



Inexact Newton Methods for Solving Generalized Equations on Riemannian Manifolds

Mauricio S. Louzeiro^{1,2} · Gilson N. Silva³ · Jinyun Yuan^{1,5} · Daoping Zhang⁴

Received: 11 October 2023 / Revised: 8 February 2025 / Accepted: 21 March 2025

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2025

Abstract

The convergence of inexact Newton methods is studied for solving generalized equations on Riemannian manifolds by using the metric regularity property, which is also explored. Under appropriate conditions and without any additional geometric assumptions, local convergence results with linear and quadratic rates, as well as a semi-local convergence result, are obtained for the proposed method. Finally, the theory is applied to the problem of finding a singularity for the sum of two vector fields. In particular, the KKT system for the constrained Riemannian center of mass on the sphere is explored numerically.

Keywords Generalized equation · Inexact Newton method · Metric regularity · Riemannian manifolds · Vector fields

Mathematics Subject Classification 90C33 · 49M37 · 65K05

✉ Mauricio S. Louzeiro
mauriciolouzeiro@ufg.br

Gilson N. Silva
gilson.silva@ufpi.edu.br

Jinyun Yuan
yuanjy@gmail.com

Daoping Zhang
daopingzhang@nankai.edu.cn

¹ School of Computer Science and Technology, Dongguan University of Technology, Dongguan, Guangdong, China

² IME/UFG, Avenida Esperança, s/n, Campus Samambaia, Goiânia, GO 74690-900, Brazil

³ Departamento de Matemática, Universidade Federal do Piauí, Teresina, PI 64049-550, Brazil

⁴ School of Mathematical Sciences and LPMC, Nankai University, Tianjin 300071, China

⁵ Department of Mathematics and Statistics, Henan Normal University, Xinxiang 453007, China

1 Introduction

In recent years, constrained generalized equations on Banach spaces, i.e., finding a solution to the inclusion

$$x \in C, \quad f(x) + F(x) \ni 0 \quad (1)$$

where \mathbb{X} and \mathbb{Y} are two Banach spaces, $C \subset \mathbb{X}$ is a nonempty, closed, and convex set, $f: \mathbb{X} \rightarrow \mathbb{Y}$ is a mapping, and $F: \mathbb{X} \rightrightarrows \mathbb{Y}$ is a set-valued mapping, have gained increased attention [7, 18]. This is attributed to the fact that the model (1) covers many well-known problems, such as constrained variational inequality and split variational inequality problem [15, 35], nonlinear equations, systems of equations and inequalities, optimality condition in mathematical programming and optimal control, and equilibrium problem. The readers are referred to [3–5, 8, 9, 17, 22–25, 27, 29, 31, 32, 34, 37, 39] for a detailed study of (1) with $C = \mathbb{X}$.

For a Riemannian manifold \mathcal{M} , a closed set $\Omega \subset \mathcal{M}$ with a nonempty relative interior, a continuously differentiable mapping $f: \mathcal{M} \rightarrow \mathbb{R}^m$, and a set-valued mapping $F: \mathcal{M} \rightrightarrows \mathbb{R}^m$, we consider the generalized equation

$$p \in \Omega, \quad f(p) + F(p) \ni 0. \quad (2)$$

Evidently, (2) covers the nonlinear equation $f(p) = 0$ ($F \equiv 0$) and the nonlinear inclusion problem $f(p) \in K$ ($F \equiv -K$) for a fixed cone $K \subset \mathbb{R}^m$ [53]. In this paper, we shall prove that problem (2) also covers the problem of finding a singularity of the sum of two vector fields of the form

$$p \in \Omega, \quad V(p) + Z(p) \ni 0_p, \quad (3)$$

where $V: \mathcal{M} \rightarrow \mathcal{TM}$ is a single-valued vector field, $Z: \mathcal{M} \rightrightarrows \mathcal{TM}$ is a set-valued vector field. In particular, we demonstrate that (2) can also be used to obtain a solution for a variational inequality problem and the Karush-Kuhn-Tucker (KKT) conditions for a constrained optimization problem on manifolds.

Problem (3) has been extensively investigated [5, 29] and solved using Newton's method for $Z \equiv 0$ [2, 6, 19, 26, 43, 52]. This problem naturally arises, for example, in the first-order optimality conditions of the minimization problem

$$\text{Minimize } \varsigma(p) + \vartheta(p), \quad p \in \text{int } \Omega, \quad (4)$$

where $\varsigma: \mathcal{M} \rightarrow \overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$ is a differentiable function defined over $\text{int } \Omega$ (the interior of Ω) and $\vartheta: \mathcal{M} \rightarrow \overline{\mathbb{R}}$ is non-differentiable. In fact, one can consider (3) with V representing the Riemannian gradient of ς ($\text{grad } \varsigma$), Z representing the Riemannian subdifferential of ϑ ($\partial\vartheta$), and $\Omega = \text{int } \Omega$, i.e.,

$$p \in \text{int } \Omega, \quad \text{grad } \varsigma(p) + \partial\vartheta(p) \ni 0_p.$$

Problem (4) is associated with several important applications, including sparse principal component analysis [33], sparse blind deconvolution [55], unsupervised feature selection [50], and image restoration [11, 12].

The extensive scope of generalized equations and the growing interest in optimization on manifolds [1, 14, 48] in recent years have motivated us to explore the integration of these two areas in this paper. Our focus is on studying a Riemannian version of the inexact Newton method proposed by [16] with the aim of solving problem (2). Given an initial point $p_0 \in \mathcal{M}$,

our method generates a sequence of iterations $\{p_k\}$ as follows:

$$(f(p_k) + \mathcal{D}f(p_k)[v_k] + F(\exp_{p_k} v_k)) \cap R_k(p_k) \neq \emptyset, \quad p_{k+1} := \exp_{p_k} v_k, \quad (5)$$

where $R_k: \mathcal{M} \rightrightarrows \mathbb{R}^m$ is a sequence of set-valued mappings with closed graphs representing the inexactness, and $\mathcal{D}f$ denotes the differential of f . In other words, the method involves: selecting an initial point p_0 on the manifold (sufficiently close to a solution), solving a subproblem to find v_k in the intersection above (under certain conditions, it is possible to guarantee that it is nonempty), and computing the next iteration p_{k+1} by applying the exponential map to (p_k, v_k) . It is noteworthy that whenever v_k is sufficiently small, (5) can be expressed as:

$$(f(p_k) + \mathcal{D}f(p_k)[\exp_{p_k}^{-1} p_{k+1}] + F(p_{k+1})) \cap R_k(p_k) \neq \emptyset. \quad (6)$$

Although it may appear that obtaining v_k requires solving the subproblem in (5) exactly, the map R_k is specifically introduced to circumvent this necessity. This is evident in the particular case of (5) where $R_k(p_k)$ is defined as the closed ball centered at 0 with a radius of $\eta_k \|f(p_k)\|_e$, with $\eta_k \in (0, 1)$, and $F \equiv 0$. In this scenario, the subproblem in (5) consists of finding a v_k that satisfies

$$\|f(p_k) + \mathcal{D}f(p_k)[v_k]\|_e \leq \eta_k \|f(p_k)\|_e, \quad k \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$$

where $\|\cdot\|_e$ denotes the Euclidean norm, which can be interpreted as an inexact Newton method for by solving $f(p) = 0$. The convergence analysis presented in this paper is conducted using (5) because it is more general and can be applied to other potential particular cases of (2).

Assuming that the set-valued mapping $f + F$ in (2) is metrically regular at $\bar{p} \in \Omega$ for 0 and that R_k satisfies a suitable boundedness condition, we demonstrate that a sequence $\{p_k\}$ generated by (6) exhibits both linear and quadratic convergence towards \bar{p} , with the exact nature of the convergence depending on the specific assumptions regarding R_k . These results can be obtained without requiring prior knowledge of the sequence of mappings R_k in terms of problem-specific data. However, it is necessary to ensure that a sequence $\{u_k\}$ in $R_k(p_k)$ converges to 0 at the same rate as the sequence $\{p_k\}$ converges to \bar{p} . This requirement is a standard assumption in the context of inexact Newton-type methods, even when applied to nonlinear equations.

Under the condition of metric regularity of a linearization of $f + F$ at \bar{p} for 0, and provided that certain additional conditions are met, we present variations of the aforementioned results. In these variations, a neighborhood of \bar{p} is assumed to be known, which allows for a more suitable choice of the initial point p_0 for the sequence $\{p_k\}$. We also provide a semi-local convergence result, where the required conditions are related to p_0 rather than \bar{p} . This result is new even in the case where \mathcal{M} is a Euclidean space.

To understand the concept of metric regularity well on Riemannian manifolds, we constructed examples of mappings that are metrically regular over the set of positive definite symmetric matrices equipped with a well-established Riemannian metric. Additionally, we present conditions that guarantee the metric regularity property for certain mappings, and in particular, we compare the metric regularity of $f + F$ with that of its linearization. Finally, some examples are given to show that our proposed concept is useful and a numerical example applied to the KKT system is given as well for the constrained Riemannian center of mass on the sphere to illustrate our theoretical results.

This work is organized as follows. In Sect. 2, some notations and basic concepts are reviewed. In Sect. 3 the metric regularity assumption is explored. In Sect. 4, local and

semi-local convergence are studied for the proposed method. In Sect. 5 the relationship is investigated between (2) and (3), and some examples of classical problems which can be viewed as generalized equations are given. In Sect. 6, a numerical example is provided to illustrate our theoretical results. Finally, conclusions are presented in the last section.

2 Preliminary

In this section we recall some notations, definitions and basic properties of Riemannian manifolds used throughout the paper, which can be found in many introductory books on Riemannian geometry, for example [21, 41, 42, 51].

Suppose that \mathcal{M} is a connected, n -dimensional smooth manifold. At each point $p \in \mathcal{M}$, the tangent space $\mathcal{T}_p\mathcal{M}$ is an n -dimensional vector space with its origin at 0_p . The disjoint union of all tangent spaces, denoted as \mathcal{TM} , is the *tangent bundle* of \mathcal{M} . We assume that \mathcal{M} is equipped with a Riemannian metric, making it a *Riemannian manifold*. At each point $p \in \mathcal{M}$, this Riemannian metric is denoted by $\langle \cdot, \cdot \rangle_p: \mathcal{T}_p\mathcal{M} \times \mathcal{T}_p\mathcal{M} \rightarrow \mathbb{R}$, and the associated norm in $\mathcal{T}_p\mathcal{M}$ is represented as $\| \cdot \|_p$. The Riemannian distance between two points p and q in \mathcal{M} , denoted as $d(p, q)$, is defined as the infimum of the lengths of all piecewise smooth curve segments connecting p to q . Additionally, the distance from a point p to a subset $\mathcal{W} \subset \mathcal{M}$ is defined as $d(p, \mathcal{W}) := \inf_{q \in \mathcal{W}} d(p, q)$ and the interior of \mathcal{W} is represented by $\text{int } \mathcal{W}$.

A *vector field* V on \mathcal{M} is a correspondence that associates to each point $p \in \mathcal{M}$ a vector $V(p) \in \mathcal{T}_p\mathcal{M}$. The point p is said to be a singularity of V if and only if $V(p) = 0_p$. The set of smooth vector fields on $\mathcal{W} \subseteq \mathcal{M}$ is denoted by $\mathcal{X}(\mathcal{W})$.

The tangent vector of a smooth curve $\gamma: I \rightarrow \mathcal{M}$ defined on some open interval $I \subseteq \mathbb{R}$ is denoted by $\dot{\gamma}(t)$. For each $a, t \in I, a < t$, the Levi-Civita connection $\nabla: \mathcal{X}(\mathcal{M}) \times \mathcal{X}(\mathcal{M}) \rightarrow \mathcal{X}(\mathcal{M})$ induces an isometry $P_{\gamma,a,t}: \mathcal{T}_{\gamma(a)}\mathcal{M} \rightarrow \mathcal{T}_{\gamma(t)}\mathcal{M}$ relative to the Riemannian metric on \mathcal{M} , given by $P_{\gamma,a,t}v = V(\gamma(t))$, where V is the unique vector field on γ such that $\nabla_{\dot{\gamma}(t)}V(\gamma(t)) = 0$ and $V(\gamma(a)) = v$. The isometry $P_{\gamma,a,t}$ is the *parallel transport* along γ joining $\gamma(a)$ to $\gamma(t)$. When the geodesic γ connecting $p = \gamma(a)$ and $q = \gamma(t)$ is unique, the notation P_{pq} will be used instead of $P_{\gamma,a,t}$. It is well-known that $P_{qp} \circ P_{pq}$ is equal to the identity map over $\mathcal{T}_p\mathcal{M}$.

The *differential* of a smooth function $f: \mathcal{M} \rightarrow \mathbb{R}$ at p is the linear map $\mathcal{D}f(p): \mathcal{T}_p\mathcal{M} \rightarrow \mathbb{R}$ which assigns to each $v \in \mathcal{T}_p\mathcal{M}$ the value

$$\mathcal{D}f(p)[v] = \dot{\gamma}(t_0)[f] = \left. \frac{d}{dt}(f \circ \gamma) \right|_{t=t_0},$$

for every smooth curve $\gamma: I \rightarrow \mathcal{M}$ satisfying $\gamma(t_0) = p$ and $\dot{\gamma}(t_0) = v$. The *gradient* at p of f , denoted as $\text{grad } f(p)$, is defined by the unique tangent vector at p such that $\langle \text{grad } f(p), v \rangle_p = \mathcal{D}f(p)[v]$ for all $v \in \mathcal{T}_p\mathcal{M}$. For a smooth multifunction $f := (f_1, \dots, f_m): \mathcal{M} \rightarrow \mathbb{R}^m$, its differential $\mathcal{D}f(p): \mathcal{T}_p\mathcal{M} \rightarrow \mathbb{R}^m$ is given by

$$\mathcal{D}f(p)[v] = (\langle \text{grad } f_1(p), v \rangle_p, \dots, \langle \text{grad } f_m(p), v \rangle_p), \quad v \in \mathcal{T}_p\mathcal{M}. \quad (7)$$

Note that $\mathcal{D}f(p)$ is a linear map from $\mathcal{T}_p\mathcal{M}$ into \mathbb{R}^m for all $p \in \mathcal{M}$. The norm of a linear map $A: \mathcal{T}_p\mathcal{M} \rightarrow \mathbb{R}^m$ is defined by $\|A\|_{\text{map}} := \sup\{\|Av\|_e: v \in \mathcal{T}_p\mathcal{M}, \|v\|_p = 1\}$ where $\| \cdot \|_e$ is the Euclidean norm on \mathbb{R}^m .

A vector field V along a smooth curve γ is said to be *parallel* if and only if $\nabla_{\dot{\gamma}}V = 0$. The curve γ is a *geodesic* when $\dot{\gamma}$ is self-parallel. When the geodesic equation $\nabla_{\dot{\gamma}}\dot{\gamma} = 0$ is

a second-order nonlinear ordinary differential equation, the geodesic $\gamma = \gamma_v(\cdot, p)$ is determined by its position p and velocity v at p . The restriction of a geodesic to a closed bounded interval is called a *geodesic segment*. Denote the unique geodesic segment $\gamma: [0, 1] \rightarrow \mathcal{M}$ satisfying $\gamma(0) = p$ and $\gamma(1) = q$ by γ_{pq} . A geodesic segment joining p to q in \mathcal{M} is said to be *minimal* if its length is equal to $d(p, q)$. A Riemannian manifold is *complete* if its geodesics are defined for all values of $t \in \mathbb{R}$. Hopf-Rinow's theorem asserts that every pair of points in a complete, connected Riemannian manifold \mathcal{M} can be joined by a (not necessarily unique) minimal geodesic segment. Due to the completeness of \mathcal{M} , the *exponential map* $\exp_p: \mathcal{T}_p\mathcal{M} \rightarrow \mathcal{M}$ is given by $\exp_p v = \gamma_v(1, p)$, for each $p \in \mathcal{M}$. In this paper, all manifolds are assumed to be connected, finite-dimensional, and complete.

The open and closed balls on \mathcal{M} of radius $r > 0$ centered at $p \in \mathcal{M}$ are defined, respectively, as $\mathcal{B}_r(p) := \{q \in \mathcal{M} : d(p, q) < r\}$ and $\mathcal{B}_r[p] := \{q \in \mathcal{M} : d(p, q) \leq r\}$. Analogously, the open and closed balls on $\mathcal{T}_p\mathcal{M}$ of radius $r > 0$ centered at $u \in \mathcal{T}_p\mathcal{M}$ are defined, respectively, as $\mathcal{B}_r(u) := \{v \in \mathcal{T}_p\mathcal{M} : \|v - u\|_p < r\}$ and $\mathcal{B}_r[u] := \{v \in \mathcal{T}_p\mathcal{M} : \|v - u\|_p \leq r\}$. It is well-known that there exists $r > 0$ such that $\exp_p: \mathcal{B}_r(0_p) \rightarrow \mathcal{B}_r(p)$ is a diffeomorphism, and $\mathcal{B}_r(p)$ is called a *normal ball* with center p and radius r . Whenever $\mathcal{B}_r(p)$ is a normal ball, it and its closure will be denoted by $\mathcal{B}_r(p)$ and $\mathcal{B}_r[p]$, respectively. The *injectivity radius of \mathcal{M} at p* , denoted by $r_{inj}(p)$, is the supremum of all $r > 0$ such that $\mathcal{B}_r(p)$ is a normal ball. The equality

$$d(q, p) = \|\exp_p^{-1} q\|_p \quad (8)$$

holds for all $q \in \mathcal{B}_r(p)$, where $\exp_p^{-1}: \mathcal{B}_r(p) \rightarrow \mathcal{T}_p\mathcal{M}$ denotes the inverse of the exponential map. Recall that there exist $r > 0$ and $\delta > 0$ such that, for every $q \in \mathcal{B}_r(p)$, $\mathcal{B}_\delta(q)$ is a normal ball and $\mathcal{B}_r(p) \subset \mathcal{B}_\delta(q)$, see [21, Theorem 3.7]. In this case, $\mathcal{B}_r(p)$ is called a *totally normal ball* of center p and radius r . When \mathcal{M} is a *Hadamard manifold*, that is, a complete, simply connected Riemannian manifold with nonpositive sectional curvature, $\mathcal{B}_r(p)$ is totally normal for all $r > 0$ and $p \in \mathcal{M}$.

A sequence $\{p_k\} \subset \mathcal{M}$ is linearly convergent to \bar{p} when there exist $\theta \in (0, 1)$ and $k_0 \in \mathbb{N}$ such that

$$d(p_{k+1}, \bar{p}) \leq \theta d(p_k, \bar{p}), \quad \text{for all } k > k_0.$$

It is said to be quadratically convergent to \bar{p} when there exist $\theta > 0$ and $k_0 \in \mathbb{N}$ such that

$$d(p_{k+1}, \bar{p}) \leq \theta d^2(p_k, \bar{p}), \quad \text{for all } k > k_0.$$

We end this section by presenting the concept of Lipschitz continuity for the differential of continuously differentiable functions in the Riemannian context. This concept will be fundamental for obtaining the quadratic convergence rate for our algorithm.

