A Riemannian BFGS method without differentiated retraction for nonconvex optimization problems

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

Abstract

In this paper, a Riemannian BFGS method for minimizing a smooth function on a Riemannian manifold is defined, based on a Riemannian generalization of a cautious update and a weak line search condition. It is proven that the Riemannian BFGS method converges (i) globally to stationary points without assuming the objective function to be convex and (ii) superlinearly to a nondegenerate minimizer. Using the weak line search condition removes the need for information from differentiated retraction. The joint matrix diagonalization problem is chosen to demonstrate the performance of the algorithms with various parameters, line search conditions and pairs of retraction and vector transport. A preliminary version can be found in [HAG16a].

1 Introduction

In the Euclidean setting, the BFGS method is a well-known quasi-Newton method that has been viewed for many years as the best quasi-Newton method for solving unconstrained optimization problems [DS83, NW06]. Its global and superlinear local convergence have been analyzed in many papers for convex problems (see [DS83] and references therein). For many years, it was not clear whether the standard BFGS method could be shown to converge globally for a nonconvex cost function. In fact, it is only recently that [Dai13] has given an example where the cost function is smooth (polynomial) and nonconvex and the BFGS method cannot converge, which means that modifications of the BFGS method are required if the global convergence for general nonconvex problems is desired. Such modifications have been proposed and shown to converge globally in [LF01a, LF01b, WLQ06, WYYL04]. In the Riemannian setting, quasi-Newton methods are also favored in many applications and much attention has been paid to generalizing the Euclidean BFGS method. So far, many Riemannian versions of the BFGS method have appeared, e.g., [Gab82, BM06, QGA10, SL10, RW12, SKH13, HGA15], and only two of them [RW12, HGA15] have general discussions with complete global and local convergence analyses rather than considering only a specific cost function or manifold. In spite of this, both of them require cost functions to satisfy a Riemannian version of convexity in their analyses.

In this paper, we adopt the approach in [LF01b] for nonconvex problems, which updates the Hessian approximation cautiously, with a weak line search condition [BN89, (3.2), (3.3)]. Global and local superlinear convergence analyses of the proposed cautious Riemannian BFGS method are given. In addition, unlike the Riemannian versions of the BFGS method in [RW12, HGA15], the global convergence analysis of the cautious Riemannian BFGS does not require a convexity assumption on the cost function.

The Riemannian BFGS method (RBFGS) in this paper has another advantage over the approaches in [RW12, HGA15]. The Riemannian versions of BFGS method in [RW12, HGA15] and this paper all rely on the notion of retraction and vector transport developed in [ADM02, AMS08]. The version in [RW12] involves

[†]Department of Computational and Applied Mathematics, 3087 Duncan Hall, Rice University, Houston, TX, USA.
[‡]Department of Mathematical Engineering, ICTEAM Institute, Université catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium.
[§]Department of Mathematics, 208 Love Building, 1017 Academic Way, Florida State University, Tallahassee FL 32306-4510, USA.
[¶]Corresponding author. E-mail: huwst08@gmail.com.

the vector transport of differentiated retraction which may not be available to users or may be too expensive. The version in [HGA15] uses less information from differentiated retraction–probably the least possible if the Wolfe condition is used in line search. The Riemannian BFGS method in this paper also allows the use of line search conditions other than the Wolfe condition and the requirement of differentiated retraction is completely avoided. For example, the Armijo-Goldstein condition is such a condition and is satisfied by step size found by the frequently-used back-tracking line search algorithm.

The joint diagonalization problem [TCA09] is used to demonstrate that the proposed Riemannian BFGS framework makes it possible to use simpler line search procedures and simpler vector transports than in earlier Riemannian BFGS methods (such as the one in [HGA15] which was found to be the best performing Riemannian method for this problem), without significantly affecting the efficiency of the algorithm.

This paper is organized as follows. Section 2 presents notation used in this paper. Section 3 defines the Riemannian version of BFGS method. Global and local convergence analyses are given in Sections 4 and 5 respectively. The corresponding limited-memory version of the RBFGS method is given in Section 6. Numerical experiments are reported in Section 7 and finally conclusions are drawn in Section 8.

2 Notation

The Riemannian concepts follow from the standard literature, e.g., [Boo86, AMS08] and the notation of this paper follows [AMS08]. Let f denote a cost function defined on a d-dimensional Riemannian manifold $\mathcal{M}$ with the Riemannian metric $g : (\eta_x, \xi_x) \mapsto g_x(\eta_x, \xi_x) \in \mathbb{R}$. $\mathrm{T}_x \mathcal{M}$ denotes the tangent space of $\mathcal{M}$ at x and $\mathrm{T}\mathcal{M}$ denotes the tangent bundle, i.e., the set of all tangent spaces. The Riemannian gradient and Hessian of f at x are denoted by $\operatorname{grad} f(x)$ and $\operatorname{Hess} f(x)$ respectively and the action of $\operatorname{Hess} f(x)$ on a tangent vector $\eta_x \in \mathrm{T}_x \mathcal{M}$ is denoted by $\operatorname{Hess} f(x)[\eta_x]$. Let $\mathcal{A}_x$ be a linear operator on $\mathrm{T}_x \mathcal{M}$. $\mathcal{A}_x^*$ denotes the adjoint operator of $\mathcal{A}_x$, i.e., $\mathcal{A}_x^*$ satisfies $g_x(\eta_x, \mathcal{A}_x \xi_x) = g_x(\mathcal{A}_x^* \eta_x, \xi_x)$ for all $\eta_x, \xi_x \in \mathrm{T}_x \mathcal{M}$. $\mathcal{A}_x$ is called self-adjoint or symmetric with respect to g if $\mathcal{A}_x^* = \mathcal{A}_x$. Given $\eta_x \in \mathrm{T}_x \mathcal{M}$, $\eta_x^\flat$ represents the flat of η_x, i.e., $\eta_x^\flat : \mathrm{T}_x \mathcal{M} \to \mathbb{R} : \xi_x \mapsto g_x(\eta_x, \xi_x)$.

A retraction is a C^1 map from the tangent bundle to the manifold such that (i) $R(0_x) = x$ for all $x \in \mathcal{M}$ (where 0_x denotes the origin of $\mathrm{T}\mathcal{M}$) and (ii) $\frac{d}{dt} R(t\xi_x)|_{t=0} = \xi_x$ for all $\xi_x \in \mathrm{T}_x \mathcal{M}$. The domain of R does not need the entire tangent bundle. However, it is usually the case in practice. In this paper, we assume that R is well-defined whenever needed. R_x denotes the restriction of R to $\mathrm{T}_x \mathcal{M}$. A vector transport $\mathcal{T} : \mathrm{T}\mathcal{M} \oplus \mathrm{T}\mathcal{M} \to \mathrm{T}\mathcal{M}, (\eta_x, \xi_x) \mapsto \mathcal{T}_{\eta_x} \xi_x$ with associated retraction R is a mapping[1] such that, for all (x, η_x) in the domain of R and all $\xi_x \in \mathrm{T}_x \mathcal{M}$, it holds that (i) $\mathcal{T}_{\eta_x} \xi_x \in \mathrm{T}_{R(\eta_x)} \mathcal{M}$, (ii) $\mathcal{T}_{\eta_x}$ is a linear map. An isometric vector transport $\mathcal{T}_{\mathrm{S}}$ additionally satisfies $g_{R_x(\eta_x)}(\mathcal{T}_{\mathrm{S}_{\eta_x}} \xi_x, \mathcal{T}_{\mathrm{S}_{\eta_x}} \zeta_x) = g_x(\xi_x, \zeta_x)$, where $\zeta_x \in \mathrm{T}_x \mathcal{M}$. The vector transport by differentiated retraction $\mathcal{T}_R$ is defined to be $\mathcal{T}_{R_{\eta_x}} \xi_x := \frac{d}{dt} R_x(\eta_x + t\xi_x)|_{t=0}$.

Coordinate expressions are denoted with a hat. Let $(\mathcal{U}, \varphi)$ be a chart of a manifold $\mathcal{M}$ and $x \in \mathcal{U}$. The coordinate expression of x is defined by $\hat{x} = \varphi(x)$. E_i, the i-th coordinate vector field of $(\mathcal{U}, \varphi)$, is defined by

$$(E_i f)(x) := \partial_i (f \circ \varphi^{-1})(\varphi(x)) = D(f \circ \varphi^{-1})(\varphi(x))[e_i].$$

These coordinate vector fields are smooth and every vector field ξ on $\mathcal{U}$ has a decomposition

$$\xi = \sum_i (\xi \varphi_i) E_i.$$

Therefore, $(E_i)_x, i = 1, \dots, d$ is a basis of $\mathrm{T}_x \mathcal{M}$ and the coordinate expression $\hat{\xi}_x$ of ξ_x with this basis is $(\xi_x \varphi_1, \dots, \xi_x \varphi_d)$. The coordinate expressions of a vector transport $\mathcal{T}$ and a linear operator $\mathcal{A}$ on a tangent space are represented by matrices $\hat{\mathcal{T}}$ and $\hat{\mathcal{A}}$ respectively.

