - \mathcal{T}_{u_k}\xi(p_k) - \mathcal{T}_{u_k}\mathbb{D}\xi(p_k)u_k\| = \|\mathcal{T}_{u_k}^{-1}\xi(p_{k+1}) - \xi(p_k) - \mathbb{D}\xi(p_k)u_k\| \leqslant \frac{1}{2}\lambda\|u_k\|^2.$$

Therefore, the third term of (4.53) can be dominated in norm by the quantity below

$$\|u_k\|^{-2}\|\mathcal{T}_{u_k}u_k\|\left(\frac{1}{2}\lambda\|u_k\|^2\right) = \frac{1}{2}\lambda\|u_k\|.$$

Taking into account $\|u_k\| = d(p_k, p_{k+1})$, we obtain (4.52) by applying the triangle inequality $d(p_k, p_{k+1}) \leqslant d(p_{k+1}, \bar{p}) + d(p_k, \bar{p})$. □

**Lemma 4.15.** *Let $\kappa$, $L$, $a$, $b$, $a'$, $b'$, $a''$ and $b''$ be given positive constants with $\kappa L < 1$. Suppose that the sequences $(L_n)$, $(\alpha_n)$ and $(\beta_n)$ are generated by the recurrent relations* (4.55a)–(4.55c) *below*

$$L_n = L + a\alpha_n + b\beta_n, \tag{4.55a}$$
$$(1 - \kappa L_n)\alpha_{n+1} = a'\alpha_n^2 + b'\alpha_n\beta_n, \tag{4.55b}$$
$$\beta_{n+1} = \beta_n + a''\alpha_{n+1} + b''\alpha_n. \tag{4.55c}$$

*Further, suppose that $\alpha_0$, $\beta_0$, $L_0$ satisfy the following conditions*

$$a'\alpha_0 + b'\beta_0 + \left(a''b'\frac{\mu}{1-\mu} + \frac{b'b''}{1-\mu}\right)\alpha_0 \leqslant \mu\left(1 - \kappa L_0 - \kappa\left(\frac{a''b\mu}{1-\mu} + \frac{bb''}{1-\mu}\right)\alpha_0\right) \tag{4.56}$$

*for some $0 < \mu < 1$. Then, the following estimation holds for each $n = 0, 1, \ldots$*

$$\alpha_{n+1} \leqslant \mu\alpha_n, a'\alpha_n + b'\beta_n \leqslant \mu(1 - \kappa L_n). \tag{4.57}$$

**Proof.** Indeed, by virtue of (4.56), we have

$$\alpha_1 = \frac{a'\alpha_0 + b'\beta_0}{1 - \kappa L_0}\,\alpha_0 \leqslant \mu\alpha_0.$$

Thus, (4.57) is valid for $n = 0$. By inductive hypothesis, assume that $\alpha_1, \ldots, \alpha_n$ are well-defined through (4.55a)–(4.55c) and that (4.57) holds for each $j = 0, \ldots, n-1$. According to (4.55c), we get

$$\begin{aligned}\beta_j &= \beta_0 + a''\sum_{i=1}^{j}\alpha_i + b''\sum_{i=0}^{j-1}\alpha_i \leqslant \beta_0 + a''\sum_{i=1}^{j}\mu^i\alpha_0 + b''\sum_{i=0}^{j-1}\mu^i\alpha_0\\ &\leqslant \beta_0 + a''\frac{\mu}{1-\mu}\alpha_0 + b''\frac{1}{1-\mu}\alpha_0.\end{aligned}$$

Invoking (4.55a), we find 2

$$
\begin{aligned}
L_n &\leqslant L + a\mu^n\alpha_0 + b\left(\beta_0 + a''\frac{\mu}{1-\mu}\alpha_0 + b''\frac{1}{1-\mu}\alpha_0\right)\\
&\leqslant L + a\alpha_0 + b\beta_0 + \left(\frac{a''b\mu}{1-\mu} + \frac{b''b}{1-\mu}\right)\alpha_0\\
&= L_0 + \left(\frac{a''b\mu}{1-\mu} + \frac{b''b}{1-\mu}\right)\alpha_0,
\end{aligned}
$$

which yields $\kappa L_n < 1$. Moreover, taking into account (4.56), it follows that

$$
\mu(1-\kappa L_n) \geqslant a'\alpha_0 + b'\beta_0 + \left(a''b'\frac{\mu}{1-\mu} + \frac{b'b''}{1-\mu}\right)\alpha_0 \geqslant a'\alpha_n + b'\beta_n.
$$

As a result, we obtain $\alpha_{n+1} = \dfrac{a'\alpha_n + b'\beta_n}{1-\kappa L_n}\alpha_n \leqslant \mu\alpha_n$.

