A Riemannian symmetric rank-one trust-region method*

Wen Huang† P.-A. Absil‡§ K. A. Gallivan†

January 15, 2014

Abstract

The well-known symmetric rank-one trust-region method—where the Hessian approximation is generated by the symmetric rank-one update—is generalized to the problem of minimizing a real-valued function over a d-dimensional Riemannian manifold. The generalization relies on basic differential-geometric concepts, such as tangent spaces, Riemannian metrics, and the Riemannian gradient, as well as on the more recent notions of (first-order) retraction and vector transport. The new method, called RTR-SR1, is shown to converge globally and $d+1$-step q-superlinearly to stationary points of the objective function. A limited-memory version, referred to as LRTR-SR1, is also introduced. In this context, novel efficient strategies are presented to construct a vector transport on a submanifold of a Euclidean space. Numerical experiments—Rayleigh quotient minimization on the sphere and a joint diagonalization problem on the Stiefel manifold—illustrate the value of the new methods.

Key words: Riemannian optimization; optimization on manifolds; symmetric rank-one update; Rayleigh quotient; joint diagonalization; Stiefel manifold

2010 Mathematics Subject Classification: 65K05, 90C48, 90C53

1 Introduction

We consider the problem

$$\min_{x \in \mathcal{M}} f(x) \tag{1}$$

of minimizing a smooth real-valued function f defined on a Riemannian manifold $\mathcal{M}$. Recently investigated application areas include image segmentation [RW12] and recognition [TVSC11], electrostatics and electronic structure calculation [WY12], finance and chemistry [Bor12], multilinear algebra [SL10, IAVD11], low-rank learning [MMBS11, BA11], and blind source separation [KS12, SAGQ12].

*This paper presents research results of the Belgian Network DYSCO (Dynamical Systems, Control, and Optimization), funded by the Interuniversity Attraction Poles Programme initiated by the Belgian Science Policy Office. This work was financially supported by the Belgian FRFC (Fonds de la Recherche Fondamentale Collective). This work was performed in part while the third author was a Visiting Professor at the Institut de mathématiques pures et appliquées (MAPA) at Université catholique de Louvain.

†Department of Mathematics, 208 Love Building, 1017 Academic Way, Florida State University, Tallahassee FL 32306-4510, USA

‡Department of Mathematical Engineering, ICTEAM Institute, Université catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium

§Corresponding author. E-mail: absil@inma.ucl.ac.be. Phone: +32-10-472597. Fax: +32-10-472180.

The wealth of applications has stimulated the development of general-purpose methods for (1)—see, e.g., [AMS08, RW12, SI13] and references therein—including the trust-region approach upon which we focus in this work. A well-known technique in optimization [CGT00], the trust-region method was extended to Riemannian manifolds in [ABG07] (or see [AMS08, Ch. 7]), and found applications, e.g., in [JBAS10, VV10, IAVD11, MMBS11, BA11]. Trust-region methods construct a quadratic model m_k of the objective function f around the current iterate x_k and produce a candidate new iterate by (approximately) minimizing the model m_k within a region where it is "trusted". Depending on the discrepancy between f and m_k at the candidate new iterate, the size of the trust region is updated and the candidate new iterate is accepted or rejected.

For lack of efficient techniques to produce a second-order term in m_k that is inexact but nevertheless guarantees superlinear convergence, the Riemannian trust-region (RTR) framework loses some of its appeal when the exact second-order term—the Hessian of f—is not available. This is in contrast with the Euclidean case, where several strategies exist to build an inexact second-order term that preserves superlinear convergence of the trust-region method. Among these strategies, the symmetric rank-one (SR1) update is favored in view of its simplicity and because it preserves symmetry without unnecessarily enforcing positive definiteness; see, e.g., [NW06, §6.2] for a more detailed discussion. The $n+1$ step q-superlinear rate of convergence of the SR1 trust-region method was shown by Byrd et al. [BKS96] using a sophisticated analysis that builds on [CGT91, KBS93].

The classical (Euclidean) SR1 trust-region method can also be viewed as a quasi-Newton method, enhanced with a trust-region globalization strategy. The idea of quasi-Newton methods on manifolds is not new [Gab82, §4.5], however, most of the literature of which we are aware restricts consideration to generalizing the Broyden–Fletcher–Goldfarb–Shanno (BFGS) quasi-Newton method combined with a line search strategy. Early work such as [BM06, SL10] used Riemannian BFGS methods for a specific application without an analysis of convergence properties, but more recently, systematic analyses of Riemannian BFGS methods based on the framework of retraction and vector transport developed in [ADM02, AMS08] have been made. Qi [Qi11] analyzed a version of Riemannian BFGS methods with retraction and vector transport restricted to exponential mapping and parallel translation and showed superlinear convergence using a Riemannian Dennis-Moré condition. Ring and Wirth [RW12] proposed and analyzed a Riemannian BFGS that avoids the restrictions on retraction and vector transport assumed by Qi but that needs to resort to the derivative of the retraction. Seibert et al. [SKH13] discussed the freedom available when generalizing BFGS to Riemannian manifolds and analyzed one generalization of BFGS method on Riemannian manifolds that are isometric to $\mathbb{R}^n$. Most recently, Huang [Hua13] developed a complete convergence theory that avoids the restrictions of Qi, Ring and Wirth, guarantees superlinear convergence for the Riemannian Broyden family of quasi-Newton methods (including a version of SR1), and facilitates efficient implementation.

In this paper, motivated by the situation described above, we introduce a generalization of the classical (i.e., Euclidean) SR1 trust-region method to the Riemannian setting (1). Besides making use of basic Riemannian geometric concepts (tangent space, Riemannian metric, gradient), the new method, called RTR-SR1, relies on the notions of retraction and vector transport introduced in [ADM02, AMS08]. A detailed global and local convergence analysis is given. A limited-memory version of RTR-SR1, referred to as LRTR-SR1, is also introduced. Numerical experiments show that the RTR-SR1 method displays the expected convergence properties. When the Hessian of f is not available, RTR-SR1 thus offers an attractive way of tackling (1) by a trust-region approach. Moreover, even when the Hessian of f *is* available, making use of it can be expensive computa-

tionally, and the numerical experiments show that ignoring the Hessian information and resorting instead to the RTR-SR1 approach can be beneficial.

Another contribution of this paper with respect to [BKS96] is an extension of the analysis to allow for inexact solutions of the trust-region subproblem—compare (10) with [BKS96, (2.4)]. This extension makes it possible to resort to inner iterations such as the Steihaug–Toint truncated CG method (see [AMS08, §7.3.2] for its Riemannian extension) while staying within the assumptions of the convergence analysis.

The paper is organized as follows. The RTR-SR1 method is stated and discussed in Section 2. The convergence analysis is carried out in Section 3. The limited-memory version is introduced in Section 4. Numerical experiments are reported in Section 5. Conclusions are drawn in Section 6.

2 The Riemannian SR1 trust-region method

The proposed Riemannian SR1 trust-region (RTR-SR1) method is described in Algorithm 1. The algorithm statement is commented in Section 2.1 and the important questions of representing tangent vectors and choosing the vector transport are discussed in Sections 2.2 and 2.3.

2.1 A guide to Algorithm 1

Algorithm 1 can be viewed as a Riemannian version of the classical (Euclidean) SR1 trust-region method (see, e.g., [NW06, Algorithm 6.2]). It can also be viewed as an SR1 version of the Riemannian trust-region framework [AMS08, Algorithm 10 p. 142]. Therefore, several pieces of information given in [AMS08, Ch. 7] remain relevant for Algorithm 1.

In particular, the algorithm statement makes use of standard Riemannian concepts that are described, e.g., in [O'N83, AMS08], such as the tangent space $\mathrm{T}_x\mathcal{M}$ to the manifold $\mathcal{M}$ at a point x, a Riemannian metric g, and the gradient $\operatorname{grad} f$ of a real-valued function f on $\mathcal{M}$. The algorithm statement also relies on the notion of retraction, introduced in [ADM02] (or see [AMS08, §4.1]). A *retraction* R on $\mathcal{M}$ is a smooth map from the tangent bundle $\mathrm{T}\mathcal{M}$ (i.e., the set of all tangent vectors to $\mathcal{M}$) onto $\mathcal{M}$ such that, for all $x \in \mathcal{M}$ and all $\xi_x \in \mathrm{T}_x\mathcal{M}$, the curve $t \mapsto R(t\xi_x)$ is tangent to ξ_x at $t = 0$. We let R_x denote the restriction of R to $\mathrm{T}_x\mathcal{M}$. The domain of R need not be the entire tangent bundle, but this is usually the case in practice, and in this work we assume throughout that R is defined wherever needed. Specific ways of constructing retractions are proposed in [ADM02, AMS08, AM12]; see also [WY12, JD13] for the important case of the Stiefel manifold.

Within the Riemannian trust-region framework, the characterizing aspect of Algorithm 1 lies in the update mechanism for the Hessian approximation $\mathcal{B}_k$. The proposed update mechanism, based on formula (3) and on Step 6 of Algorithm 1, is a rather straightforward Riemannian generalization of the classical SR1 update

$$B_{k+1} = B_k + \frac{(y_k - B_k s_k)(y_k - B_k s_k)^T}{(y_k - B_k s_k)^T s_k}.$$

Significantly less straightforward is the Riemannian generalization of the superlinear convergence result, as we will see in Section 3.4. (Observe that the local convergence result [AMS08, Theorem 7.4.11] does not apply here because the Hessian approximation condition [AMS08, (7.36)] is not guaranteed to hold.)

Algorithm 1 Riemannian trust region with symmetric rank-one update (RTR-SR1)

Input: Riemannian manifold $\mathcal{M}$ with Riemannian metric g; retraction R; isometric vector transport $\mathcal{T}_S$; differentiable real-valued objective function f on $\mathcal{M}$; initial iterate $x_0 \in \mathcal{M}$; initial Hessian approximation $\mathcal{B}_0$, symmetric with respect to g.

1: Choose $\Delta_0 > 0$, $\nu \in (0,1)$, $c \in (0, 0.1)$, $\tau_1 \in (0,1)$ and $\tau_2 > 1$; Set $k \leftarrow 0$;

2: Obtain $s_k \in \mathrm{T}_{x_k}\mathcal{M}$ by (approximately) solving

$$s_k = \arg\min_{s \in \mathrm{T}_{x_k}\mathcal{M}} m_k(s) = \arg\min_{s \in \mathrm{T}_{x_k}\mathcal{M}} f(x_k) + g(\operatorname{grad} f(x_k), s) + \frac{1}{2} g(s, \mathcal{B}_k s), \text{s.t. } \|s\| \leq \Delta_k; \quad (2)$$

3: Set $\rho_k \leftarrow \frac{f(x_k) - f(R_{x_k}(s_k))}{m_k(0) - m_k(s_k)}$;

4: Let $y_k = \mathcal{T}_{S_{s_k}}^{-1} \operatorname{grad} f(R_{x_k}(s_k)) - \operatorname{grad} f(x_k)$; If $|g(s_k, y_k - \mathcal{B}_k s_k)| < \nu \|s_k\| \|y_k - \mathcal{B}_k s_k\|$, then $\tilde{\mathcal{B}}_{k+1} = \mathcal{B}_k$, otherwise define the linear operator $\tilde{\mathcal{B}}_{k+1} : \mathrm{T}_{x_k}\mathcal{M} \to \mathrm{T}_{x_k}\mathcal{M}$ by

$$\tilde{\mathcal{B}}_{k+1} = \mathcal{B}_k + \frac{(y_k - \mathcal{B}_k s_k)(y_k - \mathcal{B}_k s_k)^\flat}{g(s_k, y_k - \mathcal{B}_k s_k)}, \quad \text{(SR1)} \quad (3)$$

where $a^\flat$ denotes the flat of $a \in \mathrm{T}_x\mathcal{M}$, i.e., $a^\flat : \mathrm{T}_x\mathcal{M} \to \mathbb{R} : v \to g(a, v)$;

5: **if** $\rho_k > c$ **then**

6: $\quad x_{k+1} \leftarrow R_{x_k}(s_k)$; $\mathcal{B}_{k+1} \leftarrow \mathcal{T}_{S_{s_k}} \circ \tilde{\mathcal{B}}_{k+1} \circ \mathcal{T}_{S_{s_k}}^{-1}$;

7: **else**

8: $\quad x_{k+1} \leftarrow x_k$; $\mathcal{B}_{k+1} \leftarrow \tilde{\mathcal{B}}_{k+1}$;

9: **end if**

10: **if** $\rho_k > \frac{3}{4}$ **then**

11: $\quad$ **if** $\|s_k\| \geq 0.8\Delta_k$ **then**

12: $\quad\quad \Delta_{k+1} \leftarrow \tau_2 \Delta_k$;

13: $\quad$ **else**

14: $\quad\quad \Delta_{k+1} \leftarrow \Delta_k$;

15: $\quad$ **end if**

16: **else if** $\rho_k < 0.1$ **then**

17: $\quad \Delta_{k+1} \leftarrow \tau_1 \Delta_k$;

18: **else**

19: $\quad \Delta_{k+1} \leftarrow \Delta_k$;

20: **end if**

21: $k \leftarrow k + 1$, goto 2 until convergence.

Instrumental in the Riemannian SR1 update is the notion of vector transport, introduced in [AMS08, §8.1] as a generalization of the classical Riemannian concept of parallel translation. A *vector transport* on a manifold $\mathcal{M}$ on top of a retraction R is a smooth mapping

$$T\mathcal{M} \oplus T\mathcal{M} \to T\mathcal{M} : (\eta_x, \xi_x) \mapsto \mathcal{T}_{\eta_x}(\xi_x) \in T\mathcal{M}$$

satisfying the following properties for all $x \in \mathcal{M}$:

1. (Associated retraction) $\mathcal{T}_{\eta_x}(\xi_x) \in \mathrm{T}_{R_x(\xi_x)}\mathcal{M}$ for all $\xi_x \in \mathrm{T}_x\mathcal{M}$;

2. (Consistency) $\mathcal{T}_{0_x}(\xi_x) = \xi_x$ for all $\xi_x \in T_x\mathcal{M}$;

3. (Linearity) $\mathcal{T}_{\eta_x}(a\xi_x + b\zeta_x) = a\mathcal{T}_{\eta_x}(\xi_x) + b\mathcal{T}_{\eta_x}(\zeta_x)$.

The Riemannian SR1 update uses tangent vectors at the current iterate to produce a new Hessian approximation at the next iterate, hence the need to perform a vector transport (see Step 6) from the current iterate to the next.

In the Input step of Algorithm 1, the requirement that the vector transport $\mathcal{T}_S$ is isometric means that, for all $x \in \mathcal{M}$ and all $\xi_x, \zeta_x, \eta_x \in \mathrm{T}_x\,\mathcal{M}$, the equation

$$g(\mathcal{T}_{S_{\eta_x}}\xi_x, \mathcal{T}_{S_{\eta_x}}\zeta_x) = g(\xi_x, \zeta_x) \tag{4}$$

holds. Techniques for constructing an isometric vector transport on submanifolds of Euclidean spaces are described in Section 2.3.

The symmetry requirement on $\mathcal{B}_0$ with respect to the Riemannian metric g means that $g(\mathcal{B}_0\xi_{x_0}, \eta_{x_0}) = g(\xi_{x_0}, \mathcal{B}_0\eta_{x_0})$ for all $\xi_{x_0}, \eta_{x_0} \in \mathrm{T}_{x_0}\,\mathcal{M}$. It is readily seen from (3) and Step 6 of Algorithm 1 that $\mathcal{B}_k$ is symmetric for all k. Note however that $\mathcal{B}_k$ is, in general, not positive definite.

A possible stopping criterion for Algorithm 1 is $\|\operatorname{grad} f(x_k)\| < \epsilon$ for some specified $\epsilon > 0$, where $\|\cdot\|$, which also appears in the statement of Algorithm 1, denotes the norm induced by the Riemannian metric g, i.e.,

$$\|\xi\| = \sqrt{g(\xi, \xi)}. \tag{5}$$

In the spirit of [RW12, Remark 4], we point out that it is possible to formulate the SR1 update (3) in the new tangent space $\mathrm{T}_{x_{k+1}}\,\mathcal{M}$; in the present case of SR1, the algorithm remains equivalent since the vector transport is isometric.

Otherwise, Algorithm 1 does not call for comments other than those made in [AMS08, Ch. 7]. In particular, we point out that the meaning of "approximately" in Step 2 of Algorithm 1 depends on the desired convergence results. We will see in the convergence analysis (Section 3) that enforcing the Cauchy decrease (9) is enough to ensure global convergence to stationary points, but another condition such as (10) is needed to guarantee superlinear convergence. The truncated CG method, discussed in [AMS08, §7.3.2] in the Riemannian context, is an inner iteration for Step 2 that returns an s_k satisfying conditions (9) and (10).

2.2 Representation of tangent vectors

Let us now consider the frequently encountered situation where the manifold $\mathcal{M}$ is described as a d-dimensional submanifold of an m-dimensional Euclidean space $\mathcal{E}$. In particular, this is the case of the sphere and the Stiefel manifold involved in the numerical experiments in Section 5.

A tangent vector in $\mathrm{T}_x\,\mathcal{M}$ can be represented either by its d-dimensional vector of coordinates in a given basis B_x of $\mathrm{T}_x\,\mathcal{M}$, or else as an m-dimensional vector in $\mathcal{E}$ since $\mathrm{T}_x\,\mathcal{M} \subset \mathrm{T}_x\,\mathcal{E} \simeq \mathcal{E}$. The latter option may be preferable when the codimension $m - d$ is small (e.g., the sphere) because building, storing and manipulating the basis B_x of $\mathrm{T}_x\,\mathcal{M}$ may be inconvenient.

Likewise, since $\mathcal{B}_k$ is a linear transformation of $\mathrm{T}_{x_k}\,\mathcal{M}$, it can be represented in the basis B_x as a $d \times d$ matrix, or as an $m \times m$ matrix restricted to act on $\mathrm{T}_{x_k}\,\mathcal{M}$. Here again, the latter approach may be computationally more efficient when the codimension $m - d$ is small.