Definition 1 Let $f: \mathcal{M} \rightarrow \mathbb{R}^m$ be continuously differentiable at \bar{p} . Then $\mathcal{D}f$ is L -Lipschitz continuous around \bar{p} if there exists $\delta_L > 0$ such that for every $p, q \in \mathcal{B}_{\delta_L}(\bar{p})$

$$\|\mathcal{D}f(q)P_{\gamma,0,1} - \mathcal{D}f(p)\|_{map} \leq L\|\dot{\gamma}(0)\|_p, \quad (9)$$

where $\gamma: [0, 1] \rightarrow \mathcal{M}$ is a geodesic connecting $p = \gamma(0)$ to $q = \gamma(1)$.

3 Metric Regularity on Riemannian Manifolds

Let $F: \mathcal{M} \rightrightarrows \mathbb{R}^m$ be a set-valued mapping, whose the graph of F is the set $\text{gph } F := \{(p, x) \in \mathcal{M} \times \mathbb{R}^m : x \in F(p)\}$ with domain $\text{dom } F := \{p \in \mathcal{M} : F(p) \neq \emptyset\}$. The inverse of F is defined as $x \mapsto F^{-1}(x) = \{p \in \mathcal{M} : x \in F(p)\}$.

A function $f: \mathcal{M} \rightarrow \mathbb{R}^m$ is said to be *Lipschitz continuous* relative to a set $\mathcal{W} \subset \text{dom } f$ if there exists a constant $\kappa \geq 0$ such that

$$|f(p') - f(p)| \leq \kappa d(p', p) \text{ for all } p', p \in \mathcal{W}.$$

Let $\bar{p} \in \text{int dom } f$. The *Lipschitz modulus* of f at \bar{p} is defined by

$$\text{lip}(f; \bar{p}) := \limsup_{\substack{p', p \rightarrow \bar{p} \\ p \neq p'}} \frac{|f(p') - f(p)|}{d(p', p)}.$$

We now introduce the following notations on \mathbb{R}^m : $\mathbb{B}_r[x]$ denotes the Euclidean closed ball of radius $r > 0$ and center x , and d_e denotes the Euclidean distance from a point to a set. The next concept plays an important role in this paper and it comes from [24, p. 279] in the metric spaces setting:

Definition 2 A set-valued mapping $\Phi: \mathcal{M} \rightrightarrows \mathbb{R}^m$ is said to be *metrically regular* at $\bar{p} \in \mathcal{M}$ for $\bar{x} \in \mathbb{R}^m$ when $\bar{x} \in \Phi(\bar{p})$, and there exist positive constants σ, a and b such that

$$\text{the set } \text{gph } \Phi \cap (\mathcal{B}_a[\bar{p}] \times \mathbb{B}_b[\bar{x}]) \text{ is closed} \quad (10)$$

and

$$d(p, \Phi^{-1}(x)) \leq \sigma d_e(x, \Phi(p)) \text{ for all } (p, x) \in \mathcal{B}_a[\bar{p}] \times \mathbb{B}_b[\bar{x}]. \quad (11)$$

The infimum of σ over all such combinations of σ, a and b is called the *regularity modulus* for Φ at \bar{p} for \bar{x} and denoted by $\text{reg}(\Phi; \bar{p}|\bar{x})$. The absence of metric regularity is signaled by $\text{reg}(\Phi; \bar{p}|\bar{x}) = \infty$.

It is important to emphasize that one of the main assumptions used to guarantee the convergence of the proposed algorithm in this paper is metric regularity. This concept, in the context of Riemannian manifolds, has been explored in recent works such as [2, 4, 29]. We will now briefly discuss this property in the case where $\mathcal{M} \equiv \mathbb{R}^n$. If $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a differentiable function, the metric regularity of f at \bar{p} for 0 is equivalent to stating that the inverse of the Jacobian matrix of f at \bar{p} exists, which is the standard regularity assumption applied to solve the nonlinear equation $f(p) = 0$. On the other hand, if we are interested in solving a nonlinear system of equations and inequalities, for example,

$$g(p) \leq x, \quad h(p) = y, \quad (12)$$

where $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^s$ are given parameters, and $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $h: \mathbb{R}^n \rightarrow \mathbb{R}^s$ are continuously differentiable functions, then we can associate problem (12) with the following generalized equation:

$$\xi \in f(p) + F(p), \quad \xi = \begin{bmatrix} x \\ y \end{bmatrix}, \quad f = \begin{bmatrix} g \\ h \end{bmatrix}, \quad F = \begin{bmatrix} \mathbb{R}_+^m \\ 0 \end{bmatrix}.$$

In this particular case, it is well known that the metric regularity of $f + F$ at \bar{p} for 0 is equivalent to the standard Mangasarian-Fromovitz constraint qualification at \bar{p} . Notably, the case where $F \equiv N_C$ is of particular interest, with N_C denoting the normal cone to a closed

convex set $C \subset \mathbb{R}^m$. The metric regularity of $f + N_C$ at \bar{p} for 0 is equivalent to the concept of *strong metric regularity*, meaning that $f + N_C$ is locally single-valued and Lipschitz continuous at \bar{p} for 0. Thus, if f and C are chosen such that $f + N_C \ni 0$ forms the KKT system for a constrained optimization problem, then strong metric regularity is equivalent to the linear independence of the gradients of the active constraints and the strong second-order sufficient condition. These details are thoroughly discussed in [25]; see also Examples 6 and 7 in Sect. 5.

We present in Subsection 3.1 some examples of set-valued mapping that satisfy the previous concept. The next lemma is a version in the Riemannian setting of Corollary 3F.3 given in [25]. Its proof is similar to the corresponding result in Euclidean space and will be omitted here.

Lemma 1 *Consider a mapping $\Phi: \mathcal{M} \rightrightarrows \mathbb{R}^m$ and a point $(\bar{p}, \bar{x}) \in \text{gph } \Phi$. Then for every $g: \mathcal{M} \rightarrow \mathbb{R}^m$ with $\text{lip}(g; \bar{p}) = 0$, one has $\text{reg}(g + \Phi; \bar{p}|g(\bar{p}) + \bar{x}) = \text{reg}(\Phi; \bar{p}|\bar{x})$.*

In the theorem below, we demonstrate that the metric regularity condition of a mapping at $\bar{p} \in \mathcal{M}$ for $\bar{x} \in \mathbb{R}^m$ is equivalent to the same condition for its linearization at \bar{p} .

Theorem 1 *Consider a normal ball $B_r(\bar{p}) \subset \mathcal{M}$. Let $f: \mathcal{M} \rightarrow \mathbb{R}^m$ be a continuously differentiable function at \bar{p} and $F: \mathcal{M} \rightrightarrows \mathbb{R}^m$ be a set-valued mapping. Then for $G: \mathcal{M} \rightrightarrows \mathbb{R}^m$ defined by*

$$G(p) = \begin{cases} f(\bar{p}) + \mathcal{D}f(\bar{p})[\exp_{\bar{p}}^{-1} p] + F(p), & p \in B_r(\bar{p}), \\ f(p) + F(p), & p \in \mathcal{M} \setminus B_r(\bar{p}), \end{cases} \quad (13)$$

one has $\text{reg}(G; \bar{p}|0) = \text{reg}(f + F; \bar{p}|0)$.

Proof Pick $\epsilon > 0$ arbitrary. By Lemma 5 given in the Appendix of this paper, there exists a totally normal ball $B_{\delta_\epsilon}(\bar{p}) \subset B_r(\bar{p})$ such that

$$\|\mathcal{D}f(p)[\exp_p^{-1} p' - \exp_p^{-1} p''] - \mathcal{D}f(\bar{p})[\exp_{\bar{p}}^{-1} p' - \exp_{\bar{p}}^{-1} p'']\|_e \leq \epsilon d(p', p''), \quad p, p', p'' \in B_{\delta_\epsilon}(\bar{p}). \quad (14)$$

Consider a function $g: \mathcal{M} \rightarrow \mathbb{R}^m$ given by

$$g(q) = \begin{cases} f(\bar{p}) - f(q) + \mathcal{D}f(\bar{p})[\exp_{\bar{p}}^{-1} q], & q \in B_r(\bar{p}), \\ 0, & q \in \mathcal{M} \setminus B_r(\bar{p}). \end{cases} \quad (15)$$

In particular, this function satisfies

$$\|g(p'') - g(p')\|_e = \|f(p') - f(p'') - \mathcal{D}f(\bar{p})[\exp_{\bar{p}}^{-1} p' - \exp_{\bar{p}}^{-1} p'']\|_e, \quad p', p'' \in B_{\delta_\epsilon}(\bar{p}). \quad (16)$$

The first part of Proposition 4 guarantees that there exists $\delta \in (0, \delta_\epsilon)$ such that

$$f(p') - f(p'') = \int_0^1 \mathcal{D}f(\gamma_{p''p'}(t))[\dot{\gamma}_{p''p'}(t)] dt, \quad p', p'' \in B_\delta(\bar{p}),$$

where $\gamma_{p''p'}: [0, 1] \rightarrow \mathcal{M}$ is the geodesic satisfying $\gamma_{p''p'}(0) = p''$ and $\gamma_{p''p'}(1) = p'$. Adding $-\mathcal{D}f(\bar{p})[\exp_{\bar{p}}^{-1} p' - \exp_{\bar{p}}^{-1} p'']$ to both sides, it follows that

$$\begin{aligned}
& f(p') - f(p'') - \mathcal{D}f(\bar{p})[\exp_{\bar{p}}^{-1} p' - \exp_{\bar{p}}^{-1} p''] \\
&= \int_0^1 \mathcal{D}f(\gamma_{p''p'}(t))[\dot{\gamma}_{p''p'}(t)] - \mathcal{D}f(\bar{p})[\exp_{\bar{p}}^{-1} p' - \exp_{\bar{p}}^{-1} p''] dt, \quad p', p'' \in B_\delta(\bar{p}).
\end{aligned}$$

Applying norm on both sides, we conclude that

$$\begin{aligned}
& \|f(p') - f(p'') - \mathcal{D}f(\bar{p})[\exp_{\bar{p}}^{-1} p' - \exp_{\bar{p}}^{-1} p'']\|_e \\
& \leq \int_0^1 \|\mathcal{D}f(\gamma_{p''p'}(t))[\dot{\gamma}_{p''p'}(t)] - \mathcal{D}f(\bar{p})[\exp_{\bar{p}}^{-1} p' - \exp_{\bar{p}}^{-1} p'']\|_e dt, \quad p', p'' \in B_\delta(\bar{p}).
\end{aligned} \tag{17}$$

On the other hand, from Proposition 5 with $p = p''$ and $q = p'$ we have

$$\dot{\gamma}_{p''p'}(t) = \exp_{\gamma_{p''p'}(t)}^{-1} p' - \exp_{\gamma_{p''p'}(t)}^{-1} p'', \quad t \in [0, 1].$$

From this equality and (14) with $p = \gamma_{p''p'}(t)$, (17) yields

$$\|f(p') - f(p'') - \mathcal{D}f(\bar{p})(\exp_{\bar{p}}^{-1} p' - \exp_{\bar{p}}^{-1} p'')\|_e \leq \epsilon d(p', p''), \quad p', p'' \in B_\delta(\bar{p}).$$

Then (16) implies that $\|g(p'') - g(p')\|_e \leq \epsilon d(p', p'')$ for all $p', p'' \in B_\delta(\bar{p})$. Since $\epsilon > 0$ was chosen arbitrarily, we can conclude that $\text{lip}(g; \bar{p}) = 0$. Finally, to obtain the desired equality, just use Lemma 1 with g as in (15) and $\Phi = f + F$. \square

Next, we define the Aubin property for a set-valued mapping $\mathfrak{N}: \mathbb{R}^m \rightrightarrows \mathcal{M}$. To this end, it is necessary to introduce the concept of excess between subsets of \mathcal{M} . For sets \mathcal{W}_1 and \mathcal{W}_2 in \mathcal{M} , the *excess of \mathcal{W}_1 beyond \mathcal{W}_2* is defined as

$$e(\mathcal{W}_1, \mathcal{W}_2) = \sup_{p \in \mathcal{W}_1} d(p, \mathcal{W}_2),$$

with the convention that $e(\emptyset, \mathcal{W}_2) = 0$ for $\mathcal{W}_2 \neq \emptyset$ and $e(\emptyset, \mathcal{W}_2) = \infty$ otherwise.

Definition 3 A set-valued mapping $\mathfrak{N}: \mathbb{R}^m \rightrightarrows \mathcal{M}$ is said to have the Aubin property at $\bar{x} \in \mathbb{R}^m$ for $\bar{p} \in \mathcal{M}$ if $\bar{p} \in \mathfrak{N}(\bar{x})$, and there exist positive constants σ , a , and b such that

$$e(\mathfrak{N}(x) \cap \mathcal{B}_a[\bar{p}], \mathfrak{N}(x')) \leq \sigma \|x - x'\| \quad \text{for all } x, x' \in \mathbb{B}_b[\bar{x}].$$

The infimum of σ over all such combinations of σ , a , and b is called the Lipschitz modulus of \mathfrak{N} at \bar{x} for \bar{p} and is denoted by $\text{lip}(\mathfrak{N}; \bar{x}|\bar{p})$. The absence of this property is indicated by $\text{lip}(\mathfrak{N}; \bar{x}|\bar{p}) = \infty$.

The next result provides a relationship between the metric regularity of a set-valued mapping $\Phi: \mathcal{M} \rightrightarrows \mathbb{R}^m$ and the Aubin property for its inverse. Its proof follows from a straightforward adaptation of the demonstration in [25, Theorem 3E.6] and the concepts introduced above, and hence will be omitted.

Theorem 2 A set-valued mapping $\Phi: \mathcal{M} \rightrightarrows \mathbb{R}^m$ is metrically regular at \bar{p} for \bar{x} with a constant σ if and only if its inverse $\Phi^{-1}: \mathbb{R}^m \rightrightarrows \mathcal{M}$ has the Aubin property at \bar{x} for \bar{p} with constant σ . Thus, $\text{lip}(\Phi^{-1}; \bar{x}|\bar{p}) = \text{reg}(\Phi; \bar{p}|\bar{x})$.

3.1 Example in the SPD Matrices Cone

In this section, some examples of metrically regular mappings Φ on a particular manifold \mathcal{M} are presented.

Denote the set of symmetric matrices of size $n \times n$ for some $n \in \mathbb{N}$ by $\mathcal{P}(n)$, and the cone of symmetric positive definite matrices by $\mathcal{M} = \mathcal{P}_+(n)$. The latter is endowed with the affine invariant Riemannian metric given by

$$\langle v, u \rangle := \text{tr}(vp^{-1}up^{-1}), \quad p \in \mathcal{M}, \quad v, u \in \mathcal{T}_p\mathcal{M}, \quad (18)$$

where tr denotes the trace of a matrix. The tangent space $\mathcal{T}_p\mathcal{M}$ can be identified with $\mathcal{P}(n)$. Moreover, \mathcal{M} is a Hadamard manifold, see, for example, [40, Theorem 1.2. p. 325]. The Riemannian distance is given by

$$d(p, q) = \text{tr}^{\frac{1}{2}}(\ln^2(p^{-\frac{1}{2}}qp^{-\frac{1}{2}})), \quad p, q \in \mathcal{M}, \quad (19)$$

where \ln denotes the matrix logarithm. The Riemannian gradient of $f: \mathcal{M} \rightarrow \mathbb{R}$ at p is given by

$$\text{grad } f(p) = pf'(p)p, \quad (20)$$

where $f'(p)$ represents the Euclidean gradient of f at p , see [49]. The identity matrix of size $n \times n$ will be denoted by id .

In the following two examples, we verify the property of metric regularity for single-valued functions.

Example 1 Consider the function $\Phi: \mathcal{M} \rightarrow \mathbb{R}$ defined by $\Phi(p) = \ln(\text{tr}(p))$. For every $x \in \mathbb{R}$, the following equality holds:

$$\Phi(e^{x-\ln(\text{tr}(p))}p) = \ln(\text{tr}(e^{x-\ln(\text{tr}(p))}p)) = \ln(e^{x-\ln(\text{tr}(p))} \text{tr}(p)) = x, \quad p \in \mathcal{M}. \quad (21)$$

This implies that $\{e^{x-\ln(\text{tr}(p))}p: p \in \mathcal{M}\} \subset \Phi^{-1}(x)$ for all $x \in \mathbb{R}$. Thus, using (19) and simple algebraic manipulations, we obtain

$$\begin{aligned} d(\Phi^{-1}(x), q) &\leq d(e^{x-\ln(\text{tr}(q))}q, q), \\ &= \text{tr}^{\frac{1}{2}}(|x - \ln(\text{tr}(q))|^2 \text{id}), \\ &= \sqrt{n} |x - \ln(\text{tr}(q))| = \sqrt{n} |x - \Phi(q)|, \quad x \in \mathbb{R}, \quad q \in \mathcal{M}. \end{aligned}$$

On the other hand, since Φ is Euclidean differentiable, it follows from (20) that Φ will also be Riemannian differentiable with respect to metric (18). Therefore, Φ is continuous and, by Proposition 6, it has closed graph. Given this and the last inequality above, we conclude that Φ is metrically regular at p for $\Phi(p)$ for all $p \in \mathcal{M}$.

Example 2 Define the function $\Phi: \mathcal{M} \rightarrow \mathbb{R}$ by $\Phi(p) = 1/\text{tr}(p)$. Note that $\{(x \text{tr}(p))^{-1}p: p \in \mathcal{M}\} \subset \Phi^{-1}(x)$ for all $x \in \mathbb{R} \setminus \{0\}$. Using this and (19), we obtain

$$\begin{aligned} d(\Phi^{-1}(x), p) &\leq d((x \text{tr}(p))^{-1}p, p) \\ &= \text{tr}^{\frac{1}{2}}(\ln^2(x \text{tr}(p) \text{id})) = \sqrt{n} |\ln(x \text{tr}(p))|, \quad x \in \mathbb{R} \setminus \{0\}, \quad p \in \mathcal{M}. \quad (22) \end{aligned}$$

Now recall that

$$\text{tr}(p) = \lambda_1(p) + \dots + \lambda_n(p), \quad (23)$$

where $\lambda_1(p), \dots, \lambda_n(p)$ are the eigenvalues of p . Pick a constant $a > 0$. From (19), for all $i \in \{1, \dots, n\}$, one has

$$|\ln(\lambda_i(p))| < (\ln^2(\lambda_1(p)) + \dots + \ln^2(\lambda_n(p)))^{\frac{1}{2}} = \text{tr}^{\frac{1}{2}}(\ln^2(p)) = d(p, \text{id}) < a, \quad p \in B_a(\text{id}).$$

Then $e^{-a} < \lambda_i(p) < e^a$ holds for all $p \in B_a(\text{id})$ and $i = 1, \dots, n$. Consequently

$$ne^{-a} < \lambda_1(p) + \dots + \lambda_n(p) < ne^a, \quad p \in B_a(\text{id}).$$

With (23), we conclude that $1/\text{tr}(p) \in (e^{-a}/n, e^a/n)$ for all $p \in B_a(\text{id})$. Since the function \ln is Lipschitz on the interval $(e^{-a}/n, e^a/n)$ (its gradient is bounded on this interval), there exists $\sigma > 0$ such that

$$|\ln(x \text{tr}(p))| = |\ln(x) - \ln(1/\text{tr}(p))| \leq \sigma |x - 1/\text{tr}(p)|, \quad x \in (e^{-a}/n, e^a/n), \quad p \in B_a(\text{id}). \quad (24)$$

Therefore, (22) and (24) yield

$$d(\Phi^{-1}(x), p) \leq \sqrt{n} \sigma |x - \Phi(p)|, \quad x \in (e^{-a}/n, e^a/n), \quad p \in B_a(\text{id}).$$

Furthermore, taking into account the argument presented at the end of Example 1, it can be concluded that $\text{gph } \Phi$ is closed. Then Φ is metrically regular at p for $\Phi(p)$ for all $p \in B_a(\text{id})$.