The lemma is therefore proved. □

**Proof of Theorem 4.10.** Since $\operatorname{reg}\Upsilon(\bar p, 0_{\bar p}) < +\infty$, there are some parameters $r_1 > 0$, $r_2 > 0$ and $\kappa > 0$ such that the following statement is fulfilled: if $d(q,\bar p) < r_1$ and $v \in T_q\mathbb{M}$ with $\|v\| < r_2$, then it holds that

$$
d\Big(q, \Big\{q' \;\Big|\; \Upsilon(q') \ni \mathcal{T}_{\exp_q^{-1}(q')}v\Big\}\Big) \leqslant \kappa d\big(\mathcal{T}_{\exp_q^{-1}(q')}v, \Upsilon(q')\big). \tag{4.58}
$$

By making $r_1$, $r_2$ smaller if necessary, we can assume that the ball $\mathbb{B}_{\mathbb{M}}(\bar p, r_1)$ with center $\bar p$ and radius $r_1$ is totally normal neighborhood, while Assumption 4.3 is valid with respect to $\Omega = \mathbb{B}_{\mathbb{M}}(\bar p, r_1)$. Further, since the map $q \longmapsto i_{\mathbb{M}}(q)$ is continuous (cf. [37]), we suppose

$$
\iota = \inf\big\{i_{\mathbb{M}}(q) : d(q,\bar p) < r_1\big\} > 0.
$$

Let us fix $\epsilon \in (0,1)$ and $\beta = \frac{(1-\epsilon)\sigma_1}{\kappa(1+\sigma_2)}$. Let $\alpha$ such that $0 < \alpha \leqslant \min\{r_1, r_2, \iota\}$ fulfilling the following inequality

$$
\kappa L + a'\alpha + (b' + \kappa b)\beta < 1, \tag{4.59}
$$

where $L = (\sigma_1^{-1} + \sigma_1^{-2})\alpha\sqrt{C}$, $b = \sigma_1^{-1}$, $a' = \frac{\sigma\kappa\lambda}{2}$, $b' = \sigma\kappa$ and $\sigma = \sigma_2/\sigma_1 \geqslant 1$. By letting $\alpha$ be smaller if necessary, we can assume that the constraint (4.56) of Lemma 4.15 holds for some $\mu \in (\epsilon, 1)$ and $a = 0$, $a'' = b'' = \frac{3\lambda}{2}$, $\alpha_0 = \alpha$, $\beta_0 = \beta$, $L_0 = L + b\beta$. Applying Lemma 4.15, the recurrent relations (4.55a)–(4.55c) generate three sequences $(L_k)$, $(\alpha_k)$, $(\beta_k)$ such that $\alpha_{k+1} \leqslant \mu\alpha_k$.

Let $p_0 \neq \bar p \in U = \mathbb{B}_{\mathbb{M}}(\bar p, \alpha)$ and let $B_0 : T_{p_0}\mathbb{M} \longrightarrow T_{p_0}\mathbb{M}$ be a linear map such that

$$
\beta_0' = \|\mathbb{D}\xi(p_0) - B_0\| \leqslant \beta_0.
$$

To build the tangent vector $u_0$, we consider the following inclusion

$$
0_{p_0} \in \xi(p_0) + B_0(u) + \mathcal{T}_u^{-1}\varPi\left(\exp_{p_0}(u)\right), u \in T_{p_0}\mathbb{M}. \tag{4.60}
$$

We apply Proposition 4.8 with $A = B_0$. Indeed, let $\gamma : [0,1] \longrightarrow \mathbb{M}$ be a geodesic such that $\gamma([0,1]) \subset U$. Invoking Remark 4.13, the local vector field

$$
\zeta_0 : q \in U \longmapsto \xi(q) - \tau_{p_0,q}B_0\left(\exp_{p_0}^{-1}(q)\right) - \tau_{p_0,q}\xi(p_0)
$$

possesses the following property

$$\left\|P_\gamma^{1,0}\zeta_0(\gamma(1)) - \zeta_0(\gamma(0))\right\| \leqslant L_0\ell_0^1(\gamma).$$

If we set $\bar{u}_0 = \exp_{p_0}^{-1}(\bar{p})$ and $\delta_0 = \left\|\xi(p_0) + B_0(\bar{u}_0) - \mathcal{T}_{\bar{u}_0}^{-1}\xi(\bar{p})\right\|$, then $\alpha_0' = d(p_0, \bar{p}) = \|\bar{u}_0\| \leqslant \alpha_0$. Therefore, the following estimation is valid due to Lemma 4.11

$$\delta_0 \leqslant \frac{1}{2}\lambda \|\bar{u}_0\|^2 + \beta_0' \|\bar{u}_0\| = \frac{1}{2}\lambda (\alpha_0')^2 + \beta_0'\alpha_0'.$$