A related choice has to be made for the representation of the vector transport, since $\mathcal{T}_{\eta_x}$ is a linear map from $\mathrm{T}_x\,\mathcal{M}$ to $\mathrm{T}_{R_x(\eta_x)}\,\mathcal{M}$. This question is addressed in Section 2.3.

2.3 Isometric vector transport

We present two ways of constructing an isometric vector transport on a d-dimensional submanifold $\mathcal{M}$ of an m-dimensional Euclidean space $\mathcal{E}$.

2.3.1 Vector transport by parallelization

An open subset $\mathcal{U}$ of $\mathcal{M}$ is termed *parallelizable* if it admits a smooth field of tangent bases, i.e., a smooth function $B : \mathcal{U} \to \mathbb{R}^{m\times d} : z \mapsto B_z$ where B_z is a basis of $\mathrm{T}_z\,\mathcal{M}$. The whole manifold $\mathcal{M}$ itself may not be parallelizable; in particular, every manifold of nonzero Euler characteristic is not parallelizable [Sti35, §1.6], and it is also known that the only parallelizable spheres are those of dimension 1, 3, and 7 [Boo03, p. 116]. However, for the global convergence analysis carried out in Section 3.2, the vector transport is not required to be smooth or even continuous, and for the local convergence analysis in Section 3.4, we only need a parallelizable neighborhood $\mathcal{U}$ of the limit point x^*. Such a neighborhood always exists (take for example a coordinate neighborhood [AMS08, p. 37]).

Given an orthonormal smooth field of tangent bases B, i.e., such that $B_x^\flat B_x = I$ for all x (where I stands for the identity matrix of adequate size), the proposed isometric vector transport from $\mathrm{T}_x\,\mathcal{M}$ to $\mathrm{T}_y\,\mathcal{M}$ is given by

$$\mathcal{T} = B_y B_x^\flat. \tag{6}$$

The $d \times d$ matrix representation of this vector transport in the pair of bases (B_x, B_y) is simply the identity. This considerably simplifies the implementation of Algorithm 1.

2.3.2 Vector transport by rigging

If $\mathcal{M}$ is described as a d-dimensional submanifold of an m-dimensional Euclidean space $\mathcal{E}$ and the codimension $(m - d)$ is much smaller than the dimension d, then the vector transport by rigging, introduced next, may be preferable. For generality, we do not assume that $\mathcal{M}$ is a *Riemannian* submanifold of $\mathcal{E}$; in other words, the Riemannian metric g on $\mathcal{M}$ may not be the one induced by the metric of $\mathcal{E}$. A motivation for this generality is to be able to handle the canonical metric of the Stiefel manifold [EAS98, (2.22)]. For simplicity of the exposition, we work in an orthonormal basis of $\mathcal{E}$ and, for $x \in \mathcal{M}$, we let G_x denote a matrix expression of g_x, i.e., $g_x(\xi_x, \eta_x) = \xi_x^T G_x \eta_x$ for all $\xi_x, \eta_x \in \mathrm{T}_x\,\mathcal{M}$.

An open subset $\mathcal{U}$ of $\mathcal{M}$ is termed *rigged* if it admits a smooth field of normal bases, i.e., a smooth function $N : \mathcal{U} \to \mathbb{R}^{m\times d} : z \mapsto N_z$ where N_z is a basis of the normal space $\mathrm{N}_z\mathcal{M}$. The whole manifold $\mathcal{M}$ itself may not be rigged, but it is always locally rigged.

Given a smooth field of normal bases N, the proposed isometric vector transport $\mathcal{T}$ from $\mathrm{T}_x\,\mathcal{M}$ to $\mathrm{T}_y\,\mathcal{M}$ is defined as follows. Compute $(I - N_x(N_x^T N_x)^{-1} N_x^T)N_y$ (i.e., the orthogonal projection of N_y onto $\mathrm{T}_x\,\mathcal{M}$) and observe that its column space is $\mathrm{T}_x\,\mathcal{M} \ominus (\mathrm{T}_x\,\mathcal{M} \cap \mathrm{T}_y\,\mathcal{M})$. Obtain an orthonormal matrix Q_x by Gram-Schmidt orthonormalizing $(I - N_x(N_x^T N_x)^{-1} N_x^T)N_y$. Proceed likewise with x and y interchanged to get Q_y. Finally, let

$$\mathcal{T} = G_y^{-\frac{1}{2}}(I - Q_x Q_x^T - Q_y Q_x^T)G_x^{\frac{1}{2}}. \tag{7}$$

While it is clear that $\mathcal{T}$ satisfies the three properties of vector transport mentioned in Section 2.1, proving that $\mathcal{T}$ is (locally) smooth remains an open question. Moreover, the column space of

$(I - N_x(N_x^T N_x)^{-1} N_x^T) N_y$ gets more sensitive to numerical errors as the distance between x and y decreases. Nevertheless, there is evidence that $\mathcal{T}$ is smooth indeed, and we have observed that using vector transport by rigging in Algorithm 1 is a worthy alternative in large-scale low-codimension problems.

3 Convergence analysis of RTR-SR1

3.1 Notation and standing assumptions

Throughout the convergence analysis, unless otherwise specified, we let $\{x_k\}$, $\{\mathcal{B}_k\}$, $\{\tilde{\mathcal{B}}_k\}$, $\{s_k\}$, $\{y_k\}$, and $\{\Delta_k\}$ be infinite sequences generated by Algorithm 1, and we make use of the notation introduced in that algorithm. We let Ω denote the sublevel set of x_0, i.e.,

$$\Omega = \{x \in \mathcal{M} : f(x) \leq f(x_0)\}.$$

The global and local convergence analyses each make standing assumptions at the beginning of their respective sections. The numbered assumptions introduced below are not standing assumptions and will be invoked specifically whenever needed. Note that, apart from Assumption 3.6, all the numbered assumptions are Riemannian generalizations of assumptions made in [BKS96] for the analysis of the Euclidean SR1 trust-region method.

3.2 Global convergence analysis

In some results, we will assume for the retraction R that there exists $\mu > 0$ and $\delta_\mu > 0$ such that

$$\|\xi\| \geq \mu \operatorname{dist}(x, R_x(\xi)) \quad \text{for all } x \in \Omega, \text{ for all } \xi \in \mathrm{T}_x\mathcal{M}, \|\xi\| \leq \delta_\mu. \tag{8}$$

This corresponds to [AMS08, (7.25)] restricted to the sublevel set Ω. Such a condition is instrumental in the global convergence analysis of Riemannian trust-region schemes. Note that, in view of [RW12, Lemma 6], condition (8) can be shown to hold globally under the condition that R has equicontinuous derivatives.

The next assumption corresponds to [BKS96, (A3)].

Assumption 3.1. *The sequence of linear operators $\{\mathcal{B}_k\}$ is bounded by a constant M such that $\|\mathcal{B}_k\| \leq M$ for all k.*

We will often require that the trust-region subproblem (2) is solved accurately enough that, for some positive constants σ_1 and σ_2,

$$m_k(0) - m_k(s_k) \geq \sigma_1 \|\operatorname{grad} f(x_k)\| \min\{\Delta_k, \sigma_2 \frac{\|\operatorname{grad} f(x_k)\|}{\|\mathcal{B}_k\|}\}, \tag{9}$$

and that

$$\mathcal{B}_k s_k = -\operatorname{grad} f(x_k) + \delta_k \text{ with } \|\delta_k\| \leq \|\operatorname{grad} f(x_k)\|^{1+\theta}, \quad \text{whenever } \|s_k\| \leq 0.8\Delta_k, \tag{10}$$

where $\theta > 0$ is a constant. These conditions are generalizations of [BKS96, (2.3–4)]. Observe that, even if we restrict to the Euclidean case, condition (10) remains weaker than condition [BKS96, (2.4)]. The purpose of introducing δ_k in (10) is to encompass stopping criteria such as [AMS08,

(7.10)] that do not require the computation of an exact solution of the trust-region subproblem. We point out in particular that (9) and (10) hold if the approximate solution of the trust-region subproblem (2) is obtained from the truncated CG method, described in [AMS08, §7.3.2] in the Riemannian context.

We can now state and prove the main global convergence results. Point (iii) generalizes [BKS96, Theorem 2.1] while points (i) and (ii) are based on [AMS08, §7.4.1].

Theorem 3.1 (convergence)**.** *(i) If $f \in C^2$ is bounded below on the sublevel set Ω, Assumption 3.1 holds, condition* (9) *holds, and* (8) *is satisfied then* $\lim_{k\to\infty} \operatorname{grad} f(x_k) = 0$*. (ii) If $f \in C^2$, $\mathcal{M}$ is compact, Assumption 3.1 holds, and* (9) *holds then* $\lim_{k\to\infty} \operatorname{grad} f(x_k) = 0$*, $\{x_k\}$ has at least one limit point, and every limit point of $\{x_k\}$ is a stationary point of f. (iii) If $f \in C^2$, the sublevel set Ω is compact, f has a unique stationary point x^* in Ω, Assumption 3.1 holds, condition* (9) *holds, and* (8) *is satisfied then $\{x_k\}$ converges to x^*.*

Proof. (i) Observe that the proof of [AMS08, Theorem 7.4.4] still holds when condition [AMS08, (7.25)] is weakened to its restriction (8) to Ω. Indeed, since the trust-region method is a descent iteration, it follows that all iterates are in Ω. The assumptions thus allow us to conclude, by [AMS08, Theorem 7.4.4], that $\lim_{k\to\infty} \operatorname{grad} f(x_k) = 0$. (ii) It follows from [AMS08, Proposition 7.4.5] and [AMS08, Corollary 7.4.6] that all the assumptions of [AMS08, Theorem 7.4.4] hold. Hence $\lim_{k\to\infty} \operatorname{grad} f(x_k) = 0$, and every limit point is thus a stationary point of f. Since $\mathcal{M}$ is compact, $\{x_k\}$ is guaranteed to have at least one limit point. (iii) Again by [AMS08, Theorem 7.4.4], we get that $\lim_{k\to\infty} \operatorname{grad} f(x_k) = 0$. Since $\{x_k\}$ belongs to the compact set Ω and cannot have limit points other than x^*, it follows that $\{x_k\}$ converges to x^*. □

3.3 More notation and standing assumptions

For the purpose of conducting a local convergence analysis, we now assume that $\{x_k\}$ converges to a point x^*. Moreover, we assume throughout that $f \in C^2$.

We let $\mathcal{U}_{\text{trn}}$ be a *totally retractive neighborhood* of x^*, a concept inspired from the notion of totally normal neighborhood (see [dC92, §3.3]). By this, we mean that there is $\delta_{\text{trn}} > 0$ such that, for each $y \in \mathcal{U}_{\text{trn}}$, we have that $R_y(\mathbb{B}(0_y, \delta_{\text{trn}})) \supseteq \mathcal{U}_{\text{trn}}$ and $R_y(\cdot)$ is a diffeomorphism on $\mathbb{B}(0_y, \delta_{\text{trn}})$, where $\mathbb{B}(0_y, \delta_{\text{trn}})$ denotes the ball of radius δ_{trn} in $\mathrm{T}_y\mathcal{M}$ centered at the origin 0_y. The existence of a totally retractive neighborhood can be shown along the lines of [dC92, Theorem 3.3.7]. We assume without loss of generality that $\{x_k\} \subset \mathcal{U}_{\text{trn}}$. Whenever we consider an inverse retraction $R_x^{-1}(y)$, we implicitly assume that $x, y \in \mathcal{U}_{\text{trn}}$.

3.4 Local convergence analysis

The purpose of this section is to obtain a superlinear convergence result for Algorithm 1, stated in Theorem 3.18. The analysis can be viewed as a Riemannian generalization of the local analysis in [BKS96, §2]. As we proceed, we will point out the main hurdles that had to be overcome in the generalization. The analysis makes use of several preparation lemmas, independent of Algorithm 1, that are of potential interest in the broader context of Riemannian optimization. These preparation lemmas become trivial or well known in the Euclidean context.

The next assumption corresponds to a part of [BKS96, (A1)].

Assumption 3.2. *The point x^* is a nondegenerate local minimizer of f. In other words,* $\operatorname{grad} f(x^*) = 0$ *and* $\operatorname{Hess} f(x^*)$ *is positive definite.*

The next assumption generalizes the assumption, contained in [BKS96, (A1)], that the Hessian of f is Lipschitz continuous near x^*. (Recall that $\mathcal{T}_S$ is the vector transport invoked in Algorithm 1.) Note that the assumption holds if $f \in C^3$; see Lemma 3.5.

Assumption 3.3. *There exists a constant c_0 such that for all $x, y \in \mathcal{U}_{\mathrm{trn}}$,*

$$\| \operatorname{Hess} f(y) - \mathcal{T}_{S_\eta} \operatorname{Hess} f(x) \mathcal{T}_{S_\eta}^{-1} \| \leq c \operatorname{dist}(x, y),$$

where $\operatorname{Hess} f(x)$ *is the Riemannian Hessian of f at x (see, e.g., [AMS08, §5.5]), $\eta = R_x^{-1}(y)$, and $\| \cdot \|$ is also used to denote the operator norm induced by the Riemannian norm* (5).

The next assumption is introduced to handle the Riemannian case; in the classical Euclidean setting, Assumption 3.4 follows from Assumption 3.3. Assumption 3.4 is mild since it holds if $f \in C^3$, as shown in Lemma 3.5.

Assumption 3.4. *There exists a constant c_0 such that for all $x, y \in \mathcal{U}_{\mathrm{trn}}$, all $\xi_x \in \mathrm{T}_x \mathcal{M}$ with $R_x(\xi_x) \in \mathcal{U}_{\mathrm{trn}}$, and all $\xi_y \in \mathrm{T}_y \mathcal{M}$ with $R_y(\xi_y) \in \mathcal{U}_{\mathrm{trn}}$, the inequality*

$$\| \operatorname{Hess} \hat{f}_y(\xi_y) - \mathcal{T}_{S_\eta} \operatorname{Hess} \hat{f}_x(\xi_x) \mathcal{T}_{S_\eta}^{-1} \| \leq c_0 (\|\xi_y\| + \|\xi_x\| + \|\eta\|)$$

holds, where $\eta = R_x^{-1}(y)$, $\hat{f}_x = f \circ R_x$, and $\hat{f}_y = f \circ R_y$.

The next assumption corresponds to [BKS96, (A2)]. It implies that no updates of $\mathcal{B}_k$ are skipped. In the Euclidean case, Khalfan et al. [KBS93] show that this is usually the case in practice.

Assumption 3.5. *The inequality*

$$|g(s_k, y_k - \mathcal{B}_k s_k)| \geq \nu \|s_k\| \|y_k - \mathcal{B}_k s_k\|$$

holds.

The next assumption is introduced to handle the Riemannian case. It states that the iterates eventually *continuously stay* in the totally retractive neighborhood $\mathcal{U}_{\mathrm{trn}}$ (the terminology is borrowed from [ATV13, Definition 2.8]). The assumption is needed, in particular, for Lemma 3.6. Note that, whereas in the Euclidean setting the assumption follows from the standing assumption that $\{x_k\}$ converges to x^*, this is no longer the case on some Riemannian manifolds, where $\{x_k\}$ may converge to x^* while the connecting segments $\{R_{x_k}(ts_k) : t \in [0, 1]\}$ do not. Assumption 3.6 is thus invoked to ensure that we are in a position to carry out a *local* convergence analysis.

Assumption 3.6. *There exists N such that, for all $k \geq N$ and all $t \in [0, 1]$, $R_{x_k}(ts_k) \in \mathcal{U}_{\mathrm{trn}}$.*

The next lemma is proved in [GQA12, Lemma 14.1].

Lemma 3.2. *Let $\mathcal{M}$ be a Riemannian manifold, let $\mathcal{U}$ be a compact coordinate neighborhood in $\mathcal{M}$, and let the hat denote coordinate expressions. Then there are $c_2 > c_1 > 0$ such that, for all $x, y \in \mathcal{U}$, we have*

$$c_1 \|\hat{x} - \hat{y}\|_2 \leq \operatorname{dist}(x, y) \leq c_2 \|\hat{x} - \hat{y}\|_2,$$

where $\| \cdot \|_2$ denotes the Euclidean norm, i.e., $\|\hat{x}\|_2 = \sqrt{\hat{x}^T \hat{x}}$.

Lemma 3.3. *Let $\mathcal{M}$ be a Riemannian manifold endowed with a retraction R and let $\bar{x} \in \mathcal{M}$. Then there exist $a_0 > 0$, $a_1 > 0$, and $\delta_{a_0,a_1} > 0$ such that for all x in a sufficiently small neighborhood of $\bar{x}$ and all $\xi, \eta \in \mathrm{T}_x \mathcal{M}$ with $\|\xi\| \leq \delta_{a_0,a_1}$ and $\|\eta\| \leq \delta_{a_0,a_1}$, the inequalities*

$$a_0\|\xi - \eta\| \leq \operatorname{dist}(R_x(\eta), R_x(\xi)) \leq a_1\|\xi - \eta\|$$

hold.

Proof. Since R is smooth, we can choose a neighborhood small enough such that R satisfies the condition of [RW12, Lemma 6], and the result follows from that lemma. □

The following lemma follows from Lemma 3.3 by setting $\eta = 0$. We state it separately for convenience as we will frequently invoke it in the analysis.

Lemma 3.4. *Let $\mathcal{M}$ be a Riemannian manifold endowed with retraction R and let $\bar{x} \in \mathcal{M}$. Then there exist $a_0 > 0$, $a_1 > 0$, and $\delta_{a_0,a_1} > 0$ such that for all x in a sufficiently small neighborhood of $\bar{x}$ and all $\xi \in \mathrm{T}_x \mathcal{M}$ with $\|\xi\| \leq \delta_{a_0,a_1}$, the inequalities*

$$a_0\|\xi\| \leq \operatorname{dist}(x, R_x(\xi)) \leq a_1\|\xi\|$$

hold.

