Tilt Stability on Riemannian Manifolds with Application to Convergence Analysis of Generalized Riemannian Newton Method

Zijian Shi[1] and Miantao Chao[1*]

[1*]College of Mathematics and Information Science, Guangxi University, Nanning, 530004, Guangxi, P. R. China.

*Corresponding author(s). E-mail(s): chaomiantao@126.com;
Contributing authors: zijianshi@foxmail.com;

Abstract

We generalize tilt stability, a fundamental concept in perturbation analysis of optimization problems in Euclidean spaces, to the setting of Riemannian manifolds. We prove the equivalence of the following conditions: Riemannian tilt stability, Riemannian variational strong convexity, Riemannian uniform quadratic growth, local strong monotonicity of Riemannian subdifferential, strong metric regularity of Riemannian subdifferential, and positive definiteness of generalized Riemannian Hessian. For Riemannian nonlinear programming, we provide a characterization of Riemannian tilt stability under a weak constraint qualification. Leveraging these results, we propose a generalized Riemannian Newton method and establish its superlinear convergence under Riemannian tilt stability.

Keywords: Riemannian tilt stability, generalized differentiation on Riemannian manifolds, Riemannian nonlinear programming, generalized Riemannian Newton methods

MSC Classification: 49J53 , 90C31 , 90C46

1 Introduction

Let $M \subset \mathbb{R}^n$ be an embedded Riemannian submanifold and consider an optimization problem constrained on M:

$$\min_{x \in M} f(x), \tag{1.1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is C^2. To find a solution (a local minimizer or at least a stationary point), in Riemannian optimization, one typically runs an iterative algorithm which generates a sequence $\{x_k \in M\}_{k \geq 0}$ of points on M and stops when the Riemannian gradient $v_K := \operatorname{grad} f(x_K)$ at x_K is small enough. However, a small gradient norm does not necessarily guarantee proximity to a stationary point,[1] so it is important to ask how close x_K actually is to a solution $\bar{x}$. This problem is exactly the focus of local convergence rate analysis. On the other hand, one can approach it from the perspective of perturbation analysis. Indeed, consider the parametric optimization problem defined by a tilt (i.e., linear) perturbation:

$$(P_v) \quad \min_{x \in M} f(x) - \langle v, x - \bar{x} \rangle . \tag{1.2}$$

One can show that $\bar{x}$ is a stationary point of (P_0) and x_K is a stationary point of (P_{v_K}). Then the problem of whether x_K is close to $\bar{x}$ when v_K is small is translated into the problem of whether the solution $S(v_K)$ of (P_{v_K}) is close to the solution $S(0) = \bar{x}$ of (P_0) when v_K is a small tilt perturbation. Therefore, just as in Euclidean spaces [2], numerical methodology and algorithm analysis in Riemannian optimization are inherently linked to stability analysis of tilt perturbations.

Our goal in this paper is to study tilt stability of general optimization problems (not necessarily smooth) on general Riemannian manifolds (not necessarily embedded in Euclidean spaces). The first difficulty we encounter is the absence of a global linear structure on a Riemannian manifold, which obstructs a straightforward definition of tilt perturbation as in (1.2). We address this issue by noting that tilt stability is a local concept and can thus be defined by pullback to tangent spaces via the exponential map. To characterize Riemannian tilt stability, we propose appropriate extensions of several basic concepts in stability analysis (such as strong metric regularity, strong monotonicity, and uniform quadratic growth) and generalized differentiation (such as generalized Hessian) to the Riemannian setting. We prove that the following conditions are equivalent: Riemannian tilt stability, Riemannian variational strong convexity, Riemannian uniform quadratic growth, local strong monotonicity of the Riemannian subdifferential, strong metric regularity of the Riemannian subdifferential, and positive definiteness of the generalized Riemannian Hessian. A notable feature of our results is that we do not require subdifferential continuity (in its Riemannian form). Next we obtain explicit characterization of Riemannian tilt stability for nonlinear programming problems on Riemannian manifolds under the assumption of Riemannian metric subregularity constraint qualification, which is weaker than other Riemannian constraint qualifications such as Riemannian Mangasarian-Fromovitz constraint qualification. To

[1] This phenomenon already arises in Euclidean spaces [1]. Consider a function $f : \mathbb{R}^n \to \mathbb{R}$ that is nearly flat over a large region containing x_K, yet decreases sharply towards a minimizer $\bar{x}$ located outside that region. In this case both $\|x_K - \bar{x}\|$ and $|f(x_K) - f(\bar{x})|$ are large despite that $\|\nabla f(x_K)\|$ is small.

illustrate the utility of Riemannian tilt stability in convergence analysis, we propose a generalized Riemannian Newton method for nonsmooth Riemannian optimization problems and establish its superlinear convergence under Riemannian tilt stability.

Now we briefly review related work in the literature.

Tilt stability in Euclidean spaces. Tilt stability is proposed in [2] for extended-real-valued functions on Euclidean spaces, as a strong form of optimality condition that supports computational practice. Under both prox-regularity and subdifferential continuity, it is proved in [2] that the following conditions are equivalent: tilt stability, variational strong convexity (though not explicitly named as such), local maximal strong monotonicity of the subdifferential, and positive definiteness of the generalized Hessian (introduced in [3, 4]). In [5], it is proved for prox-regular and C^2-partly smooth functions that tilt stability, strong criticality, and quadratic growth are equivalent. In [6], under the assumption of both prox-regularity and subdifferential continuity, it is shown that tilt stability, uniform quadratic growth, and strong metric regularity of the subdifferential are equivalent. It is also revealed in [6] that prox-regularity is essential for establishing characterizations of tilt stability since it is implied by either uniform quadratic growth or strong metric regularity of the subdifferential. On the other hand, as indicated by the results in [5], subdifferential continuity does not seem to be essential for characterizing tilt stability. In [7, 8], as a byproduct of results established for variational strong convexity, it is proved, under prox-regularity of f and without assuming subdifferential continuity, that tilt stability, variational strong convexity, f-attentive uniform quadratic growth, local strong monotonicity of the f-attentive subdifferential, strong metric regularity of the f-attentive subdifferential, and positive definiteness of the f-attentive generalized Hessian are all equivalent. In [9], it is proved, under prox-regularity alone, that tilt stability is equivalent to uniform positive definiteness of the quadratic bundle (introduced in [10–12]). To apply these general results to problems with specific structures one needs calculations of second-order information. In [13], calculus results for generalized Hessian are developed and applied to second-order characterizations of tilt stable minimizers for important classes of constrained optimization problems. In particular, for nonlinear programming, it is shown in [13] that under the linear independence constraint qualification (LICQ), tilt stability is equivalent to the strong second-order sufficient condition (SSOSC). Thus, under LICQ, tilt stability for nonlinear programming is equivalent to Robinson's strong regularity [14] of the KKT system. However, without assuming LICQ (which is too restrictive), tilt stability is weaker than strong regularity. Under both Mangasarian-Fromovitz constraint qualification (MFCQ) and constant rank constraint qualification (CRCQ), it is shown that SSOSC is sufficient for tilt stability [15] but not necessary [16]. The uniform second-order sufficient condition (USOSC) is introduced in [16] and is shown to be equivalent to tilt stability under MFCQ and CRCQ. In [17], point-based sufficient conditions for tilt stability are obtained under some weak constraint qualifications. In [18], under the metric subregularity constraint qualification (MSCQ, which is weaker than either MFCQ or CRCQ), it is proved that tilt stability for nonlinear programming is equivalent to the relaxed uniform second-order sufficient condition (RUSOSC). Tilt stability of other structured optimization problems have also been characterized under various assumptions [19–23].

Stability analysis on Riemannian manifolds. As Riemannian optimization methods have become increasingly popular and successful in various fields [24–26], stability analysis on Riemannian manifolds has begun to attract attention [27, 28]. To our best knowledge, tilt stability on Riemannian manifolds has not been studied in the literature. In [29], the concept of manifold augmented tilt stability is introduced. It refers to stability of the solution map of certain Lagrangian function with respect to both tilt perturbation and dual variable. Thus it is different from the Riemannian tilt stability introduced in our paper and is also more specialized.

Nonsmooth Newton methods on Riemannian manifolds. Problems from diverse applications can be formulated as (first/second-order) nonsmooth optimization problems on Riemannian manifolds [30] and many algorithms have been proposed [31–38]. In [39], a nonsmooth Riemannian Newton method for finding a zero of a locally Lipschitz continuous vector field on a Riemannian manifold is proposed. The algorithm of [39] is globalized in [40]. In [38], a globalized semismooth Riemannian Newton method for minimizing a $C^{1,1}$ function (i.e., a function whose Riemannian gradient vector field is locally Lipschitz continuous) on a Riemannian manifold is proposed. The major difference between our proposed generalized Riemannian Newton method and those in [38–40] is that our algorithm can deal with more general problems where the Riemannian gradient vector field may not exist (in which case it is replaced by the Riemannian subdifferential) or may fail to be locally Lipschitz continuous (when it does exist). We note that in these more general situations, Riemannian CD-regularity, the regularity condition used in [38–40] to guarantee local superlinear convergence, is not even well-defined. Nonetheless, under the assumption of Riemannian tilt stability, we are able to establish local superlinear convergence of the proposed generalized Riemannian Newton method, which underscores the importance of Riemannian tilt stability for algorithm analysis on Riemannian manifolds.

The remainder of the paper is organized as follows. In Section 2 preliminaries from Riemannian geometry and variational analysis are recalled. In Section 3 we define and characterize Riemannian tilt stability by extending some key concepts in stability analysis to Riemannian manifolds. For Riemannian nonlinear programming, we obtain explicit characterization of Riemannian tilt stability in terms of the initial data. In Section 4 we propose a generalized Riemannian Newton method and establish its superlinear convergence under Riemannian tilt stability. We draw some conclusions and discuss future directions in Section 5.

2 Preliminaries

We recall some preliminaries in Riemannian geometry and variational analysis.

2.1 Riemannian geometry

We recall some concepts in Riemannian geometry [24, 25, 41, 42].

A *smooth manifold* of dimension n is a Hausdorff and second-countable topological space that is locally homeomorphic to $\mathbb{R}^n$ via an atlas of charts whose transition maps are smooth (i.e., C^∞). The tangent space at $x \in M$ is the linear space T_xM of derivations of the algebra of germs at x of C^∞ functions on M. The disjoint union of all tangent spaces of M is called the *tangent bundle* of M and is denoted by TM. The associated projection map is denoted by $\pi : TM \to M, (x, v) \mapsto \pi(x, v) := x$. A *Riemannian manifold* is a smooth manifold with a Riemannian metric, i.e., a family of inner products $\langle \cdot, \cdot \rangle_x : T_xM \times T_xM \to \mathbb{R}$ that varies smoothly with $x \in M$. The associated norm at $x \in M$ is written $\| \cdot \|_x$.

Let M be a Riemannian manifold and let $\mathfrak{X}(M)$ denote the set of smooth vector fields on M. The *Levi–Civita connection* on a Riemannian manifold M is the unique map $\nabla : TM \times \mathfrak{X}(M), (u, V) \mapsto \nabla_u V$ satisfying some nice properties. Let $f : M \to \mathbb{R}$ a differentiable function. The *Riemannian gradient* $\operatorname{grad} f(x) \in T_xM$ of f at $x \in M$ is defined by

$$Df(x)(v) = \langle \operatorname{grad} f(x), v \rangle_x, \quad \forall\, v \in T_xM, \tag{2.1}$$

where $Df(x) : T_xM \to \mathbb{R}$ is the derivative of f at $\bar{x}$. The *Riemannian Hessian* $\operatorname{Hess} f(x) : T_xM \to T_xM$ of f at x is a linear map defined by

$$\operatorname{Hess} f(x)(u) := \nabla_u \operatorname{grad} f, \tag{2.2}$$

where ∇ is the Levi-Civita connection on M.

For $x, y \in M$, the Riemannian distance between x and y is defined as $d(x, y) := \inf_c \int_a^b \|c'(t)\|_{c(t)} dt$ where the infimum is taken over all piecewise regular curves on M which connect x to y. The Riemannian distance function defines a metric on each connected component of M. A *geodesic* on M is a curve on M whose acceleration (with respect to the Levi-Civita connection) is zero. Geodesics can also be characterized as locally minimizing (with respect to the Riemannian distance) curves.