[1]This mapping is not required to be continuous. We make further assumptions when needed. Note that in the global convergence analysis (Section 4), the only constraints on $\mathcal{T}_{\mathrm{S}}$ are given in Assumption 4.2.

3 Riemannian BFGS Method with Cautious Update

The proposed Riemannian BFGS method with cautious update is stated in Algorithm 1.

Algorithm 1 Cautious RBFGS method

Input: Riemannian manifold $\mathcal{M}$ with Riemannian metric g; retraction R; isometric vector transport $\mathcal{T}_{\mathrm{S}}$, with R as the associated retraction; continuously differentiable real-valued function f on $\mathcal{M}$, bounded below; initial iterate $x_0 \in \mathcal{M}$; initial Hessian approximation $\mathcal{B}_0$ that is symmetric positive definite with respect to the metric g; convergence tolerance $\varepsilon > 0$;

1: $k \leftarrow 0$;
2: **while** $\|\operatorname{grad} f(x_k)\| > \varepsilon$ **do**
3: Obtain $\eta_k \in \mathrm{T}_{x_k} \mathcal{M}$ by solving $\mathcal{B}_k \eta_k = -\operatorname{grad} f(x_k)$;
4: Set

$$x_{k+1} = R_{x_k}(\alpha_k \eta_k), \tag{1}$$

where $\alpha_k > 0$ is computed by a line-search procedure that guarantees the existence of $0 < \chi_1 < 1$ and $0 < \chi_2 < 1$, independent of k such that

$$h_k(\alpha_k) - h_k(0) \leq -\chi_1 \frac{h_k'(0)^2}{\|\eta_k\|^2} \tag{2}$$

or

$$h_k(\alpha_k) - h_k(0) \leq \chi_2 h_k'(0), \tag{3}$$

where $h_k(t) = f(R_{x_k}(t\eta_k))$.
5: Define the linear operator $\mathcal{B}_{k+1} : \mathrm{T}_{x_{k+1}} \mathcal{M} \to \mathrm{T}_{x_{k+1}} \mathcal{M}$ by (7);
6: $k \leftarrow k + 1$;
7: **end while**

When $\mathcal{M}$ is a Euclidean space, the line search condition in Step 4 of Algorithm 1 is weak since it has been shown in [BN89, Sections 3 and 4] and references therein that many line search conditions, e.g., the Curry-Altman condition, the Goldstein condition, the Wolfe condition and the Armijo-Goldstein condition, imply either (2) or (3) if the gradient of the function is Lipschitz continuous. In the Riemannian setting, note that function $f \circ R_x : \mathrm{T}_x \mathcal{M} \to \mathbb{R}$ is defined on a linear space. It follows that the Euclidean results about line search are applicable, i.e., the conditions above also imply either (2) or (3) if the gradient of the function satisfies the Riemannian Lipschitz continuous condition in Definition 3.1:

Definition 3.1. *[AMS08, Definition 7.4.1] The function $\hat{f} = f \circ R$ is radially L-C^1 function for all $x \in \Omega$ if there exists a positive constant L_2 such that for all $x \in \Omega$ and all $\eta_x \in \mathrm{T}_x \mathcal{M}$, it holds that*

$$\left|\frac{d}{d\tau}\hat{f}(\tau\eta_x)|_{\tau=t} - \frac{d}{d\tau}\hat{f}(\tau\eta_x)|_{\tau=0}\right| \leq L_2 t \|\eta\|^2 \tag{4}$$

where t and η_x satisfy that $R_x(t\eta_x) \in \Omega$.

In an RBFGS method, the search direction η_k is given by solving

$$\eta_k = -\mathcal{B}_k^{-1} \operatorname{grad} f(x_k), \tag{5}$$

where $\mathcal{B}_k^{-1}$, a linear operator on $\mathrm{T}_{x_k} \mathcal{M}$, approximates the action of the inverse Hessian along $\operatorname{grad} f(x_k)$ direction. It remains to define the update formula for $\mathcal{B}_k$ to be used in line 5 of Algoriithm 1. The classical (Euclidean) BFGS update admits several Riemannian generalizations; see [HGA15, Section 6]. We will start from the following one:

$$\mathcal{B}_{k+1} = \tilde{\mathcal{B}}_k - \frac{\tilde{\mathcal{B}}_k s_k (\tilde{\mathcal{B}}_k^* s_k)^\flat}{(\tilde{\mathcal{B}}_k^* s_k)^\flat s_k} + \frac{y_k y_k^\flat}{y_k^\flat s_k}, \tag{6}$$

where $\tilde{\mathcal{B}}_k = \mathcal{T}_{S_{\alpha_k \eta_k}} \circ \mathcal{B}_k \circ \mathcal{T}_{S_{\alpha_k \eta_k}}^{-1}$, $y_k = \beta_k^{-1} \operatorname{grad} f(x_{k+1}) - \mathcal{T}_{S_{\alpha_k \eta_k}} \operatorname{grad} f(x_k)$, $s_k = \mathcal{T}_{S_{\alpha_k \eta_k}} \alpha_k \eta_k$, and β_k is arbitrary number satisfying that $|\beta_k - 1| \leq L_\beta \|\alpha_k \eta_k\|$, $|\beta_k^{-1} - 1| \leq L_\beta \|\alpha_k \eta_k\|$ and $L_\beta > 0$ is a constant. The motivation for introducing β_k is to make this update subsume the update in [HGA15], which uses $\beta_k = \frac{\|\alpha_k \eta_k\|}{\|\mathcal{T}_{R_{\alpha_k \eta_k}} \alpha_k \eta_k\|}$. In Algorithm 1, β_k can be chosen as 1 for all k. Note that it is well-known that there exists an update formula for $\mathcal{B}_k^{-1}$, which is given later in (50).

If $y_k^\flat s_k > 0$, then the symmetric positive definiteness of $\tilde{\mathcal{B}}_k$ implies the symmetric positive definiteness of $\mathcal{B}_{k+1}$ [HGA15]. The positive definiteness of the sequence $\{\mathcal{B}_k\}$ is important in the sense that it guarantees that the search direction (5) is a descent direction. However, not all line search conditions imply $y_k^\flat s_k > 0$. In [RW12] and [HGA15], the Wolfe condition with information about the differential of the retraction R, $\mathcal{T}_R$, is used to guarantee $y_k^\flat s_k > 0$. In this paper, instead of enforcing $y_k^\flat s_k > 0$, we guarantee the symmetric positive definiteness of $\mathcal{B}_{k+1}$ by resorting to the following cautious update formula

$$\mathcal{B}_{k+1} = \begin{cases} \tilde{\mathcal{B}}_k - \frac{\tilde{\mathcal{B}}_k s_k (\tilde{\mathcal{B}}_k^* s_k)^\flat}{(\tilde{\mathcal{B}}_k^* s_k)^\flat s_k} + \frac{y_k y_k^\flat}{y_k^\flat s_k}, & \text{if } \frac{y_k^\flat s_k}{\|s_k\|^2} \geq \vartheta(\|\operatorname{grad} f(x_k)\|) \\ \tilde{\mathcal{B}}_k, & \text{otherwise,} \end{cases} \tag{7}$$

where ϑ is a monotone increasing function satisfying $\vartheta(0) = 0$ and ϑ is strictly increasing at 0. Formula (7) reduces to the cautious update formula of [LF01b] when $\mathcal{M}$ is a Euclidean space. Since $y_k^\flat s_k > 0$ is not longer enforced, we have more flexibility in the line search procedure, which only needs to satisfy Line 4 of Algorithm 1. It is pointed out that one also can reset $\mathcal{B}_{k+1}$ to be any given matrix, e.g., id, rather than $\tilde{\mathcal{B}}_k$ when $\frac{y_k^\flat s_k}{\|s_k\|^2} \ngeq \vartheta(\|\operatorname{grad} f(x_k)\|)$. Using either approach does not affect the theoretical results given later.

Note that it is important to ensure that $\vartheta(0) = 0$ and ϑ is strictly increasing at 0. If ϑ is not strictly increasing at 0, then the update (7) still guarantees the positive definiteness of the Hessian approximation sequence $\{\mathcal{B}_k\}$. However, the global convergence of Algorithm 1 for nonconvex problems may not hold (let $\vartheta \equiv 0$ and see the example in [Dai13]). If $\vartheta(0) > 0$, then the global convergence of Algorithm 1 holds for nonconvex problems but the local superlinear convergence does not hold generally (see the proofs in Theorem 5.1). If $\vartheta(0) < 0$, then Algorithm 1 may not be well defined in the sense that the search direction (5) may not be descent.

4 Global Convergence Analysis

The global convergence analysis of Algorithm 1 is generalized from the Euclidean analysis in [BN89, LF01b] and the differences for the Riemannian setting are highlighted.