Lemma 3.5. *If $f \in C^3$, then Assumptions 3.3 and 3.4 hold.*

Proof. First, we prove that Assumption 3.3 holds. Define a function $h : \mathcal{M} \times \mathcal{M} \times \mathrm{T}\mathcal{M} \to \mathrm{T}\mathcal{M}, (x, y, \xi_y) \to \mathcal{T}_{S_\eta} \operatorname{Hess} f(x) \mathcal{T}_{S_\eta}^{-1} \xi_y$, where $\eta = R_x^{-1}(y)$. Since $f \in C^3$, we know that $h(x, y, \xi_y)$ is C^1. Then there exists b_0 such that for all $x, y \in \mathcal{U}_{\mathrm{trn}}$, $\xi_y \in \mathrm{T}_y \mathcal{M}, \|\xi_y\| = 1$,

$$\begin{aligned}
\|h(y, y, \xi_y) - h(x, y, \xi_y)\| &\leq b_0 \operatorname{dist}(\{y, y, \xi_y\}, \{x, y, \xi_y\}) \\
&\leq b_1 \|\{\hat{y}, \hat{y}, \hat{\xi}_y\} - \{\hat{x}, \hat{y}, \hat{\xi}_y\}\|_2 \text{ (by Lemma 3.2)} \\
&= b_1 \|\hat{y} - \hat{x}\|_2 \\
&\leq b_2 \operatorname{dist}(y, x), \text{ (by Lemma 3.2)}
\end{aligned}$$

where b_0, b_1 and b_2 are some constants. So we have

$$\begin{aligned}
b_2 \operatorname{dist}(y, x) &\geq \|h(y, y, \xi_y) - h(x, y, \xi_y)\| \\
&= \|(\operatorname{Hess} f(y) - \mathcal{T}_{S_\eta} \operatorname{Hess} f(x) \mathcal{T}_{S_\eta}^{-1}) \xi_y\|
\end{aligned}$$

Given any linear operator A on $\mathrm{T}_y \mathcal{M}$, we have $\|A\|$ by definition is $\sup_{\|\xi\|=1} \|A\xi\|$. Note that $\|\xi\| = 1$ is a compact set. Hence, there exists $\|\xi^*\| = 1$ such that $\|A\| = \|A\xi^*\|$. Therefore, we can choose $\xi_y, \|\xi_y\| = 1$ such that

$$\|(\operatorname{Hess} f(y) - \mathcal{T}_{S_\eta} \operatorname{Hess} f(x) \mathcal{T}_{S_\eta}^{-1}) \xi_y\| = \|(\operatorname{Hess} f(y) - \mathcal{T}_{S_\eta} \operatorname{Hess} f(x) \mathcal{T}_{S_\eta}^{-1})\|.$$

We obtain

$$\|\operatorname{Hess} f(y) - \mathcal{T}_{S_\eta} \operatorname{Hess} f(x) \mathcal{T}_{S_\eta}^{-1}\| \leq b_2 \operatorname{dist}(y, x).$$

To prove Assumption 3.4, we redefine h as $h(y, x, \xi_x) = \mathcal{T}_{S_\eta} \operatorname{Hess} \hat{f}_x(\xi_x)\mathcal{T}_{S_\eta}^{-1}$. If we use orthonormal vector fields to obtain the coordinate expression of h, denoted by $\hat{h}$, then the manifold norm and the Euclidean norm of coordinate expressions are the same and we have

$$\| \operatorname{Hess} \hat{f}_y(\xi_y) - \mathcal{T}_{S_\eta} \operatorname{Hess} \hat{f}_x(\xi_x)\mathcal{T}_{S_\eta}^{-1}\| = \| \operatorname{Hess} \hat{f}_y(\hat{\xi}_y) - \hat{\mathcal{T}}_{S_\eta} \operatorname{Hess} \hat{f}_x(\hat{\xi}_x)\hat{\mathcal{T}}_{S_\eta}^{-1}\|_2. \tag{11}$$

Since $f \in C^3$, we know that $\hat{h}$ is also in C^1. Hence there exists a constant b_3 such that

$$\|\hat{h}(\hat{y}, \hat{y}, \hat{\xi}_y) - \hat{h}(\hat{y}, \hat{x}, \hat{\xi}_x)\|_2 \leq b_3\|\{\hat{y}, \hat{y}, \hat{\xi}_y\} - \{\hat{y}, \hat{x}, \hat{\xi}_x\}\|_2.$$

Therefore

$$\begin{aligned}
\| \operatorname{Hess} \hat{f}_y(\hat{\xi}_y) - \hat{\mathcal{T}}_{S_\eta} \operatorname{Hess} \hat{f}_x(\hat{\xi}_x)\hat{\mathcal{T}}_{S_\eta}^{-1}\|_2 &= \|\hat{h}(\hat{y}, \hat{y}, \hat{\xi}_y) - \hat{h}(\hat{y}, \hat{x}, \hat{\xi}_x)\|_2 \\
&\leq b_3\|\{\hat{y}, \hat{y}, \hat{\xi}_y\} - \{\hat{y}, \hat{x}, \hat{\xi}_x\}\|_2 \\
&\leq b_4(\|\hat{y} - \hat{x}\|_2 + \|\hat{\xi}_y\|_2 + \|\hat{\xi}_x\|_2) \\
&\leq b_5(\operatorname{dist}(x, y) + \|\hat{\xi}_y\|_2 + \|\hat{\xi}_x\|_2) \text{ (by Lemma 3.2)} \\
&\leq b_6(\|\eta\| + \|\xi_y\| + \|\xi_x\|) \text{ (by Lemma 3.4)}
\end{aligned}$$

This and (11) give us Assumption 3.4. □

The next lemma generalizes [BKS96, Lemma 2.2]. The key difference with the Euclidean case is the following: in the Euclidean case, when s_k is accepted, we simply have $\|s_k\| = \|x_{k+1} - x_k\|$, while in the Riemannian generalization, we invoke Assumption 3.6 and Lemma 3.4 to deduce that $\|s_k\| \leq \frac{1}{a_0} \operatorname{dist}(x_{k+1}, x_k)$. Note that Assumption 3.6 cannot be removed. To see this, consider for example the unit sphere with the exponential retraction, where we can have $x_k = x_{k+1}$ with $\|s_k\| = 2\pi$.

Lemma 3.6. *Suppose Assumption 3.6 holds. Then either*

$$\Delta_k \to 0 \tag{12}$$

or there exist $K > 0$ and $\Delta > 0$ such that for all $k > K$

$$\Delta_k = \Delta. \tag{13}$$

In either case $s_k \to 0$.

Proof. Let $\Delta = \liminf \Delta_k$ and suppose first that $\Delta > 0$. From line 11 of Algorithm 1, if Δ_k is increased, then $\|s_k\| \geq 0.8\Delta_k$ and $x_{k+1} = R_{x_k}(s_k)$, which implies by Lemma 3.4 and Assumption 3.6 that $\operatorname{dist}(x_k, x_{k+1}) \geq a_0 0.8\Delta_k$. The latter inequality cannot hold for infinitely many values of k since $x_k \to x^*$ and $\liminf \Delta_k > 0$. Hence, there exists $K \geq 0$ such that Δ_k is not increased for any $k \geq K$. Since $\Delta > 0$, this implies that $\Delta_k \geq \Delta$ for all $k \geq K$. In view of the trust-region update mechanism in Algorithm 1 and since $\Delta = \liminf \Delta_k$, we also know that, for some $K_1 > K$, $\Delta_{K_1} < \frac{1}{\tau_1}\Delta$. If the trust region radius were to be decreased we would have $\Delta_{K_1+1} < \Delta$, which we have ruled out. Since neither increase nor decrease can occur, we must have $\Delta_k = \Delta$ for all $k \geq K_1$.

Suppose now that $\Delta = 0$. Since $x_k \to x^*$, for every $\epsilon > 0$ there exists $K_\epsilon \geq 0$ such that $\operatorname{dist}(x_{k+1}, x_k) < \epsilon$ for all $k \geq K_\epsilon$. Since $\liminf \Delta_k = 0$, there exists $j \geq K_\epsilon$ such that $\Delta_j < \epsilon$. But

since Δ_k is increased only if $\Delta_k \leq \frac{1}{0.8}\|s_k\| \leq \frac{1}{0.8a_0}\operatorname{dist}(x_{k+1}, x_k) < \frac{\epsilon}{0.8a_0}$, and the increase factor is τ_2, we have that $\Delta_k < \frac{\tau_2\epsilon}{0.8a_0}$ for all $k \geq j$. Therefore (12) follows.

To show that $\|s_k\| \to 0$, note that if (12) is true, then clearly $\|s_k\| \to 0$. If (13) is true, then for all $k > K$, the step s_k is accepted and $\|s_k\| \leq \frac{1}{a_0}\operatorname{dist}(x_{k+1}, x_k)$ (by Lemma 3.4), hence $\|s_k\| \to 0$ since $\{x_k\}$ converges. □

Lemma 3.7. *Let $\mathcal{M}$ be a Riemannian manifold endowed with two vector transports $\mathcal{T}_1$ and $\mathcal{T}_2$, and let $\bar{x} \in \mathcal{M}$. Then there exist a constant a_4 and a neighborhood $\mathcal{U}$ of $\bar{x}$ such that for all $x, y \in \mathcal{U}$ and all $\xi \in \mathrm{T}_y\mathcal{M}$,*

$$\|\mathcal{T}_{1_\eta}^{-1}\xi - \mathcal{T}_{2_\eta}^{-1}\xi\| \leq a_4\|\xi\|\|\eta\|,$$

where $\eta = R_x^{-1}(y)$.

Proof. We use the hat to denote coordinate expressions. Let $T_1(\hat{x}, \hat{\eta})$ and $T_2(\hat{x}, \hat{\eta})$ denote the coordinate expression of $\mathcal{T}_{1_\eta}^{-1}$ and $\mathcal{T}_{2_\eta}^{-1}$, respectively. Then

$$\begin{aligned}
\|\mathcal{T}_{1_\eta}^{-1}\xi - \mathcal{T}_{2_\eta}^{-1}\xi\| &\leq b_0\|(T_1(\hat{x}, \hat{\eta}) - T_2(\hat{x}, \hat{\eta}))\hat{\xi}\|_2 \\
&\leq b_0\|\hat{\xi}\|_2\|T_1(\hat{x}, \hat{\eta}) - T_2(\hat{x}, \hat{\eta})\|_2 \\
&\leq b_1\|\hat{\xi}\|_2\|\hat{\eta}\|_2 \text{ (since } T_1(\hat{x}, 0) = T_2(\hat{x}, 0) \text{ and both } T_1 \text{ and } T_2 \text{ are smooth)} \\
&\leq b_2\|\xi\|\|\eta\|
\end{aligned}$$

for some constants b_0, b_1, and b_2. □

The next lemma is proved in [GQA12, Lemma 14.5].

Lemma 3.8. *Let F be a C^1 vector field on a Riemannian manifold $\mathcal{M}$ and let $\bar{x} \in \mathcal{M}$ be a nondegenerate zero of F. Then there exist a neighborhood $\mathcal{U}$ of $\bar{x}$ and $a_5, a_6 > 0$ such that for all $x \in \mathcal{U}$,*

$$a_5 \operatorname{dist}(x, \bar{x}) \leq \|F(x)\| \leq a_6 \operatorname{dist}(x, \bar{x}).$$

In the Euclidean case, the next lemma holds with $\tilde{a}_7 = 0$ and reduces to the Fundamental Theorem of Calculus.

Lemma 3.9. *Let F be a C^1 vector field on a Riemannian manifold $\mathcal{M}$, let R be a retraction on $\mathcal{M}$, and let $\bar{x} \in \mathcal{M}$. Then there exist a neighborhood $\mathcal{U}$ of $\bar{x}$ and a constant $\tilde{a}_7$ such that for all $x, y \in \mathcal{U}$,*

$$\|P_\gamma^{0\leftarrow 1}F(y) - F(x) - [\int_0^1 P_\gamma^{0\leftarrow t}\mathbb{D}F(\gamma(t))P_\gamma^{t\leftarrow 0}dt]\eta\| \leq \tilde{a}_7\|\eta\|^2,$$

where $\eta = R_x^{-1}(y)$ and P_γ is the parallel translation along the curve γ given by by $\gamma(t) = R_x(t\eta)$.

Proof. Define $G : [0, 1] \to \mathrm{T}_x\mathcal{M} : t \mapsto G(t) = P_\gamma^{0\leftarrow t}F(\gamma(t))$. Observe that $G(0) = F(x)$ and $G(1) = P_\gamma^{0\leftarrow 1}F(y)$. We have

$$\begin{aligned}
G'(t) &= \frac{d}{d\epsilon}G(t + \epsilon)|_{\epsilon=0} \\
&= P_\gamma^{0\leftarrow t}\frac{d}{d\epsilon}P_\gamma^{t\leftarrow t+\epsilon}F(\gamma(t + \epsilon))|_{\epsilon=0} \\
&= P_\gamma^{0\leftarrow t}\mathbb{D}F(\gamma(t))[\frac{d}{d\epsilon}\gamma(t + \epsilon)]|_{\epsilon=0} \\
&= P_\gamma^{0\leftarrow t}\mathbb{D}F(\gamma(t))[\mathcal{T}_{R_(t\eta)}\eta],
\end{aligned}$$

where we have used an expression of the covariant derivative $\mathbb{D}$ in terms of the parallel translation P (see, e.g., [Cha06, theorem I.2.1]), and where $\mathcal{T}_{R(t\eta)}\eta = \frac{d}{dt}(R(t\eta))$. Since $G(1) - G(0) = \int_0^1 G'(t)dt$, we obtain

$$\begin{aligned}
&\|P_\gamma^{0\leftarrow 1}F(y) - F(x) - \int_0^1 P_\gamma^{0\leftarrow t}\mathbb{D}F(\gamma(t))P_\gamma^{t\leftarrow 0}\eta dt\| \\
&= \|\int_0^1 P_\gamma^{0\leftarrow t}\mathbb{D}F(\gamma(t))(\mathcal{T}_{R(t\eta)}\eta - P_\gamma^{t\leftarrow 0}\eta)dt\| \\
&\leq \int_0^1 \|P_\gamma^{0\leftarrow t}\mathbb{D}F(\gamma(t))P_\gamma^{t\leftarrow 0}\|\|(P_\gamma^{0\leftarrow t}\mathcal{T}_{R_(t\eta)}\eta - \eta)\|dt \\
&\leq \int_0^1 \|P_\gamma^{0\leftarrow t}\mathbb{D}F(\gamma(t))P_\gamma^{t\leftarrow 0}\|\|(P_\gamma^{0\leftarrow t}\mathcal{T}_{R(t\eta)}\eta - \mathcal{T}_{R(t\eta)}^{-1}\mathcal{T}_{R(t\eta)}\eta)\|dt \\
&\leq b_0\|\eta\|^2 \text{ (by Lemma 3.7)}
\end{aligned}$$

where b_0 is some constant. □

Lemma 3.10. *Suppose Assumptions 3.2 and 3.3 hold. Then there exist a neighborhood $\mathcal{U}$ and a constant a_7 such that for all x_1, $\tilde{x}_1$, x_2, and $\tilde{x}_2 \in \mathcal{U}$, we have*

$$|g(\mathcal{T}_{S_\zeta}\xi_1, y_2) - g(\mathcal{T}_{S_\zeta}y_1, \xi_2)| \leq a_7 \max\{\mathrm{dist}(x_1, x^*), \mathrm{dist}(x_2, x^*), \mathrm{dist}(\tilde{x}_1, x^*), \mathrm{dist}(\tilde{x}_2, x^*)\}\|\xi_1\|\|\xi_2\|,$$

where $\zeta = R_{x_1}^{-1}(x_2)$, $\xi_1 = R_{x_1}^{-1}(\tilde{x}_1)$, $\xi_2 = R_{x_2}^{-1}(\tilde{x}_2)$, $y_1 = \mathcal{T}_{S_{\xi_1}}^{-1}\operatorname{grad} f(\tilde{x}_1) - \operatorname{grad} f(x_1)$, and $y_2 = \mathcal{T}_{S_{\xi_2}}^{-1}\operatorname{grad} f(\tilde{x}_2) - \operatorname{grad} f(x_2)$.