We now present two examples involving set-valued mappings. In these examples, we demonstrate that the given mappings are metrically regular at specific points in \mathcal{M} .

Example 3 Define the set-valued mapping $\Phi: \mathcal{M} \rightrightarrows \mathbb{R}$ by

$$\Phi(p) = \begin{cases} \ln(\text{tr}(p)), & p \in \mathcal{M} \setminus (1/n \text{id}), \\ \{0\} \cup [1, 2], & p = 1/n \text{id}. \end{cases} \quad (25)$$

By (25) and the calculations in (21), if $x \in \mathbb{R}$ and $q \in \mathcal{M}$ satisfy $e^{x - \ln(\text{tr}(q))} q \in \mathcal{M} \setminus (1/n \text{id})$ then $\Phi(e^{x - \ln(\text{tr}(q))} q) \ni \ln(\text{tr}(e^{x - \ln(\text{tr}(q))} q)) = x$. Now, consider the case $e^{x - \ln(\text{tr}(q))} q = 1/n \text{id}$. By applying tr on both sides, it follows that $e^{x - \ln(\text{tr}(q))} \text{tr}(q) = 1$. Clearly, this implies that $x = 0$. Therefore, by (25), we get $\Phi(e^{x - \ln(\text{tr}(q))} q) \ni x$ for this case as well. Thus, $\{e^{x - \ln(\text{tr}(q))} q : q \in \mathcal{M}\} \subset \Phi^{-1}(x)$ holds for all $x \in \mathbb{R}$, which implies

$$\begin{aligned} d(\Phi^{-1}(x), q) &\leq d(e^{x - \ln(\text{tr}(q))} q, q), \\ &= \text{tr}^{\frac{1}{2}}(|x - \ln(\text{tr}(q))|^2 \text{id}) = \sqrt{n} |x - \ln(\text{tr}(q))|, \quad x \in \mathbb{R}, \quad q \in \mathcal{M}. \end{aligned}$$

On the other hand, it follows from (25) that

$$|x - \ln(\text{tr}(q))| = d_e(x, \Phi(q)), \quad x \in (-\infty, 1/2), \quad q \in \mathcal{M}.$$

Overall, we conclude

$$d(\Phi^{-1}(x), q) \leq \sqrt{n} d_e(x, \Phi(q)), \quad x \in (-\infty, 1/2), \quad q \in \mathcal{M}.$$

With a justification analogous to the one presented in Example 1, we state that Φ has closed graph. Therefore, Φ is metrically regular at p for $\ln(\text{tr}(p))$ for all $p \in \{q \in \mathcal{M} : \ln(\text{tr}(q)) \in (-\infty, 1/2)\}$.

Example 4 Consider the set-valued mapping $\Phi: \mathcal{M} \rightrightarrows \mathbb{R}$ by

$$\Phi(p) = \begin{cases} 1/\text{tr}(p), & p \in \mathcal{M} \setminus \text{id}, \\ \{1/n\} \cup [2, 3], & p = \text{id}. \end{cases} \quad (26)$$

Following the steps of Example 3, we can show that $\Phi(1/(x \operatorname{tr}(q))q) \ni 1/\operatorname{tr}(1/(x \operatorname{tr}(q))q) = x$ holds for all $x \in \mathbb{R} \setminus \{0\}$, $q \in \mathcal{M}$. This means that $\{1/(x \operatorname{tr}(q))q : q \in \mathcal{M}\} \subset \Phi^{-1}(x)$ for all $x \in \mathbb{R} \setminus \{0\}$, and hence

$$\begin{aligned} d(\Phi^{-1}(x), q) &\leq d((x \operatorname{tr}(q))^{-1}q, q) \\ &= \operatorname{tr}^{\frac{1}{2}}(\ln^2(x \operatorname{tr}(q) \operatorname{id})) = \sqrt{n} |\ln(x \operatorname{tr}(q))|, \quad x \in \mathbb{R} \setminus \{0\}, \quad q \in \mathcal{M}. \end{aligned}$$

From (24), for all $a > 0$ there exists $\sigma > 0$ such that

$$d(\Phi^{-1}(x), q) \leq \sigma |x - 1/\operatorname{tr}(q)|, \quad x \in (e^{-a}/n, e^a/n), \quad q \in B_a(\operatorname{id}).$$

Pick $a > 0$. Exploiting the definition of Φ in (26), it follows that

$$d(\Phi^{-1}(x), q) \leq \sigma d_e(x, \Phi(q)), \quad x \in (e^{-a}/n, \min\{e^a/n, 1\}), \quad q \in B_a(\operatorname{id}).$$

Combining this with the fact that $\operatorname{gph} \Phi$ is closed, we conclude that Φ is metrically regular at p for $1/\operatorname{tr}(p)$ for all $p \in \{q \in \mathcal{M} : 1/\operatorname{tr}(q) \in (e^{-a}/n, \min\{e^a/n, 1\})\}$.

4 Convergence

In this section, we will establish three convergence theorems for the sequence generated by the inexact Newton method described in (6). The proofs presented here rely on a Riemannian version of Theorem 5G.3 from [25], which is presented in the following lemma. The proof of this lemma is a straightforward adaptation of its Euclidean version and will be omitted.

Lemma 2 Consider a set-valued mapping $F: \mathcal{M} \rightrightarrows \mathbb{R}^m$ and a point $(\bar{p}, \bar{x}) \in \operatorname{gph} F$ at which F is metrically regular with positive constants σ , a and b satisfying (10)–(11). Let $v > 0$ be such that $\sigma v < 1$ and $\sigma' > \sigma/(1 - \sigma v)$. Then for every positive α and β such that

$$\alpha \leq a/2, \quad v\alpha + 2\beta \leq b, \quad 2\sigma'\beta \leq \alpha$$

and for every function $g: \mathcal{M} \rightarrow \mathbb{R}^m$ satisfying

$$\|g(\bar{p})\|_e \leq \beta, \quad \|g(q) - g(q')\|_e \leq v d(q, q'), \quad q, q' \in B_\alpha[\bar{p}],$$

the mapping $g + F$ has the following property: for every $x, x' \in B_\beta[\bar{x}]$ and every $p \in (g + F)^{-1}(x) \cap B_\alpha[\bar{p}]$ there exists $p' \in (g + F)^{-1}(x')$ such that $d(p, p') \leq \sigma' \|x - x'\|_e$.

Our first convergence result is a local analysis of the inexact Newton method (6) for solving (2). This approach makes assumptions around a solution \bar{p} of (2).

Theorem 3 Let \bar{p} be a solution of (2). Suppose that the following conditions hold:

- (i) the function $f: \mathcal{M} \rightarrow \mathbb{R}^m$ is continuously differentiable at \bar{p} and $\operatorname{reg}(f + F; \bar{p}|0) = \sigma$;
- (ii) the sequence $\{R_k\}$ satisfies

$$\limsup_{\substack{p \rightarrow \bar{p} \\ p \neq \bar{p}}} \left\{ \frac{1}{d(p, \bar{p})} \sup_{k \in \mathbb{N}_0} \sup_{x \in R_k(p)} \|x\|_e \right\} < \frac{1}{\sigma}. \quad (27)$$

Then there exist $\theta \in (0, 1)$ and a totally normal ball $B_r(\bar{p}) \subset \mathcal{M}$ such that for every $p \in B_r(\bar{p}) \setminus \{\bar{p}\}$, every $k \in \mathbb{N}_0$, and every $u_k \in R_k(p)$ there exists $q' \in B_r(\bar{p})$ satisfying

$$f(p) + \mathcal{D}f(p)[\exp_p^{-1} q'] + F(q') \ni u_k \quad (28)$$

and

$$d(q', \bar{p}) \leq \theta d(p, \bar{p}). \quad (29)$$

Consequently, for any starting point $p_0 \in B_r(\bar{p})$, there exists a sequence $\{p_k\}$ generated by (6) that converges linearly to \bar{p} .

Proof First, (27) implies that there exists $\iota > 0$ satisfying the following inequality:

$$\limsup_{\substack{p \rightarrow \bar{p} \\ p \neq \bar{p}}} \left\{ \frac{1}{d(p, \bar{p})} \sup_{k \in \mathbb{N}_0} \sup_{x \in R_k(p)} \|x\|_e \right\} < \iota < \frac{1}{\sigma}. \quad (30)$$

Choose ι satisfying (30) and positive constants $\mu > 0, \kappa > \sigma, \epsilon > 0$ and $\theta \in (0, 1)$ such that

$$\mu\kappa < 1, \quad \kappa(\epsilon + \iota) < \theta(1 - \mu\kappa). \quad (31)$$

Pick $\Theta > 0$ and $\tau \in (\sigma, \kappa)$ such that

$$\frac{\sigma}{1 - \mu\sigma} < \Theta < \frac{\kappa}{1 - \mu\kappa}, \quad \frac{\tau}{1 - \mu\tau} < \Theta. \quad (32)$$

From the first inequality in (30), there exists $\delta > 0$ such that

$$\|x\|_e < \iota d(p, \bar{p}) \quad \text{whenever} \quad p \in B_\delta(\bar{p}) \setminus \{\bar{p}\}, \quad x \in R_k(p), \quad k \in \mathbb{N}_0. \quad (33)$$

Choose a totally normal ball $B_\delta(\bar{p}) \subset B_\delta(\bar{p})$. Note that the function $g_p: \mathcal{M} \rightarrow \mathbb{R}^m$ defined by

$$g_p(q) = \begin{cases} \mathcal{D}f(p)[\exp_p^{-1} q - \exp_p^{-1} \bar{p}] - \mathcal{D}f(\bar{p})[\exp_{\bar{p}}^{-1} q], & q \in B_\delta(\bar{p}), \\ 0, & q \in \mathcal{M} \setminus B_\delta(\bar{p}). \end{cases} \quad (34)$$

satisfies $g_p(\bar{p}) = 0$ for all $p \in B_\delta(\bar{p})$. On the other hand, since f and F satisfy (i), it follows from Theorem 1 that $\text{reg}(G; \bar{p}|0) = \sigma$, where $G: \mathcal{M} \rightrightarrows \mathbb{R}^m$ is the function defined in (13) with $r = \bar{\delta}$. Hence, by the definition of reg in Definition 2, there exist $a > 0$ and $b > 0$ such that (10) and (11) are satisfied for $\Phi = G$ and $\sigma = \tau$. Moreover, by Lemma 5 and (34), there exists $\alpha < \min\{\bar{\delta}, a/2, b/\mu\}$ such that

$$\|g_p(q) - g_p(q')\|_e = \|\mathcal{D}f(p)[\exp_p^{-1} q - \exp_p^{-1} q'] - \mathcal{D}f(\bar{p})[\exp_{\bar{p}}^{-1} q - \exp_{\bar{p}}^{-1} q']\|_e \leq \mu d(q, q'),$$

for all $p, q, q' \in B_\alpha(\bar{p})$. Therefore, for every $p \in B_\alpha(\bar{p})$ one has

$$g_p(\bar{p}) = 0, \quad \|g_p(q) - g_p(q')\|_e \leq \mu d(q, q'), \quad q, q' \in B_\alpha[\bar{p}].$$

Overall, there exists $\beta > 0$ (independent of p) such that

$$\alpha \leq a/2, \quad \mu\alpha + 2\beta \leq b, \quad 2\Theta\beta \leq \alpha,$$

and

$$\|g_p(\bar{p})\|_e \leq \beta, \quad \|g_p(q) - g_p(q')\|_e \leq \mu d(q, q'), \quad p, q, q' \in B_\alpha[\bar{p}].$$

Applying Lemma 2 with $g = g_p, F = G, v = \mu, \sigma' = \Theta$ and $\bar{x} = 0$, we conclude that for every $x, x' \in B_\beta[0]$ and every $q \in (g_p + G)^{-1}(x) \cap B_\alpha[\bar{p}]$ there exists $q' \in (g_p + G)^{-1}(x')$ such that $d(q, q') \leq \Theta\|x - x'\|_e$. Taking $x = \bar{x} = 0$ and $q = \bar{p}$, we conclude that for every $x' \in B_\beta[0]$ there exists $q' \in (g_p + G)^{-1}(x')$ such that

$$d(\bar{p}, q') \leq \Theta\|x'\|_e. \quad (35)$$

Using the second part of Proposition 4 with $\epsilon > 0$ chosen as in (31), we have the existence of a normal ball $B_{\delta_\epsilon}(\bar{p})$ such that

$$\|f(p) - f(\bar{p}) - \mathcal{D}f(\bar{p})[\exp_{\bar{p}}^{-1} p]\|_e \leq \epsilon d(p, \bar{p}), \quad p \in B_{\delta_\epsilon}(\bar{p}), \quad (36)$$

Fixed r such that

$$0 < r < \min \left\{ \frac{\beta}{\epsilon + \iota}, \bar{\delta}, \delta_\epsilon \right\} \quad (37)$$

and $p \in B_r(\bar{p}) \setminus \{\bar{p}\}$, by choosing $k \in \mathbb{N}_0$ and $u_k \in R_k(p)$, it follows from (33) that u_k satisfies $\|u_k\|_e \leq \iota d(p, \bar{p})$. Denote

$$y_k := f(p) - f(\bar{p}) + \mathcal{D}f(p)[\exp_p^{-1} \bar{p}] - u_k. \quad (38)$$

If $y_k = 0$, then $q' := \bar{p}$ satisfies (28) because $-f(\bar{p}) \in F(\bar{p})$, and (29) holds trivially. For $y_k \neq 0$, it follows from (38), (36) and (33) that

$$\|y_k\|_e \leq \|f(p) - f(\bar{p}) + \mathcal{D}f(p)[\exp_p^{-1} \bar{p}]\|_e + \|u_k\|_e \leq (\epsilon + \iota)d(p, \bar{p}). \quad (39)$$

Since $d(p, \bar{p}) < r < \beta/(\epsilon + \iota)$, it follows from (39) that $\|y_k\|_e < \beta$. Applying (35) with $x' = -y_k$, we obtain that there exists $q' \in (g_p + G)^{-1}(-y_k)$ such that $d(q', \bar{p}) \leq \Theta \|y_k\|_e$. Hence, from the upper bound for Θ given in (32), (39), and the last inequality in (31), it follows that

$$d(q', \bar{p}) < \frac{\kappa}{1 - \mu\kappa} \|y_k\|_e < \frac{\kappa(\epsilon + \iota)}{1 - \mu\kappa} d(p, \bar{p}) < \theta d(p, \bar{p}).$$

Furthermore, it comes from $-y_k \in (g_p + G)(q')$, (38), (34) and (13) that

$$-f(p) + f(\bar{p}) - \mathcal{D}f(p)[\exp_p^{-1} \bar{p}] + u_k \in f(\bar{p}) + \mathcal{D}f(p)[\exp_p^{-1} q' - \exp_p^{-1} \bar{p}] + F(q'),$$

which means that $u_k \in f(p) + \mathcal{D}f(p)[\exp_p^{-1} q'] + F(q')$. Thus, q' satisfies (28) and (29).

Now choose $p_0 \in B_r(\bar{p})$. If $p_k = \bar{p}$, then $p_{k+1} := \bar{p}$ verifies (6) because $0 \in R_k(\bar{x})$. If $p_k \neq \bar{p}$, applying (28) and (29) with $p = p_0$ and $q' = p_1$ we have $f(p_0) + \mathcal{D}f(p_0)[\exp_{p_0}^{-1} p_1] + F(p_1) \ni u_0$ and

$$d(p_1, \bar{p}) \leq \theta d(p_0, \bar{p}).$$

Repeating this argument one can conclude that there exists a sequence $\{p_k\}$ in $B_r(\bar{p})$ which satisfies (6) and converges linearly to \bar{p} . \square

The following result provides a class of sequences $\{R_k\}$ that satisfy (27).

Proposition 1 Consider the generalized equation given in (2). Let $\varkappa > 1$ be a constant, and let $\zeta: \mathcal{M} \rightarrow \mathbb{R}^s$ be a locally Lipschitz function that satisfies $\|\zeta(p)\|_e = 0$ for all $p \in (f + F)^{-1}(0)$. Then, the sequence $\{R_k\}$, defined as

$$R_k(p) := \mathbb{B}_{\|\zeta^{\varkappa}(p)\|_e}[0], \quad k \in \mathbb{N}_0, \quad (40)$$

satisfies (27).

Proof Take $\bar{p} \in (f + F)^{-1}(0)$. Based on the assumptions regarding the function ζ , we can ensure the existence of two positive constants L and δ such that

$$\|\zeta(p)\|_e = \|\zeta(p) - \zeta(\bar{p})\|_e \leq L d(p, \bar{p}),$$

for all $p \in \mathcal{B}_\delta(\bar{p})$. Using this and (40), it follows that

$$\|x\|_e \leq \|\zeta^\varkappa(p)\|_e \leq L^\varkappa d^\varkappa(p, \bar{p}),$$

for all $p \in \mathcal{B}_\delta(\bar{p})$, $k \in \mathbb{N}_0$, and $x \in R_k(p)$. Consequently, we can guarantee that

$$\limsup_{\substack{p \rightarrow \bar{p} \\ p \neq \bar{p}}} \left\{ \frac{1}{d(p, \bar{p})} \sup_{k \in \mathbb{N}_0} \sup_{x \in R_k(p)} \|x\|_e \right\} \leq \limsup_{\substack{p \rightarrow \bar{p} \\ p \neq \bar{p}}} \{L^\varkappa d^{\varkappa-1}(p, \bar{p})\} = 0,$$

which completes the proof. \square

Remark 1 By following the ideas presented in the previous proof, it can be shown that if the Lipschitz constant of ζ around $\bar{p} \in (f + F)^{-1}(0)$ is less than $1/\sigma$, where $\sigma := \text{reg}(f + F; \bar{p} \mid 0)$, then Proposition 1 also holds for $\varkappa = 1$.