We now claim that
$$\frac{\kappa}{1-\kappa L_0}\,\delta_0 < \min\left\{r_1, r_2\right\}. \tag{4.61}$$

Indeed, notice that $\sigma\kappa\delta_0 \leqslant a'\alpha_0^2 + b'\alpha_0\beta_0 = (1-\kappa L_0)\alpha_1$, so

$$\frac{\kappa}{1-\kappa L_0}\delta_0 \leqslant \sigma^{-1}\alpha_1 \leqslant \sigma^{-1}\mu\alpha_0 < \min\left\{r_1, r_2\right\}.$$

In view of Proposition 4.8, the inclusion (4.60) admits a solution, says $u_0$, such that

$$\sigma_1 \|u_0 - \bar{u}_0\| \leqslant \frac{\kappa}{1-\kappa L_0}\delta_0 \leqslant \frac{\kappa}{1-\kappa L_0}\left(\frac{1}{2}\lambda\big(\alpha_0'\big)^2 + \beta_0\alpha_0'\right). \tag{4.62}$$

By setting $p_1 = \exp_{p_0}(u_0)$, Assumption 4.1 and (4.62) give us

$$\begin{aligned} d(p_1, \bar{p}) &\leqslant \sigma_2 \|u_0 - \bar{u}_0\| \leqslant \sigma\frac{\kappa}{1-\kappa L_0}\left(\frac{1}{2}\lambda\big(\alpha_0'\big)^2 + \beta_0'\alpha_0'\right) \\ &\leqslant \frac{a'\alpha_0 + b'\beta_0}{1-\kappa L_0}\alpha_0' \leqslant \mu\alpha_0'. \end{aligned} \tag{4.63}$$

and the new point $p_1$ is still in $U$, because of $\mu < 1$.

We proceed to the induction step. Suppose the iterations $p_0, p_1, \ldots, p_k \in U$ are given such that, for each $j = 0, \ldots, k-1$:

- $u_j = \exp_{p_j}^{-1}(p_{j+1})$ solves the inclusion

  $$0_{p_j} \in \xi(p_j) + B_j(u_j) + \mathcal{T}_{u_j}^{-1}\Pi\big(\exp_{p_j}(u_j)\big),$$

  in which the operators $B_{j+1}$ is updated via the formula

  $$B_{j+1}(\cdot) = \mathcal{T}_{u_j}B_j\mathcal{T}_{u_j}^{-1}(\cdot) + \frac{\langle \mathcal{T}_{u_j}u_j, \cdot\rangle}{\|u_j\|^2}\big[\xi(p_{j+1}) - \mathcal{T}_{u_j}B_j(u_j) - \mathcal{T}_{u_j}\xi(p_k)\big];$$

- $d(p_{j+1}, \bar{p}) \leqslant \mu d\big(p_j, \bar{p}\big)$;
- $\alpha_j' = d\big(p_j, \bar{p}\big) \leqslant \alpha_j$.

To generate $p_{k+1}$, let's consider the following inclusion

$$0_{p_k} \in \xi(p_k) + B_k(u_{p_k}) + \mathcal{T}_{u_{p_k}}^{-1}\Pi\big(\exp_{p_k}(u_{p_k})\big), \quad u_{p_k} \in T_{p_k}\mathbb{M}. \tag{4.64}$$

Define $\zeta_k : q \in U \longmapsto \xi(q) + \tau_{p_k,q}\big(B_k \log_{p_k}(q) + \xi(p_k)\big)$ and let $C_j = B_j - \mathbb{D}\xi(p_j)$.

Then we deduce after invoking Lemma 4.14

$$\|C_{j+1}\| \leqslant \|C_j\| + \frac{3\lambda}{2}(\alpha_{j+1} + \alpha_j), \quad j \leqslant k-1.$$