Let O be the subset of TM consisting of $(x, v) \in TM$ such that the domain of $\gamma_{x,v}$ contains $[0, 1]$, where $\gamma_{x,v}$ the unique maximal geodesic with initial point x and initial velocity v. The *exponential map* $\exp : O \to M$ is defined by $(x, v) \mapsto \exp(x, v) = \exp_x(v) := \gamma_{x,v}(1)$. The exponential map is smooth, $\exp_x(0) = x$, and $D\exp_x(0) = \mathrm{Id}_{T_xM}$. If $V \subset T_xM$ is a sufficiently small open neighborhood of $0 \in T_xM$, then $\exp_x : V \to M$ is a diffeomorphism from V onto $U := \exp_x(V) \subset M$. In this case, U is called a *normal neighborhood* of x and the inverse $\exp_x^{-1} : U \to T_xM$ is a well-defined diffeomorphism from U onto V. For any $x \in M$, there exists a neighborhood W of x such that W is a normal neighborhood of x' for all $x' \in W$, i.e., for all $x' \in W$, $\exp_{x'}^{-1} : W \to T_{x'}M$ is well-defined. In this case W is called a *uniformly normal neighborhood* of x. Moreover, in a uniformly normal neighborhood W of x, for any two $x_1, x_2 \in W$, there exists a unique minimizing geodesic connecting x_1 to x_2.

Let c be a smooth curve on M. The parallel transport from $T_{c(t_0)}M$ to $T_{c(t_1)}M$ along c is the map $\mathrm{PT}^c_{c(t_0),c(t_1)} : T_{c(t_0)}M \to T_{c(t_1)}M$ defined by $\mathrm{PT}^c_{c(t_0),c(t_1)}(v) = Z(t_1)$ where Z is the unique parallel vector field with $Z(t_0) = v$. Here "parallel" is defined with respect to the Levi-Civita connection on M. When W is a uniformly normal neighborhood of $x \in M$, for any two $x_0, x_1 \in W$, we write PT_{x_0,x_1} for $\mathrm{PT}^c_{x_0,x_1}$ where c is the unique minimizing geodesic connecting x_0 to x_1.

2.2 Variational analysis

We recall some concepts in variational analysis [43–45].

Let E be a Euclidean space (i.e., a Hilbert space of finite dimension). The Fréchet/regular normal cone to $\Omega \subset E$ at $\bar{x} \in \Omega$ is

$$\widehat{N}_\Omega(\bar{x}) := \Big\{ v \in E : \limsup_{x \xrightarrow{\Omega} \bar{x}} \frac{\langle v, x - \bar{x} \rangle}{\|x - \bar{x}\|} \leq 0 \Big\}, \tag{2.3}$$

where $x \xrightarrow{\Omega} \bar{x}$ indicates that $x \to \bar{x}$ with $x \in \Omega$. The Mordukhovich/limiting normal cone to $\Omega \subset E$ at $\bar{x} \in \Omega$ is

$$N_\Omega(\bar{x}) = \{ v \in \mathbb{R}^n : \exists x_k \xrightarrow{\Omega} \bar{x}, v_k \to v \text{ as } k \to \infty \text{ with } v_k \in \widehat{N}_\Omega(x_k) \}. \tag{2.4}$$

Let E, F be Euclidean spaces and $S : E \rightrightarrows F$ a set-valued map with graph $\operatorname{gph} S \subset E \times F$. The Fréchet/regular coderivative and the Mordukhovich/limiting coderivative of S at $(\bar{x}, \bar{y}) \in \operatorname{gph} S$ are

$$\widehat{D}^* S(\bar{x}, \bar{y})(v) := \{ u \in E : (u, -v) \in \widehat{N}_{\operatorname{gph} S}(\bar{x}, \bar{y}) \}, \quad v \in F, \tag{2.5}$$

$$D^* S(\bar{x}, \bar{y})(v) := \{ u \in E : (u, -v) \in N_{\operatorname{gph} S}(\bar{x}, \bar{y}) \}, \quad v \in F. \tag{2.6}$$

When S is single-valued, we drop $\bar{y} = S(\bar{x})$ from the notation. If S is single-valued and C^1-smooth around $\bar{x}$, then by [43, Example 8.34], $\widehat{D}^* S(\bar{x})(v) = D^* S(\bar{x})(v) = \{ DS(\bar{x})^*(v) \}$, where $DS(\bar{x}) : E \to F$ is the derivative of S at $\bar{x}$ and $DS(\bar{x})^* : F \to E$ is the adjoint of $DS(\bar{x})$.

Let $f : E \to \overline{\mathbb{R}} := \mathbb{R} \cup \{\infty\}$ be an extended-real-valued function with effective domain $\operatorname{dom} f := \{ x \in E \,|\, f(x) \in \mathbb{R} \}$. The Fréchet/regular subdifferential and the Mordukhovich/limiting subdifferential of f at $\bar{x} \in \operatorname{dom} f$ are

$$\widehat{\partial} f(\bar{x}) := \{ v \in E \mid (v, -1) \in \widehat{N}_{\operatorname{epi} f}(\bar{x}, f(\bar{x})) \}, \tag{2.7}$$

$$\partial f(\bar{x}) := \{ v \in E \mid (v, -1) \in N_{\operatorname{epi} f}(\bar{x}, f(\bar{x})) \}. \tag{2.8}$$

For $\rho > f(\bar{x})$, the ρ-truncated f-attentive subdifferential of f is

$$\operatorname{gph} \partial_\rho f := \{ (x, v) \in \operatorname{gph} \partial f \mid f(x) < \rho \}. \tag{2.9}$$

Let $f : E \to \overline{\mathbb{R}}$ be an extended-real-valued function and $\bar{x} \in \operatorname{dom} f$. The generalized Hessian (or second-order subdifferential) of f at $(\bar{x}, \bar{v}) \in \operatorname{gph} \partial f$ is

$$\partial^2 f(\bar{x}, \bar{v}) := D^*(\partial f)(\bar{x}, \bar{v}). \tag{2.10}$$

For $\rho > f(\bar{x})$, the ρ-truncated f-attentive generalized Hessian of f at $(\bar{x}, \bar{v}) \in \operatorname{gph} \partial f$ is

$$\partial^2_\rho f(\bar{x}, \bar{v}) := D^*(\partial_\rho f)(\bar{x}, \bar{v}). \tag{2.11}$$

Definition 1 (tilt stability [2]) Let E be a Euclidean space. A point $\bar{x} \in E$ is called a tilt stable local minimizer of the function $f : E \to \overline{\mathbb{R}}$ if $f(\bar{x}) \in \mathbb{R}$ and there exist a neighborhood

$U \subset E$ of $\bar{x}$ and a neighborhood $V \subset E$ of 0 such that the following map is single-valued and Lipschitz continuous:

$$S_U : V \to E,\ v \mapsto S_U(v) := \operatorname*{argmin}_{x \in U} \{f(x) - \langle v, x - \bar{x}\rangle\}. \tag{2.12}$$

We will abbreviate "lower semicontinuous" as "lsc".

Definition 2 (prox-regularity) Let E be a Euclidean space. A function $f : E \to \overline{\mathbb{R}}$ is said to be prox-regular at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$ if it is lsc at $\bar{x}$ and there exist $\epsilon > 0$ and $r > 0$ such that

$$\begin{gathered} f(x') > f(x) + \langle v, x' - x\rangle - \frac{r}{2}\|x' - x\|^2 \quad \text{whenever} \\ \|x' - \bar{x}\| < \epsilon,\ \|x - \bar{x}\| < \epsilon,\ x \neq x',\ |f(x) - f(\bar{x})| < \epsilon,\ v \in \partial f(x). \end{gathered} \tag{2.13}$$

Definition 3 (subdifferential continuity) Let E be a Euclidean space. A function $f : E \to \overline{\mathbb{R}}$ is said to be subdifferentially continuous at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$ if for every $\delta > 0$, there exists $\epsilon > 0$ such that $|f(x) - f(\bar{x})| < \delta$ whenever $v \in \partial f(x)$ with $\|x - \bar{x}\| < \epsilon$ and $\|v - \bar{v}\| < \epsilon$.

An extended-real-valued function $f : M \to \overline{\mathbb{R}}$ on a Riemannian manifold is called *proper* if its effective domain $\operatorname{dom} f := \{x \in M \mid f(x) \in \mathbb{R}\}$ is nonempty. For $C \subset M$, its *indicator function* is the function $\delta_C : M \to \mathbb{R}$ with $\delta_C(x) = 0$ for $x \in C$ and $\delta_C(x) = \infty$ for $x \notin C$. A function $f : M \to \overline{\mathbb{R}}$ on a Riemannian manifold is called lower semicontinuous (lsc) at $\bar{x} \in M$ if $\liminf_{x \xrightarrow{M} \bar{x}} = f(\bar{x})$ where $\liminf_{x \xrightarrow{M} \bar{x}} := \sup_{V \in \mathcal{N}(\bar{x})} \left[\inf_{x \in V} f(x)\right]$. It follows that f is lsc at $\bar{x}$ if and only if $f \circ \exp_{\bar{x}} : T_{\bar{x}}M \to \overline{\mathbb{R}}$ is lsc at $0 \in T_{\bar{x}}M$.

3 Riemannian Tilt Stability

In this section we define and characterize tilt stability on Riemannian manifolds.

Let M be a Riemannian manifold. Consider an optimization problem on M

$$\min_{x \in M} f(x), \tag{3.1}$$

where $f : M \to \overline{\mathbb{R}}$ is a proper, extended-real-valued function. This formulation allows us to handle constraints $C \subset M$ implicitly via the indicator function $\delta_C : M \to \overline{\mathbb{R}}$.

Tilt stability in Euclidean spaces (see Definition 1) is a local concept, i.e., it is completely determined by the behavior of f around $\bar{x}$. This motivates us to define Riemannian tilt stability as follows.

Definition 4 (Riemannian tilt stability) A point $\bar{x} \in M$ is called a Riemannian tilt stable local minimizer of (3.1) if $f(\bar{x}) \in \mathbb{R}$ and there exist a neighborhood $U \subset M$ of $\bar{x}$ and a neighborhood $V \subset T_{\bar{x}}M$ of $0 \in T_{\bar{x}}M$ such that the following map is single-valued and Lipschitz continuous:

$$S_U : V \to M,\ v \mapsto S_U(v) := \operatorname*{argmin}_{x \in U} \left\{f(x) - \langle v, \exp_{\bar{x}}^{-1}(x)\rangle_{\bar{x}}\right\}, \tag{3.2}$$

where $\exp_{\bar{x}} : \mathcal{O}_{\bar{x}} \subset T_{\bar{x}}M \to M$ is the exponential map at $\bar{x}$,[2] $\exp_{\bar{x}}^{-1} : U \subset M \to T_{\bar{x}}M$ is the locally defined inverse of $\exp_{\bar{x}}$, and $\langle \cdot, \cdot \rangle_{\bar{x}}$ is the inner product on $T_{\bar{x}}M$. The Lipschitz continuity of S_U is understood with respect to the Riemannian distance function d on M: there exists a constant $\kappa \geq 0$ such that $d(S_U(v), S_U(v')) \leq \kappa \|v - v'\|$ for all $v, v' \in V$. The infimum of all such κ is called the tilt stability modulus of f at $\bar{x}$.

We list the following observations as evidence that the proposed Definition 4 is a sensible generalization of tilt stability [2] to Riemannian manifolds.

- When M is a Euclidean space, Definition 4 reduces to the usual (i.e., Euclidean) tilt stability [2] (see Definition 1) since $\exp_{\bar{x}}^{-1}(x) = x - \bar{x}$ in this case.
- When M is embedded in a Euclidean space E so that a global linear structure is available, Riemannian tilt perturbation is equivalent to Euclidean tilt perturbation (defined via the global linear structure). More precisely, when $f : E \to \mathbb{R}$ is C^2, we prove (in Proposition A.1) that $\bar{x} \in M$ is a Riemannian tilt stable minimizer of $f|_M : M \to \mathbb{R}$ if and only if $\bar{x}$ is a Euclidean tilt stable minimizer of $\tilde{f} := f + \delta_M : E \to \overline{\mathbb{R}}$ where δ_M is the indicator function of M.
- In Euclidean spaces, an arbitrary (smooth, additive) perturbation to the objective function can be approximated, up to first-order, by a tilt perturbation. This remains true for Riemannian tilt perturbations. To see this, suppose a smooth perturbation $h : P \times M \to \mathbb{R}, (p, x) \mapsto h(p, x)$ (where P is a Riemannian manifold of parameters) is added to f in (3.1) to get the parametrized functions $x \mapsto f_p(x) := f(x) + h(p, x)$ and we are interested in the stability of the minimizer(s) $S(p)$ of f_p with respect to p. By [25, Proposition 5.44], $h(p, x) = h(p, \bar{x}) + \langle \operatorname{grad}_x h(p, \bar{x}), \exp_{\bar{x}}(x) \rangle + o(\|\exp_{\bar{x}}(x)\|) = h(p, \bar{x}) + \langle v, \exp_{\bar{x}}(x) \rangle + o(\|\exp_{\bar{x}}(x)\|)$ where $v := \operatorname{grad}_x h(p, \bar{x})$. Thus, up to first-order, the general perturbation $h(p, x)$ can be approximated by the tilt perturbation $\langle v, \exp_{\bar{x}}(x) \rangle$. This means that understanding Riemannian tilt stability is a first step towards more general perturbation analysis in Riemannian optimization.
- In Euclidean spaces, tilt stability is closely related to numerical methodology and convergence analysis for optimization algorithms. This remains true for Riemannian tilt stability, as we now illustrate. Suppose a Riemannian optimization algorithm generates $\{x_k \in M\}_{k \geq 0}$. A common stopping criterion is $\|\operatorname{grad} f(x_K)\|_{x_K} < \epsilon$. In general, this condition cannot guarantee that x_K is actually close to a minimizer $\bar{x}$. However, if $\bar{x}$ is a Riemannian tilt stable minimizer, we can ensure that $d(x_K, \bar{x}) < C\epsilon$ for some $C > 0$. See Proposition A.2 for details.