Next, we discuss the inexactness of method (5) and the previous proposition.

Remark 2 For practical purposes, it is desirable that the sequence $\{R_k\}$ satisfies the following conditions:

- (a) It is independent of any $\bar{p} \in (f + F)^{-1}(0)$;
- (b) $R_k(p) = 0$ if and only if $p \in (f + F)^{-1}(0)$, for all $k \in \mathbb{N}_0$.

Condition (a) ensures that the method in (5) can be applied even in cases where no $\bar{p} \in (f + F)^{-1}(0)$ is known, while condition (b) guarantees that the set $\{R_k(p_k)\}$ contains elements with positive norm whenever $p_k \notin (f + F)^{-1}(0)$, and also provides a stopping criterion for the method. It is worth noting that the inexactness of the inexact Newton method introduced in [20] for solving a specific case of (2), where $F \equiv 0$ and $\Omega = \mathcal{M} = \mathbb{R}^n$, satisfies both conditions (a) and (b). In the case where (2) represents a general generalized equation, Proposition 1 helps in identifying a sequence $\{R_k\}$ that satisfies these properties. Indeed, based on Proposition 1, we can conclude that if the function ζ satisfies its assumptions and is independent of any $\bar{p} \in (f + F)^{-1}(0)$, then the sequence $\{R_k\}$ defined in (40) satisfies conditions (a) and (b). In Proposition 3, we provide an example of such a function ζ for the case where (2) represents the generalized equation associated with the Riemannian KKT system.

In the following remark, we provide a condition, independent of the solution set of (2), that guarantees the sequence R_k , as defined in (40), satisfies (27) as well as conditions (a) and (b) in Remark 2.

Remark 3 Let $\tilde{\zeta}: \mathcal{M} \rightarrow \mathbb{R}$ be the function defined as $\tilde{\zeta}(p) = d_e(0, (f + F)(p))$ for all $p \in \mathcal{M}$. Since $\tilde{\zeta}(p) = 0$ for every $p \in (f + F)^{-1}(0)$, it follows from Proposition 1 that if $\tilde{\zeta}$ is locally Lipschitz then the sequence $\{R_k\}$ given in (40), with $\zeta = \tilde{\zeta}$, satisfies (27).

An upper bound for the radius r of the ball mentioned in the statement of Theorem 3 is given in (37). This bound depends on certain constants that, while not known a priori, are guaranteed to exist due to conditions i) and ii) in Lemma 2, as well as additional results presented in the Appendix of this paper.

Below, we present a version of Theorem 3 with more technical conditions, where it is assumed that the constants mentioned above are known. Under these new conditions, we can explicitly determine a radius of convergence for the sequence p_k generated by (6). The proof of this new theorem follows a similar structure to the proof of Theorem 3 and will therefore be omitted.

Theorem 4 Let \bar{p} be a solution to the equation (2). We assume that the following conditions are satisfied:

- (i) $B_{\bar{\delta}}(\bar{p})$ is a totally normal ball, and the mapping $G: \mathcal{M} \rightrightarrows \mathbb{R}^m$ defined in (13) with $r = \bar{\delta}$ is metrically regular at \bar{p} for 0 with the constant $\tau > 0$ and neighborhoods $B_a[\bar{p}]$ and $B_b[0]$.
- (ii) $\iota \in (0, 1/\tau)$ is chosen such that $\|x\|_e < \iota d(p, \bar{p})$ holds for all $p \in B_{\bar{\delta}}(\bar{p}) \setminus \{\bar{p}\}$, $x \in R_k(p)$, $k \in \mathbb{N}_0$.
- (iii) $\mu > 0$, $\kappa > \tau$, $\epsilon > 0$ and $\Theta > 0$ satisfy

$$\mu\kappa < 1, \quad \kappa(\epsilon + \iota) < 1 - \mu\kappa, \quad \frac{\tau}{1 - \mu\tau} < \Theta < \frac{\kappa}{1 - \mu\kappa}.$$

- (iv) $\alpha < \min\{\bar{\delta}, a/2, b/\mu\}$ and

$$\begin{aligned} & \|\mathcal{D}f(p)[\exp_p^{-1} q - \exp_p^{-1} q'] - \mathcal{D}f(\bar{p})[\exp_{\bar{p}}^{-1} q - \exp_{\bar{p}}^{-1} q']\|_e \\ & \leq \mu d(q, q') \text{ for all } p, q, q' \in B_{\alpha}(\bar{p}). \end{aligned}$$

- (v) $\beta > 0$ is such that $\mu\alpha + 2\beta \leq b$ and $2\Theta\beta \leq \alpha$.
- (vi) $\delta_{\epsilon} > 0$ satisfies $\|f(p) - f(\bar{p}) - \mathcal{D}f(\bar{p})[\exp_{\bar{p}}^{-1} p]\|_e \leq \epsilon d(p, \bar{p})$ for all $p \in B_{\delta_{\epsilon}}(\bar{p})$.

Then, for every starting point $p_0 \in B_r(\bar{p})$ with $0 < r < \min\{\beta/(\epsilon + \iota), \bar{\delta}, \delta_{\epsilon}\}$, there exists a sequence $\{p_k\}$ generated by (6) that is well defined and converges linearly to \bar{p} .

Next, we modify the assumption on the multifunction R_k and introduce a stronger condition for the differentiability of the function f in order to establish quadratic convergence of the sequence (6).

Theorem 5 Let \bar{p} be a solution of (2). Suppose the following conditions hold:

- (i) the function $f: \mathcal{M} \rightarrow \mathbb{R}^m$ is continuously differentiable at \bar{p} , and $\text{reg}(f + F; \bar{p}|0) = \sigma$;
- (ii) $\mathcal{D}f$ is L -Lipschitz continuous around \bar{p} ;
- (iii) the sequence $\{R_k\}$ satisfies

$$\limsup_{\substack{p \rightarrow \bar{p} \\ p \neq \bar{p}}} \left\{ \frac{1}{d^2(p, \bar{p})} \sup_{k \in \mathbb{N}_0} \sup_{x \in R_k(p)} \|x\|_e \right\} < \frac{1}{\sigma}. \quad (41)$$

Then there exist $\theta > 0$ and a totally normal ball $B_r(\bar{p}) \subset \mathcal{M}$ such that for every $p \in B_r(\bar{p}) \setminus \{\bar{p}\}$, every $k \in \mathbb{N}_0$, and every $u_k \in R_k(p)$, there exists $q' \in B_r(\bar{p})$ satisfying

$$f(p) + \mathcal{D}f(p)[\exp_p^{-1} q'] + F(q') \ni u_k \quad (42)$$

and

$$d(q', \bar{p}) \leq \theta d^2(p, \bar{p}). \quad (43)$$

Consequently, for any starting point $p_0 \in B_r(\bar{p})$, there exists a sequence $\{p_k\}$ generated by (6) that converges quadratically to \bar{p} .

Proof This proof is analogous to the proof of Theorem 3. Using (41), we can find $\iota < 1/\sigma$ and a totally normal ball $B_{\bar{\delta}}(\bar{p})$ such that

$$\|x\|_e < \iota d^2(p, \bar{p}) \text{ whenever } p \in B_{\bar{\delta}}(\bar{p}) \setminus \{\bar{p}\}, \quad k \in \mathbb{N}_0, \quad x \in R_k(p). \quad (44)$$

Choose $\mu > 0$, $\kappa > \sigma$, $\Theta > 0$ and $\tau \in (\sigma, \kappa)$ satisfying

$$\mu\kappa < 1, \quad \frac{\sigma}{1 - \mu\sigma} < \Theta < \frac{\kappa}{1 - \mu\kappa}, \quad \frac{\tau}{1 - \mu\tau} < \Theta. \quad (45)$$

Note that $\mu\tau < 1$. From (i) and Theorem 1, we obtain $\text{reg}(G; \bar{p}|0) = \sigma$, where $G: \mathcal{M} \rightrightarrows \mathbb{R}^m$ is the function defined in (13) with $r = \delta$. Thus, there exist $a > 0$ and $b > 0$ such that G is metrically regular at \bar{p} for 0 with the constant τ and neighborhoods $\mathcal{B}_a[\bar{p}]$ and $\mathbb{B}_b[0]$. Now, consider the auxiliary functions $g_p: \mathcal{M} \rightarrow \mathbb{R}^m$, $p \in \mathcal{B}_\delta(\bar{p})$, defined in (34). Recall that $g_p(\bar{p}) = 0$ for all $p \in \mathcal{B}_\delta(\bar{p})$. Make $\delta > 0$ smaller, if necessary, to ensure

$$\|g_p(q) - g_p(q')\|_e \leq \mu d(q, q'), \quad p, q, q' \in \mathcal{B}_\delta[\bar{p}].$$

Hence, we can apply Lemma 2 with $g = g_p$, $F = G$, $\sigma = \tau$, $v = \mu$, $\sigma' = \Theta$, $x = \bar{x} = 0$, and $p = \bar{p}$. This yields the existence of $\beta > 0$ (independent of the point $p \in \mathcal{B}_\delta(\bar{p})$ that determines the function g_p) such that for each $x' \in \mathbb{B}_\beta[0]$, there exists $q' \in (g_p + G)^{-1}(x')$ such that

$$d(q', \bar{p}) \leq \Theta \|x'\|_e. \quad (46)$$

With (ii), the last part of Proposition 4 implies that there exists $\delta_L > 0$ such that

$$\|f(p) - f(\bar{p}) - \mathcal{D}f(\bar{p})[\exp_{\bar{p}}^{-1} p]\|_e \leq L d^2(p, \bar{p}), \quad p \in \mathcal{B}_{\delta_L}(\bar{p}). \quad (47)$$

Fixed r such that

$$0 < r < \min \left\{ \left(\frac{\beta}{L + \iota} \right)^{1/2}, \frac{1 - \mu\kappa}{\kappa(L + \iota)}, \delta, \delta_L \right\} \quad (48)$$

and $p \in \mathcal{B}_r(\bar{p}) \setminus \{\bar{p}\}$ fixed, by choosing $k \in \mathbb{N}_0$ and $u_k \in R_k(p)$ it follows from (44) that u_k satisfies $\|u_k\|_e \leq \iota d^2(p, \bar{p})$. Denote

$$y_k := f(p) - f(\bar{p}) + \mathcal{D}f(p)[\exp_p^{-1} \bar{p}] - u_k. \quad (49)$$

If $y_k = 0$, then $q' := \bar{p}$ satisfies (42) because $-f(\bar{p}) \in F(\bar{p})$ and (43) hold trivially. Assume that $y_k \neq 0$. Using (47) and (49), we obtain

$$\|y_k\|_e \leq \|f(p) - f(\bar{p}) + \mathcal{D}f(p)[\exp_p^{-1} \bar{p}]\|_e + \|u_k\|_e \leq (L + \iota) d^2(p, \bar{p}). \quad (50)$$

Since $d(p, \bar{p}) < r < (\beta/(L + \iota))^{1/2}$, it follows from (50) that $\|y_k\|_e < \beta$. Applying (46) with $x' = -y_k$, we obtain that there exists $q' \in (g_p + G)^{-1}(-y_k)$ such that $d(q', \bar{p}) \leq \Theta \|y_k\|_e$. Hence, utilizing the upper bound for Θ given in (45) and (50), it follows that

$$d(q', \bar{p}) \leq \frac{\kappa}{1 - \mu\kappa} \|y_k\|_e \leq \theta d^2(p, \bar{p}) \quad \text{where} \quad \theta := \frac{\kappa(L + \iota)}{1 - \mu\kappa}.$$

Furthermore, it comes from $-y_k \in (g_p + G)(q')$, (49), (34) and (13) that

$$-f(p) + f(\bar{p}) - \mathcal{D}f(p)[\exp_p^{-1} \bar{p}] + u_k \in f(\bar{p}) + \mathcal{D}f(p)[\exp_p^{-1} q' - \exp_p^{-1} \bar{p}] + F(q'),$$

which means that $u_k \in f(p) + \mathcal{D}f(p)[\exp_p^{-1} q'] + F(q')$. Thus, q' satisfies (42) and (43).

To finish the proof, choose any $p_0 \in \mathcal{B}_r(\bar{p})$. If $p_0 = \bar{p}$, then $p_1 := \bar{p}$ verifies (6) because $0 \in R_0(\bar{p})$. If $p_0 \neq \bar{p}$, applying (42) and (43) we obtain that for every $u_0 \in R_0(p_0)$ there exists p_1 such that

$$f(p_0) + \mathcal{D}f(p_0)[\exp_{p_0}^{-1} p_1] + F(p_1) \ni u_0 \quad \text{and} \quad d(p_1, \bar{p}) \leq \theta d^2(p_0, \bar{p}).$$

By considering the definition of r and θ , we obtain from the above inequality that $p_1 \in B_r(\bar{p})$. Repeating the previous argument it is possible to construct a sequence $\{p_k\}$ in $B_r(\bar{p})$ that satisfies (6) and converges quadratically to \bar{p} . \square

Under suitable conditions, it is possible to determine the radius r mentioned in Theorem 5. Details are given in the following result. The proof of this result is along the same lines as the proof of Theorem 5 and will therefore be omitted.

Theorem 6 *Let \bar{p} be a solution of (2). Suppose that the following conditions hold:*

- (i) $B_\delta(\bar{p})$ is a totally normal ball, and the mapping $G: \mathcal{M} \rightrightarrows \mathbb{R}^m$ defined in (13) with $r = \delta$ is metrically regular at \bar{p} for 0 with the constant $\tau > 0$ and neighborhoods $B_a[\bar{p}]$ and $B_b[0]$.
- (ii) $\iota \in (0, 1/\tau)$ satisfies $\|x\|_e < \iota d^2(p, \bar{p})$ for all $p \in B_\delta(\bar{p}) \setminus \{\bar{p}\}$, $x \in R_k(p)$, $k \in \mathbb{N}_0$.
- (iii) $\mu > 0$, $\kappa > \tau$ and $\Theta > 0$ satisfy

$$\mu\kappa < 1, \quad \frac{\tau}{1 - \mu\tau} < \Theta < \frac{\kappa}{1 - \mu\kappa}.$$

- (iv) $\alpha < \min\{\delta, a/2, b/\mu\}$ and

$$\begin{aligned} & \|\mathcal{D}f(p)[\exp_p^{-1}q - \exp_p^{-1}q'] - \mathcal{D}f(\bar{p})[\exp_{\bar{p}}^{-1}q - \exp_{\bar{p}}^{-1}q']\|_e \\ & \leq \mu d(q, q') \quad \text{for all } p, q, q' \in B_\alpha(\bar{p}). \end{aligned}$$

- (v) $\beta > 0$ satisfies $\mu\alpha + 2\beta \leq b$ and $2\Theta\beta \leq \alpha$.
- (vi) $L > 0$ and $\delta_L > 0$ satisfy $\|f(p) - f(\bar{p}) - \mathcal{D}f(\bar{p})[\exp_{\bar{p}}^{-1}p]\|_e \leq Ld^2(p, \bar{p})$ for all $p \in B_{\delta_L}(\bar{p})$.

Then for every r satisfying (48) and starting point $p_0 \in B_r(\bar{p})$ there exists a sequence $\{p_k\}$ generated by (6) that is well defined and converges linearly to \bar{p} .

Some comments about the previous results are in order. First, Theorems 3 and 5 establish conditions for ensuring that the inexact Newton method (6) converges with linear and quadratic rates, respectively. Second, a similar result to Theorem 3 is presented in [17, Theorem 2.1] by considering \mathcal{M} as Banach spaces. Thus, if $\mathcal{M} = X$, where X is a Banach space, and the derivative $\mathcal{D}f(\cdot)$ is replaced by a suitable approximation, then Theorems 3 and [17, Theorem 2.1] are equivalent. Third, to ensure superlinear and quadratic convergence, the authors in [17] assume a stronger condition, namely, the strong metric regularity (see [17, Theorem 2.3]). We recall that a set-valued mapping F is strongly metrically regular at \bar{x} for \bar{y} if and only if its inverse F^{-1} has a single-valued graphical localization around \bar{y} for \bar{x} which is Lipschitz continuous around \bar{y} with Lipschitz modulus at \bar{y} equal to $\text{reg}(F; \bar{x}|\bar{y})$, see [17, p. 1007]. Fourth, we establish quadratic convergence of (6) by only assuming the metric regularity condition, which is clearly a weaker assumption than strong metric regularity, see the previous comment. Thus, the result obtained in Theorem 5 is stronger than [17, Theorem 2.3]. Finally, Theorems 4 and 6 refine Theorems 3 and 5, respectively, in the sense that they provide guidance on how to find the neighborhood to start the proposed method in (6).

It is also important to mention that in [17] is introduced a version of the Dennis-Moré theorem for the sequence generated by (6). We do not go further in this topic because we do not propose a quasi-Newton method in this paper.

We conclude this section by presenting a semi-local convergence result for the inexact Newton method proposed in this paper to solve (2). This result makes no assumptions about the unknown solution to the problem under investigation; instead, the assumptions are made about the starting point p_0 . It is worth mentioning that this result is novel, even in the Euclidean context.

Theorem 7 Assume that for $(p_0, y_0) \in \Omega \times \mathbb{R}^m$ with $y_0 \in f(p_0) + F(p_0)$ and $u_0 \in R_0(p_0)$, the following conditions hold:

(i) $B_\delta(p_0) \subset \Omega$ is a totally normal ball, and the mapping $G: \mathcal{M} \rightrightarrows \mathbb{R}^m$ defined by

$$G(p) = \begin{cases} f(p_0) + \mathcal{D}f(p_0)[\exp_{p_0}^{-1} p] + F(p), & p \in B_\delta(p_0) \\ f(p) + F(p), & p \in \mathcal{M} \setminus B_\delta(p_0) \end{cases} \quad (51)$$

is metrically regular at p_0 for y_0 with the constant $\sigma > 0$ and neighborhoods $\mathcal{B}_a[p_0]$ and $\mathbb{B}_b[y_0]$.

(ii) the positive constants $\mu, \alpha, \beta, \Theta, \epsilon$ and ι satisfy

$$\begin{aligned} \mu\sigma < 1, \quad \alpha &\leq a/2, \quad \mu\alpha + 2\beta \leq b, \\ \sigma/(1 - \mu\sigma) < \Theta &\leq \alpha/(2\beta), \quad \epsilon + \iota < 2\beta/\alpha. \end{aligned}$$

(iii) $\|u_0\|_e \leq \iota\|y_0\|_e$ and for all $p \in B_\delta(p_0)$ there holds

$$\|y_0\|_e \leq \min \left\{ \frac{\beta}{\Theta(1 + \iota)}, \frac{\beta}{1 + \iota}, \frac{\beta(1 - \hat{\alpha})}{\hat{\alpha} - \hat{\alpha}^2 + 1}, \frac{\delta(1 - \hat{\alpha})}{\Theta(1 + \iota)} \right\}, \quad \hat{\alpha} := \Theta(\epsilon + \iota). \quad (52)$$

(iv) $\|x\|_e + \|y\|_e \leq \iota d(\tilde{p}, \tilde{q})$ holds for all $\tilde{p}, \tilde{q} \in B_\delta(p_0)$, $x \in R_k(\tilde{p})$, $y \in R_{k-1}(\tilde{q})$, $k \in \mathbb{N}$ and $\tilde{p} \neq \tilde{q}$.