The convergence analysis is built on the next two assumptions, which are not blanket assumptions and invoked only when needed. Assumptions 4.1 and 4.2 generalize the assumptions of Euclidean setting [LF01b, Assumption A].

Assumption 4.1. *The level set $\Omega = \{x \in \mathcal{M} \mid f(x) \leq f(x_0)\}$ is compact.*

Multiple Riemannian versions of Lipschitz continuous differentiable functions have been defined, e.g., [AMS08, Definitions 7.4.1 and 7.4.3]. Definition 4.1 gives another version which is slightly more general than [AMS08, Definition 7.4.3] in the sense that Definition 4.1 reduces to [AMS08, Definition 7.4.3] when $\mathcal{T}$ is chosen to be the parallel translation along the shortest geodesic. Assumption 4.2 assumes that the function f satisfies the Lipschitz continuous differentiability in Definition 4.1. Note that if f further satisfies [AMS08, Definitions 7.4.1] or equivalently Definition 3.1, then the line search conditions e.g., the Wolfe condition and the Armijo-Goldstein condition imply condition either (2) or (3).

Definition 4.1. *Let $\mathcal{T}$ be a vector transport associated with a retraction R. A function $\tilde{f}$ on $\mathcal{M}$ is said to be Lipschitz continuously differentiable with respect to $\mathcal{T}$ on $\mathcal{U} \subset \mathcal{M}$ if there exists $L_1 > 0$ such that*

$$\|\mathcal{T}_\eta \operatorname{grad} \tilde{f}(x) - \operatorname{grad} \tilde{f}(R_x(\eta))\| \leq L_1 \|\eta\|$$

for all $x \in \mathcal{U}$ and η such that $R_x(\eta) \in \mathcal{U}$.

Assumption 4.2. *The function f is Lipschitz continuously differentiable with respect to the isometric vector transport $\mathcal{T}_{\mathrm{S}}$ on Ω.*

Theorem 4.1 gives a sufficient condition for global convergence of Algorithm 1. This theorem is generalized from [LF01b, Theorem 3.1] but the proof is different since a different line search condition is used.

Theorem 4.1. *Let $\{x_k\}$ and η_k be sequences generated by Algorithm 1. If there are positive constants κ_1, κ_2 and κ_3 such that the inequalities*

$$\|\mathcal{B}_k \eta_k\| \leq \kappa_1 \|\eta_k\|, \quad \kappa_2 \|\eta_k\|^2 \leq \eta_k^\flat \mathcal{B}_k \eta_k \leq \kappa_3 \|\eta_k\|^2 \tag{8}$$

hold for infinitely many k's, then $\liminf_{k\to\infty} \|\operatorname{grad} f(x_k)\| = 0$.

Proof. Let $\mathcal{K}$ denote the index set such that (8) holds. Using (5) and (8) yields

$$\kappa_2 \|\eta_k\| \leq \|\operatorname{grad} f(x_k)\| \leq \kappa_1 \|\eta_k\|. \tag{9}$$

It follows from either (2) or (3) that

$$\begin{aligned}
\infty > f(x_0) - f(x_{k+1}) &= \sum_{i=0}^{k} (f(x_i) - f(x_{i+1})) = \sum_{i=0}^{k} (h_i(0) - h_i(\alpha_i)) \\
&\geq \sum_{i=0}^{k} \min(\chi_1 \frac{h_i'(0)^2}{\|\eta_i\|^2}, -\chi_2 h_i'(0)) = \sum_{i=0}^{k} \min(\chi_1 \frac{g(\operatorname{grad} f(x_i), \eta_i)^2}{\|\eta_i\|^2}, -\chi_2 g(\operatorname{grad} f(x_i), \eta_i)) \\
&= \sum_{i=0}^{k} \min(-\chi_1 \frac{g(\mathcal{B}_k \eta_k, \eta_k)}{\|\eta_i\|^2} g(\operatorname{grad} f(x_i), \eta_i), -\chi_2 g(\operatorname{grad} f(x_i), \eta_i)) \\
&\geq \min(\chi_1 \kappa_2, \chi_2) \sum_{i=0, i\in\mathcal{K}}^{k} (-g(\operatorname{grad} f(x_i), \eta_i)).
\end{aligned}$$

Therefore, it holds that

$$\liminf_{k\to\infty} -g(\operatorname{grad} f(x_k), \eta_k) = 0. \tag{10}$$

Combining (10) with (9), (8) yields

$$\|\operatorname{grad} f(x_k)\|^2 \leq \kappa_1^2 \|\eta_k\|^2 \leq \kappa_1^2 \kappa_2^{-1} g(\eta_k, \mathcal{B}_k \eta_k) = \kappa_1^2 \kappa_2^{-1} g(\eta_k, -\operatorname{grad} f(x_k)),$$

which implies $\liminf_{k\to\infty} \|\operatorname{grad} f(x_k)\| = 0$. □

Theorem 4.1 shows that the global convergence is ensured if there exist three positive constants $\kappa_1, \kappa_2, \kappa_3$ such that (8) holds for infinitely many k's. Lemma 4.1 proves that if (11) holds, then such three constants exist. Lemma 4.1 is generalized from [BN89, Theorem 2.1]. The main difference is that in the Euclidean setting, $\mathcal{B}_{k+1}$ is equal to $\mathcal{B}_k$ if the update is skipped. However, in the Riemannian setting, even the update is skipped, $\mathcal{B}_{k+1}$ is still different from $\mathcal{B}_k$ due to the existence of the vector transport $\mathcal{T}_{\mathrm{S}}$.

Lemma 4.1. *Let $\tilde{\mathcal{I}} = \{j_0, j_1, j_2, \ldots\}$ be an infinite index set, and $\{\mathcal{B}_k\}$ be the sequence generated by the Riemannian BFGS update*

$$\mathcal{B}_{k+1} = \begin{cases} \tilde{\mathcal{B}}_k - \frac{\tilde{\mathcal{B}}_k s_k (\tilde{\mathcal{B}}_k^* s_k)^\flat}{(\tilde{\mathcal{B}}_k^* s_k)^\flat s_k} + \frac{y_k y_k^\flat}{y_k^\flat s_k}, & \text{if } k \in \tilde{\mathcal{I}}; \\ \tilde{\mathcal{B}}_k, & \text{otherwise.} \end{cases}$$

Suppose $\mathcal{B}_0$ is symmetric and positive definite and there are positive constants $a_0 \leq a_1$ such that for all $k \geq 0$, y_{j_k} and s_{j_k} satisfy

$$\frac{y_{j_k}^\flat s_{j_k}}{\|s_{j_k}\|^2} \geq a_0, \frac{\|y_{j_k}\|^2}{y_{j_k}^\flat s_{j_k}} \leq a_1. \tag{11}$$

Then for any $p \in (0,1)$ there exist constants $\tilde{\kappa}_1, \tilde{\kappa}_2, \tilde{\kappa}_3 > 0$ such that, for any positive integer k,

$$\cos\theta_{j_i} \geq \tilde{\kappa}_1, \tag{12}$$

$$\tilde{\kappa}_2 \leq q_{j_i} \leq \tilde{\kappa}_3, \tag{13}$$

$$\tilde{\kappa}_2 \leq \frac{\|\mathcal{B}_{j_i}\eta_{j_i}\|}{\|\eta_{j_i}\|} \leq \frac{\tilde{\kappa}_3}{\tilde{\kappa}_1}, \tag{14}$$

hold for at least $\lceil (k+1)p \rceil$ values of $i \in \{0, 1, \dots, k\}$, where $\cos\theta_{j_i} = \frac{g(\eta_{j_i}, \mathcal{B}_{j_i}\eta_{j_i})}{\|\eta_{j_i}\|\|\mathcal{B}_{j_i}\eta_{j_i}\|}$, $q_{j_i} = \frac{g(\eta_{j_i}, \mathcal{B}_{j_i}\eta_{j_i})}{\|\eta_{j_i}\|^2}$ and $\lceil a \rceil$ denote the greatest integer that is smaller than a.