Riemannian tilt stability is related to Euclidean tilt stability as explained in the lemma below, which follows from the fact that $\exp_{\bar{x}}^{-1}$ is a diffeomorphism (see, e.g., [25, Corollary 10.25]).

Lemma 3.1 Let M be a Riemannian manifold and $f : M \to \overline{\mathbb{R}}$. For a point $\bar{x} \in \operatorname{dom} f$, a neighborhood $U \subset M$ of $\bar{x}$ on which $\exp_{\bar{x}}^{-1}$ is defined, and $v \in T_{\bar{x}}M$, consider the following two optimization problems:

$$\min_{x \in U \subset M} f_v^m(x) := f(x) - \langle v, \exp_{\bar{x}}^{-1}(x) \rangle, \tag{3.3}$$

[2] Here $\mathcal{O}_{\bar{x}} := \mathcal{O} \cap T_{\bar{x}}M$ is the domain of $\exp_{\bar{x}}$.

and

$$\min_{s\in\exp_{\bar x}^{-1}(U)\subset T_{\bar x}M} f_v^e(s) := f(\exp_{\bar x}(s)) - \langle v, s\rangle\,. \tag{3.4}$$

Then $x_v \in M$ is a minimizer of (3.3) if and only if $s = \exp_{\bar x}^{-1}(x_v)$ is a minimizer of (3.4).

We recall the definition of Riemannian subdifferentials.

Definition 5 (Riemannian subdifferentials, [46]) Let M be a Riemannian manifold and $f : M \to \overline{\mathbb{R}}$ be lsc at $\bar x \in \operatorname{dom} f$. The Riemannian Fréchet/regular subdifferential of f at $\bar x$ is

$$\widehat{\partial}_R f(\bar x) := \widehat{\partial}(f\circ \exp_{\bar x})(0),$$

where $\widehat{\partial}(f\circ\exp_{\bar x})(0) \subset T_{\bar x}M$ is the Fréchet/regular subdifferential of $f\circ\exp_{\bar x} : T_{\bar x}M \to \overline{\mathbb{R}}$ at 0. The Riemannian Mordukhovich/limiting subdifferential of f at $\bar x$ is

$$\partial_R f(\bar x) := \{v \in T_{\bar x}M \mid \exists x_k \in M, \exists v_k \in \widehat{\partial}_R f(x_k) \subset T_{x_k}M,\ \text{s.t. } x_k \to \bar x, \mathrm{PT}_{x_k,\bar x}(v_k) \to v\}.$$

The Riemannian Mordukhovich/limiting normal cone of $C \subset M$ at $\bar x \in C$ is $N_C^R(\bar x) := \partial_R\delta_C(\bar x) \subset T_{\bar x}M$ where δ_C is the indicator function of C.

We define the f-attentive Riemannian subdifferential of f, which will be used in our later characterizations of Riemannian tilt stability.

Definition 6 (f-attentive Riemannian subdifferential) Let M be a Riemannian manifold and $f : M \to \overline{\mathbb{R}}$ be lsc at $\bar x \in \operatorname{dom} f$. Given $\rho > f(\bar x)$, the f-attentive Riemannian subdifferential $\partial_{R,\rho} f$ is defined by

$$\operatorname{gph}\partial_{R,\rho} f := \{(x, v) \in \operatorname{gph}\partial_R f \mid f(x) < \rho\}. \tag{3.5}$$

The following special chain rule for Riemannian subdifferential plays an important role in our later characterizations of Riemannian tilt stability.

Proposition 3.2 (special chain rule for Riemannian subdifferential) *Let M be a Riemannian manifold, $f : M \to \overline{\mathbb{R}}$ be lsc at $\bar x \in \operatorname{dom} f$, and $V \subset T_{\bar x}M$ an open neighborhood of $0 \in T_{\bar x}M$ such that $\exp_{\bar x} : V \to M$ a diffeomorphism from V onto $\exp_{\bar x}(V) \subset M$. For $s \in T_{\bar x}M$ small,*

$$\partial(f\circ\exp_{\bar x})(s) = D\exp_{\bar x}(s)^*(\partial_R f(\exp_{\bar x}(s))), \tag{3.6}$$

where $D\exp_{\bar x}(s)^* : T_{\exp_{\bar x}(s)}M \to T_{\bar x}M$ is the adjoint map of $D\exp_{\bar x}(s) : T_{\bar x}M \to T_{\exp_{\bar x}(s)}M$ and $\partial_R f(\exp_{\bar x}(s)) \subset T_{\exp_{\bar x}(s)}M$ is the Riemannian subdifferential of f at $\exp_{\bar x}(s)$.

Proof We first prove $\partial(f\circ\exp_{\bar x})(s) \subset D\exp_{\bar x}(s)^*(\partial_R f(\exp_{\bar x}(s)))$. Given $v \in \partial(f\circ\exp_{\bar x})(s)$, by the definition of (Euclidean) limiting subdifferential, there exist $s_k \in T_{\bar x}M$ with $s_k \to s$ and $v_k \in \widehat{\partial}(f\circ\exp_{\bar x})(s_k)$ with $v_k \to v$. By the definition of $\widehat{\partial}(f\circ\exp_{\bar x})(s_k)$, we have

$$\liminf_{s_k' \xrightarrow{T_{\bar x}M} s_k} \frac{f(\exp_{\bar x}(s_k')) - f(\exp_{\bar x}(s_k)) - \langle v_k, s_k' - s_k\rangle_{\bar x}}{\|s_k' - s_k\|_{\bar x}} \geq 0. \tag{3.7}$$

Define $x_k := \exp_{\bar{x}}(s_k)$ and $v'_k := D\exp_{\bar{x}}(s_k)^{*,-1}(v_k)$. Then $x_k \to \exp_{\bar{x}}(s)$ (since $s_k \to s$) and $\mathrm{PT}_{x_k,\exp_{\bar{x}}(s)}(v'_k) = \mathrm{PT}_{x_k,\exp_{\bar{x}}(s)}(D\exp_{\bar{x}}(s_k)^{*,-1}(v_k)) \to D\exp_{\bar{x}}(s)^{*,-1}(v)$ since the exponential and parallel transport maps are smooth (and, in particular, continuous). Now we only need to show that $v'_k \in \widehat{\partial}_R f(x_k)$ to conclude that $D\exp_{\bar{x}}(s)^{*,-1}(v) \in \partial_R f(\exp_{\bar{x}}(s))$, which is equivalent to $v \in D\exp_{\bar{x}}(s)^*(\partial_R f(\exp_{\bar{x}}(s)))$. Thus we need to show that

$$\liminf_{s' \xrightarrow{T_{x_k}M} 0} \frac{f(\exp_{x_k}(s')) - f(\exp_{x_k}(0)) - \langle v'_k, s'\rangle_{x_k}}{\|s'\|_{x_k}} \geq 0. \tag{3.8}$$

To see this, note that by defining $s'_k := \exp_{\bar{x}}^{-1}(\exp_{x_k}(s'))$ we have $s' \to 0 \iff s'_k \to s_k$ (since $s_k = \exp_{\bar{x}}^{-1}(x_k)$) and

$$\begin{aligned}
&f(\exp_{x_k}(s')) - f(\exp_{x_k}(0)) - \langle v'_k, s'\rangle_{x_k}\\
=&f(\exp_{\bar{x}}(s'_k)) - f(\exp_{\bar{x}}(s_k)) - \langle v'_k, \exp_{x_k}^{-1}(\exp_{\bar{x}}(s'_k))\rangle_{x_k}\\
=&f(\exp_{\bar{x}}(s)) - f(\exp_{\bar{x}}(s_k)) - \langle v'_k, D\exp_{\bar{x}}(s_k)(s'_k - s_k)\rangle_{x_k} - \langle v'_k, o(\|s'_k - s_k\|_{\bar{x}})\rangle_{x_k}\\
=&f(\exp_{\bar{x}}(s)) - f(\exp_{\bar{x}}(s_k)) - \langle D\exp_{\bar{x}}(s_k)^*(v'_k), s'_k - s_k\rangle_{\bar{x}} + o(\|s'_k - s_k\|_{\bar{x}})\\
=&f(\exp_{\bar{x}}(s)) - f(\exp_{\bar{x}}(s_k)) - \langle v_k, s'_k - s_k\rangle_{\bar{x}} + o(\|s'_k - s_k\|_{\bar{x}})\\
\geq& o(\|s'_k - s_k\|_{\bar{x}})\\
=&o(\|s'\|_{x_k}),
\end{aligned} \tag{3.9}$$

where the second equality follows from Lemma A.3, the fourth equality follows from the definition of v'_k, and the inequality follows from (3.7).

We next prove $D\exp_{\bar{x}}(s)^*(\partial_R f(\exp_{\bar{x}}(s))) \subset \partial(f \circ \exp_{\bar{x}})(s)$. Given $v = D\exp_{\bar{x}}(s)^*(w)$ with $w \in \partial_R f(\exp_{\bar{x}}(s))$, by Definition 5, there exist $x_k \in M$ with $x_k \to \exp_{\bar{x}}(s)$ and $v'_k \in \widehat{\partial}_R f(x_k) \subset T_{\bar{x}_k}M$ with $\mathrm{PT}_{x_k,\exp_{\bar{x}}(s)}(v'_k) \to w$. By the definition of $\widehat{\partial}_R f(x_k)$, we have $v'_k \in \widehat{\partial}(f \circ \exp_{x_k})(0)$, which means that

$$\liminf_{s' \xrightarrow{T_{x_k}M} 0} \frac{f(\exp_{x_k}(s')) - f(\exp_{x_k}(0)) - \langle v'_k, s'\rangle_{x_k}}{\|s'\|_{x_k}} \geq 0. \tag{3.10}$$

Define $s_k := \exp_{\bar{x}}^{-1}(x_k) \in T_{\bar{x}}M$ and $v_k := D\exp_{\bar{x}}(s_k)^*(v'_k) \in T_{\bar{x}}M$. Then $s_k \to s$ (since $x_k \to \exp_{\bar{x}}(s)$) and $v_k = D\exp_{\bar{x}}(s_k)^*(v'_k) = D\exp_{\bar{x}}(s_k)^*(\mathrm{PT}_{\exp_{\bar{x}}(s),x_k}(\mathrm{PT}_{x_k,\exp_{\bar{x}}(s)}(v'_k))) \to D\exp_{\bar{x}}(s)^*(w) = v$ since $\mathrm{PT}_{x_k,\exp_{\bar{x}}(s)}(v'_k) \to w$ and $\mathrm{PT}_{\exp_{\bar{x}}(s),x_k} \to \mathrm{PT}_{\exp_{\bar{x}}(s),\exp_{\bar{x}}(s)} = \mathrm{Id}$ and $D\exp_{\bar{x}}(s_k) \to D\exp_{\bar{x}}(s)$. Now we only need to show that $v_k \in \widehat{\partial}(f \circ \exp_{\bar{x}})(s_k)$ to conclude that $v \in \partial(f \circ \exp_{\bar{x}})(s)$. Thus we need to show that

$$\liminf_{s'_k \xrightarrow{T_{\bar{x}}M} s_k} \frac{f(\exp_{\bar{x}}(s'_k)) - f(\exp_{\bar{x}}(s_k)) - \langle v_k, s'_k - s_k\rangle_{\bar{x}}}{\|s'_k - s_k\|_{\bar{x}}} \geq 0. \tag{3.11}$$

To see this, note that by defining $s' := \exp_{x_k}^{-1}(\exp_{\bar{x}}(s'_k))$ we have $s' \to 0 \iff s'_k \to s_k$ and

$$\begin{aligned}
&f(\exp_{\bar{x}}(s'_k)) - f(\exp_{\bar{x}}(s_k)) - \langle v_k, s'_k - s_k\rangle_{\bar{x}}\\
=&f(\exp_{x_k}(s')) - f(x_k) - \langle D\exp_{\bar{x}}(s_k)^*(v'_k), s'_k - s_k\rangle_{\bar{x}}\\
=&f(\exp_{x_k}(s')) - f(\exp_{x_k}(0)) - \langle v'_k, D\exp_{\bar{x}}(s_k)(s'_k - s_k)\rangle_{x_k}\\
=&f(\exp_{x_k}(s')) - f(\exp_{x_k}(0)) - \langle v'_k, \exp_{x_k}^{-1}(\exp_{\bar{x}}(s'_k))\rangle_{x_k} - \langle v'_k, o(\|s'_k - s_k\|_{\bar{x}})\rangle_{x_k}\\
=&f(\exp_{x_k}(s')) - f(\exp_{x_k}(0)) - \langle v'_k, s'\rangle_{x_k} + o(\|s'_k - s_k\|_{\bar{x}})\\
\geq& o(\|s'\|_{x_k}) + o(\|s'_k - s_k\|_{\bar{x}})\\
=&o(\|s'_k - s_k\|_{\bar{x}}),
\end{aligned} \tag{3.12}$$

where the third equality follows from Lemma A.3 and the inequality follows from (3.10). $\square$

A corollary of Proposition 3.2 is that the Riemannian limiting subdifferential can be computed via pullback.