Recalling $\|C_0\| \leqslant \beta_0$ we conclude $\beta'_k = \|C_k\| \leqslant \beta_k$. Therefore, by an analogous argument as in the case $k = 0$, one has

$$\left\|P_\gamma^{1,0}\zeta_k(\gamma(1)) - \zeta_k(\gamma(0))\right\| \leqslant L_k \ell_0^1(\gamma) \tag{4.65}$$

by virtue of (4.51). Here, $\gamma$ stands for a geodesic on $U$. With the aim of using Proposition 4.8, we set $\bar{u}_k = \exp_{p_k}^{-1}(\bar{p})$ and $\delta_k = \left\|\xi(p_k) + B_k(\bar{u}_k) - \mathcal{T}_{\bar{u}_k}^{-1}\xi(\bar{p})\right\|$. By applying Lemma 4.33

$$\delta_k = \left\|\mathcal{T}_{\bar{u}_k}^{-1}\xi(\bar{p}) - B_k(\bar{u}_k) - \xi(p_k)\right\| \leqslant \frac{\lambda}{2}\|\bar{u}_k\|^2 + \|C_k\|\|\bar{u}_k\| = \frac{\lambda}{2}\left(\alpha'_k\right)^2 + \beta'_k\alpha'_k,$$

which permits us to write

$$\kappa\sigma\delta_k \leqslant \left(a'\alpha_k + b'\beta_k\right)\alpha_k = (1 - \kappa L_k)\alpha_{k+1}.$$

Consequently, we deduce

$$\frac{\kappa}{1-\kappa L_k}\delta_k \leqslant \sigma^{-1}\alpha_{k+1} < \min\left\{r_1, r_2\right\}.$$

Hence, by invoking again Proposition 4.8, the inclusion (4.64) has a solution $u_k$ such that

$$\sigma_1\|u_k - \bar{u}_k\| \leqslant \frac{\kappa}{1-\kappa L_k}\delta_k \leqslant \sigma^{-1}\frac{1}{1-\kappa L_k}\left(a'\alpha'_k + b'\beta'_k\right)\alpha'_k.$$

Define the next iteration $p_{k+1} = \exp_{p_k}(u_k)$, we obtain

$$\begin{aligned} d\left(p_{k+1}, \bar{p}\right) &\leqslant \sigma_2\|u_k - \bar{u}_k\| \leqslant \sigma_2\sigma_1^{-1}\sigma^{-1}\frac{1}{1-\kappa L_k}\left(a'\alpha'_k + b'\beta'_k\right)\alpha'_k \\ &\leqslant \frac{a'\alpha_k + b'\beta_k}{1-\kappa L_k}\alpha'_k \leqslant \mu\alpha'_k = \mu d\left(p_k, \bar{p}\right). \end{aligned}$$

Furthermore, since $\alpha'_k \leqslant \alpha_k$, it follows that

$$d\left(p_{k+1}, \bar{p}\right) \leqslant \frac{a'\alpha_k + b'\beta_k}{1-\kappa L_k}\alpha_k = \alpha_{k+1}.$$

The proof is then completed by induction. □

By making use of the Dennis-Moré type condition in the previous section, we establish the following superlinear convergence of the Broyden update. To prove this result, the notion of Hilbert-Schmidt norm will be useful. Let $\left(e_i : i \in \mathcal{I}\right)$ form an orthonormal basis of some Hilbert space $\mathcal{H}$ endowed with the inner product $\langle\cdot,\cdot\rangle$ (see, e.g. [8]). The associated Hilbert-Schmidt norm $\|\cdot\|_{HS}$ to the linear and continuous map $L : \mathcal{H} \longrightarrow \mathcal{H}$ is defined as follows

$$\|L\|_{HS}^2 := \sum_{i\in\mathcal{I}}\langle Le_i, Le_i\rangle. \tag{4.66}$$

In finite dimensional space, such a notion reduces the usual Frobenius norm.

**Theorem 4.16.** (Locally superlinear convergence) *Let the assumptions in the statement of Theorem* 4.10 *hold. Suppose further* $\Upsilon = \xi + \Pi$ *is* $\mathcal{T}$*-strongly subregular in the sense of Definition* 2.1*. Then, the sequence* $(p_k)$ *generated in the proof of Theorem* 4.10 *converges superlinearly to the solution* $\bar{p}$*.*

**Proof.** It is sufficient to establish the relation

$$\limsup_{k\to\infty} \frac{\|C_k u_k\|}{\|u_k\|} = 0, \quad C_k = B_k - \mathbb{D}\xi(p_k). \tag{4.67}$$

Indeed, if (4.67) is verified, then we just apply Theorem 3.7 with $\mathcal{H}_k = \mathbb{D}\xi(p_k)$. To do this, it is necessary to estimate the quantity $\|C_k u_k\|$. We have

$$C_k u_k = \big(B_k - \mathcal{T}_{u_k}^{-1} B_{k+1} \mathcal{T}_{u_k}\big) u_k + \big[\mathcal{T}_{u_k}^{-1} B_{k+1} \mathcal{T}_{u_k} u_k - \mathbb{D}\xi(p_k) u_k\big]. \tag{4.68}$$