Proof. Since both $\mathrm{tr}(\hat{\mathcal{B}}_k)$ and $\det(\hat{\mathcal{B}}_k)$ are independent of the choice of basis, they are well-defined. Since

$$\mathcal{B}_{j_k} = \mathcal{T}_{\mathrm{S}_{\alpha_{j_k-1}\eta_{j_k-1}}} \circ \dots \circ \mathcal{T}_{\mathrm{S}_{\alpha_{j_{k-1}+1}\eta_{j_{k-1}+1}}} \circ \mathcal{B}_{j_{k-1}+1} \circ \mathcal{T}^{-1}_{\mathrm{S}_{\alpha_{j_{k-1}+1}\eta_{j_{k-1}+1}}} \circ \dots \circ \mathcal{T}^{-1}_{\mathrm{S}_{\alpha_{j_k-1}\eta_{j_k-1}}},$$

we have $\mathrm{tr}(\hat{\mathcal{B}}_{j_k}) = \mathrm{tr}(\hat{\mathcal{B}}_{j_{k-1}+1})$ and $\det(\hat{\mathcal{B}}_{j_k}) = \det(\hat{\mathcal{B}}_{j_{k-1}+1})$. Using the update formula of (6) yields that

$$\mathrm{tr}(\hat{\mathcal{B}}_{j_{k+1}}) = \mathrm{tr}(\hat{\mathcal{B}}_{j_k}) - \frac{\|\tilde{\mathcal{B}}_{j_k}s_{j_k}\|^2}{g(s_{j_k}, \tilde{\mathcal{B}}_{j_k}s_{j_k})} + \frac{\|y_{j_k}\|^2}{g(y_{j_k}, s_{j_k})}.$$

$$\det(\hat{\mathcal{B}}_{j_{k+1}}) = \det(\hat{\mathcal{B}}_{j_k})\frac{g(y_{j_k}, s_{j_k})}{g(s_{j_k}, \tilde{\mathcal{B}}_{j_k}s_{j_k})},$$

which have been used in [HGA15, pp. 5 and 13]. The Euclidean versions of these equations can be found in [NW06, (6.44) and (6.45)]. Define $\psi(\hat{\mathcal{B}}_{j_k}) = \mathrm{tr}(\hat{\mathcal{B}}_{j_k}) - \ln(\det(\hat{\mathcal{B}}_{j_k}))$ and it follows from the approach in [BN89, Section 2] or [NW06, Section 6.4] that

$$\begin{aligned}\psi(\hat{\mathcal{B}}_{j_{k+1}}) =& \psi(\hat{\mathcal{B}}_{j_k}) + \frac{\|y_{j_k}\|^2}{g(y_{j_k}, s_{j_k})} - 1 - \ln\frac{g(y_{j_k}, s_{j_k})}{\|s_{j_k}\|^2} + \ln\cos^2\theta_{j_k} \\ &+ 1 - \frac{q_{j_k}}{\cos^2\theta_{j_k}} + \ln\frac{q_{j_k}}{\cos^2\theta_{j_k}}.\end{aligned} \tag{15}$$

Using (11) and (15) yields

$$\begin{aligned}\psi(\hat{\mathcal{B}}_{j_{k+1}}) &\leq \psi(\hat{\mathcal{B}}_{j_k}) + a_1 - 1 - \ln a_0 + \ln\cos^2\theta_{j_k} + 1 - \frac{q_{j_k}}{\cos^2\theta_{j_k}} + \ln\frac{q_{j_k}}{\cos^2\theta_{j_k}} \\ &\leq \psi(\hat{\mathcal{B}}_0) + (a_1 - 1 - \ln a_0)(k+1) + \sum_{i=0}^{k}(\ln\cos^2\theta_{j_i} + 1 - \frac{q_{j_i}}{\cos^2\theta_{j_i}} + \ln\frac{q_{j_i}}{\cos^2\theta_{j_i}}).\end{aligned}$$

Define $\tau_{j_i} = -\ln\cos^2\theta_{j_i} - (1 - \frac{q_{j_i}}{\cos^2\theta_{j_i}} + \ln\frac{q_{j_i}}{\cos^2\theta_{j_i}})$. Note that the function $\ell(t) = 1 - t + \ln(t)$ is nonpositive for all t and $\ln\cos^2\theta_{j_i} \leq 0$ and therefore $\tau_{j_i} \geq 0$. Also note that $\psi(\hat{\mathcal{B}}_{j_{k+1}}) = \sum_{i=1}^{d}(\lambda_i - \ln\lambda_i) > 0$, where d is the dimension of $\mathcal{M}$ and λ_i is the eigenvalue of $\hat{\mathcal{B}}_{k+1}$. It follows that

$$\frac{1}{k+1}\sum_{i=0}^{k}\tau_{j_i} < \frac{\psi(\hat{\mathcal{B}}_0)}{k+1} + (a_1 - 1 - \ln a_0). \tag{16}$$

Define $\mathcal{J}_k$ to be the set of the $\lceil (k+1)p \rceil$ indices corresponding to $\lceil (k+1)p \rceil$ smallest values of τ_{j_i} for $i \leq k$. Let τ_{m_k} denote the largest of the τ_{j_i} for $i \in \mathcal{J}_k$. It follows that

$$\frac{1}{k+1}\sum_{i=0}^{k}\tau_{j_i} \geq \frac{1}{k+1}(\tau_{m_k} + \sum_{i=0, i\notin\mathcal{J}_k}^{k}\tau_{j_i}) \geq \tau_{m_k}(1-p). \tag{17}$$

Combining (16) with (17) yields

$$\tau_{j_i} < \frac{1}{1-p}(\psi(\hat{\mathcal{B}}_0) + a_1 - 1 - \ln a_0) := \omega$$

for all $i \in \mathcal{J}_k$. By the definition of τ_{j_i}, we have that

$$-\ln\cos^2\theta_{j_i} < \omega \tag{18}$$

$$1 - \frac{q_{j_i}}{\cos^2\theta_{j_i}} + \ln\frac{q_{j_i}}{\cos^2\theta_{j_i}} > -\omega \tag{19}$$

for all $i \in \mathcal{J}_k$. Inequality (18) implies $\cos\theta_{j_i} > \exp(-\omega/2) := \tilde{\kappa}_1$ which completes the proof of (12).

Using (19) and noting $\ell(t) \to -\infty$ both as $t \to 0$ and as $t \to \infty$, it follows that there exist positive constants $\tilde{\kappa}_3$ and b for all $i \in \mathcal{J}_k$ $0 < b \leq \frac{q_{j_i}}{\cos^2\theta_{j_i}} \leq \tilde{\kappa}_3$. Additionally using (12) yields $\tilde{\kappa}_2 := \tilde{\kappa}_1^2 b \leq q_{j_i} \leq \tilde{\kappa}_3$ for all $i \in \mathcal{J}_k$ which completes the proof of (13).

Finally, inequalities $\tilde{\kappa}_2 \leq \frac{\|\mathcal{B}_{j_i}\eta_{j_i}\|}{\|\eta_{j_i}\|} \leq \frac{\tilde{\kappa}_3}{\tilde{\kappa}_1}$ follow from $\|\mathcal{B}_{j_i}\eta_{j_i}\|/\|\eta_{j_i}\| = q_{j_i}/\cos\theta_{j_i}$. □

The desired global convergence result for a nonconvex function f is stated in Theorem 4.2.

Theorem 4.2. *Let $\{x_k\}$ be sequences generated by Algorithm 1. If Assumptions 4.1, and 4.2 hold, then*

$$\liminf_{k\to\infty} \|\operatorname{grad} f(x_k)\| = 0. \tag{20}$$

Proof. Define the index set from Algorithm 1

$$\mathcal{I} = \{k | \frac{y_k^\flat s_k}{\|s_k\|^2} \geq \vartheta(\|\operatorname{grad} f(x_k)\|)\}.$$

If $\mathcal{I}$ is finite, then there exists a $k_0 > 0$ such that $\mathcal{B}_k$ has the same eigenvalues as $\mathcal{B}_{k_0}$ for all $k \geq k_0$. Since $\mathcal{B}_{k_0}$ is symmetric positive definite, it is obvious that (8) holds. Therefore, Theorem 4.1 yields the desired result.

Contradiction is used to prove the result when $\mathcal{I}$ is infinite. Suppose (20) does not hold. Therefore, there exists a constant $\delta > 0$ such that $\|\operatorname{grad} f(x_k)\| \geq \delta$ for all k. It follows that from the definition of $\mathcal{I}$ that

$$\frac{y_k^\flat s_k}{\|s_k\|^2} \geq \vartheta(\delta) \tag{21}$$

holds for all $k \in \mathcal{I}$. Note Assumption 4.2, the definition of β_k, and the compactness of Ω, we have

$$\begin{aligned}
\|y_k\| =& \|\beta_k^{-1}\operatorname{grad} f(x_{k+1}) - \mathcal{T}_{S_{\alpha_k\eta_k}}\operatorname{grad} f(x_k)\| \\
\leq& \|\beta_k^{-1}\operatorname{grad} f(x_{k+1}) - \operatorname{grad} f(x_{k+1})\| + \|\operatorname{grad} f(x_{k+1}) - \mathcal{T}_{S_{\alpha_k\eta_k}}\operatorname{grad} f(x_k)\| \\
\leq& L_\beta\|s_k\|\|\operatorname{grad} f(x_{k+1})\| + L_1\|s_k\| \leq L_3\|s_k\|
\end{aligned}$$

where L_3 is a constant. It follows that $\frac{\|y_k\|^2}{y_k^\flat s_k} \leq \frac{L_3^2}{\vartheta(\delta)}$ Therefore, (20) follows from Lemma 4.1 and Theorem 4.1, a contradiction. □

5 Local Convergence Analysis

Section 5.1 lists the assumptions and definitions used in the local convergence analyses. Section 5.2 builds the connections between the global and local convergence analyses and shows that Algorithm 1 eventually reduces to the ordinary RBFGS method, i.e., Algorithm 1 without skipping updates. Since Algorithm 1 is equivalent to an ordinary RBFGS locally, the R-linear and superlinear convergence analyses are presented for the ordinary RBFGS in Sections 5.3 and 5.4 respectively.