Proposition 3.3 *Let M be a Riemannian manifold and $f : M \to \overline{\mathbb{R}}$ be lsc at $\bar{x} \in \operatorname{dom} f$. Then one has*

$$\partial_R f(\bar{x}) = \partial(f \circ \exp_{\bar{x}})(0), \tag{3.13}$$

Proof This follows from Proposition 3.2 with $s = 0$ since $D\exp_{\bar{x}}(0) = \mathrm{Id}_{T_{\bar{x}}M}$. $\square$

To obtain characterizations of Riemannian tilt stability, we need to extend several key concepts in variational analysis to Riemannian manifolds. We first introduce set-valued sections. The Riemannian subdifferential is a notable example.

Definition 7 (set-valued section) A set-valued section on a Riemannian manifold M is a set-valued map $S : M \rightrightarrows TM$ such that $\pi(S(x)) = \{x\}$ whenever $S(x) \neq \emptyset$, i.e., $S(x) \subset T_x M$ for all $x \in M$. The graph of S is $\operatorname{gph} S := \{(x, v) \in TM \mid v \in S(x)\}$.

Note that when $M = E$ is a Euclidean space, a set-valued section is just a set-valued map $S : E \rightrightarrows E$. In this case, the inverse of S is defined as $S^{-1} : E \rightrightarrows E, v \mapsto S^{-1}(v) := \{x \in E \mid v \in S(x)\}$. However, when M is a Riemannian manifold and $S : M \rightrightarrows TM$ is a set-valued section, if we strictly follow the definition of inverse in Euclidean spaces, we would get $S^{-1} = \pi$ for all S, which makes the concept useless. Instead we propose the concept of localized inverse, which is motivated by our later characterizations of Riemannian tilt stability.

Definition 8 (localized inverse of set-valued section) Let M be a Riemannian manifold and $S : M \rightrightarrows TM$ a set-valued section. The localized inverse of S at $\bar{x} \in \operatorname{dom} S$ is:

$$S^{-1,\bar{x}} : T_{\bar{x}}M \to M, v \mapsto S^{-1,\bar{x}}(v) := \left\{x \in M \mid v \in D\exp_{\bar{x}}(\exp_{\bar{x}}^{-1}(x))^*(S(x))\right\}. \tag{3.14}$$

Note that $x \in S^{-1,\bar{x}}(v)$ implicitly requires that x is close to $\bar{x}$ so that $\exp_{\bar{x}}^{-1}(x)$ is well-defined.

Definition 9 (localization of set-valued sections) Let M be a Riemannian manifold, $S : M \rightrightarrows TM$ a set-valued section, and $(\bar{x}, \bar{v}) \in \operatorname{gph} S$. Let $U \subset M$ be a neighborhood of $\bar{x}$ and $V \subset T_{\bar{x}}M$ a neighborhood of $\bar{v}$. We assume that U is small enough so that $\exp_{\bar{x}}^{-1}$ is well-defined on U. The neighborhood $U \times_{\bar{x}} V \subset TM$ of $(\bar{x}, \bar{v})$ in TM is defined as follows:

$$U \times_{\bar{x}} V := \left\{D\exp_{\bar{x}}(s)^{*,-1}(v) \mid v \in V,\, s = \exp_{\bar{x}}^{-1}(x),\, x \in U\right\}. \tag{3.15}$$

A localization of S at $(\bar{x}, \bar{v})$ is a set-valued section with graph $\operatorname{gph} S \cap (U \times_{\bar{x}} V)$.

Now we extend some key concepts in variational analysis to Riemannian manifolds.

Definition 10 (local monotonicity of set-valued section) Let M be a Riemannian manifold. A set-valued section $S : M \rightrightarrows TM$ is locally monotone at $\bar{x} \in M$ if there exists a neighborhood U of $\bar{x}$ such that for all $x, x' \in U$ and $v \in S(x), v' \in S(x')$,

$$\langle D\exp_{\bar{x}}(s')^*(v') - D\exp_{\bar{x}}(s)^*(v),\ s' - s\rangle_{\bar{x}} \geq 0, \tag{3.16}$$

where $s' := \exp_{\bar{x}}^{-1}(x')$ and $s := \exp_{\bar{x}}^{-1}(x)$. S is locally strongly monotone with constant $\sigma > 0$ at $\bar{x}$ if there exists a neighborhood U of $\bar{x}$ such that for all $x, x' \in U$ and $v \in S(x), v' \in S(x')$,

$$\langle D\exp_{\bar{x}}(s')^*(v') - D\exp_{\bar{x}}(s)^*(v),\ s' - s\rangle_{\bar{x}} \geq \sigma\|s' - s\|_{\bar{x}}^2, \tag{3.17}$$

where $s' := \exp_{\bar{x}}^{-1}(x')$ and $s := \exp_{\bar{x}}^{-1}(x)$. The supremum over all such σ is called the modulus of local strong monotonicity of S at $\bar{x}$.

Definition 11 (Riemannian strong metric regularity) Let M be a Riemannian manifold. A set-valued section $S : M \rightrightarrows TM$ is strongly metrically regular with constant $\kappa \geq 0$ at $(\bar{x}, \bar{v}) \in \operatorname{gph} S$, if $S^{-1,\bar{x}} : T_{\bar{x}} \rightrightarrows M$, the localized inverse of S at $\bar{x}$, has a localization at $(\bar{v}, \bar{x}) \in \operatorname{gph} S^{-1,\bar{x}}$ that is single-valued and Lipschitz continuous with constant κ. In other words, there exist neighborhoods $V \subset T_{\bar{x}}M$ of $\bar{v}$ and $U \subset M$ of $\bar{x}$ such that for all $v \in V$, $S^{-1,\bar{x}}(v) \cap U$ is a singleton and the single-valued map $V \to U : v \mapsto S^{-1,\bar{x}}(v) \cap U$ is Lipschitz continuous with κ. The infimum of all such κ is called the strong metric regularity modulus of S at $(\bar{x}, \bar{v})$.

Definition 12 (Riemannian uniform quadratic growth condition) An lsc function $f : M \to \overline{\mathbb{R}}$ on a Riemannian manifold M is said to satisfy the Riemannian uniform quadratic growth condition with constant $\sigma > 0$ at $\bar{x} \in \operatorname{dom} f$ if there exist $\rho > f(\bar{x})$ and neighborhoods $U \subset M$ of $\bar{x}$ and $V \subset T_{\bar{x}}M$ of $0 \in T_{\bar{x}}M$ such that for all $(x, v) \in \operatorname{gph} \partial_R f \cap (U \times_{\bar{x}} V)$ with $f(x) < \rho$ and for all $x' \in U$,

$$f(x') \geq f(x) + \langle D\exp_{\bar{x}}(\exp_{\bar{x}}^{-1}(x))^*(v), \exp_{\bar{x}}^{-1}(x') - \exp_{\bar{x}}^{-1}(x)\rangle_{\bar{x}} + \sigma\|\exp_{\bar{x}}^{-1}(x') - \exp_{\bar{x}}^{-1}(x)\|_{\bar{x}}^2.$$

The following concept is introduced in [29]. Our formulation is slightly different.

Definition 13 (Riemannian variational strong convexity) Let M be a Riemannian manifold. An lsc function $f : M \to \overline{\mathbb{R}}$ satisfies Riemannian variational strong convexity at $\bar{x} \in M$ for $\bar{v} \in \partial_R f(\bar{x})$ if there exists an lsc function $h : M \to \overline{\mathbb{R}}$ such that (1) $h \leq f$ in a neighborhood of $\bar{x}$; (2) $\widehat{h} := h \circ \exp_{\bar{x}} : T_{\bar{x}}M \to \overline{\mathbb{R}}$ is locally strongly convex with modulus σ around $0 \in T_{\bar{x}}M$; (3) there exist $\rho > f(\bar{x})$ and neighborhoods $U \subset M$ of $\bar{x}$ and $V \subset T_{\bar{x}}M$ of 0 such that $\operatorname{gph} \partial_R h \cap (U \times_{\bar{x}} V) = \operatorname{gph} \partial_R f \cap (U_\rho \times_{\bar{x}} V)$ and $h(x) = f(x)$ for all $x \in \pi(\operatorname{gph} \partial_R h \cap (U \times_{\bar{x}} V))$, where $U_\rho := \{x \in U \mid f(x) < \rho\}$ and $\pi : TM \to M$ is the projection.

Now we extend the generalized Hessian [3] to Riemannian manifolds.

Definition 14 (generalized Riemannian Hessian) Let M be a Riemannian manifold and $f : M \to \overline{\mathbb{R}}$ be lsc at $\bar{x} \in \operatorname{dom} f$. The generalized Riemannian Hessian of f at $(\bar{x}, \bar{v}) \in \operatorname{gph} \partial_R f$ is the set-valued map $\partial_R^2 f(\bar{x}|\bar{v}) : T_{\bar{x}}M \rightrightarrows T_{\bar{x}}M$ defined as follows:

$$\forall u \in T_{\bar{x}}M,\ \partial_R^2 f(\bar{x}|\bar{v})(u) := \partial^2(f \circ \exp_{\bar{x}})(0|\bar{v})(u) = D^*\left(\partial(f \circ \exp_{\bar{x}})\right)(0|\bar{v})(u), \tag{3.18}$$

where $\partial(f \circ \exp_{\bar{x}})$ is the Euclidean subdifferential of $f \circ \exp_{\bar{x}} : T_{\bar{x}}M \to \overline{\mathbb{R}}$. Given $\rho > f(\bar{x})$, the ρ-truncated f-attentive generalized Riemannian Hessian of f at $(\bar{x}, \bar{v}) \in \operatorname{gph} \partial_R f$ is the set-valued map $\partial^2_{R,\rho} f(\bar{x}|\bar{v}) : T_{\bar{x}}M \rightrightarrows T_{\bar{x}}M$ defined as follows:

$$\forall u \in T_{\bar{x}}M,\ \partial^2_{R,\rho} f(\bar{x}|\bar{v})(u) := \partial^2_{\rho}(f \circ \exp_{\bar{x}})(0|\bar{v})(u) = D^* \left(\partial_{\rho}(f \circ \exp_{\bar{x}})\right)(0|\bar{v})(u), \quad (3.19)$$

where $\partial_{\rho}(f \circ \exp_{\bar{x}})$ is the Euclidean ρ-truncated f-attentive subdifferential of $f \circ \exp_{\bar{x}}$.

Definition 15 (Riemannian prox-regularity) An lsc function $f : M \to \overline{\mathbb{R}}$ on a Riemannian manifold M is called prox-regular at $\bar{x}$ for $\bar{v} \in \partial_R f(\bar{x})$ if $f \circ \exp_{\bar{x}} : T_{\bar{x}}M \to \overline{\mathbb{R}}$ is prox-regular at $0 \in T_{\bar{x}}M$ for $\bar{v}$.

Definition 16 (Riemannian subdifferential continuity) An lsc function $f : M \to \overline{\mathbb{R}}$ on a Riemannian manifold M is called subdifferentially continuous at $\bar{x}$ for $\bar{v} \in \partial_R f(\bar{x})$ if $f \circ \exp_{\bar{x}} : T_{\bar{x}}M \to \overline{\mathbb{R}}$ is subdifferentially continuous at $0 \in T_{\bar{x}}M$ for $\bar{v}$.

The following proposition follows from the definitions.

Proposition 3.4 *Let M be a Riemannian manifold and $f : M \to \overline{\mathbb{R}}$ be subdifferentially continuous at $(\bar{x}, \bar{v}) \in \operatorname{gph} \partial_R f$. Then for any $\rho > f(\bar{x})$, $\partial_{R,\rho} f = \partial_R f$ and $\partial^2_{R,\rho} f(\bar{x}, \bar{v}) = \partial^2_R f(\bar{x}, \bar{v})$.*

We now characterize Riemannian tilt stability in a quantitative way.