Proof. Define $\bar{y}_1 = P_{\gamma_1}^{0\leftarrow 1}\operatorname{grad} f(\tilde{x}_1) - \operatorname{grad} f(x_1)$ and $\bar{y}_2 = P_{\gamma_2}^{0\leftarrow 1}\operatorname{grad} f(\tilde{x}_2) - \operatorname{grad} f(x_2)$, where P is the parallel transport, $\gamma_1(t) = R_{x_1}(t\xi_1)$, and $\gamma_2(t) = R_{x_2}(t\xi_2)$. From Lemma 3.9, we have

$$\|\bar{y}_1 - \bar{H}_1(x_1, \tilde{x}_1)\xi_1\| \leq b_0\|\xi_1\|^2 \quad \text{and} \quad \|\bar{y}_2 - \bar{H}_2(x_2, \tilde{x}_2)\xi_2\| \leq b_0\|\xi_2\|^2, \tag{14}$$

where $\bar{H}_1(x_1, \tilde{x}_1) = \int_0^1 P_{\gamma_1}^{0\leftarrow t}\operatorname{Hess} f(\gamma_1(t))P_{\gamma_1}^{t\leftarrow 0}dt$, $\bar{H}_2(x_2, \tilde{x}_2) = \int_0^1 P_{\gamma_2}^{0\leftarrow t}\operatorname{Hess} f(\gamma_2(t))P_{\gamma_2}^{t\leftarrow 0}dt$, and b_0 is a constant. It follows that

$$\begin{aligned}
&|g(\mathcal{T}_{S_\zeta}\xi_1, y_2) - g(\mathcal{T}_{S_\zeta}y_1, \xi_2)| \\
&\leq |g(\mathcal{T}_{S_\zeta}\xi_1, \bar{y}_2) - g(\mathcal{T}_{S_\zeta}\bar{y}_1, \xi_2)| + |g(\mathcal{T}_{S_\zeta}\xi_1, y_2 - \bar{y}_2) - g(\mathcal{T}_{S_\zeta}(y_1 - \bar{y}_1), \xi_2)| \\
&\leq |g(\mathcal{T}_{S_\zeta}\xi_1, \bar{H}_2(x_2, \tilde{x}_2)\xi_2) - g(\mathcal{T}_{S_\zeta}\bar{H}_1(x_1, \tilde{x}_1)\xi_1, \xi_2)| + b_1(\|\xi_1\| + \|\xi_2\|)\|\xi_1\|\|\xi_2\| \text{ (by (14))} \\
&\quad + |g(\mathcal{T}_{S_\zeta}\xi_1, \mathcal{T}_{S_{\zeta_2}}^{-1}\operatorname{grad} f(\tilde{x}_2) - P_{\gamma_2}^{0\leftarrow 1}\operatorname{grad} f(\tilde{x}_2))| + |g(\mathcal{T}_{S_\zeta}(\mathcal{T}_{S_{\zeta_1}}^{-1}\operatorname{grad} f(\tilde{x}_1) - P_{\gamma_1}^{0\leftarrow 1}\operatorname{grad} f(\tilde{x}_1)), \xi_2)| \\
&\leq |g(\mathcal{T}_{S_\zeta}\xi_1, \bar{H}_2(x_2, \tilde{x}_2)\xi_2) - g(\mathcal{T}_{S_\zeta}\bar{H}_1(x_1, \tilde{x}_1)\xi_1, \xi_2)| + b_1(\|\xi_1\| + \|\xi_2\|)\|\xi_1\|\|\xi_2\| \\
&\quad + b_2\|\xi_1\|\|\xi_2\|\|\operatorname{grad} f(\tilde{x}_2)\| + b_3\|\xi_1\|\|\xi_2\|\|\operatorname{grad} f(\tilde{x}_1)\| \text{ (by Lemma 3.7)} \\
&\leq |g(\bar{H}_2(x_2, \tilde{x}_2)\mathcal{T}_{S_\zeta}\xi_1, \xi_2) - g(\mathcal{T}_{S_\zeta}\bar{H}_1(x_1, \tilde{x}_1)\xi_1, \xi_2)| \text{ (average Hessian is also self-adjoint)} \\
&\quad + b_4\|\xi_1\|\|\xi_2\|(\mathrm{dist}(x_1, \tilde{x}_1) + \mathrm{dist}(x_2, \tilde{x}_2) + \mathrm{dist}(\tilde{x}_2, x^*) + \mathrm{dist}(\tilde{x}_1, x^*)) \text{ (by Lemmas 3.4 and 3.8)} \\
&\leq b_5\|\xi_1\|\|\xi_2\| \max\{\mathrm{dist}(x_1, x^*), \mathrm{dist}(x_2, x^*), \mathrm{dist}(\tilde{x}_1, x^*), \mathrm{dist}(\tilde{x}_2, x^*)\} \text{ (by triangle inequality of distance)} \\
&\quad + |g(\bar{H}_2(x_2, \tilde{x}_2)\mathcal{T}_{S_\zeta}\xi_1, \xi_2) - g(\mathcal{T}_{S_\zeta}\bar{H}_1(x_1, \tilde{x}_1)\xi_1, \xi_2)| \qquad (15)
\end{aligned}$$

where b_1, b_2, b_3, b_4 and b_5 are some constants. Using hat to denote coordinate expressions, $T(\hat{x}_1, \hat{x}_2)$ to denote $\mathcal{T}_\zeta$ and $\hat{G}(\hat{x}_2)$ to denote the matrix expression of the Riemannian metric at x_2, we have

$$\begin{aligned}
&|g(\bar{H}_2(x_2, \tilde{x}_2)\mathcal{T}_{S_\zeta}\xi_1, \xi_2) - g(\mathcal{T}_{S_\zeta}\bar{H}_1(x_1, \tilde{x}_1)\xi_1, \xi_2)| \\
&= |\hat{\xi}_1^T T(\hat{x}_1, \hat{x}_2)^T \hat{\bar{H}}_2(\hat{x}_2, \hat{\tilde{x}}_2)^T \hat{G}(\hat{x}_2)\hat{\xi}_2 - \hat{\xi}_1^T \hat{\bar{H}}_1(\hat{x}_1, \hat{\tilde{x}}_1)^T T(\hat{x}_1, \hat{x}_2)^T \hat{G}(\hat{x}_2)\hat{\xi}_2| \\
&\leq \|\hat{\xi}_1\|_2 \|T(\hat{x}_1, \hat{x}_2)^T \hat{\bar{H}}_2(\hat{x}_2, \hat{\tilde{x}}_2)^T - \hat{\bar{H}}_1(\hat{x}_1, \hat{\tilde{x}}_1)^T T(\hat{x}_1, \hat{x}_2)^T\|_2 \|\hat{G}(\hat{x}_2)\|_2 \|\hat{\xi}_2\|_2 \qquad (16)
\end{aligned}$$

where $\|\cdot\|_2$ denotes the Euclidean norm. Define a function

$$J(\hat{x}_1, \hat{\tilde{x}}_1, \hat{x}_2, \hat{\tilde{x}}_2) = T(\hat{x}_1, \hat{x}_2)^T \hat{\bar{H}}_2(\hat{x}_2, \hat{\tilde{x}}_2)^T - \hat{\bar{H}}_1(\hat{x}_1, \hat{\tilde{x}}_1)^T T(\hat{x}_1, \hat{x}_2)^T.$$

We can see that when $(\hat{x}_1^T, \hat{\tilde{x}}_1^T) = (\hat{x}_2^T, \hat{\tilde{x}}_2^T)$, $J = 0$. Since, in view of Assumption 3.3, J is Lipschitz continuous, it follows that (16) becomes

$$\begin{aligned}
|g(\bar{H}_2\mathcal{T}_{S_\zeta}\xi_1, \xi_2) - g(\mathcal{T}_{S_\zeta}\bar{H}_1\xi_1, \xi_2)| &\leq b_6 \|(\hat{x}_1^T, \hat{\tilde{x}}_1^T) - (\hat{x}_2^T, \hat{\tilde{x}}_2^T)\|_2 \|\hat{\xi}_1\|_2 \|\hat{\xi}_2\|_2 \\
&\leq b_7 \|\xi_1\| \|\xi_2\| \max\{\operatorname{dist}(x_1, x_2), \operatorname{dist}(\tilde{x}_1, \tilde{x}_2)\},
\end{aligned}$$

where b_6, b_7 are some constants. Combining this equation with (15), we obtain

$$|g(\mathcal{T}_{S_\zeta}\xi_1, y_2) - g(\mathcal{T}_{S_\zeta}y_1, \xi_2)| \leq b_8 \|\xi_1\| \|\xi_2\| \max\{\operatorname{dist}(x_1, x^*), \operatorname{dist}(x_2, x^*), \operatorname{dist}(\tilde{x}_1, x^*), \operatorname{dist}(\tilde{x}_2, x^*)\},$$

where b_8 is a constant. □

Lemma 3.11. *Let $\mathcal{M}$ be a Riemannian manifold endowed with a vector transport $\mathcal{T}$ with associated retraction R, and let $\bar{x} \in \mathcal{M}$. Then there is a neighborhood $\mathcal{U}$ of $\bar{x}$ and a_8 such that for all $x, y \in \mathcal{U}$,*

$$\begin{aligned}
\|\operatorname{id} - \mathcal{T}_\xi^{-1}\mathcal{T}_\eta^{-1}\mathcal{T}_\zeta\| &\leq a_8 \max(\operatorname{dist}(x, \bar{x}), \operatorname{dist}(y, \bar{x})), \\
\|\operatorname{id} - \mathcal{T}_\zeta^{-1}\mathcal{T}_\eta\mathcal{T}_\xi\| &\leq a_8 \max(\operatorname{dist}(x, \bar{x}), \operatorname{dist}(y, \bar{x})),
\end{aligned}$$

where $\xi = R_{\bar{x}}^{-1}(x)$, $\eta = R_x^{-1}(y)$, $\zeta = R_{\bar{x}}^{-1}(y)$, id *is the identity operator, and $\|\cdot\|$ is the operator norm induced by the Riemannian metric.*

Proof. Let the hat denote coordinate expressions, chosen such that the matrix expression of the Riemannian metric at $\bar{x}$ is the identity. Let $L(x, y)$ denote $\mathcal{T}_{R_x^{-1}(y)}$. We have

$$\|\operatorname{id} - \mathcal{T}_\xi^{-1}\mathcal{T}_\eta^{-1}\mathcal{T}_\zeta\| = \|I - L(\bar{x}, x)^{-1}L(x, y)^{-1}L(\bar{x}, y)\|.$$

Define a function $J(\bar{x}, \xi, \zeta) = I - L(\bar{x}, R_{\bar{x}}(\xi))^{-1}L(R_{\bar{x}}(\xi), R_{\bar{x}}(\zeta))^{-1}L(\bar{x}, R_{\bar{x}}(\zeta))$. Notice that J is a smooth function and $J(\bar{x}, 0_{\bar{x}}, 0_{\bar{x}}) = 0$. So

$$\begin{aligned}
\|J(\bar{x}, \xi, \zeta)\| &= \|J(\bar{x}, \xi, \zeta) - J(\bar{x}, 0_{\bar{x}}, 0_{\bar{x}})\| \\
&= \|\hat{J}(\hat{\bar{x}}, \hat{\xi}, \hat{\zeta}) - \hat{J}(\hat{\bar{x}}, \hat{0}_{\bar{x}}, \hat{0}_{\bar{x}})\|_2 \\
&\leq b_0(\|\hat{\xi}\|_2 + \|\hat{\zeta}\|_2) \text{ (smoothness of } J) \\
&\leq b_1(\operatorname{dist}(x, \bar{x}) + \operatorname{dist}(y, \bar{x})) \text{ (by Lemma 3.4)} \\
&\leq b_2 \max(\operatorname{dist}(x, \bar{x}), \operatorname{dist}(y, \bar{x})),
\end{aligned}$$

where b_0, b_1 and b_2 are some constants and $\|\cdot\|_2$ denotes the Euclidean norm. So

$$\|\operatorname{id} - \mathcal{T}_\xi^{-1}\mathcal{T}_\eta^{-1}\mathcal{T}_\zeta\| \leq b_2 \max(\operatorname{dist}(x, \bar{x}), \operatorname{dist}(y, \bar{x})).$$

This concludes the first part of the proof. The second part of the result follows from a similar argument. □

The next lemma generalizes [CGT91, Lemma 1]. It is instrumental in the proof of Lemma 3.14 below. In the Euclidean setting, it is possible to give an expression for a_9 and a_{10} in terms of c of Assumption 3.3 and ν of Assumption 3.5. In the Riemannian setting, we could not obtain such an expression, in part because the constant b_2 that appears in the proof below is no longer zero. However, the existence of a_9 and a_{10} can still be shown, under the assumption that $\{x_k\}$ converges to x^*, and this is all we need in order to carry on with Lemma 3.14.

Lemma 3.12. *Suppose Assumptions 3.1, 3.2, 3.3, and 3.5 hold. Then*

$$y_j - \tilde{\mathcal{B}}_{j+1} s_j = 0 \tag{17}$$

for all j. Moreover, there exist constants a_9 and a_{10} such that

$$\|y_j - (\mathcal{B}_i)_j s_j\| \leq a_9 a_{10}^{i-j-2} \epsilon_{i,j} \|s_j\| \tag{18}$$

for all j, $i \geq j+1$, where $\epsilon_{i,j} = \max_{j \leq k \leq i} \operatorname{dist}(x_k, x^*)$ and

$$(\mathcal{B}_i)_j = \mathcal{T}_{S_{\zeta_{j,i}}}^{-1} \mathcal{B}_i \mathcal{T}_{S_{\zeta_{j,i}}}$$

with $\zeta_{j,i} = R_{x_j}^{-1}(x_i)$.

Proof. From (3), we have

$$\tilde{\mathcal{B}}_{j+1} s_j = (\mathcal{B}_j + \frac{(y_j - \mathcal{B}_j s_j)(y_j - \mathcal{B}_j s_j)^\flat}{g(s_j, y_j - \mathcal{B}_j s_j)}) s_j = y_j.$$

This yields (17), as well as (18) with $i = j+1$. The proof of (18) for $i > j+1$ is by induction. We choose $k \geq j+1$ and assume that (18) holds for all $i = j+1, \ldots, k$. Let $r_k = y_k - \mathcal{B}_k s_k$. We have

$$\begin{aligned}
&|g(r_k, \mathcal{T}_{S_{\zeta_{j,k}}} s_j)| = |g(y_k - \mathcal{B}_k s_k, \mathcal{T}_{S_{\zeta_{j,k}}} s_j)| \\
&\leq |g(y_k, \mathcal{T}_{S_{\zeta_{j,k}}} s_j) - g(s_k, \mathcal{T}_{S_{\zeta_{j,k}}} y_j)| + |g(s_k, \mathcal{T}_{S_{\zeta_{j,k}}} (y_j - (\mathcal{B}_k)_j s_j))| + |g(s_k, \mathcal{T}_{S_{\zeta_{j,k}}}((\mathcal{B}_k)_j s_j)) - g(\mathcal{B}_k s_k, \mathcal{T}_{S_{\zeta_{j,k}}} s_j)| \\
&\leq |g(y_k, \mathcal{T}_{S_{\zeta_{j,k}}} s_j) - g(s_k, \mathcal{T}_{S_{\zeta_{j,k}}} y_j)| + \|\mathcal{T}_{S_{\zeta_{j,k}}} (y_j - (\mathcal{B}_k)_j s_j)\| \|s_k\| + |g(s_k, \mathcal{B}_k \mathcal{T}_{S_{\zeta_{j,k}}} s_j) - g(\mathcal{B}_k s_k, \mathcal{T}_{S_{\zeta_{j,k}}} s_j)| \\
&\leq |g(y_k, \mathcal{T}_{S_{\zeta_{j,k}}} s_j) - g(s_k, \mathcal{T}_{S_{\zeta_{j,k}}} y_j)| + b_0 a_9 a_{10}^{k-j-2} \epsilon_{k,j} \|s_j\| \|s_k\| \text{ ($\mathcal{B}_k$ self-adjoint and induction assumption)} \\
&\leq b_0 a_9 a_{10}^{k-j-2} \epsilon_{k,j} \|s_j\| \|s_k\| + b_1 \epsilon_{k+1,j} \|s_k\| \|s_j\|, \text{ (by Lemma 3.10)}
\end{aligned}$$

where b_0 and b_1 are some constants. It follows that

$$\begin{aligned}
&\|y_j - (\mathcal{B}_{k+1})_j s_j\| \\
&= \|y_j - \mathcal{T}_{S_{\zeta_{j,k+1}}}^{-1} \mathcal{B}_{k+1} \mathcal{T}_{S_{\zeta_{j,k+1}}} s_j\| \\
&= \|y_j - \mathcal{T}_{S_{\zeta_{j,k+1}}}^{-1} \mathcal{T}_{S_{s_k}} \tilde{\mathcal{B}}_{k+1} \mathcal{T}_{S_{s_k}}^{-1} \mathcal{T}_{S_{\zeta_{j,k+1}}} s_j\| \\
&\leq \|y_j - \mathcal{T}_{S_{\zeta_{j,k}}}^{-1} \tilde{\mathcal{B}}_{k+1} \mathcal{T}_{S_{\zeta_{j,k}}} s_j\| + \|\mathcal{T}_{S_{\zeta_{j,k}}}^{-1} \tilde{\mathcal{B}}_{k+1} \mathcal{T}_{S_{\zeta_{j,k}}} s_j - \mathcal{T}_{S_{\zeta_{j,k+1}}}^{-1} \mathcal{T}_{S_{s_k}} \tilde{\mathcal{B}}_{k+1} \mathcal{T}_{S_{s_k}}^{-1} \mathcal{T}_{S_{\zeta_{j,k+1}}} s_j\| \\
&\leq \|y_j - ((\mathcal{B}_k)_j + \mathcal{T}_{S_{\zeta_{j,k}}}^{-1} \frac{(r_k)(r_k)^\flat}{g(s_k, r_k)} \mathcal{T}_{S_{\zeta_{j,k}}}) s_j\| + b_2 \epsilon_{k+1,j} \|s_j\| \text{ (by Lemma 3.11, Assumption 3.1, and (3))} \\
&\leq \|y_j - (\mathcal{B}_k)_j s_j\| + b_3 \frac{|g(r_k, \mathcal{T}_{S_{\zeta_{j,k}}} s_j)|}{\|s_k\|} + b_2 \epsilon_{k+1,j} \|s_j\| \text{ (by Assumption 3.5)} \\
&\leq a_9 a_{10}^{k-j-2} \epsilon_{k,j} \|s_j\| + b_3 b_0 a_9 a_{10}^{k-j-2} \epsilon_{k,j} \|s_j\| + b_3 b_1 \epsilon_{k,j} \|s_j\| + b_2 \epsilon_{k+1,j} \|s_j\| \\
&\leq (a_9 a_{10}^{k-j-2} + b_3 b_0 a_9 a_{10}^{k-j-2} + b_3 b_1 + b_2) \epsilon_{k+1,j} \|s_j\|, \text{ (note that } \epsilon_{k,j} \leq \epsilon_{k+1,j})
\end{aligned}$$

where b_2, b_3 are some constant. Because b_0, b_1, b_2 and b_3 are independent of a_9 and a_{10}, we can choose a_9 and a_{10} large enough such that

$$(a_9 a_{10}^{k-j-2} + b_3 b_0 a_9 a_{10}^{k-j-2} + b_3 b_1 + b_2) \leq a_9 a_{10}^{k+1-j-2}.$$

for all j, $k \geq j+1$. Take for example, $a_9 > 1$ and $a_{10} > 1 + b_3 b_0 + b_3 b_1 + b_2$. Therefore

$$\|y_j - (\mathcal{B}_{k+1})_j s_j\| \leq a_9 a_{10}^{k+1-j-2} \epsilon_{k+1,j} \|s_j\|.$$

This concludes the argument by induction. □

Lemma 3.13. *If Assumption 3.3 holds then there exist a neighborhood $\mathcal{U}$ of x^* and a constant a_{11} such that for all $x_1, x_2 \in \mathcal{U}$, the inequality*

$$\|y - \mathcal{T}_{S_{\zeta_1}} \operatorname{Hess} f(x^*) \mathcal{T}_{S_{\zeta_1}}^{-1} s\| \leq a_{11} \|s\| \max\{\operatorname{dist}(x_1, x^*), \operatorname{dist}(x_2, x^*)\}$$

holds, where $\zeta_1 = R_{x^*}^{-1}(x_1)$, $s = R_{x_1}^{-1}(x_2)$, $y = \mathcal{T}_{S_s}^{-1} \operatorname{grad} f(x_2) - \operatorname{grad} f(x_1)$.