5.1 Basic Assumptions and Definitions

Throughout the local convergence analyses, let x^* denote a nondegenerate local minimizer, i.e., Hess $f(x^*)$ is nonsingular. Three blanket assumptions are made in Assumption 5.1. Note that it follows that $f \circ R$ is a twice continuously differentiable function.

Assumption 5.1. *(i) The objective function f is twice continuously differentiable in the level set Ω; (ii) the retraction R is twice continuously differentiable; (iii) the isometric vector transport $\mathcal{T}_S$ with associated retraction R is continuous, satisfies $\mathcal{T}_{\mathrm{S}_{0_x}} \xi_x = \xi_x$ for all $\xi_x \in \mathrm{T}_x \mathcal{M}$. Additionally, given any $x \in \mathcal{M}$, there exists a neighborhood $\mathcal{U}$ of x such that $\mathcal{T}_\mathrm{S}$ satisfies $\|\mathcal{T}_{\mathrm{S}_\eta} - \mathcal{T}_{R_\eta}\| \leq \tilde{L}\|\eta\|$ and $\|\mathcal{T}_{\mathrm{S}_\eta}^{-1} - \mathcal{T}_{R_\eta}^{-1}\| \leq \tilde{L}\|\eta\|$, where $R_x(t\eta) \in \mathcal{U}$ for all $t \in [0,1]$ and $\tilde{L}$ is a positive constant.*

Since $\eta = 0$ implies $\|\mathcal{T}_{\mathrm{S}_\eta} - \mathcal{T}_{R_\eta}\| = 0$ and $\|\mathcal{T}_{\mathrm{S}_\eta}^{-1} - \mathcal{T}_{R_\eta}^{-1}\| = 0$, we have that $\mathcal{T}_S \in C^1$ implies $\|\mathcal{T}_{\mathrm{S}_\eta} - \mathcal{T}_{R_\eta}\| \leq \tilde{L}\|\eta\|$ and $\|\mathcal{T}_{\mathrm{S}_\eta}^{-1} - \mathcal{T}_{R_\eta}^{-1}\| \leq \tilde{L}\|\eta\|$. It follows that Assumption 5.1 (iii) is weaker than $\mathcal{T}_S \in C^1$. This assumption has been used in [HGA15] and therefore the analysis framework of [HGA15] can be applied here.

As in the Euclidean setting, we also have the result that a C^2 function on a compact set implies its gradient is Lipschitz continuous on the set, i.e., $f \in C^2$ implies the function f satisfies Assumptions 4.2 and Definition 3.1 (see [AMS08, Section 7.4] and [Hua13, Section 5.2.2] for details).

Definition 5.1, given in [HGA15, Definition 3.1], generalizes the convexity of a function on $\mathcal{S} \subseteq \mathcal{M}$ from the Euclidean setting to the Riemannian setting.

Definition 5.1. *For a function $f : \mathcal{M} \to \mathbb{R} : x \mapsto f(x)$ on a Riemannian manifold $\mathcal{M}$ with retraction R, define $m_{x,\eta}(t) = f(R_x(t\eta))$ for $x \in \mathcal{M}$ and $\eta \in \mathrm{T}_x \mathcal{M}$. The function f is retraction-convex with respect to the retraction R in a set $\mathcal{S}$ if for all $x \in \mathcal{S}$, all $\eta \in \mathrm{T}_x \mathcal{M}$ and $\|\eta\| = 1$, $m_{x,\eta}(t)$ is convex for all t which satisfy $R_x(\tau\eta) \in \mathcal{S}$ for all $\tau \in [0,t]$. Moreover, f is strongly retraction-convex in $\mathcal{S}$ if $m_{x,\eta}(t)$ is strongly convex, i.e., there exist two constants $0 < a_7 < a_8$ such that $a_7 \leq \frac{d^2 m_{x,\eta}}{dt^2}(t) \leq a_8$, for all $x \in \mathcal{S}$, all $\|\eta\| = 1$ and all t such that $R_x(\tau\eta) \in \mathcal{S}$ for all $\tau \in [0,t]$.*

It has been shown in [HGA15, Lemma 3.1] that such a neighborhood, in which the function f is strongly retraction-convex, always exists around a nondegenerate minimizer. In addition, for any neighborhood $\mathcal{W}$ of $x \in \mathcal{M}$, it can be shrunk such that it is an R-star shaped neighborhood of x, i.e., $R_x(tR_x^{-1}(z)) \in \mathcal{W}$ for all $z \in \mathcal{W}$ and $t \in [0,1]$. Therefore, there exists a neighborhood $\tilde{\mathcal{W}}$ of x^* satisfying

A.1 $\tilde{\mathcal{W}}$ is a R-star shape of x^*;

A.2 The object function f is strongly retraction-convex in $\tilde{\mathcal{W}}$;

A.3 For any $x, \tilde{x} \in \tilde{\mathcal{W}}$, inequalities (23) hold. (This can be seen from Lemma 5.1 given later)

Assumption 5.2 is used in the later proofs with $\tilde{\Omega}$ satisfying some or all of *A.1–A.3*.

Assumption 5.2. *There exists $K > 0$ such that the iterates x_k stay continuously in a neighborhood $\tilde{\Omega}$ of x^* for all $k \geq K$, meaning that $R_{x_k}(t\eta_k) \in \tilde{\Omega} \subseteq \Omega$ for all $t \in [0, \alpha_k]$.*

Note that the iterates x_k must stay in $\tilde{\Omega}$ **continuously**. The "continuously" assumption cannot be removed. To see this, consider the unit sphere with the exponential retraction, where we can have $x_{k+1} \approx x_k$ with $\alpha_k \eta_k \approx 2\pi$. (A similar comment was made in [HAG15] before Assumption 6 and in [HGA15] before Assumption 3.3.)

Assumption 5.3 is used to guarantee that $R_x^{-1}(y)$ is well-defined and has been used in existing papers, e.g., [HAG15, HGA15]. This assumption is also a blanket assumption for the local convergence analysis.

Assumption 5.3. *There exists $r > 0$ such that for each $x \in \tilde{\Omega}$, $R_x(\mathbb{B}(0_x, r)) \supset \tilde{\Omega}$ and $R_x(\cdot)$ is a diffeomorphism on $\mathbb{B}(0_x, r)$, where $\tilde{\Omega}$ is defined in Assumption 5.2.*

Assumptions 5.2 and 5.3 are used in Section 5.3 to show R-linear convergence. Assumptions 5.4 and 5.5 are used in Section 5.4 to prove superlinear convergence.

When a retraction is considered, a generalization of the Euclidean triangle inequality in $\tilde{\Omega}$ must be assumed in the proofs below. As shown in [Hua13, Lemma 6.2.1], choosing the exponential mapping for the retraction R implies Assumption 5.4.

Assumption 5.4. *There is a constant a_9 such that for all $x, y \in \tilde{\Omega}$,*

$$\max_{t\in[0,1]} \mathrm{dist}(R_x(t\eta), x^*) \leq a_9 \max(\mathrm{dist}(x, x^*), \mathrm{dist}(y, x^*)),$$

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

Assumption 5.5 generalizes the Euclidean property of twice Hölder continuously differentiability of f at x^* to a Riemannian manifold.

Assumption 5.5. *There exist positive constants a_{10} and a_{11} such that for all $y \in \tilde{\Omega}$,*

$$\| \mathrm{Hess}\, f(y) - \mathcal{T}_{\mathrm{S}_\eta} \mathrm{Hess}\, f(x^*) \mathcal{T}_{S_\eta}^{-1} \| \leq a_{10} \|\eta\|^{a_{11}},$$

where $\eta = R_{x^*}^{-1} y$.

It can be shown that Assumptions 5.3, 5.4, and 5.5 hold if $\tilde{\Omega}$ is sufficiently small. It follows that if $f, R \in C^2$, $\mathcal{T} \in C^1$, and the series $\{x_k\}$ converges to x^*, then all the assumptions in this section hold for $\{x_k\}_{k=K}^{\infty}$ with K sufficiently large.

5.2 Preliminaries

Before giving the local convergence analysis, we first state an important property that holds in the Euclidean setting but may not hold in the Riemannian setting. The property is used in Theorem 5.1 and Section 5.1.