Theorem 3.5 (characterizations of Riemannian tilt stability) *Let $f : M \to \overline{\mathbb{R}}$ be lsc on a Riemannian manifold M and $\bar{x} \in M$ with $0 \in \partial_R f(\bar{x})$. Let $\kappa \geq 0$ be a real number and write $\widehat{f} := f \circ \exp_{\bar{x}} : T_{\bar{x}}M \to \overline{\mathbb{R}}$. Then the following statements are equivalent:*

(i) *f satisfies Riemannian variational strong convexity at $(\bar{x}, 0)$ with modulus $\frac{1}{\kappa}$.*

(ii) *f satisfies Riemannian uniform quadratic growth condition at $\bar{x}$ with modulus $\frac{1}{\kappa}$.*

(iii) *There exists $\rho > f(\bar{x})$ such that $\partial_{R,\rho} f$ has a localization $\operatorname{gph} \partial_{R,\rho} f \cap (U \times_{\bar{x}} V)$ at $(\bar{x}, 0)$ that is locally strongly monotone at $\bar{x}$ with modulus $\frac{1}{\kappa}$.*

(iv) *There exists $\rho > f(\bar{x})$ such that $\partial_{R,\rho} f$ is strongly metrically regular at $(\bar{x}, 0)$ with modulus κ;*

(v) *f satisfies Riemannian prox-regularity at $(\bar{x}, 0)$ and $\bar{x}$ is a Riemannian tilt stable minimizer of f with modulus κ;*

(vi) *f satisfies Riemannian prox-regularity at $(\bar{x}, 0)$ and there exists $\rho > f(\bar{x})$ such that $\partial^2_{R,\rho} f(\bar{x}|0)$ is positive definite with modulus $\frac{1}{\kappa}$:*

$$\langle z, w \rangle_{\bar{x}} \geq \frac{1}{\kappa} \|w\|^2_{\bar{x}} \quad \text{whenever} \quad z \in \partial^2_{R,\rho} f(\bar{x}|0)(w). \quad (3.20)$$

Proof We first prove the following claims.

Claim 1: $f : M \to \overline{\mathbb{R}}$ satisfies (i) if and only if $\widehat{f}$ is variationally strongly convex at $(0, 0)$. Suppose that (i) holds. We need to show that $\operatorname{gph} \widehat{h} \cap \left(\exp_{\bar{x}}^{-1}(U) \times V\right) = \operatorname{gph} \widehat{f} \cap \left(\exp_{\bar{x}}^{-1}(U)_{\rho} \times V\right)$ with $\widehat{h}(s) = \widehat{f}(s)$ for all $s \in \pi_1(\operatorname{gph} \widehat{h} \cap \left(\exp_{\bar{x}}^{-1}(U) \times V\right))$ where π_1 :

$T_{\bar{x}}M \times T_{\bar{x}}M \to T_{\bar{x}}M, (s, v) \mapsto \pi_1(s, v) := s$ is the projection to the first component. By the definition of localization, for $x \in U$ and $v \in D\exp_{\bar{x}}(\exp_{\bar{x}}^{-1}(x))^{*,-1}[V]$, $v \in \partial_R h(x)$ if and only if $v \in \partial_R f(x)$. Write $s := \exp_{\bar{x}}^{-1}(x)$. By Proposition 3.2, we have $v \in \partial_R h(x)$ if and only $v \in D\exp_{\bar{x}}(s)^{-1}\widehat{\partial} \widehat{h}(s)$ and similarly for f. Then the conclusion follows. The converse implication can be proved similarly.

Claim 2: f satisfies (ii) if and only if $\widehat{f}$ satisfies the uniform quadratic growth condition at $0 \in T_{\bar{x}}M$. To see this, suppose that f satisfies (ii). Write $s := \exp_{\bar{x}}^{-1}(x)$. Then for $x, x' \in U$ with $f(x) < \rho$ and $v \in \partial_R f(x) \cap D\exp_{\bar{x}}(s)^{*,-1}[V]$, we have

$$f(x') \geq f(x) + \langle D\exp_{\bar{x}}(s)^*[v], s' - s\rangle + \frac{1}{2\kappa}\|s' - s\|_{\bar{x}}^2. \tag{3.21}$$

By Proposition 3.2, $D\exp_{\bar{x}}(s)^*[v] \in \partial \widehat{f}(s)$. Then (3.21) implies that $\widehat{f}$ satisfies the uniform quadratic growth condition at $0 \in T_{\bar{x}}M$. The converse implication can be proved similarly.

Claim 3: $\partial_{R,\rho} f$ satisfies (iii) if and only if $\partial_\rho \widehat{f}$ has a localization $\operatorname{gph} \partial_\rho \widehat{f} \cap (\exp_{\bar{x}}^{-1}(U) \times V)$ at $(0, 0) \in \operatorname{gph} \partial \widehat{f}$ that is locally strongly monotone with modulus $\frac{1}{\kappa}$. To see this, suppose first that $\partial_{R,\rho} f$ satisfies (iii). Write $s = \exp_{\bar{x}}^{-1}(x)$. Then for $x, x' \in U$ and $v \in \partial_{R,\rho} f(x) \cap D\exp_{\bar{x}}(s)^{*,-1}[V]$, $v' \in \partial_{R,\rho} f(x') \cap D\exp_{\bar{x}}(s')^{*,-1}[V]$, we have

$$\langle D\exp_{\bar{x}}(s')^*[v'] - D\exp_{\bar{x}}(s)^*[v],\ s' - s\rangle_{\bar{x}} \geq \frac{1}{\kappa}\|s' - s\|_{\bar{x}}^2. \tag{3.22}$$

By Proposition 3.2, we have $D\exp_{\bar{x}}(\exp_{\bar{x}}^{-1}(x'))^*[v'] \in \partial_\rho(f \circ \exp_{\bar{x}})(s')$. Thus (3.22) implies that the localization $\operatorname{gph} \partial_\rho \widehat{f} \cap (\exp_{\bar{x}}^{-1}(U) \times V)$ is locally strongly monotone with modulus $\frac{1}{\kappa}$ at $0 \in T_{\bar{x}}M$. The converse implication can be proved similarly.

Claim 4: $\partial_{R,\rho} f$ satisfies (iv) if and only if $\partial_\rho \widehat{f}$ is strongly metrically regular with modulus κ at $(0, 0) \in \operatorname{gph} \partial \widehat{f}$. To see this, suppose that $\partial_{R,\rho} f$ satisfies (iv). By definition, $v \mapsto (\partial_{R,\rho} f)^{-1,\bar{x}}(v) \cap U$ is a single-valued and Lipschitz continuous with constant κ. Write $s = \exp_{\bar{x}}^{-1}(x)$. By Proposition 3.2, $(\partial_{R,\rho} f)^{-1,\bar{x}}(v) = \{x \in M \mid v \in D\exp_{\bar{x}}(s)^*[\partial_{R,\rho} f(x)] = \partial_\rho \widehat{f}(s)\} = \exp_{\bar{x}}((\partial_\rho \widehat{f})^{-1}(v))$. Then $(\partial_\rho \widehat{f})^{-1} = \exp_{\bar{x}}^{-1} \circ (\partial_{R,\rho} f)^{-1,\bar{x}}$. Since $\exp_{\bar{x}}^{-1}$ is Lipschitz continuous, it follows that $(\partial_\rho \widehat{f})^{-1}$ is single-valued and Lipschitz continuous, i,e., $\partial_\rho \widehat{f}$ is strongly metrically regular at $(0, 0)$. Moreover, since the exact Lipschitz modulus of $\exp_{\bar{x}}^{-1}$ is 1 (see [46, Theorem 1]), we conclude that the Lipschitz modulus of $(\partial_\rho \widehat{f})^{-1}$ coincides with that of $(\partial_{R,\rho} f)^{-1,\bar{x}}$. The converse implication is similar.

Claim 5: $\bar{x}$ satisfies (v) if and only if $\widehat{f}$ is prox-regular at $(0, 0) \in \operatorname{gph} \partial \widehat{f}$ and $0 \in T_{\bar{x}}M$ is a Riemannian tilt stable minimizer of $\widehat{f}$. This follows from Lemma 3.1.

Claim 6: $\partial_{R,\rho}^2 f(\bar{x}|0)$ satisfies (vi) if and only if $\partial_\rho^2 \widehat{f}(0|0)$ is positive definite. This follows from the definition of $\partial_{R,\rho}^2 f(\bar{x}|0)$.

In view of Claims 1-6, it follows from [8, Theorem 5.1, Proposition 3.5] that (i) $\iff$ (v) $\iff$ (vi). It follows from [8, Theorem 5.2] that (iv) $\iff$ (v). If we can verify that $0 \in \widehat{\partial} \widehat{f}(0)$, then it follows from [7, Theorem 2] that (i) $\iff$ (ii) $\iff$ (iii). To verify the condition $0 \in \widehat{\partial} \widehat{f}(0)$, we note that by [47, Theorem 1], prox-regularity of $\widehat{f}$ at 0 for $0 \in \partial \widehat{f}(0)$ is implied by each one of the conditions (i)(ii)(iii). This implies, by the definition of prox-regularity, that 0 is a proximal subgradient of $\widehat{f}$ at 0 and therefore 0 is a regular subgradient of $\widehat{f}$ at 0, i.e., $0 \in \widehat{\partial} \widehat{f}(0)$. The proof is thus completed. □

Theorem 3.6 (Riemannian tilt stability under Riemannian subdifferential continuity) *Let $f : M \to \overline{\mathbb{R}}$ be lsc on a Riemannian manifold M and $\bar{x} \in M$ with $0 \in \partial_R f(\bar{x})$. Assume that f is subdifferentially continuous at $(\bar{x}, 0)$. Let $\kappa \geq 0$ be a real number and write $\widehat{f} := f \circ \exp_{\bar{x}} : T_{\bar{x}}M \to \overline{\mathbb{R}}$. Then the following statements are equivalent:*

(i) f *satisfies Riemannian variational strong convexity at* $(\bar{x},0)$ *with modulus* $\frac{1}{\kappa}$.

(ii) f *satisfies Riemannian uniform quadratic growth condition at* $\bar{x}$ *with modulus* $\frac{1}{\kappa}$.

(iii) $\partial_R f$ *has a localization* $\operatorname{gph}\partial_R f \cap (U \times_{\bar{x}} V)$ *at* $(\bar{x},0)$ *that is locally strongly monotone at* $\bar{x}$ *with modulus* $\frac{1}{\kappa}$.

(iv) $\partial_R f$ *is strongly metrically regular at* $(\bar{x},0)$ *with modulus* κ*;*

(v) f *satisfies Riemannian prox-regularity at* $(\bar{x},0)$ *and* $\bar{x}$ *is a Riemannian tilt stable minimizer of* f *with modulus* κ*;*

(vi) f *satisfies Riemannian prox-regularity at* $(\bar{x},0)$ *and* $\partial_R^2 f(\bar{x}|0)$ *is positive definite with modulus* $\frac{1}{\kappa}$*:*

$$\langle z,w\rangle_{\bar{x}} \geq \frac{1}{\kappa}\|w\|_{\bar{x}}^2 \quad \textit{whenever} \quad z\in\partial_R^2 f(\bar{x}|0)(w). \tag{3.23}$$

Proof This follows from Theorem 3.5 and Proposition 3.4. □

Next we study Riemannian tilt stability for nonlinear programming problems on Riemannian manifolds. Let M be a Riemannian manifold and consider the Riemannian nonlinear programming problem

$$\min_{x\in M} f(x), \quad \text{s.t.} \quad g(x)\in\Theta, \tag{3.24}$$

where $f: M\to\mathbb{R}$, $g: M\to\mathbb{R}^l$ are C^2 functions and $\Theta := \{0\}^{l_1}\times\mathbb{R}_-^{l_2}\subset\mathbb{R}^l$ (with $l_1+l_2=l$). Write $\Gamma := \{x\in M \,|\, g(x)\in\Theta\}$ and $\tilde{f} := f+\delta_\Gamma : M\to\overline{\mathbb{R}}$ where δ_Γ is the indicator function of Γ. Problem (3.24) is the same as $\min\limits_{x\in M}\tilde{f}(x)$. A point $\bar{x}\in M$ is called a Riemannian tilt stable minimizer of (3.24) if it is a Riemannian tilt stable local minimizer of $\tilde{f}$.

Definition 17 (Riemannian MSCQ) For the nonlinear programming problem (3.24), the Riemannian metric subregularity constraint qualification (Riemannian MSCQ) is said to hold at $\bar{x}\in\Gamma$ if the set-valued map $G: M\rightrightarrows\mathbb{R}^l, x\mapsto G(x) := g(x)-\Theta$ is metrically subregular at $(\bar{x},0)\in\operatorname{gph}G$, i.e., there exist a neighborhood $U\subset M$ of $\bar{x}$ and a constant $\kappa\geq 0$ such that for all $x\in U$,

$$d(x;\Gamma) = d(x;G^{-1}(0)) \leq \kappa\, d(0;G(x)) = \kappa\, d(g(x);\Theta), \tag{3.25}$$

where d is the Riemannian distance on M and $d(x;\Gamma) := \inf\limits_{x'\in\Gamma} d(x,x')$.