Proof. Define $\bar{y} = P_\gamma^{0\leftarrow 1} \operatorname{grad} f(x_2) - \operatorname{grad} f(x_1)$, where P is the parallel transport along the curve γ defined by $\gamma(t) = R_{x_1}(ts)$. From Lemma 3.9, we have

$$\|\bar{y} - \bar{H} s\| \leq b_0 \|s\|^2, \tag{19}$$

where $\bar{H} = \int_0^1 P_\gamma^{0\leftarrow t} \operatorname{Hess} f(\gamma(t)) P_\gamma^{t\leftarrow 0} dt$ and b_0 is a constant. We then have

$$\begin{aligned}
\|y - \mathcal{T}_{S_{\zeta_1}} \operatorname{Hess} f(x^*) \mathcal{T}_{S_{\zeta_1}}^{-1} s\| &\leq \|y - \bar{y}\| + \|\bar{y} - \bar{H} s\| + \|\bar{H} s - \mathcal{T}_{S_{\zeta_1}} \operatorname{Hess} f(x^*) \mathcal{T}_{S_{\zeta_1}}^{-1} s\| \\
&= \|\mathcal{T}_{S_\zeta}^{-1} \operatorname{grad} f(x_2) - P_\gamma^{0\leftarrow 1} \operatorname{grad} f(x_2)\| + b_0 \|s\|^2 + \|\bar{H} - \mathcal{T}_{S_{\zeta_1}} \operatorname{Hess} f(x^*) \mathcal{T}_{S_{\zeta_1}}^{-1}\| \|s\| \\
&\leq b_1 \|s\| \max\{\operatorname{dist}(x_1, x^*), \operatorname{dist}(x_2, x^*)\} + b_0 \|s\|^2 \text{ (by Lemma 3.7)} \\
&\quad + (\| \int_0^1 P_\gamma^{0\leftarrow t} \operatorname{Hess} f(\gamma(t)) P_\gamma^{t\leftarrow 0} dt - \operatorname{Hess} f(x_1)\| \\
&\quad + \| \operatorname{Hess} f(x_1) - \mathcal{T}_{S_{\zeta_1}} \operatorname{Hess} f(x^*) \mathcal{T}_{S_{\zeta_1}}^{-1}\|) \|s\| \\
&\leq b_2 \|s\| \max\{\operatorname{dist}(x_1, x^*), \operatorname{dist}(x_2, x^*)\}, \text{ (by Assumption 3.3)}
\end{aligned}$$

where b_1 and b_2 are some constants. □

With these technical lemmas in place, we now start the Riemannian generalization of the sequence of lemmas in [BKS96] that leads to the main result [BKS96, Theorem 2.7], generalized here as Theorem 3.18. For an easier comparison with [BKS96], in the rest of the convergence analysis, we let n (instead of d) denote the dimension of the manifold $\mathcal{M}$.

The next lemma generalizes [BKS96, Lemma 2.3], itself a slight variation of [KBS93, Lemma 3.2]. The proof of [BKS96, Lemma 2.3] involves considering the span of a few s_j's. In the Riemannian setting, a difficulty arises from the fact that the s_j's are not in the same tangent space. We overcome this difficulty by transporting the s_j's to $\mathrm{T}_{x^*} \mathcal{M}$.

Lemma 3.14. *Let s_k be such that $R_{x_k}(s_k) \to x^*$. If Assumptions 3.1, 3.2, 3.3, and 3.5 hold then there exists $K \geq 0$ such that for any set of $n+1$ steps $S = \{s_{k_j} : K \leq k_1 < \ldots < k_{n+1}\}$, there exists an index k_m with $m \in \{2, 3, \ldots, n+1\}$ such that*

$$\frac{\|(\mathcal{B}_{k_m} - H_{k_m}) s_{k_m}\|}{\|s_{k_m}\|} < (a_{12} a_{10}^{k_{n+1} - k_1 - 2} + \bar{a}_{12}) \epsilon_S^{\frac{1}{n}},$$

where $\epsilon_S = \max_{1\le j\le n+1}\{\mathrm{dist}(x_{k_j}, x^*), \mathrm{dist}(R_{x_{k_j}}(s_{k_j}), x^*)\}$, $H_{k_m} = \mathcal{T}_{S_{\zeta_{k_m}}} \mathrm{Hess}\, f(x^*) \mathcal{T}^{-1}_{S_{\zeta_{k_m}}}$, $\zeta_{k_m} = R^{-1}_{x^*}(x_{k_m})$, $a_{12}, \bar{a}_{12}$ *are some constants, and* n *is the dimension of the manifold.*

Proof. Given S, for $j = 1, 2, \ldots, n+1$, define

$$S_j = \left[\frac{\bar{s}_{k_1}}{\|\bar{s}_{k_1}\|}, \frac{\bar{s}_{k_2}}{\|\bar{s}_{k_2}\|}, \ldots, \frac{\bar{s}_{k_j}}{\|\bar{s}_{k_j}\|}\right],$$

where $\bar{s}_{k_i} = \mathcal{T}^{-1}_{S_{\zeta_{k_i}}} s_{k_i}, i = 1, 2, \ldots, j$. The proof is organized as follows. We will first obtain in (28) that there exists $m \in [2, n+1]$ and $u \in \mathbb{R}^{m-1}$, $w \in \mathrm{T}_{x^*}\mathcal{M}$ such that $\bar{s}_{k_m}/\|\bar{s}_{k_m}\| = S_{m-1}u - w$, S_{m-1} has full column rank and is well conditioned, and $\|w\|$ is small. We will also obtain in (30) that $(\mathcal{T}^{-1}_{S_{\zeta_{k_m}}} \mathcal{B}_{k_m} \mathcal{T}_{S_{\zeta_{k_m}}} - \mathrm{Hess}\, f(x^*))S_{m-1}$ is small due to the Hessian approximating properties of the SR1 update given in Lemma 3.13 above. The conclusion follows from these two results.

Let G_* denote the matrix expression of inner product of $\mathrm{T}_{x^*}\mathcal{M}$ and $\hat{S}_j$ denote the coordinate expression of S_j, for $j \in \{1, \ldots, n\}$. Let κ_j be the smallest singular value of $G_*^{1/2}\hat{S}_j$ and define $\kappa_{n+1} = 0$. We have

$$1 = \kappa_1 \ge \kappa_2 \ldots \ge \kappa_{n+1} = 0.$$

Let m be the smallest integer for which

$$\frac{\kappa_m}{\kappa_{m-1}} < \epsilon_S^{\frac{1}{n}}. \tag{20}$$

Since $m \le n+1$ and $\kappa_1 = 1$, we have

$$\kappa_{m-1} = \kappa_1(\frac{\kappa_2}{\kappa_1}) \ldots (\frac{\kappa_{m-1}}{\kappa_{m-2}}) > \epsilon_S^{(m-2)/n} > \epsilon_S^{(n-1)/n}. \tag{21}$$

Since $x_k \to x^*$ and $R_{x_k}(s_k) \to x^*$, we can assume that $\epsilon_S \in (0, (\frac{1}{4})^n)$ for all k. Now, we choose $z \in \mathbb{R}^m$ such that

$$\|G_*^{1/2}\hat{S}_m z\|_2 = \kappa_m \|z\|_2 \tag{22}$$

and

$$z = \begin{pmatrix} u \\ -1 \end{pmatrix},$$

where $u \in \mathbb{R}^{m-1}$. (The last component of z is nonzero due to that m is the smallest such that (20) is true.) Let $w = S_m z$ and its coordinate expression $\hat{w} = \hat{S}_m z$. From the definition of $G_*^{1/2}\hat{S}_m$ and z, we have

$$G_*^{1/2}\hat{S}_{m-1}u - G_*^{1/2}\hat{w} = \frac{G_*^{1/2}\hat{\bar{s}}_{k_m}}{\|G_*^{1/2}\hat{\bar{s}}_{k_m}\|_2}, \tag{23}$$

where $\hat{\bar{s}}_{k_m}$ is the coordinate expression of $\bar{s}_{k_m}$. Since κ_{m-1} is the smallest singular value of $G_*^{1/2}\hat{S}_{m-1}$, we have that

$$\|u\|_2 \le \frac{1}{\kappa_{m-1}}\|G_*^{1/2}\hat{S}_{m-1}u\|_2 = \frac{1}{\kappa_{m-1}}\|G_*^{1/2}\hat{w} + \frac{G_*^{1/2}\hat{\bar{s}}_{k_m}}{\|G_*^{1/2}\hat{\bar{s}}_{k_m}\|_2}\|_2 \le \frac{\|G_*^{1/2}\hat{w}\|_2 + 1}{\kappa_{m-1}} = \frac{\|w\| + 1}{\kappa_{m-1}} \tag{24}$$

$$< \frac{\|G_*^{1/2}\hat{w}\|_2 + 1}{\epsilon_S^{(n-1)/n}} = \frac{\|w\| + 1}{\epsilon_S^{(n-1)/n}}. \text{ (by (21))} \tag{25}$$

Using (22) and (24), we have that

$$\begin{aligned}\|w\|^2 = \|G_*^{1/2}\hat{w}\|_2^2 = \|G_*^{1/2}\hat{S}_m z\|_2^2 = \kappa_m^2\|z\|_2^2 = \kappa_m^2(1+\|u\|_2^2)\\ \leq \kappa_m^2 + (\frac{\kappa_m}{\kappa_{m-1}})^2(\|G_*^{1/2}\hat{w}\|_2+1)^2 = \kappa_m^2 + (\frac{\kappa_m}{\kappa_{m-1}})^2(\|w\|+1)^2.\end{aligned}$$

Therefore, since (20) implies that $\kappa_m < \epsilon_S^{1/n}$, using (20),

$$\|w\|^2 < \epsilon_S^{2/n} + \epsilon_S^{2/n}(\|w\|+1)^2 < 4\epsilon_S^{2/n}(\|w\|+1)^2. \tag{26}$$

This implies

$$\|w\|(1-2\epsilon_S^{1/n}) < 2\epsilon_S^{1/n},$$

and hence $\|w\| < 1$, since $\epsilon_S < (\frac{1}{4})^n$. Therefore, (25) and (26) imply that

$$\|u\|_2 < \frac{2}{\epsilon_S^{(n-1)/n}}, \tag{27}$$

$$\|w\| < 4\epsilon_S^{1/n}. \tag{28}$$

Equation (28) is the announced result that w is small. The bound (27) will also be invoked below.

Now we show that $\|(\mathcal{T}_{S_{\zeta_{k_j}}}^{-1}\mathcal{B}_{k_j}\mathcal{T}_{S_{\zeta_{k_j}}} - \operatorname{Hess} f(x^*))S_{j-1}\|$ is small for all $j \in [2, n+1]$ (and thus in particular for $j = m$). By Lemma 3.12, we have

$$\|y_i - (\mathcal{B}_{k_j})_i s_i\| \leq a_9 a_{10}^{k_j-i-2}\epsilon_{k_j,i}\|s_i\| \leq a_9 a_{10}^{k_{n+1}-k_1-2}\epsilon_S\|s_i\|, \tag{29}$$

for all $i \in \{k_1, k_2, \ldots, k_{j-1}\}$. Therefore,

$$\begin{aligned}&\|(\mathcal{T}_{S_{\zeta_{k_j}}}^{-1}\mathcal{B}_{k_j}\mathcal{T}_{S_{\zeta_{k_j}}} - \operatorname{Hess} f(x^*))\frac{\bar{s}_i}{\|\bar{s}_i\|}\|\\ &\leq \|\frac{\mathcal{T}_{S_{\zeta_i}}^{-1}y_i - \mathcal{T}_{S_{\zeta_{k_j}}}^{-1}\mathcal{B}_{k_j}\mathcal{T}_{S_{\zeta_{k_j}}}\bar{s}_i}{\|\bar{s}_i\|}\| + \|\frac{\mathcal{T}_{S_{\zeta_i}}^{-1}y_i - \operatorname{Hess} f(x^*)\bar{s}_i}{\|\bar{s}_i\|}\|\\ &\leq \|\frac{\mathcal{T}_{S_{\zeta_i}}^{-1}y_i - \mathcal{T}_{S_{\zeta_{k_j}}}^{-1}\mathcal{B}_{k_j}\mathcal{T}_{S_{\zeta_{k_j}}}\bar{s}_i}{\|\bar{s}_i\|}\| + b_1\epsilon_S \text{ (by Lemma 3.13)}\\ &= \|\frac{\mathcal{T}_{S_{\zeta_i}}^{-1}(y_i - \mathcal{T}_{S_{\zeta_i}}\mathcal{T}_{S_{\zeta_{k_j}}}^{-1}\mathcal{B}_{k_j}\mathcal{T}_{S_{\zeta_{k_j}}}\mathcal{T}_{S_{\zeta_i}}^{-1}s_i)}{\|\bar{s}_i\|}\| + b_1\epsilon_S\\ &\leq b_2\frac{\|(y_i - (\mathcal{B}_{k_j})_i s_i)\|}{\|s_i\|} + b_3\epsilon_S \text{ (by Lemma 3.11 and Assumption 3.1)}\\ &\leq (b_4 a_{10}^{k_{n+1}-k_1-2} + b_3)\epsilon_S \text{ (by (29))}\end{aligned}$$

where b_2, b_3 and b_4 are some constants. Therefore, we have that for any $j \in [2, n+1]$,

$$\|(\mathcal{T}_{S_{\zeta_{k_j}}}^{-1}\mathcal{B}_{k_j}\mathcal{T}_{S_{\zeta_{k_j}}} - \operatorname{Hess} f(x^*))S_{j-1}\|_{g,2} \leq b_5\epsilon_S, \tag{30}$$

where $b_5 = \sqrt{n}(b_4 a_{10}^{k_{n+1}-k_1-2} + b_3)$ and $\|\cdot\|_{g,2}$ is the norm induced by the Riemannian metric g and the Euclidean norm, i.e., $\|A\|_{g,2} = \sup\|Av\|/\|v\|_2$ with $\|\cdot\|$ defined in (5).

We can now conclude the proof as follows. Using (23) and (30) with $j = m$, (27) and (28), we have

$$\begin{aligned}
&\frac{\|(\mathcal{T}_{S_{\zeta_{k_m}}}^{-1}\mathcal{B}_{k_m}\mathcal{T}_{S_{\zeta_{k_m}}} - \operatorname{Hess} f(x^*))\bar{s}_m\|}{\|\bar{s}_m\|} \\
&= \|(\mathcal{T}_{S_{\zeta_{k_m}}}^{-1}\mathcal{B}_{k_m}\mathcal{T}_{S_{\zeta_{k_m}}} - \operatorname{Hess} f(x^*))(S_{m-1}u - w)\| \\
&\leq \|(\mathcal{T}_{S_{\zeta_{k_m}}}^{-1}\mathcal{B}_{k_m}\mathcal{T}_{S_{\zeta_{k_m}}} - \operatorname{Hess} f(x^*))S_{m-1}\|_{g,2}\|u\|_2 + \|(\mathcal{T}_{S_{\zeta_{k_m}}}^{-1}\mathcal{B}_{k_m}\mathcal{T}_{S_{\zeta_{k_m}}} - \operatorname{Hess} f(x^*))\|\|w\| \\
&\leq b_5\epsilon_S\frac{2}{\epsilon_S^{(n-1)/n}} + (M + \operatorname{Hess} f(x^*))4\epsilon_S^{1/n} \text{ (by Assumption 3.1)} \\
&\leq (2b_5 + b_6)\epsilon_S^{1/n}
\end{aligned}$$

where b_6 is some constant. Finally,

$$\begin{aligned}
\frac{\|(\mathcal{B}_{k_m} - H_{k_m})s_{k_m}\|}{\|s_{k_m}\|} &= \frac{\|(\mathcal{B}_{k_m} - \mathcal{T}_{S_{\zeta_{k_m}}}\operatorname{Hess} f(x^*)\mathcal{T}_{S_{\zeta_{k_m}}}^{-1})s_{k_m}\|}{\|s_{k_m}\|} \\
&= \frac{\|(\mathcal{T}_{S_{\zeta_{k_m}}}^{-1}\mathcal{B}_{k_m}\mathcal{T}_{S_{\zeta_{k_m}}} - \operatorname{Hess} f(x^*))\bar{s}_{k_m}\|}{\|\bar{s}_{k_m}\|} \\
&\leq (2b_5 + b_6)\epsilon_S^{1/n}.
\end{aligned}$$

□

The next lemma generalizes [BKS96, Lemma 2.4]. Its proof is a translation of the proof of [BKS96, Lemma 2.4], where we invoke two manifold-specific results: the equality of $\operatorname{Hess} f(x^*)$ and $\operatorname{Hess}(f \circ R_{x^*})(0_{x^*})$ (which holds in view of [AMS08, Proposition 5.5.6] since x^* is a critical point of f), and the bound in Lemma 3.4 on the retraction R.