(v) $f: \mathcal{M} \rightarrow \mathbb{R}^m$ is continuously differentiable, and the inequalities

$$\|f(q) - f(p) - \mathcal{D}f(p)[\exp_p^{-1} q]\|_e \leq \epsilon d(p, q)$$

and

$$\|\mathcal{D}f(p)[\exp_p^{-1} q - \exp_p^{-1} q'] - \mathcal{D}f(p_0)[\exp_{p_0}^{-1} q - \exp_{p_0}^{-1} q']\|_e \leq \mu d(q, q')$$

hold for all $p, q, q' \in B_\delta(p_0)$.

Then there exist sequences $\{p_k\}_{k \in \mathbb{N}}$ and $\{u_k\}_{k \in \mathbb{N}}$ satisfying:

$$(A1) \quad d(p_k, p_0) \leq \frac{1 - \hat{\alpha}^k}{1 - \hat{\alpha}} \Theta(1 + \iota) \|y_0\|_e.$$

$$(A2) \quad d(p_k, p_{k-1}) \leq \hat{\alpha}^{k-1} \Theta(1 + \iota) \|y_0\|_e.$$

$$(A3) \quad u_{k-1} \in f(p_{k-1}) + \mathcal{D}f(p_{k-1})[\exp_{p_{k-1}}^{-1} p_k] + F(p_k).$$

In particular, $\{p_k\}_{k \in \mathbb{N}}$ remains in $B_\delta(p_0)$, converges to a solution \bar{p} of (2), and satisfies

$$d(p_k, \bar{p}) \leq \frac{\hat{\alpha}^k}{1 - \hat{\alpha}} \Theta(1 + \iota) \|y_0\|_e \text{ for all } k \in \mathbb{N}.$$

Proof Note that the function $g_p: \mathcal{M} \rightarrow \mathbb{R}^m$, $p \in B_\delta(p_0)$, defined by

$$g_p(q) = \begin{cases} \mathcal{D}f(p)[\exp_p^{-1} q - \exp_p^{-1} p_0] - \mathcal{D}f(p_0)[\exp_{p_0}^{-1} q], & q \in B_\delta(p_0) \\ 0, & q \in \mathcal{M} \setminus B_\delta(p_0), \end{cases} \quad (53)$$

satisfies $g_p(p_0) = 0$. From (53) and the second inequality in (v), it follows that

$$\begin{aligned} \|g_p(q) - g_p(q')\|_e &= \|\mathcal{D}f(p)[\exp_p^{-1} q - \exp_p^{-1} q'] - \mathcal{D}f(p_0)[\exp_{p_0}^{-1} q \\ &\quad - \exp_{p_0}^{-1} q']\|_e \leq \mu d(q, q'), \end{aligned}$$

for all $p, q, q' \in B_\delta(p_0)$. Thus, by using (ii), one has

$$\alpha \leq a/2, \quad \mu\alpha + 2\beta \leq b, \quad 2\Theta\beta \leq \alpha,$$

and

$$\|g_p(p_0)\|_e < \beta, \quad \|g_p(q) - g_p(q')\|_e \leq \mu d(q, q'), \quad p, q, q' \in \mathcal{B}_\alpha(p_0).$$

Hence, we can apply Lemma 2 with $g = g_p$, $F = G$, $v = \mu$, $\sigma' = \Theta$, $\bar{x} = y_0$ and $\bar{p} = p_0$, to obtain the following statement:

(S) for every $x, x' \in \mathbb{B}_\beta[y_0]$ and every $\tilde{q} \in (g_p + G)^{-1}(x) \cap \mathcal{B}_\alpha[p_0]$ there exists $\tilde{q}' \in (g_p + G)^{-1}(x')$ such that $d(\tilde{q}, \tilde{q}') \leq \Theta \|x - x'\|_e$.

Due to (iv), for every $k \in \mathbb{N}$, the following can be stated:

$$p_k \neq p_{k-1} \in \mathcal{B}_\delta(p_0), u_k \in R_k(p_k) \implies \|u_k\|_e + \|u_{k-1}\|_e \leq \iota d(p_k, p_{k-1}). \quad (54)$$

We are now able to prove (A1)–(A3) by induction. By using statement (S) with $x = y_0$, $\tilde{q} = p = p_0$, and $x' = u_0$, we conclude that there exists $p_1 := \tilde{q}' \in (g_{p_0} + G)^{-1}(u_0)$ such that

$$d(p_1, p_0) \leq \Theta \|y_0 - u_0\|_e \leq \Theta(1 + \iota) \|y_0\|_e,$$

where the last inequality follows from (iii). Moreover, the inclusion $p_1 \in (g_{p_0} + G)^{-1}(u_0)$ is equivalent to

$$u_0 \in f(p_0) + \mathcal{D}f(p_0)[\exp_{p_0}^{-1} p_1] + F(p_1).$$

Hence, (A1)–(A3) hold for $k = 1$. For $k > 1$, assume the induction hypothesis: there exist $p_j \in \mathcal{B}_\delta(p_0)$ and $u_j \in R_j(p_j)$ such that (A1)–(A3) hold for every $j \in \{1, 2, \dots, k\}$. Denote

$$z_k := f(p_0) - f(p_k) - \mathcal{D}f(p_k)[\exp_{p_k}^{-1} p_0] + u_k, \quad (55)$$

$$w_k := f(p_0) - f(p_{k-1}) - \mathcal{D}f(p_k)[\exp_{p_k}^{-1} p_0] - \mathcal{D}f(p_{k-1})[\exp_{p_{k-1}}^{-1} p_k] + u_{k-1}. \quad (56)$$

By (A3), (53), (51) we find $p_k \in (g_{p_k} + G)^{-1}(w_k)$. On the other hand, the first inequality in (v), (54), (A1), (A2) and (52) yield

$$\begin{aligned} \|z_k - y_0\|_e &\leq \|f(p_0) - f(p_k) - \mathcal{D}f(p_k)[\exp_{p_k}^{-1} p_0]\|_e + \|u_k\|_e + \|y_0\|_e \\ &\leq \epsilon d(p_k, p_0) + \iota d(p_k, p_{k-1}) + \|y_0\|_e \\ &\leq \epsilon d(p_k, p_0) + (\epsilon + \iota) d(p_k, p_{k-1}) + \|y_0\|_e \end{aligned}$$

and

$$\begin{aligned} \|w_k - y_0\|_e &\leq \|f(p_0) - f(p_k) - \mathcal{D}f(p_k)[\exp_{p_k}^{-1} p_0]\|_e \\ &\quad + \|f(p_k) - f(p_{k-1}) - \mathcal{D}f(p_{k-1})[\exp_{p_{k-1}}^{-1} p_k]\|_e + \|u_{k-1}\|_e + \|y_0\|_e \\ &\leq \epsilon d(p_k, p_0) + (\epsilon + \iota) d(p_k, p_{k-1}) + \|y_0\|_e \end{aligned}$$

and

$$\begin{aligned} \epsilon d(p_k, p_0) + (\epsilon + \iota) d(p_k, p_{k-1}) + \|y_0\|_e &\leq \frac{\epsilon \Theta}{1 - \hat{\alpha}} \|y_0\| + (\epsilon + \iota) \Theta \|y_0\| + \|y_0\|_e \\ &\leq \frac{(\epsilon + \iota) \Theta}{1 - \hat{\alpha}} \|y_0\| + (\epsilon + \iota) \Theta \|y_0\| + \|y_0\|_e \\ &= \frac{\hat{\alpha} - \hat{\alpha}^2 + 1}{1 - \hat{\alpha}} \|y_0\|_e \leq \beta, \end{aligned}$$

which implies $w_k, z_k \in \mathbb{B}_\beta[y_0]$. Therefore, we can apply (S) with $x' = z_k$, $\tilde{q} = p = p_k$ and $x = w_k$ to conclude that there exists $q' \in (g_{p_k} + G)^{-1}(z_k)$ such that $d(q', p_k) \leq \Theta \|z_k - w_k\|_e$.

Putting $q' = p_{k+1}$ in this inequality and taking into account (55), (56), the first inequality in (v), (54) and (A2), we get

$$\begin{aligned} d(p_{k+1}, p_k) &\leq \Theta \|z_k - w_k\|_e \\ &\leq \Theta \left(\| -f(p_k) + f(p_{k-1}) + \mathcal{D}f(p_{k-1})[\exp_{p_{k-1}}^{-1} p_k] \|_e + \|u_k\|_e + \|u_{k-1}\|_e \right) \\ &\leq \Theta(\epsilon + \iota) d(p_k, p_{k-1}) = \hat{\alpha} d(p_k, p_{k-1}) \leq \hat{\alpha}^k \Theta \|y_0\|_e \leq \hat{\alpha}^k \Theta(1 + \iota) \|y_0\|_e. \end{aligned}$$

In view of this and (A1), we also have

$$\begin{aligned} d(p_{k+1}, p_0) &\leq d(p_{k+1}, p_k) + d(p_k, p_0) \\ &\leq \hat{\alpha}^k \Theta(1 + \iota) \|y_0\|_e + \frac{1 - \hat{\alpha}^k}{1 - \hat{\alpha}} \Theta(1 + \iota) \|y_0\|_e = \frac{1 - \hat{\alpha}^{k+1}}{1 - \hat{\alpha}} \Theta(1 + \iota) \|y_0\|_e. \end{aligned}$$

Furthermore, it comes from $p_{k+1} \in (g_{p_k} + G)^{-1}(z_k)$ that

$$u_k \in f(p_k) + \mathcal{D}f(p_k)[\exp_{p_k}^{-1} p_{k+1}] + F(p_{k+1}).$$

Overall, we can conclude that (A1), (A2) and (A3) hold for all $k \geq 1$.

Now our focus is to show the convergence of $\{p_k\}$. Using (A1) and (A2) in a simple induction procedure, we can show that

$$d(p_m, p_n) \leq \hat{\alpha}^m \frac{1 - \hat{\alpha}^{n-m}}{1 - \hat{\alpha}} \Theta(1 + \iota) \|y_0\|_e$$

holds for all $m < n$, $m, n \in \mathbb{N}$. Hence, since $\hat{\alpha} \in (0, 1)$, $\{p_k\}$ is a Cauchy sequence, and consequently, there exists $\bar{p} \in \mathcal{M}$ such that $\lim_{k \rightarrow \infty} p_k = \bar{p}$. From (A1) and (52), it follows that $p_k \in B_\delta(p_0)$ for all \mathbb{N} , which implies that $\bar{p} \in B_\delta[p_0] \subset \Omega$. As f is continuous and F has closed graph, by letting $k \rightarrow +\infty$ in (54) and (A3), we conclude that \bar{p} is a solution of (2). Finally, using (A2), we get

$$\begin{aligned} d(p_k, \bar{p}) &= \lim_{m \rightarrow \infty} d(p_k, p_{k+m}) \leq \lim_{m \rightarrow \infty} \sum_{i=k}^{k-1+m} d(p_i, p_{i+1}) \\ &\leq \lim_{m \rightarrow \infty} \sum_{i=k}^{k-1+m} \hat{\alpha}^i \Theta(1 + \iota) \|y_0\|_e = \lim_{m \rightarrow \infty} \frac{\hat{\alpha}^k (1 - \hat{\alpha}^m)}{1 - \hat{\alpha}} \Theta(1 + \iota) \|y_0\|_e \\ &= \frac{\hat{\alpha}^k}{1 - \hat{\alpha}} \Theta(1 + \iota) \|y_0\|_e, \end{aligned}$$

for all $k \in \mathbb{N}$. □

We conclude this section with some remarks about Theorem 7. First, it is a result to guarantee the existence of a neighborhood which the inexact Newton's method is well-defined and, hence its convergence to a solution of (2). Thus, in practice, it can be hard to evaluate all the parameters in items (i) – (v). Second, it states that if the initial point $p_0 \in \mathcal{M}$ satisfies the inclusion $y_0 \in (f + F)(p_0)$ where $y_0 \in B_\chi(0)$ and

$$0 < \chi := \min \left\{ \frac{\beta}{\Theta(1 + \iota)}, \frac{\beta}{1 + \iota}, \frac{\beta(1 - \hat{\alpha})}{\hat{\alpha} - \hat{\alpha}^2 + 1}, \frac{\delta(1 - \hat{\alpha})}{\Theta} \right\}$$

then the inexact Newton's method finds a solution of (2) in $B_{\tilde{\chi}}(p_0)$, where

$$\tilde{\chi} = \frac{\Theta(1 + \iota)}{1 - \hat{\alpha}} \|y_0\|_e.$$

We can interpret this remark as follows: although we cannot evaluate χ in practice, if we choose y_0 close to 0 then the proposed method will work. Third, the assumption $\|u_0\| \leq \epsilon \|y_0\|_e$ in (iii) ensures that the first call (and the subsequent calls) to the inexact Newton's method for solving (2) address the inexactness. Fourth, conditions in (i)-(ii) are usual in the context of metric regularity, and both inequalities in (v) are related to the smoothness assumption of f and the continuity of the exponential map. Thus, if ϵ and η are small enough, this assumption holds.

5 Application

In this section, we investigate three well-known problems that can be formulated as generalized equations on Riemannian manifolds.

Example 5 (System of Inequalities and Equalities) Consider the generalized equation (2) with $F \equiv -K$, where $K \subset \mathbb{R}^m$ is the fixed cone

$$\mathbb{R}_-^s \times \{0\}^{m-s} := \{x \in \mathbb{R}^m : x_i \leq 0 \text{ for } i = 1, \dots, s \text{ and } x_i = 0 \text{ for } i = s+1, \dots, m\}.$$

It is easy to see that this generalized equation is equivalent to the following system of equalities and inequalities:

$$\begin{aligned} f_i(\bar{p}) &\leq 0 & \text{for } i = 1, \dots, s, \\ f_i(\bar{p}) &= 0 & \text{for } i = s+1, \dots, m. \end{aligned} \quad (57)$$

Let \bar{p} solve (57) and let each f_i be continuously differentiable around \bar{p} for all $i = 1, \dots, m$. As defined in [10, Definition 3.12], the Mangasarian–Fromovitz condition for system (57) is as follows:

$$\exists v \in \mathcal{T}_{\bar{p}}\mathcal{M} \text{ such that } \begin{cases} \langle \text{grad } f_i(\bar{p}), v \rangle < 0 & \text{for } i \in \{1, \dots, s\} \text{ with } f_i(\bar{p}) = 0, \\ \langle \text{grad } f_i(\bar{p}), v \rangle = 0 & \text{for } i \in \{s+1, \dots, m\}. \end{cases} \quad (58)$$

After making simple adaptations of [25, Example 4D.3] to the Riemannian context, we can assert that condition (58) is equivalent to the Aubin property of $(-f + K)^{-1}$ at \bar{x} for \bar{p} , which, in turn, by Theorem 2, is equivalent to the metric regularity of $-f + K$ at \bar{p} for \bar{x} .

The following proposition serves as a prerequisite for discussing the last two examples in this section. This result establishes an equivalence between problems (2) and (3).

Proposition 2 Let $\Omega \subseteq \mathcal{M}$ and let $\{E_1, \dots, E_n\}$ be a basis for $\mathcal{X}(\Omega)$. Suppose $V : \Omega \rightarrow \mathcal{TM}$ is a single-valued vector field, and $Z : \Omega \rightrightarrows \mathcal{TM}$ is a set-valued vector field. Then, a point \bar{p} is a solution to (3) if and only if it is a solution to the generalized equation (2) with $f : \Omega \rightarrow \mathbb{R}^n$ and $F : \Omega \rightrightarrows \mathbb{R}^n$ defined, respectively, by

$$f(p) := (\langle V(p), E_1(p) \rangle, \dots, \langle V(p), E_n(p) \rangle) \text{ and } F(p) := \bigcup_{v \in Z(p)} (\langle v, E_1(p) \rangle, \dots, \langle v, E_n(p) \rangle). \quad (59)$$

Proof If \bar{p} is a solution of (3), then there exists $\bar{v} \in Z(\bar{p})$ such that $V(\bar{p}) + \bar{v} = 0_{\bar{p}}$. Consequently,

$$\langle V(\bar{p}), E_i(\bar{p}) \rangle + \langle \bar{v}, E_i(\bar{p}) \rangle = \langle V(\bar{p}) + \bar{v}, E_i(\bar{p}) \rangle = 0 \text{ for all } i = 1, \dots, n.$$

Using the definitions of f and F in (59), we find that \bar{p} satisfies (2). Conversely, if \bar{p} is a solution to (2) with f and F defined in (59), then there exists $\bar{v} \in Z(\bar{p})$ such that

$$(\langle V(\bar{p}), E_1(\bar{p}) \rangle, \dots, \langle V(\bar{p}), E_n(\bar{p}) \rangle) + (\langle \bar{v}, E_1(\bar{p}) \rangle, \dots, \langle \bar{v}, E_n(\bar{p}) \rangle) = 0.$$

Since $\{E_1(\bar{p}), \dots, E_n(\bar{p})\}$ forms a basis for $\mathcal{T}_{\bar{p}}\mathcal{M}$, it follows that $\langle V(\bar{p}) + \bar{v}, v \rangle = 0$ for all $v \in \mathcal{T}_{\bar{p}}\mathcal{M}$. Thus, $0_{\bar{p}} = V(\bar{p}) + \bar{v}$, implying that \bar{p} is a solution to (3). \square

In the following example, we discuss the variational inequality problem proposed in [45], which extends the problem introduced in [47].

Example 6 (Variational Inequality Problem) Let $V: \Omega \subseteq \mathcal{M} \rightarrow \mathcal{TM}$ be a single-valued vector field. Consider the variational inequality problem

$$p \in \Omega, \quad \langle V(p), \dot{\gamma}(0) \rangle \geq 0 \quad \text{for all } \gamma \in \Gamma_{p,q}^\Omega, \quad (60)$$

where $\Gamma_{p,q}^\Omega$ is the set of all geodesics $\gamma: [0, 1] \rightarrow \mathcal{M}$ satisfying $\gamma(0) = p$, $\gamma(1) = q$, and $\gamma(t) \in \Omega$ for all $t \in [0, 1]$. Since the normal cone associated with Ω is the set-valued vector field $N_\Omega: \Omega \rightrightarrows \mathcal{TM}$ defined by

$$N_\Omega(p) := \left\{ v \in \mathcal{T}_p\mathcal{M} : \langle v, \dot{\gamma}(0) \rangle \leq 0 \text{ for each } \gamma \in \Gamma_{p,q}^\Omega \right\},$$

for all $p \in \Omega$ (see [44]), the problem (60) is equivalent to

$$p \in \Omega, \quad V(p) + N_\Omega(p) \ni 0_p,$$

which, in turn, by Proposition 2, is equivalent to generalized equation (2) with $f: \Omega \rightarrow \mathbb{R}^n$ and $F: \Omega \rightrightarrows \mathbb{R}^n$ as in (59), with $Z = N_\Omega$ and $\{E_1, \dots, E_n\}$ being any basis for $\mathcal{X}(\Omega)$.