Consider the following nonlinear programming problem in Euclidean spaces:

$$\min_{s\in V\subset T_{\bar{x}}M} f(\exp_{\bar{x}}(s)), \quad \text{s.t.} \quad g(\exp_{\bar{x}}(s))\in\Theta, \tag{3.26}$$

where V is a neighborhood of $0\in T_{\bar{x}}M$ on which $\exp_{\bar{x}}$ is well-defined.

Proposition 3.7 *The Riemannian MSCQ holds at* $\bar{x}$ *for problem* (3.24) *if and only if the (Euclidean) MSCQ holds at* $0\in T_{\bar{x}}M$ *for problem* (3.26).

Proof Suppose that (Euclidean) MSCQ holds at $0 \in T_{\bar{x}}M$ for problem (3.26). Write $\widehat{G}(s) := G(\exp_{\bar{x}}(s)) - \Theta$. This means that there exists a neighborhood $V' \subset T_{\bar{x}}M$ of 0 and a constant $\kappa \geq 0$ such that for all $s \in V'$,

$$d(s; \widehat{G}^{-1}(0)) \leq \kappa\, d(0; \widehat{G}(s)) = \kappa\, d(g(\exp_{\bar{x}}(s)); \Theta). \tag{3.27}$$

Note that $\widehat{G}^{-1}(0) = \exp_{\bar{x}}^{-1}(\Gamma)$. Then

$$d(x; \Gamma) = d(\exp_{\bar{x}}(s); \exp_{\bar{x}}(\widehat{G}^{-1}(0))) \leq Ld(s; \widehat{G}^{-1}(0)) \leq L\kappa\, d(g(\exp_{\bar{x}}(s)); \Theta), \tag{3.28}$$

where L is a Lipschitz constant for $\exp_{\bar{x}}$. The converse implication is similar. □

Proposition 3.8 *Suppose that Riemannian MSCQ holds at $\bar{x} \in \Gamma$ for (3.24). Then there exists a neighborhood $U \subset M$ of $\bar{x}$ such that for all $x \in \Gamma \cap U$,*

$$N_\Gamma^R(x) = Dg(x)^*(N_\Theta(g(x))). \tag{3.29}$$

Moreover, δ_Γ satisfies Riemannian prox-regularity and Riemannian subdifferential continuity at $(\bar{x}, \bar{v})$ for all $\bar{v} \in \partial_R \delta_\Gamma(\bar{x})$.

Proof This follows from Proposition 3.7 and [18, Lemma 4.2]. □

Proposition 3.9 (Riemannian KKT conditions) *If $\bar{x}$ is a local minimizer of (3.24) and the Riemannian MSCQ holds at $\bar{x}$, then the following KKT conditions hold: there exists $\bar{\lambda} \in \mathbb{R}^l$ with $\bar{\lambda}_{l_1+1}, \ldots, \bar{\lambda}_{l_1+l_2} \geq 0$ such that*

$$0 = \operatorname{grad} f(\bar{x}) + Dg(\bar{x})^*(\bar{\lambda}), \quad \bar{\lambda}_i = 0 \ \ \forall\, i \notin I(\bar{x}), \quad g(x) \in \Theta, \tag{3.30}$$

where $Dg(x)^* : \mathbb{R}^l \to T_{\bar{x}}M$ is the adjoint of the derivative $Dg(x) : T_{\bar{x}}M \to \mathbb{R}^l$ and $I(\bar{x}) := \{i \mid g_i(\bar{x}) = 0\}$. Denote by $L : M \times \mathbb{R}^l \to \mathbb{R}, (x, \lambda) \mapsto L(x, \lambda) := f(x) + \langle \lambda, g(x) \rangle$ the Riemannian Lagrangian function. Then the KKT conditions can be reformulated as

$$\operatorname{grad}_x L(\bar{x}, \bar{\lambda}) = 0, \quad \bar{\lambda}_i g_i(x) = 0\, (\forall i \in [l]), \quad g(x) \in \Theta. \tag{3.31}$$

Proof This follows from Proposition 3.7 and [18, (4.6)]. □

For $x \in \Gamma$ and $v \in N_\Gamma^R(x)$, we define

$$\Lambda(x, v) := \{\lambda \in \mathbb{R}^l \mid \lambda_{l_1+1}, \ldots, \lambda_{l_1+l_2} \geq 0,\ Dg(x)^*(\lambda) = v,\ \lambda_i = 0\, (\forall i \notin I(x))\}.$$

By Proposition 3.8, if Riemannian MSCQ holds at $\bar{x} \in \Gamma$, then $\Lambda(\bar{x}, \bar{v})$ is a nonempty convex polyhedral set for any $\bar{v} \in N_\Gamma^R(\bar{x})$. For each $s \in T_{\bar{x}}M$, consider the following linear program:

$$\min_\lambda - \langle s, \operatorname{Hess}_{xx}(\lambda^T g)(\bar{x})(s) \rangle_{\bar{x}}, \quad \text{s.t.} \quad \lambda \in \Lambda(\bar{x}, \bar{v}). \tag{3.32}$$

Its optimal solution set is denoted by $\Lambda(\bar{x}, \bar{v}; s)$.

Definition 18 (Riemannian relaxed uniform second-order sufficient condition) For (3.24), the Riemannian relaxed uniform second-order sufficient condition (Riemannian RUSOSC) holds at $\bar{x} \in \Gamma$ with modulus $l > 0$ if there exists $\eta > 0$ such that

$$\langle \mathrm{Hess}_{xx} L(x, \lambda)(w), w \rangle \geq l\|w\|^2, \tag{3.33}$$

whenever $(x, v) \in \mathrm{gph}\, \partial_R \tilde{f} \cap (\mathbb{B}_\eta(\bar{x}) \times_{\bar{x}} \mathbb{B}_\eta(0))$ and $\lambda \in \Lambda(x, v - \mathrm{grad}\, f(x); w)$ with w satisfying

$$\langle \mathrm{grad}\, g_i(x), w \rangle = 0, \ \ \forall i \in I^+(\lambda) \quad \text{and} \quad \langle \mathrm{grad}\, g_i(x), w \rangle \geq 0, \ \ \forall i \in I(x)/I^+(\lambda). \tag{3.34}$$

Theorem 3.10 (Riemannian tilt stability in Riemannian nonlinear programming) *Consider the Riemannian nonlinear programming problem* (3.24) *where f, g are C^2. Let $\bar{x} \in \Gamma$ be a stationary point of* (3.24) *at which the Riemannian MSCQ holds. Then $\bar{x}$ is a Riemannian tilt stable minimizer if and only if the Riemannian RUSOSC holds at $\bar{x}$.*

Proof This follows from Lemma 3.1 and [18, Theorem 4.5]. □

4 Generalized Riemannian Newton Methods

In this section we propose a generalized Riemannian Newton methods for minimizing a function $f : M \to \overline{\mathbb{R}}$ that is prox-regular and subdifferentially continuous in the Riemannian sense. This covers a large class of both unconstrained and constrained Riemannian optimization problems, including unconstrained $C^{1,1}$ minimization and Riemannian nonlinear programming problems. Under Riemannian tilt stability, we establish superlinear convergence of the algorithm .

Consider the following optimization problem:

$$\min_{x \in M} f(x), \tag{4.1}$$

where M is a Riemannian manifold and $f : M \to \overline{\mathbb{R}}$ is lsc.

Definition 19 (surrogate Riemannian Hessian) Let M be a Riemannian manifold and $f : M \to \overline{\mathbb{R}}$ an lsc function. A surrogate Riemannian Hessian of f is a map $H : (x, v) \in \mathrm{gph}\, \partial_R f \mapsto H(x, v) : T_x M \rightrightarrows T_x M$ such that for all $(x, v) \in \mathrm{gph}\, \partial_R f$, $\mathrm{gph}\, H(x, y) \subset T_x M \times T_x M$ is a cone, i.e., $0 \in H(x, v)(0)$ and $\lambda w \in H(x, v)(\lambda u)$ whenever $w \in H(x, v)(u)$ and $\lambda \geq 0$.

For applications in Newton-type methods, we require that $\partial_R f$ is g-semismooth with respect to the surrogate Riemannian Hessian H.

Definition 20 (g-semismoothness with respect to surrogate Riemannian Hessian) Let $f : M \to \overline{\mathbb{R}}$ be an lsc function on a Riemannian manifold M. We say that $\partial_R f$ is g-semismooth at $(\bar{x}, \bar{y}) \in \mathrm{gph}\, \partial_R f$ with respect to the surrogate Riemannian Hessian H if for any $\epsilon > 0$, there exists a neighborhood U of $\bar{x}$ and a neighborhood $V \subset T_{\bar{x}} M$ of $\bar{y}$ such that for all $(x, y) \in U \times_{\bar{x}} V$,

$$\|\mathrm{PT}_{y, \bar{y}}(y + w) - \bar{y}\|_{\bar{x}} \leq \epsilon\, d(x, \bar{x}), \quad \forall\, w \in H(x, y)(\exp_x^{-1}(\bar{x})). \tag{4.2}$$

Generalized derivatives in the Euclidean setting can be extended (via pullback) to the Riemannian setting to serve as surrogate Riemannian Hessian. Examples include Riemannian versions of graphical derivative $D\partial_R f(x|y) := D\partial(f \circ \exp_x)(0|y)$, subspace containing derivative $D_{\text{sc}}\partial_R f(x,y)$ (introduced in [48] for the Euclidean setting), limiting graphical derivative $D^\sharp \partial_R f(x,y)$ [48], strict graphical derivative $D_*\partial_R f(x,y)$, coderivative $D^*\partial_R f(x,y)$, and adjoint subspace containing derivative $D^*_{\text{scd}}\partial_R f(x,y)$ [48]. When f is $C^{1,1}$ (i.e., its Riemannian gradient is locally Lipschitz continuous), one can also consider Riemannian versions of B-subdifferential $\partial_B \operatorname{grad} f(x)$ and Clarke's generalized Jacobian $\partial_C \operatorname{grad} f(x)$ [38–40]. In practice, one can construct a surrogate Riemannian Hessian by exploiting the specific structure of $\partial_R f$ (such as using an upper estimate of certain generalized derivative of $\partial_R f$).

Given a surrogate Riemannian Hessian H, we define

$$\mathcal{A}_{\text{reg}} H(x,y) := \{A : T_x M \to T_x M \mid A \text{ is linear and } \forall w,\ A(w) \in H(x,y)^{-1}(w)\}.$$

For example, when f is C^2 and $H(x, \operatorname{grad} f(x)) = \operatorname{Hess} f(x)$, then $\mathcal{A}_{\text{reg}} H(x, \operatorname{grad} f(x))$ is nonempty if and only if $\operatorname{Hess} f(x)$ is invertible, in which case $\mathcal{A}_{\text{reg}} H(x, \operatorname{grad} f(x)) = \{\operatorname{Hess} f(x)^{-1}\}$.

We propose a framework of generalized Riemannian Newton methods below as Algorithm 1, which is similar in spirit to several generalized Newton methods in Euclidean spaces [49–52]. We remark that in the Newton step, we do not need to find A_k explicitly and the Newton direction d^k is computed by solving a linear system. We also note that implementation of the correction step is non-trivial in general; however, if f is $C^{1,1}$ as in [38–40], we can easily implement the correction step by setting $\hat{x}^k = x^k$ and $\hat{y}^k = \operatorname{grad} f(\hat{x}^k) = \operatorname{grad} f(x^k)$. Finaly we mention that in Riemannian optimization, iterates are updated by a retraction $R : TM \to M$ [24–26], which is more efficient than the exponential map.

Algorithm 1 A Framework of Generalized Riemannian Newton Methods

Input: problem (4.1), retraction R on M and initial iterate $x^0 \in M$.

1: **for** $k = 0, 1, \dots$ **do**
2: If $0 \in \partial_R f(x^k)$, stop.
3: **Correction Step**: Compute $(\hat{x}^k, \hat{y}^k) \in \operatorname{gph} \partial_R f$ that is close to $(x^k, 0)$ with $\mathcal{A}_{\text{reg}} H(\hat{x}^k, \hat{y}^k) \neq \emptyset$.
4: **Newton Step**: Compute $d^k = -A_k(\hat{y}^k)$ with $A_k \in \mathcal{A}_{\text{reg}} H(\hat{x}^k, \hat{y}^k)$ and set $x^{k+1} = R_{\hat{x}^k}(d^k)$.
5: **end for**

Theorem 4.1 (general convergence theorem) *Consider problem (4.1) and Algorithm 1. We assume the following conditions:*

(1) $\partial_R f$ *is g-semismooth at* $(\bar{x}, 0) \in \operatorname{gph} \partial_R f$ *with respect to* H*;*

(2) *There exist constants* $L, \kappa > 0$ *and a neighborhood* U *of* $\bar{x}$ *such that for all* $x \in U$*, the following set is nonempty:*

$$\mathcal{G}_{\bar{x}}^{L,\kappa,H}(x) := \{(\hat{x}, \hat{y}, A) \,|\, d(\hat{x}, \bar{x}) + \|\hat{y}\|_{\hat{x}} \le L d(x, \bar{x}), A \in \mathcal{A}_{\mathrm{reg}} H(\hat{x}, \hat{y}), \|A\| \le \kappa\}; \tag{4.3}$$

(3) $(\hat{x}^k, \hat{y}^k, A_k) \in \mathcal{G}_{\bar{x}}^{L,\kappa,H}(x^k)$ *for all* $k \ge 0$.