Lemma 3.15. *Suppose that Assumptions 3.1, 3.2, 3.3, 3.4, 3.5 and 3.6 hold and the trust-region subproblem* (2) *is solved accurately enough for (9) to hold. Then there exists N such that for any set of $p > n$ consecutive steps $s_{k+1}, s_{k+1}, \ldots, s_{k+p}$ with $k \geq N$, there exists a set, $\mathcal{G}_k$, of at least $p - n$ indices contained in the set $\{i : k + 1 \leq i \leq k + p\}$ such that for all $j \in \mathcal{G}_k$,*

$$\frac{\|(\mathcal{B}_j - H_j)s_j\|}{\|s_j\|} < a_{13}\epsilon_k^{\frac{1}{n}},$$

where $a_{13} = a_{12}a_{10}^{p-2} + \bar{a}_{12}$, $H_j = \mathcal{T}_{S_{\zeta_j}}\operatorname{Hess} f(x^*)\mathcal{T}_{S_{\zeta_j}}^{-1}$, $\zeta_j = R_{x^*}^{-1}(x_j)$, and

$$\epsilon_k = \max_{k+1\leq j\leq k+p}\{\operatorname{dist}(x_j, x^*), \operatorname{dist}(R_{x_j}(s_j), x^*)\}.$$

Furthermore, for k sufficiently large, if $j \in \mathcal{G}_k$, then

$$\|s_j\| < a_{14}\operatorname{dist}(x_j, x^*), \tag{31}$$

where a_{14} is a constant, and

$$\rho_j \geq 0.75. \tag{32}$$

Proof. By Lemma 3.6, $s_k \to 0$. Therefore, by Lemma 3.14, applied to the set

$$\{s_k, s_{k+1}, \ldots, s_{k+p}\}, \tag{33}$$

there exists N such that for any $k \geq N$ there exists an index l_1, with $k+1 \leq l_1 \leq k+p$ satisfying

$$\frac{\|(\mathcal{B}_{l_1} - H_{l_1})s_{l_1}\|}{\|s_{l_1}\|} < a_{13}\epsilon_k^{\frac{1}{n}},$$

where $a_{13} = a_{12}a_{10}^{p-2} + \bar{a}_{12}$. Now we can apply Lemma 3.14 to the set $\{s_k, s_{k+1}, \ldots, s_{k+p}\} - s_{l_1}$ to get l_2. Repeating this $p-n$ times, we get a set of $p-n$ indices $\mathcal{G}_k = \{l_1, l_2, \ldots, l_{p-n}\}$ such that if $j \in \mathcal{G}_k$, then

$$\frac{\|(\mathcal{B}_j - H_j)s_j\|}{\|s_j\|} < a_{13}\epsilon_k^{\frac{1}{n}}. \tag{34}$$

We show (31) next. Consider $j \in \mathcal{G}_k$. By (34), we have

$$g(s_j, (H_j - \mathcal{B}_j)s_j) \leq \|s_j\|\|(H_j - \mathcal{B}_j)s_j\| \leq a_{13}\epsilon_k^{\frac{1}{n}}\|s_j\|^2.$$

Therefore,

$$\begin{aligned} g(s_j, \mathcal{B}_j s_j) &\geq g(s_j, H_j s_j) - a_{13}\epsilon_k^{\frac{1}{n}}\|s_j\|^2 \\ &> b_0\|s_j\|^2, \text{ (choosing } k \text{ large enough)} \end{aligned}$$

where b_0 is a constant and we have

$$\begin{aligned} 0 \leq m_j(0) - m_j(s_j) &= -g(\operatorname{grad} f(x_j), s_j) - \frac{1}{2}g(s_j, \mathcal{B}_j s_j) \\ &\leq \|\operatorname{grad} f(x_j)\|\|s_j\| - \frac{1}{2}b_0\|s_j\|^2 \\ &\leq b_1 \operatorname{dist}(x_j, x^*)\|s_j\| - \frac{1}{2}b_0\|s_j\|^2, \text{ (by Lemma 3.8)} \end{aligned}$$

where b_1 is some constant. This yields (31).

Finally, we show (32). Let $j \in \mathcal{G}_k$ and define $\hat{f}_x(\eta) = f(R_x(\eta))$. It follows that

$$
\begin{aligned}
&|f(x_j) - f(R_{x_j}(s_j)) - (m_j(0) - m_j(s_j))| \\
&= |f(x_j) - f(R_{x_j}(s_j)) + g(\operatorname{grad} f(x_j), s_j) + \frac{1}{2} g(s_j, \mathcal{B}_j s_j)| \\
&= |\hat{f}_{x_j}(0_{x_j}) - \hat{f}_{x_j}(s_j) + g(\operatorname{grad} f(x_j), s_j) + \frac{1}{2} g(s_j, \mathcal{B}_j s_j)| \\
&= |\frac{1}{2} g(s_j, \mathcal{B}_j s_j) - \int_0^1 g(\operatorname{Hess} \hat{f}_{x_j}(\tau s_j)[s_j], s_j)(1-\tau) d\tau| \text{ (by Taylor's theorem)} \\
&\leq |\frac{1}{2} g(s_j, \mathcal{B}_j s_j) - \frac{1}{2} g(s_j, H_j s_j)| + |\frac{1}{2} g(s_j, H_j s_j) - \int_0^1 g(\operatorname{Hess} \hat{f}_{x_j}(\tau s_j)[s_j], s_j)(1-\tau) d\tau| \\
&= |\frac{1}{2} g(s_j, (\mathcal{B}_j - H_j) s_j)| \\
&\quad + |\int_0^1 (g(s_j, \mathcal{T}_{S_{\zeta_j}} \operatorname{Hess} f(x^*) \mathcal{T}_{S_{\zeta_j}}^{-1} s_j) - g(\operatorname{Hess} \hat{f}_{x_j}(\tau s_j)[s_j], s_j))(1-\tau) d\tau| \\
&\leq \frac{1}{2} \|s_j\| \|(\mathcal{B}_j - H_j) s_j\| \\
&\quad + \|s_j\|^2 \int_0^1 \|(\mathcal{T}_{S_{\zeta_j}} \operatorname{Hess} \hat{f}_{x^*}(0_{x^*}) \mathcal{T}_{S_{\zeta_j}}^{-1} - \operatorname{Hess} \hat{f}_{x_j}(\tau s_j))\|(1-\tau) d\tau \text{ (by [AMS08, Proposition 5.5.6])} \\
&\leq b_2 \|s_j\|^2 \epsilon_k^{\frac{1}{n}} + b_3 \|s_j\|^2 (\operatorname{dist}(x_j, x^*) + \|s_j\|) \text{ (by (34), Lemma 3.4 and Assumption 3.4)} \\
&\leq b_4 \|s_j\|^2 \epsilon_k^{\frac{1}{n}}, \text{ (by (31) and } \operatorname{dist}(x_j, x^*) \text{ is smaller than } \epsilon_k^{\frac{1}{n}} \text{ eventually)}
\end{aligned}
$$

where b_2, b_3 and b_4 are some constants. In view of (31) and Lemma 3.8, we have

$$\|s_j\| < b_5 \|\operatorname{grad} f(x_j)\|,$$

where b_5 is some constant. Combining with $\|s_j\| \leq \Delta_j$, we obtain

$$\|s_j\|^2 \leq b_5 \|\operatorname{grad} f(x_j)\| \min\{\Delta_j, b_5 \|\operatorname{grad} f(x_j)\|\}.$$

Noticing (9), we have

$$|f(x_j) - f(R_{x_j}(s_j)) - (m_j(0) - m_j(s_j))| \leq b_6 \epsilon_k^{\frac{1}{n}} (m_j(0) - m_j(s_j)),$$

where b_6 is a constant. This implies (32). □

The next result generalizes [BKS96, Lemma 2.5] in two ways: the Euclidean setting is extended to the Riemannian setting, and inexact solves are allowed by the presence of δ_k. The main hurdle that we had to overcome in the Riemannian generalization is that the equality $\operatorname{dist}(x_k + s_k, x^*) = \|s_k - \xi_k\|$ does not necessarily hold. As we will see, Lemma 3.3 comes to our rescue.

Lemma 3.16. *Suppose Assumptions 3.2 and 3.3 hold. If the quantities*

$$e_k = \operatorname{dist}(x_k, x^*) \text{ and } \frac{\|(\mathcal{B}_k - H_k) s_k\|}{\|s_k\|}$$

are sufficiently small and if $\mathcal{B}_k s_k = -\operatorname{grad} f(x_k) + \delta_k$ *with* $\|\delta_k\| \leq \|\operatorname{grad} f(x_k)\|^{1+\theta}$*, then*

$$\operatorname{dist}(R_{x_k}(s_k), x^*) \leq a_{15} \frac{\|(\mathcal{B}_k - H_k)s_k\|}{\|s_k\|} e_k + a_{16} e_k^{1+\min\{\theta,1\}}, \tag{35}$$

$$h(R_{x_k}(s_k)) \leq a_{17} \frac{\|(\mathcal{B}_k - H_k)s_k\|}{\|s_k\|} h(x_k) + a_{18} h^{1+\min\{\theta,1\}}(x_k), \tag{36}$$

and

$$a_{19} h(x_k) \leq e_k \leq a_{20} h(x_k) \tag{37}$$

where a_{15}, a_{16}, a_{17} *and* a_{18} *are some constants and* $h(x) = (f(x) - f(x^*))^{\frac{1}{2}}$*.*

Proof. By definition of s_k, we have

$$s_k = H_k^{-1}[(H_k - \mathcal{B}_k)s_k - \operatorname{grad} f(x_k) + \delta_k]. \tag{38}$$

Define $\xi_k = R_{x_k}^{-1} x^*$. Therefore, letting γ be the curve defined by $\gamma(t) = R_{x_k}(t\xi_k)$, we have

$$\begin{aligned}
&\|s_k - \xi_k\| \\
&= \|H_k^{-1}[(H_k - \mathcal{B}_k)s_k - \operatorname{grad} f(x_k) + \delta_k - H_k\xi_k]\| \\
&\leq b_0(\|(H_k - \mathcal{B}_k)s_k\| + \|\delta_k\| + \|P_\gamma^{0\leftarrow 1} \operatorname{grad} f(x^*) - \operatorname{grad} f(x_k) - (\int_0^1 P_\gamma^{0\leftarrow t} \operatorname{Hess} f(\gamma(t)) P_\gamma^{t\leftarrow 0} dt)\xi_k\| \\
&+ \|(\int_0^1 P_\gamma^{0\leftarrow t} \operatorname{Hess} f(\gamma(t)) P_\gamma^{t\leftarrow 0} dt)\xi_k - \operatorname{Hess} f(x_k)\xi_k\| + \|\operatorname{Hess} f(x_k)\xi_k - H_k\xi_k\|) \\
&\leq b_0(\|(H_k - \mathcal{B}_k)s_k\| + b_1\|\xi_k\|^{1+\min\{\theta,1\}} \text{ (by Lemmas 3.8 and 3.9)} \\
&+ \|(\int_0^1 P_\gamma^{0\leftarrow t} \operatorname{Hess} f(\gamma(t)) P_\gamma^{t\leftarrow 0} dt)\xi_k - \operatorname{Hess} f(x_k)\xi_k\| + \|\operatorname{Hess} f(x_k) - H_k\|\|\xi_k\|) \\
&\leq b_0\|(H_k - \mathcal{B}_k)s_k\| + b_0 b_1\|\xi_k\|^{1+\min\{\theta,1\}} + b_0 b_3\|\xi_k\|^2 \text{ (by Assumption 3.3)} \\
&\leq b_0\|(H_k - \mathcal{B}_k)s_k\| + b_4\|\xi_k\|^{1+\min\{\theta,1\}}
\end{aligned} \tag{39}$$

where b_1, b_2, b_3 and b_4 are some constants. From Lemma 3.3, we have

$$\operatorname{dist}(R_{x_k}(s_k), x^*) = \operatorname{dist}(R_{x_k}(s_k), R_{x_k}(\xi_k)) \leq b_5\|s_k - \xi_k\|, \tag{40}$$

where b_5 is a constant. Combining (39) and (40) and using Lemma 3.4, we obtain

$$\operatorname{dist}(R_{x_k}(s_k), x^*) \leq b_0 b_5\|(H_k - \mathcal{B}_k)s_k\| + \bar{b}_4 b_5 e_k^{1+\min\{\theta,1\}}. \tag{41}$$

From (38), for k large enough such that $\|H_k^{-1}\|\|(H_k - \mathcal{B}_k)s_k\| \leq \frac{1}{2}\|s_k\|$, we have

$$\|s_k\| \leq \frac{1}{2}\|s_k\| + \|H_k^{-1}\|(\|\operatorname{grad} f(x_k)\| + \|\operatorname{grad} f(x_k)\|^{1+\theta}).$$

Using Lemma 3.8, this yields

$$\|s_k\| \leq b_6 \operatorname{dist}(x_k, x^*),$$

where b_6 is a constant. Using the latter in (41) yields

$$\operatorname{dist}(R_{x_k}(s_k), x^*) \leq b_0 b_5 b_6 \frac{\|(H_k - \mathcal{B}_k)s_k\|}{\|s_k\|} \operatorname{dist}(x_k, x^*) + \bar{b}_4 b_5 e_k^{1+\min\{\theta,1\}},$$

which shows (35).

We show (37) next. Define $\hat{f}_x(\eta) = f(R_x(\eta))$ and let $\zeta_k = R_{x^*}^{-1}(x_k)$. We have, for some $t \in (0,1)$,

$$\begin{aligned}\hat{f}_{x^*}(\zeta_k) - \hat{f}_{x^*}(0_{x^*}) &= g(\operatorname{grad} f(x^*), \zeta_k) + g(\operatorname{Hess} \hat{f}_{x^*}(t\zeta_k)[\zeta_k], \zeta_k) \\ &= g(\operatorname{Hess} \hat{f}_{x^*}(t\zeta_k)[\zeta_k], \zeta_k),\end{aligned}$$

where we have used (Euclidean) Taylor's theorem to get the first equality and the fact that x^* is a critical point of f (Assumption 3.2) for the second one. Therefore, since $\operatorname{Hess} \hat{f}_{x^*} = \operatorname{Hess} f(x^*)$ is positive definite (in view of [AMS08, Proposition 5.5.6] and Assumption 3.2), there exist b_7 and b_8 such that

$$b_7(\hat{f}_{x^*}(\zeta_k) - \hat{f}_{x^*}(0_{x^*})) \leq \|\hat{\zeta}_k\|^2 \leq b_8(\hat{f}_{x^*}(\zeta_k) - \hat{f}_{x^*}(0_{x^*}))$$

Then, using Lemma 3.4, we obtain that there exist b_9 and b_{10} such that

$$b_9(f(x_k) - f(x^*)) \leq \operatorname{dist}(x_k, x^*)^2 \leq b_{10}(f(x_k) - f(x^*)).$$

In other words,

$$b_9 h^2(x_k) \leq e_k^2 \leq b_{10} h^2(x_k),$$

and we have shown (37). Combining it with (35), we get (36). □

With Lemmas 3.15 and 3.16 in place, the rest of the local convergence analysis is essentially a translation of the analysis in [BKS96]. The next lemma generalizes [BKS96, Lemma 2.6].

Lemma 3.17. *If Assumptions 3.1, 3.2, 3.3, 3.4, 3.5, and 3.6 hold and the subproblem is solved accurately enough for* (9) *and* (10) *to hold, then*

$$\lim_{k\to\infty} \frac{h_k}{\Delta_k} = 0,$$

where $h_k = h(x_k)$.

Proof. Let p be the smallest integer greater than $2n + n(-\ln \tau_1 / \ln \tau_2)$, where τ_1 and τ_2 are defined in Algorithm 1. Then

$$\tau_1^n \tau_2^{p-2n} \geq 1. \tag{42}$$

Applying Lemma 3.15 with this value of p, there exists N such that if $k \geq N$, then there exists a set of at least $p - n$ indices, $\mathcal{G}_k \subset \{j : k+1 \leq j \leq k+p\}$, such that if $j \in \mathcal{G}_k$, then

$$\begin{aligned}\frac{\|(\mathcal{B}_j - H_j)s_j\|}{\|s_j\|} &< c\epsilon_k^{\frac{1}{n}}, \\ \rho_j &\geq 0.75.\end{aligned} \tag{43}$$

We now show that for such steps,

$$\frac{h_{j+1}}{\Delta_{j+1}} \leq \frac{1}{\tau_2} \frac{h_j}{\Delta_j}. \tag{44}$$

If $\|s_j\| \geq 0.8\Delta_j$, then since from Step 12 of Algorithm 1, $\Delta_{j+1} = \tau_2\Delta_j$ and since $\{h_i\}$ is decreasing, (44) follows. If on the other hand $\|s_j\| < 0.8\Delta_j$, then from Step 14 of Algorithm 1, we have that $\Delta_{j+1} = \Delta_j$. Also since the trust region is inactive, by condition (10), we have that $\mathcal{B}_j s_j = -\operatorname{grad} f(x_j) + \delta_k$, $\|\delta_k\| \leq \|\operatorname{grad} f(x_j)\|^{1+\theta}$. Therefore, in view of (36) in Lemma 3.16 and of (43), if N is large enough, we have that

$$h_{j+1} \leq \frac{1}{\tau_2} h_j.$$

This implies that (44) is true for all $j \in \mathcal{G}_j$, where $k \geq N$.

In addition, note that for any j, $h_{j+1} \leq h_j$ and $\Delta_{j+1} \geq \tau_1\Delta_j$ and so

$$\frac{h_{j+1}}{\Delta_{j+1}} \leq \frac{1}{\tau_1}\frac{h_j}{\Delta_j}. \tag{45}$$

Since (44) is true for $p-n$ values of $j \in \mathcal{G}_k$ and (45) holds for all j, we have that for all $k \geq N$,

$$\frac{h_{k+p}}{\Delta_{k+p}} \leq (\frac{1}{\tau_1})^n (\frac{1}{\tau_2})^{p-n} \frac{h_k}{\Delta_k} \leq (\frac{1}{\tau_2})^n \frac{h_k}{\Delta_k},$$

where the second inequality follows from (42). Therefore, starting at $k = N$, it follows that

$$\frac{h_{N+lp}}{\Delta_{N+lp}} \to 0$$

as $l \to \infty$. Using (45) again, we complete the proof. □

The next result generalizes [BKS96, Theorem 2.7].