The following example proposes an approach based on the Riemannian extension of the analysis presented in [37].

Example 7 (KKT Conditions) Consider the constrained nonlinear optimization problem on \mathcal{M} :

$$\text{minimize } \mathbf{f}(p) \quad (61)$$

$$\text{subject to } \mathbf{g}(p) \leq 0, \quad \mathbf{h}(p) = 0, \quad (62)$$

where $\mathbf{f}: \mathcal{M} \rightarrow \mathbb{R}$, $\mathbf{g} := (\mathbf{g}_1, \dots, \mathbf{g}_{m_1}): \mathcal{M} \rightarrow \mathbb{R}^{m_1}$, and $\mathbf{h} := (\mathbf{h}_1, \dots, \mathbf{h}_{m_2}): \mathcal{M} \rightarrow \mathbb{R}^{m_2}$ are assumed to be continuously differentiable functions. The constraint $\mathbf{g}(p) \leq 0$ means that $\mathbf{g}_i(p) \leq 0$ for all $i = 1, \dots, m_1$. The Lagrangian function $\mathcal{L}: \mathcal{M} \times \mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \rightarrow \mathbb{R}$ is given by

$$\mathcal{L}(p, \mu, \lambda) := \mathbf{f}(p) + \sum_{i=1}^{m_1} \mu_i \mathbf{g}_i(p) + \sum_{i=1}^{m_2} \lambda_i \mathbf{h}_i(p),$$

where $\mu := (\mu_1, \dots, \mu_{m_1}) \in \mathbb{R}^{m_1}$ and $\lambda := (\lambda_1, \dots, \lambda_{m_2}) \in \mathbb{R}^{m_2}$. For each $(\mu, \lambda) \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$, consider the function $\mathcal{L}_{\mu,\lambda}: \mathcal{M} \rightarrow \mathbb{R}$ defined by $\mathcal{L}_{\mu,\lambda}(p) = \mathcal{L}(p, \mu, \lambda)$ for all $p \in \mathcal{M}$. Based on [10], we assert that the KKT conditions for (61) and (62) are:

$$\text{grad } \mathcal{L}_{\mu,\lambda}(p) = \text{grad } \mathbf{f}(p) + \sum_{i=1}^{m_1} \mu_i \text{grad } \mathbf{g}_i(p) + \sum_{i=1}^{m_2} \lambda_i \text{grad } \mathbf{h}_i(p) = 0_p, \quad (63)$$

$$\mu \geq 0, \quad \mathbf{g}(p) \leq 0, \quad \sum_{i=1}^{m_1} \mu_i \mathbf{g}_i(p) = 0, \quad (64)$$

$$\mathbf{h}(p) = 0. \quad (65)$$

Let $\widetilde{\mathcal{M}} := \mathcal{M} \times \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$ and consider the vector field $V: \widetilde{\mathcal{M}} \rightarrow \mathcal{T}\widetilde{\mathcal{M}} \equiv \mathcal{T}\mathcal{M} \times \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$ defined by

$$V(p, \mu, \lambda) = (\text{grad } \mathcal{L}_{\mu, \lambda}(p), \mathbf{g}(p), \mathbf{h}(p)),$$

and the set-valued vector field $Z: \widetilde{\mathcal{M}} \rightrightarrows \mathcal{T}\widetilde{\mathcal{M}}$ defined by

$$Z(p, \mu, \lambda) = \begin{cases} \{0_p\} \times \{y \in \mathbb{R}_+^{m_1} : \sum_{i=1}^{m_1} \mu_i y_i = 0\} \times \{0\}, & \text{if } \mu \geq 0; \\ \emptyset, & \text{otherwise,} \end{cases}$$

where $\mathbb{R}_+^{m_1}$ denotes the set of vectors in \mathbb{R}^{m_1} with nonnegative coordinates. Note that KKT system (63), (64) and (65) is equivalent to the problem

$$(p, \mu, \lambda) \in \widetilde{\mathcal{M}}, \quad V(p, \mu, \lambda) + Z(p, \mu, \lambda) \ni (0_p, 0, 0). \quad (66)$$

Let $\{E_1, \dots, E_n\}$ be a basis for $\mathcal{X}(\mathcal{M})$ and $\{E_{n+1}, \dots, E_{n+m_1+m_2}\}$ be the canonical basis for $\mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$. Then, $\{E_1, \dots, E_{n+m_1+m_2}\}$ forms a basis for $\mathcal{X}(\widetilde{\mathcal{M}})$ and, by Proposition 2, (66) is equivalent to the generalized equation

$$(p, \mu, \lambda) \in \widetilde{\mathcal{M}}, \quad f(p, \mu, \lambda) + F(p, \mu, \lambda) \ni 0, \quad (67)$$

where $f: \widetilde{\mathcal{M}} \rightarrow \mathbb{R}^n \times \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$ and $F: \widetilde{\mathcal{M}} \rightrightarrows \mathbb{R}^n \times \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$ are defined, respectively, by

$$(f(p, \mu, \lambda))_i = \begin{cases} \langle \text{grad } \mathcal{L}_{\mu, \lambda}(p), E_i(p) \rangle, & \text{if } i = 1, \dots, n; \\ (\mathbf{g}(p))_{i-n}, & \text{if } i = n+1, \dots, n+m_1; \\ (\mathbf{h}(p))_{i-n-m_1}, & \text{if } i = n+m_1+1, \dots, n+m_1+m_2; \end{cases} \quad (68)$$

and

$$F(p, \mu, \lambda) = \begin{cases} \{0\} \times \{y \in \mathbb{R}_+^{m_1} : \sum_{i=1}^{m_1} \mu_i y_i = 0\} \times \{0\}, & \text{if } \mu \geq 0; \\ \emptyset, & \text{otherwise.} \end{cases} \quad (69)$$

The following result provides a function ζ associated with the generalized equation discussed in the previous example, which satisfies the assumptions of Proposition 1.

Proposition 3 *Let f and F be as defined in (68) and (69), respectively. If f is locally Lipschitz, then $\zeta: \widetilde{\mathcal{M}} \rightarrow \mathbb{R}^n \times \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$, defined by*

$$(\zeta(p, \mu, \lambda))_i = \begin{cases} \langle \text{grad } \mathcal{L}_{\mu, \lambda}(p), E_i(p) \rangle, & \text{if } i = 1, \dots, n, \\ \min \{\mu_{i-n}, -(\mathbf{g}(p))_{i-n}\}, & \text{if } i = n+1, \dots, n+m_1, \\ (\mathbf{h}(p))_{i-n-m_1}, & \text{if } i = n+m_1+1, \dots, n+m_1+m_2, \end{cases} \quad (70)$$

is locally Lipschitz. Moreover, $\|\zeta(p, \mu, \lambda)\|_e = 0$ if and only if $(p, \mu, \lambda) \in (f + F)^{-1}(0)$.

Proof By utilizing (70) and (68), it can be inferred that

$$\begin{aligned} & \|\zeta(p, \mu, \lambda) - \zeta(p', \mu', \lambda')\|_e^2 \\ &= \|f(p, \mu, \lambda) - f(p', \mu', \lambda')\|_e^2 - \|\mathbf{g}(p) - \mathbf{g}(p')\|_e^2 \\ & \quad + \sum_{j=1}^{m_1} |\min \{\mu_j, -(\mathbf{g}(p))_j\} - \min \{\mu'_j, -(\mathbf{g}(p'))_j\}|^2 \\ & \leq \|f(p, \mu, \lambda) - f(p', \mu', \lambda')\|_e^2 - \|\mathbf{g}(p) - \mathbf{g}(p')\|_e^2 \end{aligned}$$

$$\begin{aligned}
& + \sum_{j=1}^{m_1} \left(\max \left\{ |\mu_j - \mu'_j|, |(\mathbf{g}(p))_j - (\mathbf{g}(p'))_j| \right\} \right)^2 \\
& \leq \|f(p, \mu, \lambda) - f(p', \mu', \lambda')\|_e^2 + \|\mu - \mu'\|_e^2 \\
& \leq \|f(p, \mu, \lambda) - f(p', \mu', \lambda')\|_e^2 + d^2((p, \mu, \lambda), (p', \mu', \lambda'))
\end{aligned}$$

for all $(p, \mu, \lambda), (p', \mu', \lambda') \in \tilde{\mathcal{M}}$. Since f is locally Lipschitz, the previous inequality implies that ζ is also locally Lipschitz. The last part of the result follows directly from the discussion in Example 7 and from (70). \square

6 Numerical Example

In this section, we present a numerical example based on a generalized equation derived from Example 7, and solve it using the inexact Newton method described in (5). All computations were performed on a MacBook Pro running macOS Sonoma 14.5, equipped with 16 GB RAM, an Apple M1 Pro CPU, and MATLAB R2022a. To ensure reproducibility, we fixed the randomness using MATLAB's built-in function *rng*(2024).

Here, we consider the Riemannian constrained optimization problem on \mathcal{M} :

$$\text{minimize } \mathbf{f}(p) := \frac{1}{N} \sum_{i=1}^N d^2(p, p^i) \quad (71)$$

$$\text{subject to } \mathbf{g}(p) := d^2(p, \tilde{p}) - r^2 \leq 0, \quad (72)$$

where $r > 0$ and $p^1, \dots, p^N, \tilde{p} \in \mathcal{M}$ are chosen such that $r < r_{\text{inj}}(\tilde{p})$ and $p^1, \dots, p^N \in \mathcal{B}_r(\tilde{p})$. We will use the fact that

$$\text{grad } \mathbf{f}(p) = -\frac{2}{N} \sum_{i=1}^N \exp_p^{-1} p^i \quad \text{and} \quad \text{grad } \mathbf{g}(p) = -2 \exp_p^{-1} \tilde{p}, \quad \forall p \in \mathcal{B}_r(\tilde{p}). \quad (73)$$

The problem defined by (71) and (72) represents a constrained version of the Riemannian center of mass, also known as the (Riemannian) mean, as proposed in [10]. This problem was first introduced in [38] and has been extensively studied in recent literature (see, for example, [13, 30, 54]).

Particularly, we are interested in problem (71) and (72) with $\mathcal{M} = \mathbb{S}^3 := \{p \in \mathbb{R}^4 \mid \|p\|_e = 1\}$, equipped with the metric of the ambient space \mathbb{R}^4 . According to [36], the skew-symmetric matrices

$$M_1 = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad M_2 = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}, \quad M_3 = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix},$$

induce an orthonormal basis of vector fields $\{E_1, E_2, E_3\}$ on \mathbb{S}^3 , defined by $E_i(p) = M_i p$ (standard matrix–vector product) for all $p \in \mathbb{S}^3$ and $i = 1, 2, 3$. By defining $\mathcal{L}_\mu: \mathbb{S}^3 \rightarrow \mathbb{R}$ by

$$\mathcal{L}_\mu(p) := \mathbf{f}(p) + \mu \mathbf{g}(p),$$

it follows from Example 7 that the KKT conditions for (71) and (72) are:

$$\text{grad } \mathcal{L}_\mu(p) = \text{grad } \mathbf{f}(p) + \mu \text{grad } \mathbf{g}(p) = 0_p, \quad (74)$$

$$\mu \geq 0, \quad \mathbf{g}(p) \leq 0, \quad \mu \mathbf{g}(p) = 0, \quad (75)$$

which is equivalent to the generalized equation

$$(p, \mu) \in \mathbb{S}^3 \times \mathbb{R}, \quad f(p, \mu) + F(p, \mu) \ni 0, \quad (76)$$

where $f: \mathbb{S}^3 \times \mathbb{R} \rightarrow \mathbb{R}^3 \times \mathbb{R}$ and $F: \mathbb{S}^3 \times \mathbb{R} \rightarrow \mathbb{R}^3 \times \mathbb{R}$ are defined, respectively, by

$$f(p, \mu) = (\langle \text{grad } \mathcal{L}_\mu(p), E_1(p) \rangle, \langle \text{grad } \mathcal{L}_\mu(p), E_2(p) \rangle, \langle \text{grad } \mathcal{L}_\mu(p), E_3(p) \rangle, \mathbf{g}(p)) \quad (77)$$

and

$$F(p, \mu) = \begin{cases} \{0\} \times \{y \in \mathbb{R}_+ : \mu y = 0\}, & \text{if } \mu \geq 0, \\ \emptyset, & \text{otherwise,} \end{cases} \quad (78)$$

for all $(p, \mu) \in \mathbb{S}^3 \times \mathbb{R}$.

To apply the inexact Newton method in (5) to solve (76), the subproblem in each iteration k involves computing $(v_k, v_k) \in \mathcal{T}_{(p_k, \mu_k)}(\mathbb{S}^3 \times \mathbb{R}) \equiv \mathcal{T}_{p_k} \mathbb{S}^3 \times \mathbb{R}$ such that

$$(f(p_k, \mu_k) + \mathcal{D}f(p_k, \mu_k)[(v_k, v_k)] + F(\exp_{p_k} v_k, \mu_k + v_k)) \cap R_k(p_k, \mu_k) \neq \emptyset.$$

To accomplish this, we need to select $u_k \in R_k(p_k, \mu_k)$. Here, let \varkappa be 2 and R_k is given in (40), with the function ζ defined in Proposition 3. Set $\rho_k = \min(\|\zeta^\varkappa(p_k, \mu_k)\|_e/4, 1/10^k)$ and we use $u_k = \rho_k \times [1, 1, 1, 1]^t$, where t denotes the transpose operation, for all k . In this way, u_k is in R_k , and $\|u_k\|_e$ is also monotonically decreasing to 0, for all k . Then solve the following optimization problem:

$$\text{minimize} \quad \frac{1}{2} \|z + f(p_k, \mu_k) + \mathcal{D}f(p_k, \mu_k)[(v, v)] - u_k\|_e^2 \quad (79)$$

$$\text{subject to } (v, v) \in \mathcal{T}_{p_k} \mathbb{S}^3 \times \mathbb{R}, \quad z = (z_1, z_2, z_3, z_4) \in F(\exp_{p_k} v, \mu_k + v). \quad (80)$$

Based on the definition of F in (78) and the fact that every $v \in \mathcal{T}_{p_k} \mathbb{S}^3$ can be expressed as a linear combination of $E_1(p_k)$, $E_2(p_k)$, and $E_3(p_k)$, we can find a solution to (79) and (80) by solving the Euclidean quadratic constraint problem:

$$\text{minimize} \quad \frac{1}{2} \|z + f(p_k, \mu_k) + \mathcal{D}f(p_k, \mu_k)[(\alpha_1 E_1(p_k) + \alpha_2 E_2(p_k) + \alpha_3 E_3(p_k), v)] - u_k\|_e^2 \quad (81)$$

$$\text{subject to } (\alpha_1, \alpha_2, \alpha_3) \in \mathbb{R}^3, \quad z_1 = z_2 = z_3 = 0, \quad z_4 \in \mathbb{R}^+, \quad \mu_k + v \in \mathbb{R}^+, \quad z_4(\mu_k + v) = 0. \quad (82)$$

Thus, the iteration $k + 1$ is obtained as follows:

$$(p_{k+1}, \mu_{k+1}) = (\exp_{p_k}(\alpha_1^k E_1(p_k) + \alpha_2^k E_2(p_k) + \alpha_3^k E_3(p_k)), \mu_k + v_k)$$

where $\alpha_1^k, \alpha_2^k, \alpha_3^k, v_k$ are a solution for (81) and (82).

For the implementation of our algorithm, we use the following expressions derived from (73) and (77):

$$f(p_k, \mu_k) = \begin{bmatrix} \left\langle -\frac{2}{N} \sum_{i=1}^N \exp_{p_k}^{-1} p^i - 2\mu_k \exp_{p_k}^{-1} \tilde{p}, E_1(p_k) \right\rangle \\ \left\langle -\frac{2}{N} \sum_{i=1}^N \exp_{p_k}^{-1} p^i - 2\mu_k \exp_{p_k}^{-1} \tilde{p}, E_2(p_k) \right\rangle \\ \left\langle -\frac{2}{N} \sum_{i=1}^N \exp_{p_k}^{-1} p^i - 2\mu_k \exp_{p_k}^{-1} \tilde{p}, E_3(p_k) \right\rangle \\ d^2(p_k, \tilde{p}) - r^2 \end{bmatrix}$$

and

$$\mathcal{D}f(p_k, \mu_k)[(v, v)] = \begin{bmatrix} \langle \text{grad}_p f_1(p_k, \mu_k), v \rangle - 2\langle \log_{p_k} \tilde{p}, E_1(p_k) \rangle v \\ \langle \text{grad}_p f_2(p_k, \mu_k), v \rangle - 2\langle \log_{p_k} \tilde{p}, E_2(p_k) \rangle v \\ \langle \text{grad}_p f_3(p_k, \mu_k), v \rangle - 2\langle \log_{p_k} \tilde{p}, E_3(p_k) \rangle v \\ -2\langle \log_{p_k} \tilde{p}, v \rangle \end{bmatrix}, \quad (v, v) \in \mathcal{T}_{p_k} \mathbb{S}^3 \times \mathbb{R},$$

where, for each $j \in \{1, 2, 3\}$, $\text{grad}_p f_j(p_k, \mu_k)$ denotes the Riemannian gradient of $f_j(\cdot, \mu_k)$ at p_k , which is the projection of the Euclidean gradient of $f_j(\cdot, \mu_k)$ at p_k onto $\mathcal{T}_{p_k} \mathbb{S}^3$. This projection can be obtained by multiplying the vector by the matrix $\text{id}_{4 \times 4} - p_k(p_k)^t$, where $\text{id}_{4 \times 4}$ denotes the identity matrix of dimension 4×4 . Additionally, we need to know the following expressions:

$$\exp_p(v) := \begin{cases} \cos(\|v\|_2)p + \sin(\|v\|_2)\frac{v}{\|v\|_2}, & v \in \mathcal{T}_p \mathbb{S}^3 \setminus \{0\}, \\ p, & v = 0, \end{cases}$$

and

$$\exp_p^{-1}(q) := \begin{cases} \frac{\arccos\langle p, q \rangle}{\sqrt{1 - \langle p, q \rangle^2}}(I - pp^T)q, & q \notin \{p, -p\}, \\ 0, & q = p. \end{cases}$$

These formulas on the sphere can be found in [28], for example.