Since $\mathcal{T}_{u_k}$ is an isometry, we deduce $\langle \mathcal{T}_{u_k} u_k, \mathcal{T}_{u_k} u_k\rangle = \langle u_k, u_k\rangle$, which yields from the updating formulae (4.1) that

$$\mathcal{T}_{u_k}^{-1} B_{k+1} \mathcal{T}_{u_k} u_k = \mathcal{T}_{u_k}^{-1} \xi(p_{k+1}) - \xi(p_k).$$

Consequently,

$$\begin{aligned} \big\|\mathcal{T}_{u_k}^{-1} B_{k+1} \mathcal{T}_{u_k} u_k - \mathbb{D}\xi(p_k) u_k\big\| &= \big\|\mathcal{T}_{u_k}^{-1} \xi(p_{k+1}) - \xi(p_k) - \mathbb{D}\xi(p_k) u_k\big\| \\ &\leqslant \frac{\lambda}{2}\|u_k\|^2. \end{aligned} \tag{4.69}$$

Moreover, it is possible to use the similar arguments as in the proof of [8, Theorem 4.8] to derive the conclusion that

$$\|\mathcal{T}_{u_k}^{-1} B_{k+1} \mathcal{T}_{u_k} - B_k\| \leqslant \|\mathcal{T}_{u_k}^{-1} B_{k+1} \mathcal{T}_{u_k} - B_k\|_{HS} \xrightarrow{k\to\infty} 0, \tag{4.70}$$

where $\|\cdot\|_{HS}$ denotes the Hilbert-Schmidt norm. In fact, let $B$ be any linear transformation on $T_{p_k}\mathbb{M}$ with $Bu_k = \mathcal{T}_{u_k}^{-1}\xi(p_{k+1}) - \xi(p_k) = y_k$. Then, writing by $\{e_i^{(k)}\}_{i=1}^m$ an orthonormal basis of $T_{p_k}\mathbb{M}$, we find

$$\begin{aligned} &\|\mathcal{T}_{u_k}^{-1} B_{k+1} \mathcal{T}_{u_k} - B_k\|_{HS}^2 = \big\|\|u_k\|^{-2}\langle \mathcal{T}_{u_k} u_k, \mathcal{T}_{u_k}(\cdot)\rangle \big(y_k - B_k u_k\big)\big\|_{HS}^2 \\ &= \frac{\sum_{i=1}^m \big\|\big\langle \mathcal{T}_{u_k} u_k, \mathcal{T}_{u_k}\big(e_i^{(k)}\big)\big\rangle (B - B_k) u_k\big\|^2}{\|u_k\|^4} = \frac{\sum_{i=1}^m \big\|\big\langle u_k, e_i^{(k)}\big\rangle (B - B_k) u_k\big\|^2}{\|u_k\|^4} \\ &= \frac{\sum_{i=1}^m \big\langle u_k, e_i^{(k)}\big\rangle^2 \|(B - B_k) u_k\|^2}{\|u_k\|^4} = \frac{\|(B - B_k) u_k\|^2}{\|u_k\|^2} \leqslant \|B - B_k\|_{HS}. \end{aligned}$$

In other words,

$$\|\mathcal{T}_{u_k}^{-1} B_{k+1} \mathcal{T}_{u_k} - B_k\|_{HS} = \min\Big\{\|B - B_k\|_{HS} \mid B : T_{p_k}\mathbb{M} \longrightarrow T_{p_k}\mathbb{M} \text{ linear, } Bu_k = y_k\Big\}.$$

Define for each $k$ the linear operator

$$A_k : v \in T_{p_k}\mathbb{M} \longmapsto \int_0^1 \Big(P_{\gamma_k}^{t,0} \nabla_{P_{\gamma_k}^{0,t} v}\, \xi\Big)\, dt, \tag{4.71}$$

where $\gamma_k(t) = \exp_{p_k}(t u_k)$ is the geodesic passing through $p_k$ with velocity $u_k$.

Recalling the Taylor's expansion for $\xi$ (see, e.g. [18, Lemma 14.4])

$$P^{1,0}_{\gamma_k}\xi(p_{k+1}) = \xi(p_k) + \int_0^1 \left(P^{t,0}_{\gamma_k}\nabla_{P^{0,t}_{\gamma_k}u_k}\xi\right)\, dt = \xi(p_k) + A_k u_k, \tag{4.72}$$

which yields
$$\mathcal{T}^{-1}_{u_k} B_{k+1}\mathcal{T}_{u_k} u_k = A_k u_k. \tag{4.73}$$

Following the proof of [8, Theorem 4.8], we reach the inequality

$$\|\mathcal{T}^{-1}_{u_k} B_{k+1}\mathcal{T}_{u_k} - A_k\|^2_{HS} + \|\mathcal{T}^{-1}_{u_k} B_{k+1}\mathcal{T}_{u_k} - B_k\|^2_{HS} \leqslant \|B_k - A_k\|^2_{HS}. \tag{4.74}$$

Let's establish now some auxiliary facts.