In the Euclidean setting, suppose f is strongly convex on $\mathcal{S} \subset \mathbb{R}^d$. It is well-known that $y_k = \bar{G}_k s_k$, where $y_k = \mathrm{grad}\, f(x_{k+1}) - \mathrm{grad}\, f(x_k)$, $s_k = x_{k+1} - x_k$ and $\bar{G}_k = \int_0^1 \mathrm{Hess}\, f(x_k + \tau s_k) d\tau$ is the average Hessian. It follows that

$$a_4 \|s_k\|^2 \leq y_k^T s_k \leq a_5 \|s_k\|^2 \text{ and } \|y_k\|^2 \leq a_6 y_k^T s_k, \tag{22}$$

where a_4, a_5 and a_6 are positive constants. Derivations can be found in, e.g., [BN89, NW06].

The RBFGS generalized by Ring and Wirth [RW12] satisfies (22) when the cost function is uniformly convex, see, [RW12, Proposition 10]. The Riemannian Broyden family of methods generalized by Huang et al. [HGA15] also satisfies (22) when the cost function is retraction-convex. However, the RBFGS in this paper does not generally imply (22) for either uniformly convex or retraction-convex cost functions. This is the main difference and probably the main difficulty compared to the existing work in [LF01b, BN89, RW12, HGA15]. Nevertheless, as shown in Lemma 5.1, we know (22) holds when iterates $\{x_k\}$ are in a neighborhood of the nondegenerate minimizer x^*.

Lemma 5.1. *Suppose Assumptions 5.1, 5.2 and 5.3 hold. Let x^* be a nondegenerate minimizer of f. Then there exists a neighborhood $\mathcal{V}$ of x^* and positive constants $\tilde{a}_4$, $\tilde{a}_5$ and $\tilde{a}_6$ such that for all $x, \tilde{x} \in \mathcal{V}$ satisfying $R_x(tR_x^{-1}(\tilde{x})) \in \mathcal{V}$ for $t \in [0,1]$, it holds that*

$$\tilde{a}_4 \|s\|^2 \leq g(y, s) \leq \tilde{a}_5 \|s\|^2 \text{ and } \|y\|^2 \leq \tilde{a}_6 g(y, s), \tag{23}$$

where $y = \frac{1}{\beta(\xi)} \mathrm{grad}\, f(\tilde{x}) - \mathcal{T}_{\mathrm{S}_\xi} \mathrm{grad}\, f(x)$, $\xi = R_x^{-1}(\tilde{x})$, $s = \mathcal{T}_{\mathrm{S}_\xi} \xi$, $\beta : \mathrm{T}\mathcal{M} \to \mathbb{R}$ satisfies $\beta(\xi) = 1 + O(\|\xi\|)$, and $O(t)$ means that $\lim_{t\to 0} O(t)/t$ is bounded.

Proof. Define $y^{\mathcal{L}}$ and $s^{\mathcal{L}}$ to be $\frac{\|\mathcal{T}_{R_\xi} \xi\|}{\|\xi\|} \mathrm{grad}\, f(\tilde{x}) - \mathcal{T}_{\mathrm{S}_\xi}^{\mathcal{L}} \mathrm{grad}\, f(x)$ and $\mathcal{T}_{\mathrm{S}_\xi}^{\mathcal{L}} \xi$ respectively, where $\mathcal{T}_{\mathrm{S}}^{\mathcal{L}}$ denotes an isometric vector transport satisfying the locking condition, i.e., $\mathcal{T}_{\mathrm{S}_\eta}^{\mathcal{L}} \eta = \frac{\|\eta\|}{\|\mathcal{T}_{R_\eta} \eta\|} \mathcal{T}_{R_\eta} \eta$, for all $\eta \in \mathrm{T}\mathcal{M}$. The existence of $\mathcal{T}_S^{\mathcal{L}}$ (at least locally) can be seen from [HGA15, Section 4]. It follows from [HGA15, Lemmas 3.3 and 3.9] that there exist positive constants b_0, b_1 and b_2 such that

$$b_0\|s^{\mathcal{L}}\|^2 \leq g(y^{\mathcal{L}}, s^{\mathcal{L}}) \leq b_1\|s^{\mathcal{L}}\|^2 \text{ and } \|y^{\mathcal{L}}\|^2 \leq b_2 g(y^{\mathcal{L}}, s^{\mathcal{L}}).$$

Using $\beta(\xi) = 1 + O(\|\xi\|)$, [HGA15, Lemmas 3.5, 3.6] and [GQA12, Lemma 14.5] yields

$$|g(s,y) - g(s^{\mathcal{L}}, y^{\mathcal{L}})| = |g\left(\left(\frac{1}{\beta(\xi)}\mathcal{T}_{S_\xi} - \frac{\|\mathcal{T}_{R_\xi}\xi\|}{\|\xi\|}\mathcal{T}_{S_\xi}^{\mathcal{L}}\right)\xi, \operatorname{grad} f(\tilde{x})\right)| = O(\|s\|^2\epsilon),$$

where $\epsilon = \max(\operatorname{dist}(\tilde{x}, x^*), \operatorname{dist}(x, x^*))$. It follows that

$$b_0\|s^{\mathcal{L}}\|^2 \leq g(y,s) + O(\|s\|^2\epsilon) \leq b_1\|s^{\mathcal{L}}\|^2 \text{ and } \|y^{\mathcal{L}}\|^2 \leq b_2 g(y,s) + O(\|s\|^2\epsilon).$$

Therefore, by choosing sufficiently small neighborhood $\mathcal{V}$ such that ϵ is small enough, we have that there exist constants $\tilde{a}_4$, $\tilde{a}_5$ and $\tilde{a}_6$ such that (23) holds. □

We can now prove Theorem 5.1, which states that if the iterates $\{x_k\}$ stay in a sufficient small neighborhood of a nondegenerate minimizer x^*, i.e., $\operatorname{Hess} f(x^*)$ is positive definite, then x_k converges to x^* and Algorithm 1 reduces to an ordinary RBFGS. This implies that the local convergence analysis of an ordinary RBFGS is equivalent to the local convergence analysis of Algorithm 1. This theorem is generalized from [LF01b, Theorem 3.5].

Theorem 5.1. *Under the assumptions of Lemma 5.1, if $s_k \to 0$ and x^* is an accumulation point of $\{x_k\}$ generated by Algorithm 1, then the sequence $\{x_k\}$ converges to x^* and Algorithm 1 reduces to the ordinary RBFGS when x_k is sufficiently close to x^*.*

Proof. The assumptions about x^* imply that x^* is an isolated minimizer of f. Since x^* is an accumulation point and $s_k \to 0$, we have $x_k \to x^*$. It follows that

$$\lim_{k\to\infty} \|\operatorname{grad} f(x_k)\| = 0. \tag{24}$$

By (23) of Lemma 5.1, there exists constant $b_2 > 0$ and an integer $K > 0$ such that $\frac{y_k^\flat s_k}{\|s_k\|^2} > b_2$ for all $k > K$. Therefore, by (24), the cautious update formula (7) reduces to the ordinary update formula (6) when $\vartheta(\|\operatorname{grad} f(x_k)\|) < b_2$ and $k > K$. □

5.3 R-Linear Convergence Analysis

Theorem 5.2 gives sufficient conditions that guarantee R-linear convergence for **all** quasi-Newton methods with a line search condition satisfying either (2) or (3). It is generalized from [BN89, Theorem 3.1].

Theorem 5.2. *Suppose Assumption 5.1 holds and Assumption 5.2 holds with $\tilde{\Omega}$ satisfy A.1 and A.2; $\{x_k\}$ is generated by (1) and (5) such that the step size α_k satisfies either (2) or (3); and $\mathcal{B}_k$ is positive definite for all $k > K$. Assume that there exists $p \in (0,1)$ and $a_{12}, a_{13} > 0$, such that for any $k \geq K$, the inequalities*

$$\cos\theta_j \geq a_{12}, \tag{25}$$

$$\frac{\|\mathcal{B}_j s_j\|}{\|s_j\|} \leq a_{13} \tag{26}$$

hold for at least $\lceil (k - K + 1)p \rceil$ values of $j \in [K, k]$, where $\cos\theta_j = \frac{g(\eta_j, \mathcal{B}_j \eta_j)}{\|\eta_j\|\|\mathcal{B}_j\eta_j\|}$. Then $\{x_k\} \to x^*$; moreover

$$\sum_{k=K}^{\infty} \|\zeta_k\| < \infty, \tag{27}$$

and there is a constant $0 \leq a_{14} < 1$ such that

$$f(x_{k+1}) - f(x^*) \leq a_{14}^{k-K+1}(f(x_K) - f(x^*)) \tag{28}$$

holds for all $k \geq K$, where $\zeta_k = R_{x^*}^{-1}(x_k)$.