Now, we are prepared for the numerical implementation. We use the Matlab built-in function *rand* to generate p^i , normalizing them to unit vectors. Each component of p^i lies within the interval $(0, 1)$. Specifically, we consider four cases:

- A1.** $N = 10, r = 2, \tilde{p} = [0, 0, 0, 1]$, and $p_0 = p^1$;
- A2.** $N = 500, r = 2, \tilde{p} = [0, 0, 0, 1]$, and $p_0 = p^1$;
- A3.** $N = 10, r = 0.1, \tilde{p} = [0, 0, 0, 1]$, and $p_0 = p^1$;
- A4.** $N = 500, r = 0.1, \tilde{p} = [0, 0, 0, 1]$, and $p_0 = p^1$;

where p_0 is the initial iteration point. For the stopping criteria, we use

$$\|\Phi(p_k, \mu_k)\|_e \leq 10^{-12} \quad \text{and} \quad \mathbf{g}(p_k) \leq 10^{-12}$$

where $\Phi := (f_1, f_2, f_3)$.

From Fig. 1 and Table 1, we can claim that for the above four cases **A1-A4**, we find a solution (p^*, μ^*) for generalized equation (76). This is because for all four cases, we achieve $\|\Phi(p^*, \mu^*)\|_e \leq 10^{-12}$, $\mathbf{g}(p^*) \leq 10^{-12}$ and $\mu^* \geq 0$. In addition, KKT system (74) and (75) is also satisfied under (p^*, μ^*) , which is asserted in theory and again ensured numerically. Furthermore, looking into cases **A1-A2**, we have $\mathbf{g}(p^*) < 0$, which means that p^* lies in the interior of the constraint region. But for cases **A3-A4**, p^* is on the boundary of the constraint region because $\mathbf{g}(p^*)$ is almost equal to 0. Due to the effect of the constraint, the convergence rate of the norm of $\Phi(p_k, \mu_k)$ for cases **A1-A2** is stable. For cases **A3-A4**, the convergence rate of the norm of $\Phi(p_k, \mu_k)$ in the first several iterations is unstable but after several iterations, the convergence rate becomes stable. This coincides with the theory that the convergence rate of the inexact Newton method becomes stable when the iteration point is close to the solution.

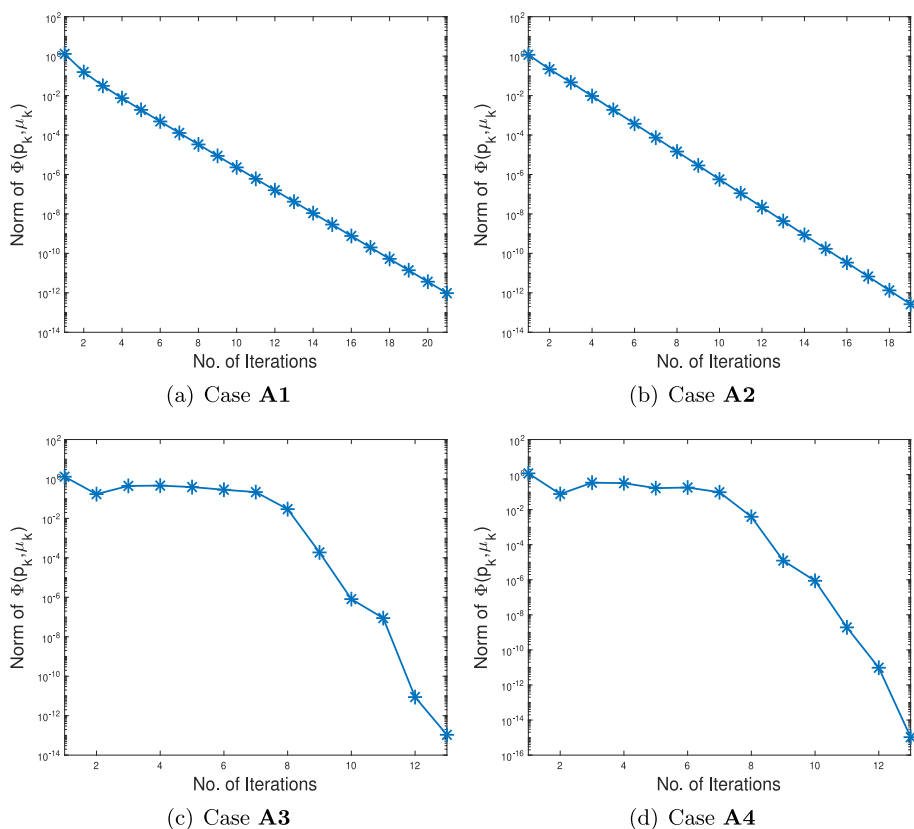


Fig. 1 Norm of $\Phi(p_k, \mu_k)$ at each iteration for four cases A1–A4

7 Conclusion

In this paper, we address the problem of finding solutions to generalized equations on Riemannian manifolds. If the manifold is a Euclidean space, then it is well-known that this problem encompasses several other contexts, such as standard nonlinear optimization, variational inequalities, and equilibrium problems. Here, we present a general inexact Newton method for solving generalized equations. Firstly, we discuss the metric regularity property and provide some examples of mappings satisfying this property. Secondly, we establish local convergence results, including both linear and quadratic convergence under suitable assumptions, along with a semi-local convergence result. Finally, we discuss the relationship between problems (2) and (3).

All results obtained here are derived under assumptions that are highly natural in comparison to their Euclidean counterparts, without requiring additional conditions related to the geometry of the manifold. From this perspective, considering a problem as a generalized equation may be a promising approach in the context of manifolds as well.

As a next step, we plan to investigate how to extend the theory presented here to cases where the exponential map is replaced by a general retraction. Moreover, we intend to apply this new concept to quasi-Newton-type methods for solving problem (2).

Table 1 The final results of four cases A1–A4

	p^*	μ^*	$g(p^*)$	$\mu g(p^*)$	$\ \text{grad}\mathcal{L}_{\mu^*}(p^*)\ _p$	No. of Iteration
Case A1	[0.4388,0.4862,0.5665,0.5001]	0	-2.9037	0	9.61×10^{-13}	21
Case A2	[0.4908,0.5100,0.4878,0.5109]	0	-2.9298	0	2.64×10^{-13}	19
Case A3	[0.0504,0.0566,0.0650,0.9950]	8.9114	1.04×10^{-17}	9.27×10^{-17}	1.08×10^{-13}	13
Case A4	[0.0570,0.0593,0.0566,0.9950]	8.7870	0	0	2.21×10^{-15}	13

A Appendix

We begin this section by presenting two supporting lemmas. The proof of the first one can be found in [46, Lemma A.1]. The second one is analogous to [46, Lemma A.2], but its proof requires some adaptations. Therefore, we provide the proof here.

Lemma 3 *Let (\bar{p}, \bar{v}) be a point on \mathcal{TM} , and $B_r(\bar{p})$ be a normal ball. Define the following vector field on $B_r(\bar{p})$:*

$$V(p) = P_{\bar{p}p}\bar{v}, \quad p \in B_r(\bar{p}).$$

Then, V is a smooth vector field on $B_r(\bar{p})$.

Lemma 4 *Let $f: \mathcal{M} \rightarrow \mathbb{R}^m$ be a continuously differentiable function at \bar{p} and $B_r(\bar{p})$ a normal ball. Then, for each $\epsilon > 0$, there exists $\delta \in (0, r)$ such that*

$$\|\mathcal{D}f(p)P_{\bar{p}p} - \mathcal{D}f(\bar{p})\|_{\text{map}} \leq \epsilon \quad \text{for all } p \in B_\delta(\bar{p}).$$

Proof It follows from (7) and parallel transport properties that

$$\begin{aligned} \mathcal{D}f(p)[P_{\bar{p}p}v] &= (\langle \text{grad } f_1(p), P_{\bar{p}p}v \rangle_p, \dots, \langle \text{grad } f_m(p), P_{\bar{p}p}v \rangle_p), \\ &= (\langle P_{p\bar{p}} \text{grad } f_1(p), v \rangle_{\bar{p}}, \dots, \langle P_{p\bar{p}} \text{grad } f_m(p), v \rangle_{\bar{p}}), \quad v \in \mathcal{T}_{\bar{p}}\mathcal{M}, \quad p \in B_r(\bar{p}). \end{aligned}$$

From norm properties, we have

$$\begin{aligned} &\|(\mathcal{D}f(p)P_{\bar{p}p} - \mathcal{D}f(\bar{p}))[v]\|_e^2 \\ &= \sum_{i=1}^m |\langle P_{p\bar{p}} \text{grad } f_i(p) - \text{grad } f_i(\bar{p}), v \rangle_{\bar{p}}|^2, \\ &\leq \sum_{i=1}^m \|P_{p\bar{p}} \text{grad } f_i(p) - \text{grad } f_i(\bar{p})\|_{\bar{p}}^2 \|v\|_{\bar{p}}^2, \quad v \in \mathcal{T}_{\bar{p}}\mathcal{M}, \quad p \in B_r(\bar{p}). \end{aligned}$$

Since $\mathcal{D}f(p)P_{\bar{p}p} - \mathcal{D}f(\bar{p}): \mathcal{T}_{\bar{p}}\mathcal{M} \rightarrow \mathbb{R}^m$ is a linear map, the above inequality implies that

$$\begin{aligned} &\|\mathcal{D}f(p)P_{\bar{p}p} - \mathcal{D}f(\bar{p})\|_{\text{map}}^2 \\ &\leq \sum_{i=1}^m \|P_{p\bar{p}} \text{grad } f_i(p) - \text{grad } f_i(\bar{p})\|_{\bar{p}}^2 \\ &= \sum_{i=1}^m \|\text{grad } f_i(p) - P_{\bar{p}p} \text{grad } f_i(\bar{p})\|_p^2, \quad p \in B_r(\bar{p}). \end{aligned}$$

Under the assumptions of the lemma, $\text{grad } f_i$ is a continuous vector field around \bar{p} for all $i = 1, \dots, m$. Hence, using Lemma 3 with $\bar{v} = \text{grad } f_i(\bar{p})$ and considering the fact that the function $\mathcal{M} \ni p \mapsto \|\cdot\|_p$ is continuous, we obtain

$$\lim_{p \rightarrow \bar{p}} \|\text{grad } f_i(p) - P_{\bar{p}p} \text{grad } f_i(\bar{p})\|_p = \|\text{grad } f_i(\bar{p}) - P_{\bar{p}\bar{p}} \text{grad } f_i(\bar{p})\|_{\bar{p}} = 0.$$

Thus, it follows from the last inequality above that $\lim_{p \rightarrow \bar{p}} \|\mathcal{D}f(p)P_{\bar{p}p} - \mathcal{D}f(\bar{p})\|_{\text{map}} = 0$, which is equivalent to what we want. \square

In the following proposition, we provide a Riemannian version of the Fundamental Theorem of Calculus and two inequalities that play a crucial role throughout this paper. We claim that the inequalities in the following result do not hold for a general retraction.

Proposition 4 Let $f: \mathcal{M} \rightarrow \mathbb{R}^m$ be a continuously differentiable function at \bar{p} . Then, there exists $\delta > 0$ such that for each $p', p \in \mathcal{B}_\delta(\bar{p})$, we have

$$f(p) - f(p') = \int_0^1 \mathcal{D}f(\gamma(t))[\dot{\gamma}(t)] dt, \quad (83)$$

where $\gamma: [0, 1] \rightarrow \mathcal{M}$ is a geodesic satisfying $\gamma(0) = p'$ and $\gamma(1) = p$. In particular, for each $\epsilon > 0$, there exists a normal ball $\mathcal{B}_{\delta_\epsilon}(\bar{p})$ such that

$$\|f(p) - f(\bar{p}) - \mathcal{D}f(\bar{p})[\exp_{\bar{p}}^{-1} p]\|_e \leq \epsilon d(p, \bar{p}), \quad p \in \mathcal{B}_{\delta_\epsilon}(\bar{p}). \quad (84)$$

Furthermore, if $\mathcal{D}f$ is L -Lipschitz continuous around \bar{p} then there exists a normal ball $\mathcal{B}_{\delta_L}(\bar{p})$ such that

$$\|f(p) - f(\bar{p}) - \mathcal{D}f(\bar{p})[\exp_{\bar{p}}^{-1} p]\|_e \leq Ld^2(p, \bar{p}), \quad p \in \mathcal{B}_{\delta_L}(\bar{p}). \quad (85)$$

Proof Choose $\delta > 0$ such that f is continuously differentiable on $\mathcal{B}_\delta(\bar{p})$. Pick $p', p \in \mathcal{B}_\delta(\bar{p})$ and consider a geodesic $\gamma: [0, 1] \rightarrow \mathcal{M}$ connecting $p' = \gamma(0)$ to $p = \gamma(1)$. Using [21, Proposition 2.7], we conclude that

$$\frac{d}{dt}(f \circ \gamma)(t) = \mathcal{D}f(\gamma(t))[\dot{\gamma}(t)], \quad t \in [0, 1].$$

Applying integral with respect to t on both sides and using the Fundamental Theorem of Calculus for the real function $f \circ \gamma$, we conclude that

$$f(p) - f(p') = f(\gamma(1)) - f(\gamma(0)) = \int_0^1 \mathcal{D}f(\gamma(t))[\dot{\gamma}(t)] dt, \quad (86)$$

which completes the proof of (83). Now, let us prove (84) and (85). Let $\mathcal{B}_r(\bar{p})$ be a normal ball. For each $p \in \mathcal{B}_r(\bar{p})$, the geodesic $\gamma_{\bar{p}p}: [0, 1] \rightarrow \mathcal{M}$ connecting \bar{p} to p can be written as $\gamma_{\bar{p}p}(t) = \exp_{\bar{p}}(t \exp_{\bar{p}}^{-1} p)$, which implies that $\dot{\gamma}_{\bar{p}p}(t) = P_{\bar{p}\gamma_{\bar{p}p}(t)} \exp_{\bar{p}}^{-1} p$ holds for all $t \in [0, 1]$. Hence, it follows from (86) with $\gamma = \gamma_{\bar{p}p}$ that

$$f(p) - f(\bar{p}) = \int_0^1 \mathcal{D}f(\gamma_{\bar{p}p}(t)) [P_{\bar{p}\gamma_{\bar{p}p}(t)} \exp_{\bar{p}}^{-1} p] dt.$$

By adding the term $-\mathcal{D}f(\bar{p})[\exp_{\bar{p}}^{-1} p]$ on both sides, it comes that

$$f(p) - f(\bar{p}) - \mathcal{D}f(\bar{p})[\exp_{\bar{p}}^{-1} p] = \int_0^1 (\mathcal{D}f(\gamma_{\bar{p}p}(t)) P_{\bar{p}\gamma_{\bar{p}p}(t)} - \mathcal{D}f(\bar{p})) [\exp_{\bar{p}}^{-1} p] dt.$$

Applying the Euclidean norm of \mathbb{R}^m , we get

$$\begin{aligned} & \|f(p) - f(\bar{p}) - \mathcal{D}f(\bar{p})[\exp_{\bar{p}}^{-1} p]\|_e \\ &= \left\| \int_0^1 (\mathcal{D}f(\gamma_{\bar{p}p}(t)) P_{\bar{p}\gamma_{\bar{p}p}(t)} - \mathcal{D}f(\bar{p})) [\exp_{\bar{p}}^{-1} p] dt \right\|_e \\ &\leq \int_0^1 \left\| (\mathcal{D}f(\gamma_{\bar{p}p}(t)) P_{\bar{p}\gamma_{\bar{p}p}(t)} - \mathcal{D}f(\bar{p})) [\exp_{\bar{p}}^{-1} p] \right\|_e dt \\ &\leq \int_0^1 \|\mathcal{D}f(\gamma_{\bar{p}p}(t)) P_{\bar{p}\gamma_{\bar{p}p}(t)} - \mathcal{D}f(\bar{p})\|_{\text{map}} \|\exp_{\bar{p}}^{-1} p\|_{\bar{p}} dt. \end{aligned}$$

It follows from (8) and simple properties of integral that

$$\begin{aligned} & \|f(p) - f(\bar{p}) - \mathcal{D}f(\bar{p})[\exp_{\bar{p}}^{-1} p]\|_e \\ & \leq \sup_{t \in [0,1]} \{ \|\mathcal{D}f(\gamma_{\bar{p}p}(t)) P_{\bar{p}\gamma_{\bar{p}p}(t)} - \mathcal{D}f(\bar{p})\|_{\text{map}} \} d(p, \bar{p}), \end{aligned} \quad (87)$$

for all $p \in B_r(\bar{p})$. For arbitrary $\epsilon > 0$, it follows from Lemma 4 and $\delta_\epsilon \in (0, r)$ that $\|\mathcal{D}f(p) P_{\bar{p}p} - \mathcal{D}f(\bar{p})\|_{\text{map}} \leq \epsilon$ for all $p \in B_{\delta_\epsilon}(\bar{p})$, which implies that $\|\mathcal{D}f(\gamma_{\bar{p}p}(t)) P_{\bar{p}\gamma_{\bar{p}p}(t)} - \mathcal{D}f(\bar{p})\|_{\text{map}} \leq \epsilon$ for all $p \in B_{\delta_\epsilon}(\bar{p})$ and $t \in [0, 1]$. Therefore, (87) yields (84). Finally, assume that $\mathcal{D}f$ is L -Lipschitz continuous around \bar{p} . By Definition 1, there exists $\delta_L \in (0, r)$ such that $\|\mathcal{D}f(p) P_{\bar{p}p} - \mathcal{D}f(\bar{p})\|_{\text{map}} \leq Ld(p, \bar{p})$ for all $p \in B_{\delta_L}(\bar{p})$, which implies that

$$\sup_{t \in [0,1]} \{ \|\mathcal{D}f(\gamma_{\bar{p}p}(t)) P_{\bar{p}\gamma_{\bar{p}p}(t)} - \mathcal{D}f(\bar{p})\|_{\text{map}} \} \leq L \sup_{t \in [0,1]} \{ d(\gamma_{\bar{p}p}(t), \bar{p}) \} = Ld(p, \bar{p}),$$

for all $p \in B_{\delta_L}(\bar{p})$. Using the previous inequality and (87), we conclude the proof of (85). \square

The following proposition provides a geometric property of the exponential map.

Proposition 5 *Let $B_r(\bar{p})$ be a totally normal ball. Then, for each p and q in $B_r(\bar{p})$, we have*

$$\dot{\gamma}_{pq}(t) = \exp_{\gamma_{pq}(t)}^{-1} q - \exp_{\gamma_{pq}(t)}^{-1} p, \quad t \in [0, 1], \quad (88)$$

where $\gamma_{pq} : [0, 1] \rightarrow \mathcal{M}$ is the geodesic connecting $p = \gamma_{pq}(0)$ to $q = \gamma_{pq}(1)$.