**Claim 1**: $\limsup\limits_{k\to\infty}\|A_k - \mathbb{D}\xi(p_k)\| = \limsup\limits_{k\to\infty}\|A_k - \mathbb{D}\xi(p_k)\|_{HS} = 0.$
Taking into account the relation $\nabla_{P^{0,t}_{\gamma_k}u}\xi = \mathbb{D}\xi\big(\gamma_k(t)\big)P^{0,t}_{\gamma_k}u$ we derive by using the assumption $\mathbb{D}\xi \in \mathrm{Lip}_\lambda(\Omega)$ that

$$\begin{aligned}\big\|\big[A_k - \mathbb{D}\xi(p_k)\big]u\big\| &= \left\|\int_0^1 \left[P^{t,0}_{\gamma_k}\mathbb{D}\xi(\gamma_k(t))P^{0,t}_{\gamma_k}u - \mathbb{D}\xi(p_k)u\right]\, dt\right\| \\ &\leqslant \int_0^1 \lambda\ell_0^t(\gamma_k)\|u\|\, dt = \|u\|\int_0^1\left(\lambda\int_0^t\|\gamma_k'(s)\|\, ds\right)dt = \frac{\lambda}{2}\|u_k\|\|u\|.\end{aligned}$$

The latter gives us $\|A_k - \mathbb{D}\xi(p_k)\| \leqslant \frac{\lambda}{2}\|u_k\| = \frac{\lambda}{2}d(p_k, p_{k+1})$, which allows us to say that $\limsup_{k\to\infty}\|A_k - \mathbb{D}\xi(p_k)\| = 0$. Since the Hilbert-Schmidt norm $\|\cdot\|_{HS}$ is equivalent to the one induced by Riemannian metric, the claim is evident.

**Claim 2**: $\lim\limits_{k\to\infty}\|B_k - A_k\|_{HS} = \lim\limits_{k\to\infty}\|C_k\|_{HS}.$
This claim is a consequence of Lemma 4.14. Indeed, the series $\sum_{k\geqslant 0} d(p_k, \bar{p})$ is convergent due to Theorem 4.10. Hence, Lemma 4.14 implies that the sequence $(\beta_k')$, for $\beta_k' = \|B_k - \mathbb{D}\xi(p_k)\|$, is Cauchy. The same conclusion is also valid for $\big(\|B_k - \mathbb{D}\xi(p_k)\|_{HS}\big)$. As a result, the second limit exists. The existence of the other one is induced from the estimation

$$\|C_k\|_{HS} - \|\mathbb{D}\xi(p_k) - A_k\|_{HS} \leqslant \|B_k - A_k\|_{HS} \leqslant \|C_k\|_{HS} + \|\mathbb{D}\xi(p_k) - A_k\|_{HS}$$

and the preceding claim. Hence, the present task has been done.

Returning back to the main proof we use the identity

$$\mathcal{T}^{-1}_{u_k}B_{k+1}\mathcal{T}_{u_k} - A_k = \big[\mathcal{T}^{-1}_{u_k}C_{k+1}\mathcal{T}_{u_k}\big] + \big[\mathcal{T}^{-1}_{u_k}\mathbb{D}\xi(p_{k+1})\mathcal{T}_{u_k} - \mathbb{D}\xi(p_k)\big] + \big[\mathbb{D}\xi(p_k) - A_k\big].$$

By the assumption that $\mathbb{D}\xi \in \mathrm{Lip}_\lambda(\Omega)$, we have

$$\|\mathcal{T}^{-1}_{u_k}\mathbb{D}\xi(p_{k+1})\mathcal{T}_{u_k} - \mathbb{D}\xi(p_k)\| \leqslant \lambda\ell_0^1(\gamma_k) = \lambda d(p_k, p_{k+1}).$$

Thus, $\|\mathcal{T}^{-1}_{u_k}\mathbb{D}\xi(p_{k+1})\mathcal{T}_{u_k} - \mathbb{D}\xi(p_k)\| \xrightarrow{k\to\infty} 0$, and then, we conclude

$$\|\mathcal{T}^{-1}_{u_k}\mathbb{D}\xi(p_{k+1})\mathcal{T}_{u_k} - \mathbb{D}\xi(p_k)\|_{HS} \xrightarrow{k\to\infty} 0.$$