Then there exists a neighborhood $U_{\bar{x}}$ *of* $\bar{x}$ *such that whenever* $x^0 \in U_{\bar{x}}$*, Algorithm 1 either stops at a stationary point after finitely many iterations or converges superlinearly to* $\bar{x}$.

Proof We only have to consider the case where Algorithm 1 produces an infinite sequence of iterates $\{x^k \in M\}_{k \ge 0}$.

We first estimate $\|d^k - \exp_{\hat{x}^k}^{-1}(\bar{x})\|_{\hat{x}^k}$. By the definition of d^k, we have

$$\begin{aligned}
&\|d^k - \exp_{\hat{x}^k}^{-1}(\bar{x})\|_{\hat{x}^k} \\
=&\| - A_k(\hat{y}^k) - \exp_{\hat{x}^k}^{-1}(\bar{x})\|_{\hat{x}^k} \\
=&\|A_k(\hat{y}^k) + A_k(A_k^{-1}(\exp_{\hat{x}^k}^{-1}(\bar{x})))\|_{\hat{x}^k} \\
\le&\|A_k\| \, \|\hat{y}^k + A_k^{-1}(\exp_{\hat{x}^k}^{-1}(\bar{x}))\|_{\hat{x}^k} \\
\le&\kappa\epsilon \, d(\hat{x}^k, \bar{x}) \\
\le&\kappa L\epsilon \, d(x^k, \bar{x}),
\end{aligned} \tag{4.4}$$

where the second inequality follows from the g-semismoothness of $\partial_R f$ with respect to H at $(\bar{x}, 0)$ and the third inequality follows from assumption (3). By [38, Lemma 4.2] and the triangle inequality, there exists $C > 0$ such that

$$d(R_{\hat{x}^k}(d^k), \bar{x}) \le d(\exp_{\hat{x}^k}(d^k), \bar{x}) + C\|d^k\|_{\hat{x}^k}^2. \tag{4.5}$$

By definition and assumption (3),

$$\|d^k\|_{\hat{x}^k} = \| - A_k(\hat{y}^k)\|_{\hat{x}^k} \le \|A_k\|\|\hat{y}^k\|_{\hat{x}^k} \le \kappa L \, d(x^k, \bar{x}). \tag{4.6}$$

By [38, Lemma 4.1, (ii)], we get

$$\begin{aligned}
&d(\exp_{\hat{x}^k}(d^k), \bar{x}) \\
=&d(\exp_{\hat{x}^k}(d^k), \exp_{\hat{x}^k}(\exp_{\hat{x}^k}^{-1}(\bar{x}))) \\
=&\|d^k - \exp_{\hat{x}^k}^{-1}(\bar{x})\|_{\hat{x}^k} \\
\le&\kappa L\epsilon \, d(x^k, \bar{x}),
\end{aligned} \tag{4.7}$$

where the second equality follows from definition of the exponential map. Thus

$$d(R_{\hat{x}^k}(d^k), \bar{x}) \le \kappa L\epsilon \, d(x^k, \bar{x}) + C\kappa^2 L^2 \, d(x^k, \bar{x}) \le \epsilon d(x^k, \bar{x}). \tag{4.8}$$

By induction on k, all claims follow. □

Now we use the above general result to establish the superlinear convergence of a specific generalized Riemannian Newton method under Riemannian tilt stability.

Theorem 4.2 (convergence under Riemannian tilt stability) *Consider problem* (4.1) *and Algorithm 1 with* $H(x, y) := D^\sharp \partial_R f(x|y) := D^\sharp \partial(f \circ \exp_x)(0|y)$ *where* $D^\sharp$ *is the limiting graphical derivative* [48]. *We assume the following conditions:*

(1) $\partial_R f$ *is g-semismooth at* $\bar{x}$ *with respect to* $H := D^\sharp \partial_R f$;

(2) f is prox-regular and subdifferentially continuous (in the Riemannian sense) at $(\bar{x},0) \in \operatorname{gph} \partial_R f$ and $\bar{x}$ is a Riemannian tilt stable minimizer of f.

Then there exists a neighborhood $U_{\bar{x}}$ of $\bar{x}$ such that when $x^0 \in U_{\bar{x}}$, Algorithm 1 with $H(x,y) := D^\sharp \partial_R f$ either stops at a stationary point after finitely many steps or converges superlinearly to $\bar{x}$.

Proof By Theorem 4.1, we only need to show that $\mathcal{G}_{\bar{x}}^{L,\kappa,H}(x) \neq \emptyset$ for x close to $\bar{x}$. Take $\hat{x}$ close to x. By Theorem 3.6, $\partial_R f$ is strongly metrically regular at $(\bar{x},0)$. By the definition of localized inverse (Definition 8), we have

$$S^{-1,\hat{x}}(v_{\hat{x}}) := \{x \mid \hat{v}_{\hat{x}} \in D\exp_{\hat{x}}(\exp_{\hat{x}}^{-1}(x))^*(S(x))\} \tag{4.9}$$

and

$$S^{-1,\bar{x}}(v_{\bar{x}}) := \{x \mid v_{\bar{x}} \in D\exp_{\bar{x}}(\exp_{\bar{x}}^{-1}(x))^*(S(x))\}. \tag{4.10}$$

We will write $S^{-1,\bar{x}} \cap U$ for the map $v \mapsto S^{-1,\bar{x}}(v) \cap U$. Since $S^{-1,\bar{x}} \cap U$ is single-valued and $v_{\hat{x}} = T_{\hat{x},x} \circ T_{\bar{x},x}^{-1}(v_{\bar{x}})$ where $T_{\hat{x},x} := D\exp_{\hat{x}}(\exp_{\hat{x}}^{-1}(x))^*$ and similarly for $T_{\bar{x},x}$, it follows that $S^{-1,\hat{x}} \cap U$ is also single-valued. Moreover,

$$d(S^{-1,\hat{x}}(v_{\hat{x}}), S^{-1,\hat{x}}(v'_{\hat{x}})) = d(S^{-1,\bar{x}}(v_{\bar{x}}), S^{-1,\bar{x}}(v'_{\bar{x}})) \leq \kappa\|v_{\bar{x}} - v'_{\bar{x}}\| \leq \kappa K\|v_{\hat{x}} - v'_{\hat{x}}\|, \tag{4.11}$$

where $K > 0$ is an upper bound of Lipschitz constants of $T_{\hat{x},x} \circ T_{\bar{x},x}^{-1}$ (which exists if we restrict to a uniformly normal neighborhood of $\bar{x}$ in which the exponential maps $\exp_x$ are local diffeomorphisms with exact Lipschitz modulus 1). Since $S^{-1,\hat{x}} \cap U$ is Lipschitz continuous, we can take $A \in \partial_B(\exp_{\hat{x}}^{-1} \circ \partial_R f^{-1,\hat{x}} \cap U)(\hat{y})$, which always exists [53]. Then $\|A\| \leq \kappa$ and for any $w \in T_{\hat{x}}M$, we get from [48, Lemma 3.10,Lemma 3.11]

$$Aw \in D^\sharp(\exp_{\hat{x}}^{-1} \circ \partial_R f^{-1,\hat{x}} \cap U)(\hat{y})(w), \tag{4.12}$$

which is equivalent to

$$w \in D^\sharp(\exp_{\hat{x}}^{-1} \circ \partial_R f^{-1,\hat{x}} \cap U)^{-1}(0,\hat{y})(Aw) \tag{4.13}$$

Moreover, we have

$$(\exp_{\hat{x}}^{-1} \circ \partial_R f^{-1,\hat{x}} \cap U)^{-1} = \partial(f \circ \exp_{\hat{x}}). \tag{4.14}$$

Thus it follows that

$$w \in D^\sharp \partial(f \circ \exp_{\hat{x}})(0,\hat{y})(Aw), \tag{4.15}$$

i.e.,

$$Aw \in D^\sharp \partial(f \circ \exp_{\hat{x}})(0,\hat{y})^{-1}(w) = D^\sharp \partial_R f(\hat{x},\hat{y})^{-1}(w) = H(\hat{x},\hat{y})^{-1}(w). \tag{4.16}$$

Thus $A \in \mathcal{A}_{\mathrm{reg}} H(\hat{x},\hat{y})$ and $\mathcal{G}_{\bar{x}}^{L,\kappa,H}(x) \neq \emptyset$. The proof is then completed. □

Remark 1 Our convergence result is local and we expect that the proposed algorithm can be globalized by using line search or other strategies. This is left as future work.

Remark 2 Since our emphasis in this section is mainly on the theoretical side, i.e., to illustrate the utility of Riemannian tilt stability for convergence analysis of generalized Riemannian Newton methods, we do not include implementations for concrete problems and corresponding numerical experiments, which will be explored in future work.

5 Conclusions

We extended tilt stability from Euclidean spaces to Riemannian manifolds. We proved comprehensive general characterizations of Riemannian tilt stability and also derived explicit conditions for Riemannian nonlinear programming under a weak constraint qualification. We proposed a generalized Riemannian Newton method and proved its superlinear convergence under Riemannian tilt stability. In the future, we will study Riemannian tilt stability in other structured optimization problems and explore applications of the proposed generalized Riemannian Newton method to nonsmooth optimization problems on Riemannian manifolds.

Appendix A Appendix

Proposition A.1 *Let M be an embedded submanifold of a Euclidean space E equipped with the induced metric and $f : E \to \mathbb{R}$ is a C^2 function. Then $\bar{x} \in M$ is a Riemannian tilt stable minimizer of $f|_M : M \to \mathbb{R}$ if and only if $\bar{x}$ if a tilt stable minimizer of $\tilde{f} := f + \delta_M$.*

Proof By Theorem 3.6, $\bar{x}$ is a Riemannian tilt stable minimizer of $f|_M$ if and only if the Riemannian Hessian $\operatorname{Hess} f(\bar{x}) : T_{\bar{x}}M \to T_{\bar{x}}M$ is positive definite. On the other hand, $\bar{x}$ is a Euclidean tilt stable minimizer of $\tilde{f}$ if and only if the generalized Hessian $\partial^2 \tilde{f}(\bar{x}|0) : E \rightrightarrows E$ is positive definite. By [5, Theorem 2.12], $\operatorname{grad} f|_M(\bar{x}) = 0$ if and only if $0 \in \partial \tilde{f}(\bar{x})$ and moreover,

$$\partial^2 \tilde{f}(\bar{x}|0)(w) = \begin{cases} \operatorname{Hess} f(\bar{x})(w) + N_{\bar{x}}M, & w \in T_{\bar{x}}M, \\ \emptyset, & w \notin T_{\bar{x}}M, \end{cases} \tag{A1}$$

where $N_{\bar{x}}M$ is the normal space of M at $\bar{x}$ (which is identical with the normal cone of M at $\bar{x}$). From this it follows that $\partial^2 \tilde{f}(\bar{x}|0)$ is positive definite if and only if $\operatorname{Hess} f(\bar{x})$ is positive definite. Then the claim follows. □

Proposition A.2 *Let M be a Riemannian manifold and $f : M \to \mathbb{R}$ a C^2 function. Suppose that a Riemannian optimization algorithm generates $\{x_k \in M\}_{k \geq 0}$. If $\bar{x}$ is a Riemannian tilt stable minimizer within a neighborhood $U \subset M$ with constant $\kappa > 0$, then there exist constants $C > 0$ and $\epsilon > 0$ such that whenever x_K enters U with $\|\operatorname{grad} f(x_K)\|_{x_K} < \epsilon$, one has $d(x_K, \bar{x}) \leq C \|\operatorname{grad} f(x_K)\|_{x_K} < C\epsilon$.*