Proof Let p and q be arbitrary points in $B_r(\bar{p})$. Note that (88) is trivially satisfied when $t = 0$ and $t = 1$. Thus, we only need to analyze the case where $t \in (0, 1)$. Considering that the geodesic starting from p with direction $\dot{\gamma}_{pq}(0)$ is unique, and that $B_r(\bar{p})$ is a totally normal ball, we have

$$\frac{1}{d(p, q)} \dot{\gamma}_{pq}(0) = \frac{1}{d(p, \gamma_{pq}(t))} \exp_p^{-1}(\gamma_{pq}(t)), \quad t \in (0, 1).$$

Applying parallel transport on both sides of the equation, we obtain

$$\begin{aligned} \frac{1}{d(p, q)} P_{p\gamma_{pq}(t)}[\dot{\gamma}_{pq}(0)] &= \frac{1}{d(p, \gamma_{pq}(t))} P_{p\gamma_{pq}(t)}[\exp_p^{-1}(\gamma_{pq}(t))] \\ &= \frac{1}{d(p, \gamma_{pq}(t))} [-\exp_{\gamma_{pq}(t)}^{-1} p], \quad t \in (0, 1). \end{aligned}$$

Since γ_{pq} is a geodesic, the field $\dot{\gamma}_{pq}$ is parallel along γ_{pq} , and, consequently, the equality $\dot{\gamma}_{pq}(t) = P_{p\gamma_{pq}(t)}[\dot{\gamma}_{pq}(0)]$ holds for all $t \in (0, 1)$. Thus, the last expression implies

$$\frac{d(p, \gamma_{pq}(t))}{d(p, q)} \dot{\gamma}_{pq}(t) = -\exp_{\gamma_{pq}(t)}^{-1}(p), \quad t \in (0, 1).$$

In a similar manner, we can show that

$$\frac{d(\gamma_{pq}(t), q)}{d(p, q)} \dot{\gamma}_{pq}(t) = \exp_{\gamma_{pq}(t)}^{-1}(q), \quad t \in (0, 1).$$

To conclude the proof, simply add the last two equalities and use the fact that $d(p, \gamma_{pq}(t)) + d(\gamma_{pq}(t), q) = d(p, q)$ for all $t \in (0, 1)$. \square

In the next result, we will establish a sufficient condition for the function $f : \mathcal{M} \rightarrow \mathbb{R}$ to have closed graph.

Proposition 6 Let $f: \mathcal{M} \rightarrow \mathbb{R}$ be a continuous function. Then, the graph of f is closed in $\mathcal{M} \times \mathbb{R}$.

Proof Since the function f is continuous, the function $\varphi: \mathcal{M} \times \mathbb{R} \rightarrow \mathbb{R}$ defined by $\varphi(p, x) = |f(p) - x|$ is also continuous. On the other hand, we have that

$$\text{gph } f := \{(p, x) \in \mathcal{M} \times \mathbb{R} : x = f(p)\} = \varphi^{-1}(0).$$

Therefore, $\text{gph } f$ is closed because it is the inverse image, by the continuous function φ , of the closed set $\{0\}$. \square

Throughout this paper, our strategy to prove some results requires the following lemma.

Lemma 5 Let $f: \mathcal{M} \rightarrow \mathbb{R}^m$ be a continuously differentiable function at \bar{p} . Then, for each $\epsilon > 0$, there exists a totally normal ball $B_{\delta_\epsilon}(\bar{p})$ such that

$$\begin{aligned} & \|\mathcal{D}f(p)[\exp_p^{-1} p' - \exp_p^{-1} p''] - \mathcal{D}f(\bar{p})[\exp_{\bar{p}}^{-1} p' - \exp_{\bar{p}}^{-1} p'']\|_e \\ & \leq \epsilon d(p', p''), \quad p, p', p'' \in B_{\delta_\epsilon}(\bar{p}). \end{aligned}$$

Proof Let $B_r(\bar{p})$ be a totally normal ball. Define the function $\Psi: B_r(\bar{p}) \times B_r(\bar{p}) \rightarrow \mathbb{R}^m$ by

$$\Psi(p, q) = \mathcal{D}f(p)[\exp_p^{-1} q] - \mathcal{D}f(\bar{p})[\exp_{\bar{p}}^{-1} q]. \quad (89)$$

Let $\epsilon > 0$, and consider $\mathcal{M} \times \mathcal{M}$ with the product metric. Since $\mathcal{D}f$ and \exp^{-1} are continuous at \bar{p} and continuously differentiable at (\bar{p}, \bar{p}) , respectively, $\mathcal{D}_2\Psi(p, q)$ (the differential of $\Psi(p, \cdot)$ at q) is continuous at (\bar{p}, \bar{p}) . Therefore, there exists $\delta \in (0, r)$ such that

$$d((p, q), (\bar{p}, \bar{p})) = (d^2(p, \bar{p}) + d^2(q, \bar{p}))^{\frac{1}{2}} < \delta \quad \text{implies} \quad \|\mathcal{D}_2\Psi(p, q) - \mathcal{D}_2\Psi(\bar{p}, \bar{p})\|_e < \epsilon.$$

From (89), it follows that $\Psi(\bar{p}, q) = 0$ for all $q \in B_r(\bar{p})$ which implies $\mathcal{D}_2\Psi(\bar{p}, \bar{p}) = 0$. Thus, the last line can be rewritten as

$$d^2(p, \bar{p}) + d^2(q, \bar{p}) < \delta^2 \quad \text{implies} \quad \|\mathcal{D}_2\Psi(p, q)\|_e < \epsilon. \quad (90)$$

Set $\delta_\epsilon := \delta/\sqrt{2}$. It follows from $d^2(p, \bar{p}) + d^2(q, \bar{p}) < \delta^2$ for all $(p, q) \in B_{\delta_\epsilon}(\bar{p}) \times B_{\delta_\epsilon}(\bar{p})$ and (90) that $\|\mathcal{D}_2\Psi(p, q)\|_e < \epsilon$ for all $(p, q) \in B_{\delta_\epsilon}(\bar{p}) \times B_{\delta_\epsilon}(\bar{p})$. Consequently,

$$\sup_{q \in B_{\delta_\epsilon}(\bar{p})} \{\|\mathcal{D}_2\Psi(p, q)\|_e\} \leq \epsilon \quad \text{for all } p \in B_{\delta_\epsilon}(\bar{p}). \quad (91)$$

For $p, p', p'' \in B_{\delta_\epsilon}(\bar{p})$, it follows from the first part of Proposition 4 and (89) that

$$\begin{aligned} & \mathcal{D}f(p)[\exp_p^{-1} p' - \exp_p^{-1} p''] - \mathcal{D}f(\bar{p})[\exp_{\bar{p}}^{-1} p' - \exp_{\bar{p}}^{-1} p''] \\ & = \Psi(p, p') - \Psi(p, p''), \\ & = \int_0^1 \mathcal{D}_2\Psi(p, \gamma(t))[\dot{\gamma}(t)] dt, \end{aligned}$$

where $\gamma: [0, 1] \rightarrow \mathcal{M}$ is a geodesic that satisfies $\gamma(0) = p''$ and $\gamma(1) = p'$. Since $B_{\delta_\epsilon}(\bar{p})$ is totally normal, $\gamma(t) \in B_{\delta_\epsilon}(\bar{p})$ for all $t \in [0, 1]$. Hence, applying the norm to both sides of the above equality and using norm properties, we obtain

$$\begin{aligned} & \|\mathcal{D}f(p)[\exp_p^{-1} p' - \exp_p^{-1} p''] - \mathcal{D}f(\bar{p})[\exp_{\bar{p}}^{-1} p' - \exp_{\bar{p}}^{-1} p'']\|_e \\ & \leq \int_0^1 \|\mathcal{D}_2\Psi(p, \gamma(t))[\dot{\gamma}(t)]\|_e dt, \end{aligned}$$

$$\begin{aligned}
&\leq \int_0^1 \|\mathcal{D}_2 \Psi(p, \gamma(t))\|_{\text{map}} \|\dot{\gamma}(t)\|_{\gamma(t)} dt, \\
&\leq \sup_{q \in \mathcal{B}_{\delta_\epsilon}(\bar{p})} \{\|\mathcal{D}_2 \Psi(p, q)\|_{\text{map}}\} d(p', p'').
\end{aligned}$$

The conclusion of this proof follows from the previous inequality and (91). \square

Acknowledgements The authors would like to thank the anonymous referees for their valuable comments and suggestions, which have significantly improved the quality of this paper.

Funding Mauricio S. Louzeiro was partially supported by National Natural Science Foundation of China (No. 12171087). Gilson N. Silva was partially supported by CNPq, Brazil (401864/2022-7 and 306593/2022-0). Jinyun Yuan was partially supported by National Natural Science Foundation of China (No. 12171087), Dongguan University of Technology, China (221110093). Daoping Zhang was supported by National Natural Science Foundation of China (No. 12201320) and the Fundamental Research Funds for the Central Universities, Nankai University (No. 63221039 and 63231144).

Data Availability Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Absil, P.-A., Mahony, R., Sepulchre, R.: Optimization algorithms on matrix manifolds. Princeton University Press, Oxford (2008)
2. Adler, R.L., Dedieu, J.-P., Margulies, J.Y., Martens, M., Shub, M.: Newton's method on Riemannian manifolds and a geometric model for the human spine. *IMA J. Numer. Anal.* **22**(3), 359–390 (2002)
3. Adly, S., Cibulka, R., van Ngai, H.: Newton's method for solving inclusions using set-valued approximations. *SIAM J. Optim.* **25**, 159–184 (2015)
4. Adly, S., van Ngai, H., Vu, N.V.: Newton-type method for solving generalized equations on Riemannian manifolds. *J. Convex Anal.* **25**(2), 341–370 (2018)
5. Adly, S., van Ngai, H., Vu, N.V.: Dennis-Moré condition for set-valued vector fields and the superlinear convergence of Broyden updates in Riemannian manifolds. *J. Convex Anal.* **29**(3), 669–701 (2022)
6. Alvarez, F., Bolte, J., Munier, J.: A unifying local convergence result for Newton's method in Riemannian manifolds. *Found. Comput. Math.* **8**(2), 197–226 (2008)
7. Andreani, R., Carvalho, R.M., Secchin, L.D., Silva, G.N.: Convergence of Quasi-Newton methods for solving constrained generalized equations. *ESAIM: Control Optim. Calc. Var.* **28**, 32 (2022)
8. Aragón Artacho, F.J., Belyakov, A., Dontchev, A.L., López, M.: Local convergence of quasi-Newton methods under metric regularity. *Comput. Optim. Appl.* **58**(1), 225–247 (2014)
9. Aragón Artacho, F.J., Dontchev, A.L., Gaydu, M., Geoffroy, M.H., Veliiov, V.M.: Metric regularity of Newton's iteration. *SIAM J. Control. Optim.* **49**(2), 339–362 (2011)
10. Bergmann, R., Herzog, R.: Intrinsic formulation of KKT conditions and constraint qualifications on smooth manifolds. *SIAM J. Optim.* **29**(4), 2423–2444 (2019)
11. Bergmann, R., Herzog, R., Silva Louzeiro, M., Tenbrinck, D., Vidal-Núñez, J.: Fenchel duality theory and a primal-dual algorithm on Riemannian manifolds. *Found. Comput. Math.* **21**(6), 1465–1504 (2021)
12. Bergmann, R., Persch, J., Steidl, G.: A parallel Douglas-Rachford algorithm for minimizing rof-like functionals on images with values in symmetric Hadamard manifolds. *SIAM J. Imag. Sci.* **9**(3), 901–937 (2016)
13. Bini, D.A., Iannazzo, B.: Computing the karcher mean of symmetric positive definite matrices. *Linear Algebra Appl.* **438**(4), 1700–1710 (2013)
14. Boumal, N.: An introduction to optimization on smooth manifolds. Cambridge University Press, Cambridge (2023)
15. Censor, Y., Gibali, A., Reich, S.: Algorithms for the split variational inequality problem. *Numer. Algorithms* **59**, 301–323 (2011)

16. Cibulka, R., Dontchev, A., Geoffroy, M.H.: Inexact Newton Methods and Dennis-Moré Theorems for Nonsmooth Generalized Equations. *SIAM J. Control. Optim.* **53**(2), 1003–1019 (2015)
17. Cibulka, R., Dontchev, A., Geoffroy, M.H.: Inexact Newton methods and Dennis-Moré theorems for nonsmooth generalized equations. *SIAM J. Control. Optim.* **53**(2), 1003–1019 (2015)
18. de Oliveira, F.R., Ferreira, O.P., Silva, G.N.: Newton's method with feasible inexact projections for solving constrained generalized equations. *Comput. Optim. Appl.* **72**, 159–177 (2019)
19. Dedieu, J.-P., Priouret, P., Malajovich, G.: Newton's method on Riemannian manifolds: covariant alpha theory. *IMA J. Numer. Anal.* **23**(3), 395–419 (2003)
20. Dembo, R.S., Eisenstat, S.C., Steihaug, T.: Inexact newton methods. *SIAM J. Numer. Anal.* **19**(2), 400–408 (1982)
21. do Carmo, M. P.: Riemannian geometry. *Mathematics: Theory & Applications*. Birkhäuser Boston, Inc., Boston, MA, (1992). Translated from the second Portuguese edition by Francis Flaherty
22. Dontchev, A. L.: Local analysis of a Newton-type method based on partial linearization. In: *The mathematics of numerical analysis* (Park City, UT, 1995), vol. 32 of *Lectures in Appl. Math.*, pp. 295–306. Amer. Math. Soc., Providence, RI, (1996)
23. Dontchev, A.L., Rockafellar, R.T.: Newton's method for generalized equations: a sequential implicit function theorem. *Math. Program.* **123**(1), 139–159 (2010)
24. Dontchev, A.L., Rockafellar, R.T.: Convergence of inexact Newton methods for generalized equations. *Math. Program.* **139**(1–2), 115–137 (2013)
25. Dontchev, A.L., Rockafellar, R.T.: *Implicit Functions and Solution Mappings: A View from Variational Analysis*, 2nd edn. Springer, New York (2014)
26. Fernandes, T.A., Ferreira, O.P., Yuan, J.: On the superlinear convergence of Newton's method on Riemannian manifolds. *J. Optim. Theory Appl.* **173**(3), 828–843 (2017)
27. Ferreira, O., Silva, G.: Local convergence analysis of Newton's method for solving strongly regular generalized equations. *J. Math. Anal. Appl.* **458**(1), 481–496 (2018)
28. Ferreira, O.P., Iusem, A.N., Németh, S.Z.: Projections onto convex sets on the sphere. *J. Global Optim.* **57**(3), 663–676 (2013)
29. Ferreira, O.P., Jean-Alexis, C., Piétrus, A.: Metrically regular vector field and iterative processes for generalized equations in Hadamard manifolds. *J. Optim. Theory Appl.* **175**(3), 624–651 (2017)
30. Ferreira, O.P., Louzeiro, M.S., Prudente, L.: Gradient method for optimization on Riemannian manifolds with lower bounded curvature. *SIAM J. Optim.* **29**(4), 2517–2541 (2019)
31. Ferreira, O.P., Silva, G.N.: Kantorovich's theorem on Newton's method for solving strongly regular generalized equation. *SIAM J. Optim.* **27**(2), 910–926 (2017)
32. Gaydu, M., Silva, G.N.: A general iterative procedure to solve generalized equations with differentiable multifunction. *J. Optim. Theory Appl.* **185**, 207–222 (2020)
33. Genicot, M., Huang, W., Trendafilov, N.T.: Weakly correlated sparse components with nearly orthonormal loadings. In: *International Conference on Geometric Science of Information*, pp. 484–490. Springer, Berlin (2015)
34. Geoffroy, M.H., Piétrus, A.: Local convergence of some iterative methods for generalized equations. *J. Math. Anal. Appl.* **290**(2), 497–505 (2004)
35. He, H., Ling, C., Xu, H.-K.: A relaxed projection method for split variational inequalities. *J. Optim. Theory Appl.* **166**, 213–233 (2015)
36. Hesselholt, L.: Vector fields on spheres. Preprint, <http://www-math.mit.edu/arsh/teaching/vectorfields.pdf>
37. Izmailov, A.F., Solodov, M.V.: Inexact Josephy-Newton framework for generalized equations and its applications to local analysis of Newtonian methods for constrained optimization. *Comput. Optim. Appl.* **46**(2), 347–368 (2010)
38. Karcher, H.: Riemannian center of mass and mollifier smoothing. *Commun. Pure Appl. Math.* **30**(5), 509–541 (1977)
39. Klatte, D., Kummer, B.: Approximations and generalized Newton methods. *Math. Program.* **168**, 673–716 (2018)
40. Lang, S.: *Fundamentals of differential geometry*. In: *Graduate Texts in Mathematics*, vol. 191. Springer-Verlag, New York (1999)
41. Lee, J.M.: *Introduction to smooth manifolds*. In: *Graduate Texts in Mathematics*, vol. 218. Springer-Verlag, New York (2003)
42. Lee, J.M.: *Riemannian manifolds: an introduction to curvature*, vol. 176. Springer Science & Business Media, Berlin (2006)
43. Li, C., Wang, J.: Convergence of the Newton method and uniqueness of zeros of vector fields on Riemannian manifolds. *Sci. China, Ser. A Math.* **48**(11), 1465–1478 (2005)

44. Li, C., Yao, J.-C.: Variational inequalities for set-valued vector fields on Riemannian manifolds: convexity of the solution set and the proximal point algorithm. *SIAM J. Control. Optim.* **50**(4), 2486–2514 (2012)
45. Li, S.-L., Li, C., Liou, Y.-C., Yao, J.-C.: Existence of solutions for variational inequalities on Riemannian manifolds. *Nonlinear Anal.: Theory, Methods Appl.* **71**(11), 5695–5706 (2009)
46. Liu, C., Boumal, N.: Simple algorithms for optimization on Riemannian manifolds with constraints. *Appl. Math. Optim.* **82**(3), 949–981 (2020)
47. Németh, S.: Variational inequalities on hadamard manifolds. *Nonlinear Anal.: Theory, Methods Appl.* **52**(5), 1491–1498 (2003)
48. Sato, H.: Riemannian optimization and its applications. Springer, Berlin (2021)
49. Sra, S., Hosseini, R.: Conic geometric optimization on the manifold of positive definite matrices. *SIAM J. Optim.* **25**(1), 713–739 (2015)
50. Tang, J., Liu, H.: Unsupervised feature selection for linked social media data. In: Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 904–912, (2012)
51. Tu, L.W.: An Introduction to Manifolds, 2nd edn. Springer, New York (2011)
52. Udriște, C.: Convex functions and optimization methods on Riemannian manifolds. In: Mathematics and its Applications, vol. 297, Kluwer Academic Publishers Group, Dordrecht (1994)
53. Wang, J.-H., Huang, S., Li, C.: Extended Newton’s method for mappings on Riemannian manifolds with values in a cone. *Taiwanese J. Math.* **13**(2B), 633–656 (2009)
54. Weber, M., Sra, S.: Riemannian optimization via frank-wolfe methods. *Math. Program.* **199**(1), 525–556 (2023)
55. Zhang, Y., Lau, Y., Kuo, H.-W., Cheung, S., Pasupathy, A., Wright, J.: On the global geometry of sphere-constrained sparse blind deconvolution. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4894–4902 (2017)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.