It is sufficient to verify $\left\|\mathcal{T}_{u_k}^{-1} C_{k+1} \mathcal{T}_{u_k}\right\|_{HS} = \left\|B_{k+1} - \mathbb{D}\xi(p_{k+1})\right\|_{HS}$. But this is clear from the definition of $\|\cdot\|_{HS}$, since

$$\begin{aligned}\left\|\mathcal{T}_{u_k}^{-1} C_{k+1} \mathcal{T}_{u_k}\right\|_{HS}^2 &= \sum_{i=0}^{m} \left\langle \mathcal{T}_{u_k}^{-1} C_{k+1} \mathcal{T}_{u_k} e_i^{(k)}, \mathcal{T}_{u_k}^{-1} C_{k+1} \mathcal{T}_{u_k} e_i^{(k)} \right\rangle \\ &= \sum_{i=0}^{m} \left\langle C_{k+1} \mathcal{T}_{u_k} e_i^{(k)}, C_{k+1} \mathcal{T}_{u_k} e_i^{(k)} \right\rangle = \|C_{k+1}\|_{HS}.\end{aligned}$$

Here, we have used the fact that $\left(\mathcal{T}_{u_k} e_i^{(k)}\right)$ forms an orthonormal basis of $T_{p_k}\mathbb{M}$.

Based on the two claims above, we obtain (4.70) by passing to the limit in (4.74). As a result, the equality (4.67) holds true by combining (4.68), (4.69) and (4.70). This completes the proof of Theorem 4.16. $\square$

## References


[1] P.-A. Absil, R. Mahony, R. Sepulchre: *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, Princeton (2008).

[2] P.-A. Absil, R. Mahony, R. Sepulchre: *Optimization on manifolds: methods and applications*, in: *Recent Advances in Optimization and its Applications in Engineering*, M. Diehl et al. (eds.), Springer, Berlin (2010) 125–144.

[3] R. L. Adler, J.-P. Dedieu, J. Y. Margulies, M. Martens, M. Shub: *Newton's method on Riemannian manifolds and a geometric model for the human spine*, IMA J. Numer. Analysis 22/3 (2002) 359–390.

[4] S. Adly, R. Cibulka, H. V. Ngai: *Newton methods for solving inclusions using set-valued approximations*, SIAM J. Optimization 25/1 (2015) 159–184.

[5] S. Adly, H. V. Ngai: *Quasi-Newton methods for solving nonsmooth equations: generalized Dennis-Moré theorem and Broyden's update*, J. Convex Analysis 25/4 (2018) 1075–1104.

[6] S. Adly, H. V. Ngai, N. V. Vu: *Newton-type method for solving generalized equations on Riemannian manifolds*, J. Convex Analysis 25/2 (2018) 341–370.

[7] Q. Ansari, F. Babu, X.-B. Li: *Variational inclusion problems in Hadamard manifolds*, J. Nonlinear Convex Analysis 19/1 (2018) 219–237.

[8] F. J. Aragón Artacho, A. Belyakov, A. L. Dontchev, M. López: *Local convergence of quasi-Newton methods under metric regularity*, Comp. Optimization Appl. 58/1 (2013) 225–247.

[9] J. X. da Cruz Neto, O. P. Ferreira, L. R. Lucambio Pérez: *Monotone point-to-set vector fields*, Balkan J. Geometry Appl. 5/1 (2000) 69–79.

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

[11] J. E. Dennis jun., J. J. Moré: *A characterization of superlinear convergence and its application to quasi-Newton methods*, Math. Computation 28/126 (1974) 549–560.

[12] J. E. Dennis jun., J. J. Moré: *Quasi-Newton methods, motivation and theory*, SIAM Review 19/1 (1977) 46–89.

[13] M. Dilshad, M. Akram: *Multi-valued variational inclusion problems in Hadamard manifolds*, Bull. Math. Analysis Appl. 12/1 (2020) 29–40.

[14] M. P. do Carmo: *Riemannian Geometry*, Mathematics: Theory & Applications, Birkhäuser, Basel (1992).

[15] A. L. Dontchev: *Generalizations of the Dennis-Moré theorem*, SIAM J. Optimization 22/3 (2012) 821–830.

[16] A. L. Dontchev, R. T. Rockafellar: *Implicit Functions and Solution Mappings: A View from Variational Analysis*, 2nd ed., Springer Series in Operations Research and Financial Engineering, Springer, New York (2014).