Proof. Let $\mathcal{J}$ denote the set of indices for which (25) and (26) hold. Consider an iterate x_j with $j \in \mathcal{J}$. Using (2), (3), (25), and (26) yields that

$$h_j(0) - h_j(\alpha_j) \geq \chi \| \operatorname{grad} f(x_j) \|^2, \tag{29}$$

where $\chi = \chi_1 a_{12}^2$ if the line search condition (2) holds, or $\chi = \chi_2 a_{12}/a_{13}$ if the condition (3) holds. Define $m_{x^*,k}(t)$ to be $f(R_{x^*}(t\zeta_k/\|\zeta_k\|))$. Taylor's Theorem gives

$$f(x_k) - f(x^*) = m_{x^*,k}(\|\zeta_k\|) - m_{x^*,k}(0) = \frac{1}{2}\frac{d^2}{dt^2} m_{x^*,k}(t)|_{t=\tau} \|\zeta_k\|^2,$$

where $\tau \in [0, \|\zeta_k\|]$. It follows from Assumption 5.2 and Definition 5.1 that $f(x_k) - f(x^*) \geq \frac{1}{2} a_7 \|\zeta_k\|^2$. Using the result from [HGA15, (3.9) in Lemma 3.7] gives $f(x_k) - f(x^*) \leq b_0 \| \operatorname{grad} f(x_k)\|^2$ for some constant b_0. Therefore, we have for all $k \geq K$,

$$\frac{1}{2} a_7 \|\zeta_k\|^2 \leq f(x_k) - f(x^*) \leq b_0 \| \operatorname{grad} f(x_k) \|^2. \tag{30}$$

By (29) and (30), we obtain that for all $j \in \mathcal{J}$, $f(x_j) - f(x_{j+1}) \geq \chi(f(x_j) - f(x^*))/b_0$, which yields

$$f(x_{j+1}) - f(x^*) \leq a_{14}^{1/p}(f(x_j) - f(x^*)),$$

where $a_{14}^{1/p} = 1 - \chi b_0^{-1}$. Since $\mathcal{J} \cap [K, k]$ has at least $\lceil (k - K + 1)p \rceil$ elements, and since $\{f_k\}$ is decreasing, it follows that

$$f(x_{k+1}) - f(x^*) \leq a_{14}^{k-K+1}(f(x_K) - f(x^*)).$$

By the lower bound of (30), we obtain

$$\sum_{k=K}^{\infty} \|\zeta_k\| \leq \sqrt{\frac{2}{a_7}} \sum_{k=K}^{\infty} \sqrt{f(x_k) - f(x^*)} \leq \sqrt{\frac{2(f(x_K) - f(x^*))}{a_7}} \sum_{k=K}^{\infty} a_{14}^{k/2} < \infty.$$

□

The R-linear convergence rate of Algorithm 1 is seen easily and is given in Corollary 5.1.

Corollary 5.1. *Suppose Assumption 5.1 holds; Assumption 5.2 holds with $\tilde{\Omega}$ satisfy A.1–A.2; and Assumption 5.3 holds. Let $\{x_k\}$ be the sequence generated by Algorithm 1. Then the iterates x_k converge to x^* and* (27) *and* (28) *hold.*

Proof. The results of Lemma 4.1 follows from Lemma 5.1 and the results of this corollary follows from Lemma 4.1 and Theorem 5.2. □

5.4 Superlinear Convergence Analysis

In this section, $\{x_k\}$, $\{\mathcal{B}_k\}$, $\{\tilde{\mathcal{B}}_k\}$, $\{\alpha_k\}$, $\{s_k\}$, $\{y_k\}$, and $\{\eta_k\}$, are infinite sequences generated by Algorithm 1. The following notation is used:

$$\begin{aligned}
&\epsilon_k = \max(\operatorname{dist}(x_{k+1}, x^*), \operatorname{dist}(x_k, x^*)), \;\; H_* = \operatorname{Hess} f(x^*), \;\; \zeta_k = R_{x^*}^{-1} x_k, \\
&H_k = \mathcal{T}_{S_{\zeta_k}} H_* \mathcal{T}_{S_{\zeta_k}}^{-1}, \quad \bar{s}_k = H_{k+1}^{1/2} s_k, \quad \bar{y}_k = H_{k+1}^{-1/2} y_k, \;\; \bar{\mathcal{B}}_k = H_k^{-1/2} \mathcal{B}_k H_k^{-1/2}, \\
&\mathcal{C}_k = H_{k+1}^{-1/2} \tilde{\mathcal{B}}_k H_{k+1}^{-1/2}, \;\; \cos \bar{\theta}_k = \frac{g(\bar{s}_k, \mathcal{C}_k \bar{s}_k)}{\|\bar{s}_k\| \|\mathcal{C}_k \bar{s}_k\|}, \;\; \bar{q}_k = \frac{g(\bar{s}_k, \mathcal{C}_k \bar{s}_k)}{\|\bar{s}_k\|^2}
\end{aligned}$$

where $H_k^{1/2} = \mathcal{T}_{S_{\zeta_k}} H_*^{1/2} \mathcal{T}_{S_{\zeta_k}}^{-1}$ denotes a linear operator on $\mathrm{T}_{x_k} \mathcal{M}$, $H_*^{1/2}$ satisfies $H_*^{1/2} H_*^{1/2} = H_*$, and $H_*^{1/2}$ is self-adjoint.

Lemma 5.2 generalizes [BN89, (3.23), (3.25)]. It is used in Theorem 5.3.

Lemma 5.2. *If Assumptions 5.1, 5.2, 5.3, 5.4 and 5.5 hold, then there exists a neighborhood* $\mathcal{U}$ *of* x^* *such that for* $x_k, x_{k+1} \in \mathcal{U}$, *it holds that*

$$\|\bar{y}_k - \bar{s}_k\| \leq a_{15}\epsilon_k^{\min(1,a_{11})}\|\bar{s}_k\| \quad \text{and} \quad g(\bar{y}_k, \bar{s}_k) \geq (1 - a_{16}\epsilon_k^{\min(1,a_{11})})\|\bar{s}_k\|^2,$$

where a_{15} *and* a_{16} *are positive constants.*

Proof. Define $y_k^{\mathrm{P}} = \operatorname{grad} f(x_{k+1}) - P_{\gamma_k}^{1\leftarrow 0} \operatorname{grad} f(x_k)$, where P is parallel transport and γ_k is the retraction line from x_k to x_{k+1}, i.e., $\gamma_k(t) = R_{x_k}(tR_{x_k}^{-1}(x_{k+1}))$. From [HAG15, Lemma 8], we have $\|P_{\gamma_k}^{0\leftarrow 1} y_k^{\mathrm{P}} - \bar{H}_k \alpha_k \eta_k\| \leq b_0\|\alpha_k\eta_k\|^2 = b_0\|s_k\|^2$, where $\bar{H}_k = \int_0^1 P_{\gamma_k}^{0\leftarrow t} \operatorname{Hess} f(\gamma_k(t)) P_{\gamma_k}^{t\leftarrow 0} dt$ and b_0 is a positive constant. It follows that

$$\begin{aligned}
\|y_k - H_{k+1}s_k\| &\leq \|y_k - y_k^{\mathrm{P}}\| \\
&\quad + \|P_{\gamma_k}^{0\leftarrow 1} y_k^{\mathrm{P}} - \bar{H}_k\alpha_k\eta_k\| + \|P_{\gamma_k}^{1\leftarrow 0}\bar{H}_k P_{\gamma_k}^{0\leftarrow 1} P_{\gamma_k}^{1\leftarrow 0}\alpha_k\eta_k - H_{k+1}\mathcal{T}_{S_{\alpha_k\eta_k}}\alpha_k\eta_k\| \\
&\leq \|\operatorname{grad} f(x_{k+1})/\beta_k - \operatorname{grad} f(x_{k+1})\| + \|P_{\gamma_k}^{1\leftarrow 0}\operatorname{grad} f(x_k) - \mathcal{T}_{S_{\alpha_k\eta_k}} \operatorname{grad} f(x_k)\| \\
&\quad + b_0\|s_k\|^2 + \|P_{\gamma_k}^{1\leftarrow 0}\bar{H}_k P_{\gamma_k}^{0\leftarrow 1} P_{\gamma_k}^{1\leftarrow 0}\alpha_k\eta_k - P_{\gamma_k}^{1\leftarrow 0}\bar{H}_k P_{\gamma_k}^{0\leftarrow 1}\mathcal{T}_{S_{\alpha_k\eta_k}}\alpha_k\eta_k\| \\
&\quad + \|P_{\gamma_k}^{1\leftarrow 0}\bar{H}_k P_{\gamma_k}^{0\leftarrow 1}\mathcal{T}_{S_{\alpha_k\eta_k}}\alpha_k\eta_k - H_{k+1}\mathcal{T}_{S_{\alpha_k\eta_k}}\alpha_k\eta_k\|
\end{aligned}$$