Theorem 3.18. *If Assumptions 3.1, 3.2, 3.3, 3.4, 3.5, and 3.6 hold and the subproblem is solved accurately enough for* (9) *and* (10) *to hold then, the sequence* $\{x_k\}$ *generated by Algorithm 1 is* $n+1$*-step q-superlinear (where* n *denotes the dimension of* $\mathcal{M}$*); i.e.,*

$$\frac{\operatorname{dist}(x_{k+n+1}, x^*)}{\operatorname{dist}(x_k, x^*)} \to 0.$$

Proof. By Lemma 3.15, there exists N such that if $k \geq N$, then the set of steps $\{s_{k+1}, \ldots, s_{k+n+1}\}$ contains at least one step s_{k+j}, $1 \leq j \leq n+1$, for which

$$\frac{\|(\mathcal{B}_j - H_j)s_j\|}{\|s_j\|} < a_{13}\epsilon_k^{\frac{1}{n}}.$$

By (31) in Lemma 3.15 and (37) in Lemma 3.16 (when checking the assumptions, recall the standing assumption made in Section 3.3 that $e_k := \operatorname{dist}(x_k, x^*) \to 0$), there exists a constant b_0 such that

$$\|s_{k+j}\| < b_0 h_{k+j}.$$

Therefore, by Lemma 3.17, if N is large enough and $k \geq N$, then $\|s_{k+j}\| < 0.8\Delta_{k+j}$. By (10), this implies $\mathcal{B}_{k+j}s_{k+j} = -\operatorname{grad} f(x_{k+j}) + \delta_{k+j}$, with $\|\delta_{k+j}\| \leq \|\operatorname{grad} f(x_{k+j})\|^{1+\theta}$. Thus by inequality (36) of Lemma 3.16, if N is large enough and $k \geq N$, then

$$h_{k+j+1} = h(R_{x_{k+j}}(s_{k+j})) \leq (a_{17}a_{13}\epsilon_k^{\frac{1}{n}} + a_{18}h_{k+j}^{\min\{\theta,1\}})h_{k+j}.$$

The first equality holds because (32) implies that the step is accepted. Since the sequence $\{h_i\}$ is decreasing, this implies that

$$h_{k+n+1} \leq (a_{17}a_{13}\epsilon_k^{\frac{1}{n}} + a_{18}h_{k+j}^{\min\{\theta,1\}})h_k$$

By (37),

$$\begin{aligned} e_{k+n+1} &\leq a_{20}h_{k+n+1} \\ &\leq a_{20}(a_{17}a_{13}\epsilon_k^{\frac{1}{n}} + a_{18}h_{k+j}^{\min\{\theta,1\}})h_k \\ &\leq a_{20}(a_{17}a_{13}\epsilon_k^{\frac{1}{n}} + a_{18}(\frac{e_k}{a_{19}})^{\min\{\theta,1\}})\frac{e_k}{a_{19}}. \end{aligned}$$

This implies $n+1$-step q-superlinear convergence. □

It is also possible to extend to the Riemannian setting the result [BKS96, Theorem 2.8] that the percentage of $\mathcal{B}_k$ being positive semidefinite approaches 1 provided that $\mathcal{B}_k$ is positive semidefinite whenever $\|s_k\| \leq 0.8\Delta_k$. In the proof of [BKS96, Theorem 2.8], replace Lemma 2.6 by Lemma 3.17, Lemma 2.4 by Lemma 3.15, (2.14) by (31), and (2.9) by (37).

4 Limited memory version of RTR-SR1

In RTR-SR1 (Algorithm 1), storing $\mathcal{B}_{k+1} = \mathcal{T}_{\eta_k} \circ \tilde{\mathcal{B}}_{k+1} \circ \mathcal{T}_{\eta_k}^{-1}$ in matrix form may be inefficient for two reasons. The first reason, which is also present in the Euclidean case, is that $\tilde{\mathcal{B}}_{k+1} = \mathcal{B}_k + \frac{(y_k - \mathcal{B}_k s_k)(y_k - \mathcal{B}_k s_k)^\flat}{g(s_k, y_k - \mathcal{B}_k s_k)}$ is a rank-one modification of $\mathcal{B}_k$. The second reason, specific to the Riemannian setting, is that when $\mathcal{M}$ is a low-codimension submanifold of a Euclidean space $\mathcal{E}$, it may be beneficial to express $\mathcal{T}_{\eta_k}$ as the restriction to $\mathrm{T}_{x_k}\mathcal{M}$ of a low-rank modification of the identity (7). Instead of storing full dense matrices, it may then be beneficial to store a few vectors that implicitly represent them. This is the purpose of the limited memory version of RTR-SR1 presented in this section.

The proposed limited memory RTR-SR1, called LRTR-SR1, is described in Algorithm 2. It relies on a Riemannian generalization of the compact representation of the classical (Euclidean) SR1 matrices presented in [BNS94, §5]. We set $\mathcal{B}_0 = \mathrm{id}$. At step $k > 0$, we first choose a basic Hessian approximation $\mathcal{B}_0^k$, which in the Riemannian setting becomes a linear transformation of $\mathrm{T}_{x_k}\mathcal{M}$. We advocate the choice

$$\mathcal{B}_0^k = \gamma_k \,\mathrm{id},$$

where

$$\gamma_k = \frac{g(y_{k-1}, y_{k-1})}{g(s_{k-1}, y_{k-1})},$$

which generalizes a choice usually made in the Euclidean case [NW06, (7.20)]. As in the Euclidean case, we let $S_{k,m}$ and $Y_{k,m}$ contain the (at most) m most recent corrections, which in the Riemannian setting must be transported to $\mathrm{T}_{x_k}\mathcal{M}$, yielding $S_{k,m} = \{s_{k-\ell}^{(k)}, s_{k-\ell+1}^{(k)}, \ldots, s_{k-1}^{(k)}\}$ and $Y_{k,m} = \{y_{k-\ell}^{(k)}, y_{k-\ell+1}^{(k)}, \ldots, y_{k-1}^{(k)}\}$, where $\ell = \min\{m, k\}$ and where $s^{(k)}$ denotes s transported to $\mathrm{T}_{x_k}\mathcal{M}$. We then have the following Riemannian generalization of the limited-memory update based on [BNS94, (5.2)]:

$$\mathcal{B}_k = \mathcal{B}_0^k + (Y_{k,m} - \mathcal{B}_0^k S_{k,m})(P_{k,m} - S_{k,m}^\flat \mathcal{B}_0^k S_{k,m})^{-1}(Y_{k,m} - \mathcal{B}_0^k S_{k,m})^\flat, \quad k > 0,$$

where $P_{k,m} = D_{k,m} + L_{k,m} + L_{k,m}^T$, $D_{k,m} = \operatorname{diag}\{g(s_{k-\ell}, y_{k-\ell}), g(x_{k-\ell+1}, y_{k-\ell+1}), \ldots, g(s_{k-1}, y_{k-1})\}$, and

$$(L_{k,m})_{i,j} = \begin{cases} g(s_{k-\ell+i-1}, y_{k-\ell+j-1}), & \text{if } i > j; \\ 0, & \text{otherwise.} \end{cases}$$

Moreover, letting $Q_{k,m}$ denote the matrix $S_{k,m}^\flat S_{k,m}$, we obtain

$$\mathcal{B}_k = \gamma_k \operatorname{id} + (Y_{k,m} - \gamma_k S_{k,m})(P_{k,m} - \gamma_k Q_{k,m})^{-1}(Y_{k,m} - \gamma_k S_{k,m})^\flat, \quad k > 0. \tag{46}$$

For all $\eta \in \mathrm{T}_{x_k}\mathcal{M}$, $\mathcal{B}_k\eta$ can thus be obtained from (46) using $Y_{k,m}$, $S_{k,m}$, $P_{k,m}$ and $Q_{k,m}$. This is how $\mathcal{B}_k$ is defined in Algorithm 2, except that the technicality that the $\mathcal{B}$ update may be skipped is also taken into account therein.

5 Numerical experiments

As an illustration, we investigate the performance of RTR-SR1 (Algorithm 1) and LRTR-SR1 (Algorithm 2) on a Rayleigh quotient minimization problem on the sphere and on a joint diagonalization (JD) problem on the Stiefel manifold.

For the Rayleigh quotient problem, the manifold $\mathcal{M}$ is the sphere

$$\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : x^T x = 1\}$$

and the objective function f is defined by

$$f(x) = x^T A x, \tag{47}$$

where A is a given n-by-n symmetric matrix. Minimizing the Rayleigh quotient of A is equivalent to computing its leftmost eigenvector (see, e.g., [AMS08, §2.1.1]). The Rayleigh quotient problem provides convenient benchmarking experiments since, except for pathological cases, there is essentially one local minimizer (specifically, there is one pair of antipodal local minimizers), which can be readily computed by standard eigenvalue software for verification. Moreover, the Hessian of f is readily available, and this allows for a comparison with RTR-Newton [AMS08, Ch. 7], which corresponds to Algorithm 1 with the exception that $\mathcal{B}_k$ is replaced by the Hessian of f at x_k (and thus the vector transport is no longer needed).

In the JD problem considered, the manifold $\mathcal{M}$ is the (compact) Stiefel manifold,

$$\mathrm{St}(p, n) = \{X \in \mathbb{R}^{n \times p} : X^T X = I_p\},$$

and the objective function f is defined by

$$f(X) = -\sum_{i=1}^{N} \| \operatorname{diag}(X^T C_i X)\|^2, \tag{48}$$

where $C_1, \ldots, C_N$ are given symmetric matrices, $\operatorname{diag}(M)$ denotes the vector formed by the diagonal entries of M, and $\| \operatorname{diag}(M)\|^2$ thus denotes the sum of the squared diagonal elements of M. This problem has applications in independent component analysis for blind source separation [TCA09].

The comparisons are performed in Matlab 7.0.0 on a 32 bit Windows platform with 2.4 GHz CPU (T8300).

Algorithm 2 Limited-memory RTR-SR1 (LRTR-SR1)

Input: Riemannian manifold $\mathcal{M}$ with Riemannian metric g; retraction R; isometric vector transports $\mathcal{T}_S$; smooth function f on $\mathcal{M}$; initial iterate $x_0 \in \mathcal{M}$;

1: Choose an integer $m > 0$ and real numbers $\Delta_0 > 0$, $\nu \in (0,1)$, $c \in (0, 0.1)$, $\tau_1 \in (0,1)$ and $\tau_2 > 1$; Set $k \leftarrow 0$, $\ell \leftarrow 0$, $\gamma_0 \leftarrow 1$;

2: Obtain $s_k \in \mathrm{T}_{x_k}\mathcal{M}$ by (approximately) solving

$$s_k = \min_{s \in \mathrm{T}_{x_k}\mathcal{M}} m_k(s) = \min_{s \in \mathrm{T}_{x_k}\mathcal{M}} f(x_k) + g(\operatorname{grad} f(x_k), s) + \frac{1}{2} g(s, \mathcal{B}_k s), \text{s.t. } \|s\| \leq \Delta_k,$$

where $\mathcal{B}_k$ is defined in accordance with (46);

3: Set $\rho_k \leftarrow \frac{f(x_k) - f(R_{x_k}(s_k))}{m_k(0) - m_k(s_k)}$;

4: Set $y_k \leftarrow \mathcal{T}_{S_{\eta_k}}^{-1} \operatorname{grad} f(R_{x_k}(s_k)) - \operatorname{grad} f(x_k)$;

5: **if** $|g(s_k, y_k - \mathcal{B}_k s_k)| \geq \nu \|s_k\| \|y_k - \mathcal{B}_k s_k\|$ **then**

6: $\gamma_{k+1} \leftarrow \frac{g(y_k, y_k)}{g(s_k, y_k)}$; Add $s_k^{(k)}$ and $y_k^{(k)}$ into storage; If $\ell \geq m$, then discard vector pair $\{s_{k-\ell}^{(k)}, y_{k-\ell}^{(k)}\}$ from storage, else $\ell \leftarrow \ell + 1$; Compute matrices $P_{k,m}$ and $Q_{k,m}$ by updating $P_{k-1,m}$ and $Q_{k-1,m}$ if available;

7: **else**

8: Set $\gamma_{k+1} \leftarrow \gamma_k$, $P_{k+1,m} \leftarrow P_{k,m}$, $Q_{k+1,m} \leftarrow Q_{k,m}$ and $\{s_k^{(k)}, y_k^{(k)}\} \leftarrow \{s_{k-1}^{(k)}, y_{k-1}^{(k)}\}, \ldots, \{s_{k-\ell+1}^{(k)}, y_{k-\ell+1}^{(k)}\} \leftarrow \{s_{k-\ell}^{(k)}, y_{k-\ell}^{(k)}\}$.

9: **end if**

10: **if** $\rho_k > c$ **then**

11: $x_{k+1} \leftarrow R_{x_k}(s_k)$; Transport $s_{k-\ell+1}^{(k)}, s_{k-\ell+2}^{(k)}, \ldots, s_k^{(k)}$ and $y_{k-\ell+1}^{(k)}, y_{k-\ell+2}^{(k)}, \ldots, y_k^{(k)}$ from $\mathrm{T}_{x_k}\mathcal{M}$ to $\mathrm{T}_{x_{k+1}}\mathcal{M}$ by $\mathcal{T}_S$;

12: **else**

13: $x_{k+1} \leftarrow x_k$;

14: **end if**

15: **if** $\rho_k > \frac{3}{4}$ **then**

16: **if** $\|\eta_k\| \geq 0.8\Delta_k$ **then**

17: $\Delta_{k+1} \leftarrow \tau_2 \Delta_k$;

18: **else**

19: $\Delta_{k+1} \leftarrow \Delta_k$;

20: **end if**

21: **else if** $\rho_k < 0.1$ **then**

22: $\Delta_{k+1} \leftarrow \tau_1 \Delta_k$;

23: **else**

24: $\Delta_{k+1} \leftarrow \Delta_k$;

25: **end if**

26: $k \leftarrow k + 1$, goto 2 until convergence.

The chosen Riemannian metric g on $\mathbb{S}^{n-1}$ is obtained by making $\mathbb{S}^{n-1}$ a Riemannian submanifold of the Euclidean space $\mathbb{R}^n$. The gradient and the Hessian of f (47) with respect to the metric are given in [AMS08, §6.4]. The chosen Riemannian metric g on $\mathrm{St}(p,n)$ is the one obtained by viewing $\mathrm{St}(p,n)$ as a Riemannian submanifold of the Euclidean space $\mathbb{R}^{n \times p}$, as in [EAS98, §2.2]. With

respect to this Riemannian metric, the gradient of the objective function (48), required in the three methods, is given in [TCA09, §2.3]. The Riemannian Hessian of (48), required for RTR-Newton, is also given therein.

The initial Hessian approximation $\mathcal{B}_0$ is set to the identity in RTR-SR1. The θ, κ parameters in the inner iteration stopping criterion [AMS08, (7.10)] of the truncated CG inner iteration [AMS08, §7.3.2] are set to 0.1, 0.9 for RTR-SR1 and LRTR-SR1 and to 1, 0.1 for RTR-Newton. The initial radius Δ_0 is set to 1, ν is the square root of machine epsilon, c is set to 0.1, τ_1 to 0.25, and τ_2 to 2.

For the retraction R on $\mathbb{S}^{n-1}$, following [AMS08, Example 4.1.1], we choose $R_x(\eta) = (x+\eta)/\|x+\eta\|_2$. For the retraction R on $\mathrm{St}(p,n)$, following [AMS08, (4.8)], we choose $R_X(\eta) = \mathrm{qf}(X+\eta)$ where qf denotes the Q factor of the QR decomposition with nonnegative elements on the diagonal of R.

The chosen isometric vector transport $\mathcal{T}$ on $\mathbb{S}^{n-1}$ is the vector transport by rigging (7), which is (locally) uniquely defined in case of submanifolds of co-dimension 1. In the case of $\mathbb{S}^{n-1}$, it turns out to be equivalent to the parallel translation along the shortest geodesic that joins the origin point x and the target point y, i.e.,

$$\mathcal{T}_{\eta_x}\xi_x = \xi_x - \frac{2y^T\xi_x}{\|x+y\|^2}(x+y), \tag{49}$$

where $y = R_x(\eta_x)$. This operation is well defined whenever x and y are not antipodal points. On $\mathrm{St}(p,n)$, since we will conduct experiments on problems of small dimension, we find it preferable to select a vector transport by parallelization (6), which amounts to selecting a smooth field of tangent bases B on $\mathrm{St}(p,n)$. To this end, we note that $\mathrm{T}_X\,\mathrm{St}(p,n) = \{X\Omega + X_\perp K : \Omega^T = -\Omega,\ K \in \mathbb{R}^{(n-p)\times p}\}$, where the columns of $X_\perp \in \mathbb{R}^{n\times(n-p)}$ form an orthonormal basis of the orthogonal complement of the column space of X (see, e.g., [AMS08, Example 3.5.2]). Hence, an orthonormal basis of $\mathrm{T}_X\,\mathrm{St}(p,n)$ is given by $\{\frac{1}{\sqrt{2}}X(e_ie_j^T - e_je_i^T) : i = 1,\ldots,p, j = i+1,\ldots,p\} \cup \{X_\perp\tilde{e}_ie_j^T, i = 1,\ldots,n-p, j = 1,\ldots,p\}$, where $(e_1,\ldots,e_p)$ is the canonical basis of $\mathbb{R}^p$ and $(\tilde{e}_1,\ldots,\tilde{e}_{n-p})$ is the canonical basis of $\mathbb{R}^{n-p}$. To ensure that the obtained field of tangent bases is smooth, we need to choose $X_\perp$ as a smooth function of X. This can be done locally by extracting the $n-p$ last columns of the Gram-Schmidt orthonormalization of $\begin{bmatrix} X & C \end{bmatrix}$ where C is a given $n \times (n-p)$ matrix. In practice, we have observed that Matlab's `null` function applied to X works adequately to produce an $X_\perp$.

The experiments on the Rayleigh quotient (47) are reported in Table 1 and Figure 1. Since the Hessian is readily available, RTR-Newton is the a priori method of choice over the proposed RTR-SR1 and LRTR-SR1 methods. Nevertheless, Table 1 illustrates that it is possible to exhibit a matrix A for which LRTR-SR1 is faster than RTR-Newton. Specifically, matrix A in the Rayleigh quotient (47) is set to UDU^T, where U is an orthonormal matrix obtained by orthonormalizing a random matrix whose elements are drawn from the standard normal distribution and $D = \mathrm{diag}(0, 0.01, \ldots, 0.01, 2, \ldots, 2)$ with 0.01 and 2 occurring $n/2-1$ and $n/2$ times, respectively. The initial iterate is generated randomly. The number of function evaluations is equal to the number of iterations (*iter*). *ng* denotes the number of gradient evaluations. The differences between *iter* and *ng* that may be observed in RTR-Newton are due to occasional rejections of the candidate new iterate as prescribed in the trust-region framework [AMS08, §7.2]; for RTR-SR1 and LRTR-SR1, *iter* and *ng* are identical because one evaluation of the gradient is required at each iterate to update $\mathcal{B}_k$ or store y_k even if the candidate is rejected. *nH* denotes the number of operations of the form $\mathrm{Hess}\, f(x)\eta$ or $\mathcal{B}\eta$. *t* denotes the run time (seconds). To obtain sufficiently stable timing values, an

average is taken over several identical runs for a total run time of at least one minute. gf_0 and gf_f denote the initial and final norm of the gradient. Two stopping criteria are tested: gf_f/gf_0 <1e-3 and gf_f/gf_0 <1e-6.