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

[18] K. A. Gallivan, C. Qi, P.-A. Absil: *A Riemannian Dennis-Moré condition*, in: *High-Performance Scientific Computing: Algorithms and Applications*, W. M. Berry et al. (eds.), Springer, London (2012) 281–293.

[19] M. Gaydu, M. H. Geoffroy: *A Newton iteration for differentiable set-valued maps*, J. Appl. Math. Analysis Appl. 399/1 (2013) 213–224.

[20] A. Griewank: *Broyden updating, the good and the bad*, Documenta Mathematica (2012) 301–315.

[21] S. Hosseini, M. R. Pouryayevali: *Nonsmooth optimization techniques on Riemannian manifolds*, J. Optimization Theory Appl. 158/2 (2013) 328–342.

[22] W. Huang, K. A. Gallivan, P.-A. Absil: *A Broyden class of quasi-Newton methods for Riemannian optimization*, SIAM J. Optimization 25/3 (2015) 1660–1685.

[23] A. F. Izmailov, M. V. Solodov: *Newton-Type Methods for Optimization and Variational Problems*, Springer Series in Operation Research and Finacial Engineering, Springer, Cham (2014).

[24] Yu. S. Ledyaev, Q. J. Zhu: *Nonsmooth analysis on smooth manifolds*, Trans. Amer. Math. Soc. 359 (2007) 3687–3732.

[25] C. Li, B. S. Mordukhovich, J. Wang, J.-C. Yao: *Weak sharp minima on Riemannian manifolds*, SIAM J. Optimization 21/4 (2011) 1523–1560.

[26] C. Li, J. Yao: *Variational inequalities for set-valued vector fields on Riemannian manifolds: convexity of the solution set and the proximal point algorithm*, SIAM J. Control Optimization 50/4 (2012) 2486–2514.

[27] S.-L. Li, C. Li, Y.-C. Liou, J.-C. Yao: *Existence of solutions for variational inequalities on Riemannian manifolds*, Nonlinear Anal. Theory Meth. Appl. 71/11 (2009) 5695–5706.

[28] J. M. Martínez: *Practical quasi-Newton methods for solving nonlinear systems*, J. Comput. Appl. Math. 124 (2000) 97–121.

[29] S. Z. Németh: *Variational inequalities on Hadamard manifolds*, Nonlinear Anal. Theory Meth. Appl. 52/5 (2003) 1491–1498.

[30] C. Qi: *Numerical Optimization Methods on Riemannian Manifolds*, PhD Thesis, Florida State University, Tallahassee (2011).

[31] C. Qi, K. A. Gallivan, P.-A. Absil: *Riemannian BFGS algorithm with applications*, in: *Recent Advances in Optimization and its Applications in Engineering*, M. Diehl et al. (eds.), Springer, Berlin (2010) 183–192.

[32] W. Ring, B. Wirth: *Optimization methods on Riemannian manifolds and their application to shape space*, SIAM J. Optimization 22/2 (2012) 596–627.

[33] S. M. Robinson: *Generalized equations and their solutions. I: Basic theory*, Math. Programming Studies 10 (1979) 128-141.

[34] S. M. Robinson: *Strongly regular generalized equations*, Math. Oper. Research 5/1 (1980) 43–62.

[35] S. M. Robinson: *Generalized equations*, in: *Mathematical Programming: The State of the Art*, A. Bachem et al. (eds.), Springer, Berlin (1982) 346–367.

[36] E. W. Sachs: *Broyden's method in Hilbert space*, Math. Programming 35/1 (1986) 71–82.

[37] T. Sakai: *Riemannian Geometry*, Translations of Mathematical Monographs 149, American Mathematical Society, Providence (1996).

[38] M. Seibert, M. Kleinstauber, K. Hüper: *Properties of the BFGS method on Riemannian manifolds*, in: *Mathematical System Theory*, K. Hüper and J. Trumpf (eds.), CreateSpace, Amazon, North Charleston (2013) 395–412.

[39] S. T. Smith: *Optimization techniques on Riemannian manifolds*, in: *Hamiltonian and Gradient Flows, Algorithms and Control*, Fields Institute Communications 3, American Mathematical Society, Providence (1994) 113–136.

[40] C. Udrişte: *Convex Functions and Optimization Methods on Riemannian Manifolds*, Mathematics and Its Applications 297, Springer, Amsterdam (1994).

[41] J.-H. Wang, C. Li: *Uniqueness of the singular points of vector fields on Riemannian manifolds under the $\gamma$-condition*, J. Complexity 22/4 (2006) 533–548.