Proof Consider the problem in (3.2) with $v = v_K := \left(D \exp_{\bar{x}}^{-1}(x_K)\right)^{*,-1} (\operatorname{grad} f(x_K))$. It follows from the chain rule that $\operatorname{grad} f_{v_K}(x_K) = 0$ where f_{v_K} is the objective function in (3.2) with $v = v_K$. Since $\bar{x}$ is a Riemannian tilt stable minimizer, by Theorem 3.6, a localization (specified by some $\epsilon > 0$) of $(\operatorname{grad} f)^{-1,\bar{x}}$ is single-valued within U. By Definition 8, we conclude that x_K is the only stationary point of f_{v_K} within U and therefore $x_K = S_U(v_K)$, i.e., x_K is the solution to the tilted problem with tilt perturbation v_K. Then $d(x_K, \bar{x}) = d(S(v_K), S(0)) \leq \kappa \|v_K - 0\|_{\bar{x}} \leq \kappa \left\| \left(D \exp_{\bar{x}}^{-1}(x_K)\right)^{*,-1} \right\| \|\operatorname{grad} f(x_K)\| < C\epsilon$ for some $C > 0$. Here C exists since $\exp_{\bar{x}}$ is Lipschitz continuous around $0 \in T_{\bar{x}}M$. □

Lemma A.3 Let M be a Riemannian manifold and $\bar{x} \in M$. Let $\bar{s} \in T_{\bar{x}}M$ is small enough so that $x := \exp_{\bar{x}}(\bar{s})$ lies in a totally normal neighborhood of $\bar{x}$. For $s \in T_{\bar{x}}M$ small,

$$\exp_x^{-1}(\exp_{\bar{x}}(s)) = D\exp_{\bar{x}}(\bar{s})[s - \bar{s}] + o(\|s - \bar{s}\|_{\bar{x}}) \text{ as } s \to \bar{s}. \tag{A2}$$

Proof We denote by g the map $s \in T_{\bar{x}}M \mapsto g(s) := \exp_x^{-1}(\exp_{\bar{x}}(s)) \in T_{\exp_{\bar{x}}(\bar{s})}M$ where $x := \exp_{\bar{x}}(\bar{s})$. By assumption, the map g is well-defined and smooth for s sufficiently small. We compute the derivative of g at $\bar{s}$ using the chain rule:

$$\begin{aligned} Dg(\bar{s}) =& D\exp_x^{-1}(\exp_{\bar{x}}(\bar{s})) \circ D\exp_{\bar{x}}(\bar{s}) \\ =& D\exp_x^{-1}(x) \circ D\exp_{\bar{x}}(\bar{s}) \\ =& (D\exp_x(0))^{-1} \circ D\exp_{\bar{x}}(\bar{s}) \\ =& \mathrm{Id}_{T_xM} \circ D\exp_{\bar{x}}(\bar{s}) \\ =& D\exp_{\bar{x}}(\bar{s}). \end{aligned} \tag{A3}$$

Then we have, by the definition of derivative,

$$\exp_x^{-1}(\exp_{\bar{x}}(s)) = g(s) = g(\bar{s}) + Dg(\bar{s})(s-\bar{s}) + o(\|s-\bar{s}\|_{\bar{x}}) = D\exp_{\bar{x}}(\bar{s}) + o(\|s-\bar{s}\|_{\bar{x}}) \text{ as } s \to \bar{s},$$

where the last equality follows from $g(\bar{s}) = \exp_x^{-1}(\exp_{\bar{x}}(\bar{s})) = \exp_x^{-1}(x) = 0$ and (A3). □

References

[1] Bertsekas, D.: Nonlinear Programming. Athena Scientific, Belmont (1999)

[2] Poliquin, R.A., Rockafellar, R.T.: Tilt stability of a local minimum. SIAM Journal on Optimization **8**(2), 287–299 (1998)

[3] Mordukhovich, B.S.: Sensitivity analysis in nonsmooth optimization. In: Field, D.A., Komkov, V. (eds.) Theoretical Aspects of Industrial Design. SIAM Volumes in Applied Mathematics, vol. 58, pp. 32–46. SIAM, Philadelphia (1992)

[4] Mordukhovich, B.: Lipschitzian stability of constraint systems and generalized equations. Nonlinear Analysis: Theory, Methods & Applications **22**(2), 173–206 (1994)

[5] Lewis, A.S., Zhang, S.: Partial smoothness, tilt stability, and generalized Hessians. SIAM Journal on Optimization **23**(1), 74–94 (2013)

[6] Drusvyatskiy, D., Lewis, A.S.: Tilt stability, uniform quadratic growth, and strong metric regularity of the subdifferential. SIAM Journal on Optimization **23**(1), 256–267 (2013)

[7] Rockafellar, R.T.: Variational convexity and the local monotonicity of subgradient mappings. Vietnam Journal of Mathematics **47**(3), 547–561 (2019)

[8] Gfrerer, H.: On second-order variational analysis of variational convexity of prox-regular functions. Set-Valued and Variational Analysis **33**(1), 8 (2025)

[9] Khanh, P.D., Mordukhovich, B.S., Phat, V.T., Viet, L.D.: Characterizations of variational convexity and tilt stability via quadratic bundles. arXiv preprint arXiv:2501.04629 (2025)

[10] Rockafellar, R.T.: Augmented Lagrangians and hidden convexity in sufficient conditions for local optimality. Mathematical Programming **198**(1), 159–194 (2023)

[11] Rockafellar, R.T.: Derivative tests for prox-regularity and the modulus of convexity. To appear in Journal of Convex Analysis (2025)

[12] Khanh, P.D., Mordukhovich, B.S., Phat, V.T., Viet, L.D.: Generalized twice differentiability and quadratic bundles in second-order variational analysis. arXiv preprint arXiv:2501.02067 (2025)

[13] Mordukhovich, B.S., Rockafellar, R.T.: Second-order subdifferential calculus with applications to tilt stability in optimization. SIAM Journal on Optimization **22**(3), 953–986 (2012)

[14] Robinson, S.M.: Strongly regular generalized equations. Mathematics of Operations Research **5**(1), 43–62 (1980)

[15] Mordukhovich, B.S., Outrata, J.V.: Tilt stability in nonlinear programming under mangasarian-fromovitz constraint qualification. Kybernetika **49**(3), 446–464 (2013)

[16] Mordukhovich, B.S., Nghia, T.T.: Second-order characterizations of tilt stability with applications to nonlinear programming. Mathematical Programming **149**(1), 83–104 (2015)

[17] Gfrerer, H., Mordukhovich, B.S.: Complete characterizations of tilt stability in nonlinear programming under weakest qualification conditions. SIAM Journal on Optimization **25**(4), 2081–2119 (2015)

[18] Chieu, N.H., Hien, L.V., Nghia, T.T.: Characterization of tilt stability via subgradient graphical derivative with applications to nonlinear programming. SIAM Journal on Optimization **28**(3), 2246–2273 (2018)

[19] Mordukhovich, B.S., Outrata, J.V., Sarabi, M.E.: Full stability of locally optimal solutions in second-order cone programs. SIAM Journal on Optimization **24**(4), 1581–1613 (2014)

[20] Mordukhovich, B.S., Nghia, T.T., Rockafellar, R.T.: Full stability in finite-dimensional optimization. Mathematics of Operations Research **40**(1), 226–252 (2015)

[21] Benko, M., Gfrerer, H., Mordukhovich, B.S.: Characterizations of tilt-stable minimizers in second-order cone programming. SIAM Journal on Optimization **29**(4), 3100–3130 (2019)

[22] Nghia, T.T.: Geometric characterizations of lipschitz stability for convex optimization problems. SIAM Journal on Optimization **35**(2), 927–958 (2025)

[23] Ding, C., Sarabi, E., Wang, S.: Characterizations of tilt-stable local minimizers of a class of matrix optimization problems. arXiv preprint arXiv:2503.03217 (2025)

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

[25] Boumal, N.: An Introduction to Optimization on Smooth Manifolds. Cambridge University Press, Cambridge (2023). https://www.nicolasboumal.net/book

[26] Hu, J., Liu, X., Wen, Z.-W., Yuan, Y.-X.: A brief introduction to manifold optimization. Journal of the Operations Research Society of China **8**(2), 199–248 (2020)

[27] Rebjock, Q., Boumal, N.: Fast convergence to non-isolated minima: four equivalent conditions for C^2 functions. Mathematical Programming, 1–49 (2024)

[28] Bao, C., Ding, C., Zhou, Y.: On the robust isolated calmness of a class of nonsmooth optimizations on Riemannian manifolds and its applications. Computational Optimization and Applications, 1–46 (2025)

[29] Zhou, Y., Ding, C., Zhang, Y.: On some perturbation properties of nonsmooth optimization on Riemannian manifolds with applications. arXiv preprint arXiv:2308.06793 (2025)

[30] Absil, P.-A., Hosseini, S.: A collection of nonsmooth Riemannian optimization problems. In: Nonsmooth Optimization and Its Applications, pp. 1–15. Springer, Berlin (2019)

[31] Ferreira, O., Oliveira, P.: Subgradient algorithm on Riemannian manifolds. Journal of Optimization Theory and Applications **97**(1), 93–104 (1998)

[32] Borckmans, P.B., Selvan, S.E., Boumal, N., Absil, P.-A.: A Riemannian subgradient algorithm for economic dispatch with valve-point effect. Journal of computational and applied mathematics **255**, 848–866 (2014)

[33] Ferreira, O., Oliveira, P.: Proximal point algorithm on Riemannian manifolds. Optimization **51**(2), 257–270 (2002)

[34] Chen, S., Ma, S., Man-Cho So, A., Zhang, T.: Proximal gradient method for nonsmooth optimization over the Stiefel manifold. SIAM Journal on Optimization **30**(1), 210–239 (2020)

[35] Huang, W., Wei, K.: Riemannian proximal gradient methods. Mathematical Programming **194**(1), 371–413 (2022)

[36] Lai, R., Osher, S.: A splitting method for orthogonality constrained problems. Journal of Scientific Computing **58**(2), 431–449 (2014)

[37] Kovnatsky, A., Glashoff, K., Bronstein, M.M.: MADMM: a generic algorithm for non-smooth optimization on manifolds. In: European Conference on Computer Vision, pp. 680–696 (2016). Springer

[38] Zhou, Y., Bao, C., Ding, C., Zhu, J.: A semismooth Newton based augmented Lagrangian method for nonsmooth optimization on matrix manifolds. Mathematical Programming **201**(1), 1–61 (2023)

[39] Oliveira, F.R., Ferreira, O.P.: Newton method for finding a singularity of a special class of locally Lipschitz continuous vector fields on Riemannian manifolds. Journal of Optimization Theory and Applications **185**(2), 522–539 (2020)

[40] Oliveira, F.R., Oliveira, F.R.: A global Newton method for the nonsmooth vector fields on Riemannian manifolds. Journal of Optimization Theory and Applications **190**(1), 259–273 (2021)

[41] Do Carmo, M.P.: Riemannian Geometry vol. 2. Springer, Berlin (1992)

[42] Lee, J.M.: Riemannian Manifolds: An Introduction to Curvature vol. 176. Springer, Berlin (2006)

[43] Rockafellar, R.T., Wets, J.-B.: Variational Analysis. Springer, Berlin (2009)

[44] Mordukhovich, B.S.: Variational Analysis and Generalized Differentiation, I: Basic Theory, II: Applications vol. 331. Springer, Berlin (2006)

[45] Mordukhovich, B.S.: Variational Analysis and Applications. Springer, Berlin (2018)

[46] Farrokhiniya, M., Barani, A.: Limiting subdifferential calculus and perturbed distance function in Riemannian manifolds. Journal of Global Optimization **77**(3), 661–685 (2020)

[47] Rockafellar, R.T.: Variational convexity and prox-regularity. J. Convex Anal (2024)

[48] Gfrerer, H., Outrata, J.V.: On (local) analysis of multifunctions via subspaces contained in graphs of generalized derivatives. Journal of Mathematical Analysis and Applications **508**(2), 125895 (2022)

[49] Gfrerer, H., Outrata, J.V.: On a semismooth* Newton method for solving generalized equations. SIAM Journal on Optimization **31**(1), 489–517 (2021)

[50] Mordukhovich, B.S., Sarabi, M.E.: Generalized Newton algorithms for tilt-stable minimizers in nonsmooth optimization. SIAM Journal on Optimization **31**(2), 1184–1214 (2021)

[51] Khanh, P.D., Mordukhovich, B.S., Phat, V.T., Tran, D.B.: Globally convergent coderivative-based generalized Newton methods in nonsmooth optimization. Mathematical Programming, 1–57 (2023)

[52] Duy Khanh, P., Mordukhovich, B., Phat, V.T.: A generalized Newton method for subgradient systems. Mathematics of Operations Research **48**(4), 1811–1845 (2023)

[53] Clarke, F.H.: Optimization and Nonsmooth Analysis. SIAM, Philadelphia (1990)

Statements and Declarations

Funding

This work was supported by National Natural Science Foundation of China (No. 12561052, No. 12061013) and Special Foundation for Guangxi Ba Gui Scholars (No. GXR-6BG2424008).

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Author Contributions

All authors contributed to the study. The first draft of the manuscript was written by Zijian Shi and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.