Unsurprisingly, Table 1 shows that RTR-Newton, which exploits the Hessian of f, requires fewer iterations than the SR1 methods, which does not use this information. However, when n gets large, the time per iteration in LRTR-SR1—with moderate memory size m—gets sufficiently smaller for the method to become faster than RTR-Newton.

Table 2 and Figure 2 present the experimental results obtained for the JD problem (48). The C_i matrices are selected as $C_i = \mathrm{diag}(n, n-1, \ldots, 1) + \epsilon_C(R_i + R_i^T)$, where the elements of $R_i \in \mathbb{R}^{n\times n}$ are independently drawn from the standard normal distribution. Table 2 and Figure 2 correspond to $\epsilon_C = 0.1$, but we have observed similar results for a wide range of values of ϵ_C. Table 2 indicates that RTR-Newton requires fewer iterations than RTR-SR1, which requires fewer iterations than LRTR-SR1. This was expected since RTR-Newton uses the Hessian of f while RTR-SR1 uses an inexact Hessian and LRTR-SR1 is further constrained by the limited memory. However, the iterations of RTR-Newton tend to be more time-consuming that those of the SR1 methods, all the more so if N gets large since the number of terms in the Hessian of f is linear in N. The experiments reported in Table 2 show that the trade-off between the number of iterations and the time per iteration is in favor of RTR-SR1 for N sufficiently large. Note also that, even though it is slower than the two other methods in the experiments presented in Table 2, LRTR-SR1 may be the method of choice in certain circumstances as it does not require the Hessian of f (unlike RTR-Newton) and it has a reduced memory usage in comparison with RTR-SR1.

Table 1: Rayleigh quotient experiments

n		RTR-Newton		RTR-SR1		LRTR-SR1					
						$m=0$		$m=2$		$m=4$	
		1e-3	1e-6	1e-3	1e-6	1e-3	1e-6	1e-3	1e-6	1e-3	1e-6
64	*iter*	3	6	4	15	4	50	4	18	4	13
	ng	3	6	4	15	4	50	4	18	4	13
	nH	4	13	6	34	0	0	0	0	0	0
	gf_f	1.48_{-4}	9.26_{-9}	1.47_{-4}	3.58_{-8}	1.47_{-4}	1.18_{-6}	1.47_{-4}	7.65_{-7}	1.47_{-4}	3.49_{-8}
	gf_f/gf_0	7.42_{-5}	4.66_{-9}	7.38_{-5}	1.80_{-8}	7.38_{-5}	5.94_{-7}	7.38_{-5}	3.85_{-7}	7.38_{-5}	1.76_{-8}
	t	1.25_{-3}	3.12_{-3}	3.42_{-3}	1.39_{-2}	2.85_{-3}	4.69_{-2}	3.27_{-3}	2.63_{-2}	4.55_{-3}	2.40_{-2}
256	*iter*	3	9	4	13	4	43	4	13	4	15
	ng	3	9	4	13	4	43	4	13	4	15
	nH	4	20	6	29	0	0	0	0	0	0
	gf_f	1.80_{-3}	9.13_{-12}	1.81_{-3}	2.28_{-7}	1.81_{-3}	1.57_{-6}	1.81_{-3}	7.57_{-8}	1.81_{-3}	1.08_{-9}
	gf_f/gf_0	9.09_{-4}	4.60_{-12}	9.14_{-4}	1.15_{-7}	9.14_{-4}	7.90_{-7}	9.14_{-4}	3.82_{-8}	9.14_{-4}	5.45_{-10}
	t	2.81_{-3}	1.58_{-2}	2.11_{-2}	6.80_{-2}	3.61_{-3}	5.24_{-2}	6.93_{-3}	2.77_{-2}	8.39_{-3}	3.64_{-2}
1024	*iter*	3	9	4	14	4	53	4	13	4	12
	ng	3	9	4	14	4	53	4	13	4	12
	nH	4	19	6	30	0	0	0	0	0	0
	gf_f	1.68_{-3}	5.28_{-12}	1.68_{-3}	1.78_{-8}	1.68_{-3}	8.26_{-8}	1.68_{-3}	3.32_{-8}	1.68_{-3}	3.90_{-8}
	gf_f/gf_0	8.47_{-4}	2.65_{-12}	8.47_{-4}	8.93_{-9}	8.47_{-4}	4.15_{-8}	8.47_{-4}	1.67_{-8}	8.47_{-4}	1.96_{-8}
	t	5.50_{-2}	2.43_{-1}	2.00_{-1}	9.96_{-1}	2.94_{-2}	4.07_{-1}	3.01_{-2}	1.06_{-1}	3.09_{-2}	1.05_{-1}

Table 2: Joint diagonalization (JD) experiments: $n = 12, p = 4$, ϵ_C = 1e-1

N		RTR-Newton		RTR-SR1		LRTR-SR1					
						$m = 2$		$m = 4$		$m = 8$	
		1e-3	1e-6	1e-3	1e-6	1e-3	1e-6	1e-3	1e-6	1e-3	1e-6
16	*iter*	10	12	58	81	80	328	61	150	57	131
	ng	10	12	58	81	80	328	61	150	57	131
	nH	63	96	160	253	0	0	0	0	0	0
	gf_f	3.03_{-1}	5.99_{-4}	1.93	2.26_{-3}	2.37	2.23_{-3}	2.37	2.21_{-3}	2.14	1.96_{-3}
	gf_f/gf_0	1.21_{-4}	2.40_{-7}	7.74_{-4}	9.07_{-7}	9.48_{-4}	8.91_{-7}	9.48_{-4}	8.85_{-7}	8.55_{-4}	7.83_{-7}
	t	4.62_{-2}	5.92_{-2}	5.48_{-2}	7.78_{-2}	1.00_{-1}	3.78_{-1}	8.75_{-2}	2.13_{-1}	1.01_{-1}	2.41_{-1}
64	*iter*	14	16	64	88	163	402	83	176	109	199
	ng	14	16	64	88	163	402	83	176	109	199
	nH	93	120	186	288	0	0	0	0	0	0
	gf_f	1.89	2.58_{-3}	5.54	5.48_{-3}	8.63	8.57_{-3}	8.79	7.39_{-3}	7.82	7.28_{-3}
	gf_f/gf_0	2.13_{-4}	2.90_{-7}	6.24_{-4}	6.18_{-7}	9.72_{-4}	9.66_{-7}	9.91_{-4}	8.33_{-7}	8.82_{-4}	8.20_{-7}
	t	1.73_{-1}	2.19_{-1}	1.01_{-1}	1.43_{-1}	3.24_{-1}	7.34_{-1}	1.74_{-1}	3.65_{-1}	2.73_{-1}	5.01_{-1}
256	*iter*	10	13	54	82	122	372	100	168	81	165
	ng	10	13	54	82	122	372	100	168	81	165
	nH	64	109	148	240	0	0	0	0	0	0
	gf_f	3.49_{1}	8.08_{-3}	3.05_{1}	3.95_{-2}	3.99_{1}	2.85_{-2}	3.83_{1}	2.77_{-2}	2.77_{1}	4.01_{-2}
	gf_f/gf_0	8.65_{-4}	2.00_{-7}	7.56_{-4}	9.79_{-7}	9.89_{-4}	7.07_{-7}	9.51_{-4}	6.88_{-7}	6.87_{-4}	9.94_{-7}
	t	4.14_{-1}	6.78_{-1}	2.31_{-1}	3.57_{-1}	6.83_{-1}	1.70	4.77_{-1}	8.06_{-1}	4.20_{-1}	8.71_{-1}

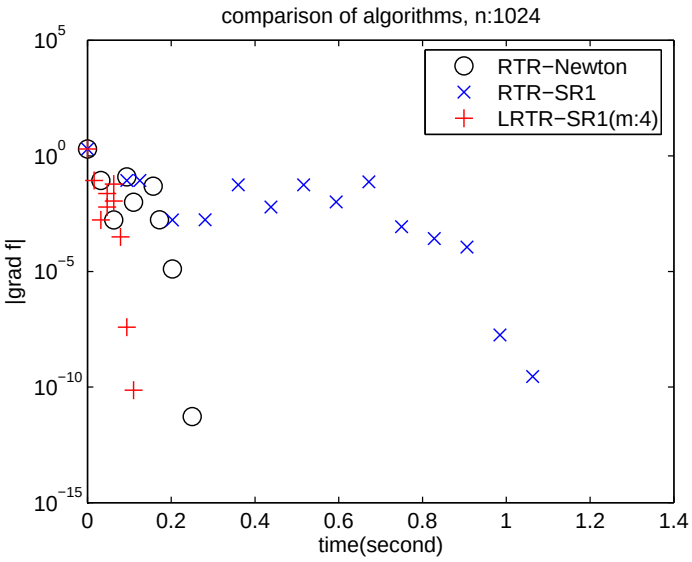

Figure 1: Comparison of RTR-Newton and the new methods RTR-SR1 and LRTR-SR1 for the Rayleigh quotient problem (47) with $n = 1024$

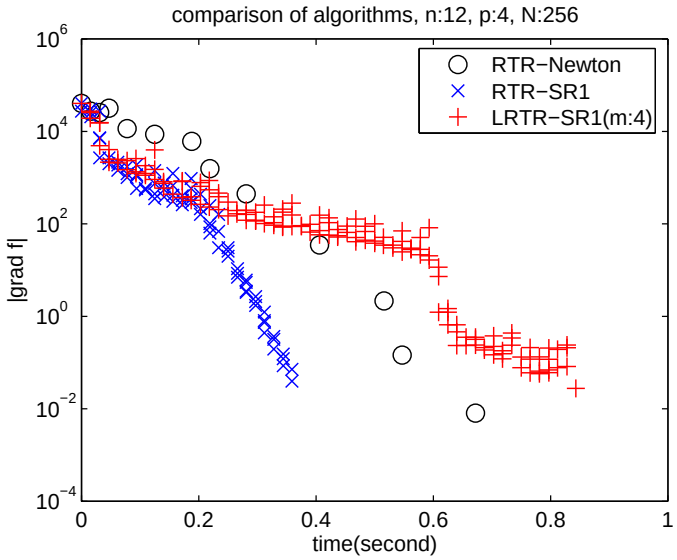

Figure 2: Comparison of RTR-Newton and the new methods RTR-SR1 and LRTR-SR1 for the joint diagonalization problem (48) with $n = 12$, $p = 4$, $N = 1024$

6 Conclusion

We have introduced a Riemannian SR1 trust-region method, where the second-order term of the model is generated using a Riemannian generalization of the classical SR1 update. Global convergence to stationary points and $d + 1$-step superlinear convergence are guaranteed, and the experiments reported here show promise. A limited-memory version of the algorithm has also been presented. The new algorithms will be made available in the Manopt toolbox [BM].

References

[ABG07] P.-A. Absil, C. G. Baker, and K. A. Gallivan. Trust-region methods on Riemannian manifolds. *Foundations of Computational Mathematics*, 7(3):303–330, 2007.

[ADM02] R. L. Adler, J.-P. Dedieu, and J. Y. Margulies. Newton's method on Riemannian manifolds and a geometric model for the human spine. *IMA Journal of Numerical Analysis*, 22(3):359–390, 2002.

[AM12] P.-A. Absil and J. Malick. Projection-like retractions on matrix manifolds. *SIAM Journal on Optimization*, 22(1):135–158, 2012.

[AMS08] P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization algorithms on matrix manifolds*. Princeton University Press, Princeton, NJ, 2008.

[ATV13] B. Afsari, R. Tron, and R. Vidal. On the convergence of gradient descent for finding the Riemannian center of mass. *SIAM Journal on Control and Optimization*, 51(3):2230–2260, 2013.

[BA11] N. Boumal and P.-A. Absil. RTRMC: A Riemannian trust-region method for low-rank matrix completion. *Advances in Neural Information Processing Systems 24 (NIPS)*, pages 406–414, 2011.

[BKS96] R. H. Byrd, H. F. Khalfan, and R. B. Schnabel. Analysis of a symmetric rank-one trust region method. *SIAM Journal on Optimization*, 6(4):1025–1039, 1996.

[BM] N. Boumal and B. Mishra. The Manopt toolbox. http://www.manopt.org.

[BM06] I. Brace and J. H. Manton. An improved BFGS-on-manifold algorithm for computing weighted low rank approximations. *Proceedings of 17th international Symposium on Mathematical Theory of Networks and Systems*, pages 1735–1738, 2006.

[BNS94] R. H. Byrd, J. Nocedal, and R. B. Schnabel. Representations of quasi-Newton matrices and their use in limited memory methods. *Mathematical Programming*, 63(1-3):129–156, 1994.

[Boo03] W. M. Boothby. *An introduction to differentiable manifolds and Riemannian geometry.* Academic Press, second edition, 2003.

[Bor12] R. Borsdorf. *Structured matrix nearness problems: theory and algorithms.* PhD thesis, The University of Manchester, 2012.

[CGT91] A. R. Conn, N. I. M. Gould, and P. L. Toint. Convergence of quasi-Newton matrices generated by the symmetric rank one update. *Mathematical Programming*, 50(1-3):177–195, March 1991.

[CGT00] A. R. Conn, N. I. M. Gould, and P. L. Toint. *Trust-region methods.* MPS/SIAM Series on Optimization. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000.

[Cha06] I. Chavel. *Riemannian geometry: a modern introduction.* Cambridge Studies in Advanced Mathematics, 2nd edition, 2006.

[dC92] M. P. do Carmo. *Riemannian geometry.* Mathematics: Theory & Applications, 1992.

[EAS98] A. Edelman, T. A. Arias, and S. T. Smith. The geometry of algorithms with orthogonality constraints. *SIAM Journal on Matrix Analysis and Applications*, 20(2):303–353, January 1998.

[Gab82] D Gabay. Minimizing a differentiable function over a differential manifold. *Journal of Optimization Theory and Applications*, 37(2):177–219, 1982.

[GQA12] K. A. Gallivan, C. Qi, and P.-A. Absil. A Riemannian Dennis-More condition. In Michael W. Berry, Kyle A. Gallivan, Efstratios Gallopoulos, Ananth Grama, Bernard Philippe, Yousef Saad, and Faisal Saied, editors, *High-Performance Scientific Computing*, pages 281–293. Springer London, 2012.

[Hua13] W. Huang. *Optimization algorithms on Riemannian manifolds with applications.* PhD thesis, Florida State University, 2013.

[IAVD11] M. Ishteva, P.-A. Absil, S. Van Huffel, and L. De Lathauwer. Best low multilinear rank approximation of higher-order tensors, based on the Riemannian trust-region scheme. *SIAM Journal on Matrix Analysis and Applications*, 32(1):115–135, 2011.

[JBAS10] M. Journée, F. Bach, P.-A. Absil, and R. Sepulchre. Low-rank optimization on the cone of positive semidefinite matrices. *SIAM Journal on Optimization*, 20(5):2327–2351, 2010.

[JD13] B. Jiang and Y.-H. Dai. A framework of constraint preserving update schemes for optimization on the Stiefel manifold, 2013. arXiv:1301.0172.

[KBS93] H. F. Khalfan, R. H. Byrd, and R. B. Schnabel. A theoretical and experimental study of the symmetric rank-one update. *SIAM Journal on Optimization*, 3(1):1–24, February 1993.

[KS12] M. Kleinsteuber and H. Shen. Blind source separation with compressively sensed linear mixtures. *IEEE Signal Processing Letters*, 19(2):107–110, 2012.

[MMBS11] B. Mishra, G. Meyer, F. Bach, and R. Sepulchre. Low-rank optimization with trace norm penalty, 2011. arXiv:1112.2318v2.

[NW06] J. Nocedal and S. J. Wright. *Numerical optimization*. Springer, second edition, 2006.

[O'N83] B. O'Neill. *Semi-Riemannian geometry*. Academic Press Incorporated [Harcourt Brace Jovanovich Publishers], 1983.

[Qi11] C. Qi. *Numerical optimization methods on Riemannian manifolds*. PhD thesis, Florida State University, 2011.

[RW12] W. Ring and B. Wirth. Optimization methods on Riemannian manifolds and their application to shape space. *SIAM Journal on Optimization*, 22(2):596–627, January 2012.

[SAGQ12] S. E. Selvan, U. Amato, K. A. Gallivan, and C. Qi. Descent algorithms on oblique manifold for source-adaptive ICA contrast. *IEEE Transactions on Neural Networks and Learning Systems*, 23(12):1930–1947, 2012.

[SI13] H. Sato and T. Iwai. Convergence analysis for the Riemannian conjugate gradient method, 2013. arXiv:1302.0125v1.

[SKH13] M. Seibert, M. Kleinsteuber, and K. Hüper. Properties of the BFGS method on Riemannian manifolds. *Mathematical System Theory Festschrift in Honor of Uwe Helmke on the Occasion of his Sixtieth Birthday*, pages 395–412, 2013.

[SL10] B. Savas and L. H. Lim. Quasi-Newton methods on Grassmannians and multilinear approximations of tensors. *SIAM Journal on Scientific Computing*, 32(6):3352–3393, 2010.

[Sti35] E. Stiefel. Richtungsfelder und Fernparallelismus in n-dimensionalen Mannigfaltigkeiten. *Commentarii Mathematici Helvetici*, 8(1):305–353, 1935.

[TCA09] F. J. Theis, T. P. Cason, and P.-A. Absil. Soft dimension reduction for ICA by joint diagonalization on the Stiefel manifold. *Proceedings of the 8th International Conference on Independent Component Analysis and Signal Separation*, 5441:354–361, 2009.

[TVSC11] P. Turaga, A. Veeraraghavan, A. Srivastava, and R. Chellappa. Statistical computations on Grassmann and Stiefel manifolds for image and video-based recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(11):2273–86, November 2011.

[VV10] B. Vandereycken and S. Vandewalle. A Riemannian optimization approach for computing low-rank solutions of Lyapunov equations. *SIAM Journal on Matrix Analysis and Applications*, 31(5):2553–2579, January 2010.

[WY12] Z. Wen and W. Yin. A feasible method for optimization with orthogonality constraints. *Mathematical Programming, Published online*, August 2012. doi:10.1007/s10107-012-0584-1.