Simplifying the right hand side yields

$$\begin{aligned}
\|y_k - H_{k+1}s_k\| &\leq \|\operatorname{grad} f(x_{k+1})\||1/\beta_k - 1| \text{ (using [GQA12, Lemma 14.5])} \\
&\quad + \|P_{\gamma_k}^{1\leftarrow 0}\operatorname{grad} f(x_k) - \mathcal{T}_{S_{\alpha_k\eta_k}}\operatorname{grad} f(x_k)\| \\
&\quad \text{(using [GQA12, Lemma 14.5] and [HGA15, Lemma 3.6])} \\
&\quad + b_0\|s_k\|^2 + \|\bar{H}_k\|\|P_{\gamma_k}^{1\leftarrow 0}\alpha_k\eta_k - \mathcal{T}_{S_{\alpha_k\eta_k}}\alpha_k\eta_k\| \text{ (using [HGA15, Lemma 3.6])} \\
&\quad + \|P_{\gamma_k}^{1\leftarrow 0}\bar{H}_k P_{\gamma_k}^{0\leftarrow 1} - H_{k+1}\|\|s_k\| \\
&\qquad\qquad \text{(using Assumptions 5.4, 5.5 and [Hua13, Lemma 6.2.5])} \\
&\leq b_1\epsilon_k\|s_k\| + b_2\epsilon_k\|s_k\| + b_3\epsilon_k\|s_k\| + b_4\epsilon_k^{\min(1,a_{11})}\|s_k\| \\
&= b_5\epsilon_k^{\min(1,a_{11})}\|s_k\|,
\end{aligned}$$

where b_1, b_2, b_3, b_4 and b_5 are positive constants. Therefore, we have

$$\|\bar{y}_k - \bar{s}_k\| \leq b_6\epsilon_k^{\min(1,a_{11})}\|\bar{s}_k\|, \tag{31}$$

where b_6 is a positive constant. It follows that $\|\bar{y}_k\| - \|\bar{s}_k\| \leq b_6\epsilon_k^{\min(1,a_{11})}\|\bar{s}_k\|$ and $\|\bar{s}_k\| - \|\bar{y}_k\| \leq b_6\epsilon_k^{\min(1,a_{11})}\|\bar{s}_k\|$, which yields

$$(1 - b_6\epsilon_k^{\min(1,a_{11})})\|\bar{s}_k\| \leq \|\bar{y}_k\| \leq (1 + b_6\epsilon_k^{\min(1,a_{11})})\|\bar{s}_k\|. \tag{32}$$

By squaring (31) and using (32), we have

$$\begin{aligned}
(1 - b_6\epsilon_k^{\min(1,a_{11})})^2\|\bar{s}_k\|^2 - 2g(\bar{y}_k, \bar{s}_k) + \|\bar{s}_k\|^2 &\leq \|\bar{y}_k\|^2 - 2g(\bar{y}_k, \bar{s}_k) + \|\bar{s}_k\|^2 \\
&\leq (b_6\epsilon_k^{\min(1,a_{11})})^2\|\bar{s}_k\|^2,
\end{aligned}$$

and therefore $g(\bar{y}_k, \bar{s}_k) \geq (1 - b_6\epsilon_k^{\min(1,a_{11})})\|\bar{s}_k\|^2$. □

Theorem 5.3 generalized from [BN89, Theorem 3.2] is the main result of this section. When Theorem 5.3 is combined with a Riemannian version of the Dennis-Moré condition, superlinear convergence follows, as shown in Corollary 5.2.

Theorem 5.3. *Suppose Assumption 5.1 holds; Assumption 5.2 holds with $\tilde{\Omega}$ satisfy A.1–A.3; Assumption 5.5 holds with $a_{11} = 1$; and Assumptions 5.3 and 5.4. Then*

$$\lim_{k\to\infty} \frac{\| \operatorname{grad} f(x_k) + \operatorname{Hess} \hat{f}_{x_k}(0_{x_k})\eta_k\|}{\|\eta_k\|} = 0. \tag{33}$$

Proof. By pre- and post- multiplying the update formula (6) by $H_{k+1}^{-1/2}$, we have

$$\bar{\mathcal{B}}_{k+1} = \mathcal{C}_k - \frac{\mathcal{C}_k \bar{s}_k (\mathcal{C}_k^* \bar{s}_k)^\flat}{(\mathcal{C}_k^* \bar{s}_k)^\flat \bar{s}_k} + \frac{\bar{y}_k \bar{y}_k^\flat}{\bar{y}_k^\flat \bar{s}_k}, \tag{34}$$

It follows from the idea used in deriving (15) that

$$\operatorname{tr}(\hat{\bar{\mathcal{B}}}_{k+1}) = \operatorname{tr}(\hat{\mathcal{C}}_k) + \frac{\|\bar{y}_k\|^2}{g(\bar{y}_k, \bar{s}_k)} - \frac{\bar{q}_k}{\cos^2 \bar{\theta}_k} \tag{35}$$

$$\ln(\det(\hat{\bar{\mathcal{B}}}_{k+1})) = \ln(\det(\hat{\mathcal{C}}_k)) + \ln \frac{g(\bar{y}_k, \bar{s}_k)}{\|\bar{s}_k\|^2} - \ln \cos^2 \bar{\theta}_k - \ln \frac{\bar{q}_k}{\cos^2 \bar{\theta}_k}. \tag{36}$$

Using the property of a determinant, $\det(M_1 M_2) = \det(M_1)\det(M_2)$, yields

$$\det(\hat{\mathcal{C}}_k) = \det(\hat{H}_{k+1}^{-1/2} \hat{\bar{\mathcal{B}}}_k \hat{H}_{k+1}^{-1/2}) = \det(\hat{H}_*^{-1/2}) \det(\hat{\mathcal{B}}_k) \det(\hat{H}_*^{-1/2}) = \det(\hat{\bar{\mathcal{B}}}_k). \tag{37}$$

It follows that

$$-\ln(\det(\hat{\mathcal{C}}_{k+1})) = -\ln(\det(\hat{\mathcal{C}}_0)) + \sum_{i=0}^{k} (-\ln \frac{g(\bar{y}_i, \bar{s}_i)}{\|\bar{s}_i\|^2} + \ln \cos^2 \bar{\theta}_i + \ln \frac{\bar{q}_i}{\cos^2 \bar{\theta}_i}). \tag{38}$$

Observing the relationship between $\operatorname{tr}(\hat{\mathcal{C}}_k)$ and $\operatorname{tr}(\hat{\bar{\mathcal{B}}}_k)$, we have

$$\begin{aligned}
\operatorname{tr}(\hat{\mathcal{C}}_k) - \operatorname{tr}(\hat{\bar{\mathcal{B}}}_k) &= \operatorname{tr}(\hat{\mathcal{T}}_{S_{\zeta_{k+1}}}^{-1} \hat{\mathcal{T}}_{S_{\alpha_k \eta_k}} \hat{\mathcal{B}}_k \hat{\mathcal{T}}_{S_{\alpha_k \eta_k}}^{-1} \hat{\mathcal{T}}_{S_{\zeta_{k+1}}} \hat{H}_*^{-1}) - \operatorname{tr}(\hat{\mathcal{T}}_{S_{\zeta_k}}^{-1} \hat{\mathcal{B}}_k \hat{\mathcal{T}}_{S_{\alpha_k \eta_k}}^{-1} \hat{\mathcal{T}}_{S_{\zeta_{k+1}}} \hat{H}_*^{-1}) \\
&\quad + \operatorname{tr}(\hat{\mathcal{T}}_{S_{\zeta_k}}^{-1} \hat{\mathcal{B}}_k \hat{\mathcal{T}}_{S_{\alpha_k \eta_k}}^{-1} \hat{\mathcal{T}}_{S_{\zeta_{k+1}}} \hat{H}_*^{-1}) - \operatorname{tr}(\hat{\mathcal{T}}_{S_{\zeta_k}}^{-1} \hat{\mathcal{B}}_k \hat{\mathcal{T}}_{S_{\zeta_k}} \hat{H}_*^{-1}) \\
&\le \|\hat{\mathcal{T}}_{S_{\zeta_{k+1}}}^{-1} \hat{\mathcal{T}}_{S_{\alpha_k \eta_k}} - \hat{\mathcal{T}}_{S_{\zeta_k}}^{-1}\|_F \|\hat{\mathcal{B}}_k \hat{\mathcal{T}}_{S_{\alpha_k \eta_k}}^{-1} \hat{\mathcal{T}}_{S_{\zeta_{k+1}}} \hat{H}_*^{-1}\|_F \\
&\quad + \|\hat{\mathcal{T}}_{S_{\alpha_k \eta_k}}^{-1} \hat{\mathcal{T}}_{S_{\zeta_{k+1}}} - \hat{\mathcal{T}}_{S_{\zeta_k}}\|_F \|\hat{H}_*^{-1} \hat{\mathcal{T}}_{S_{\zeta_k}}^{-1} \hat{\mathcal{B}}_k\|_F \\
&\le b_0 (\|\hat{\mathcal{T}}_{S_{\zeta_{k+1}}}^{-1} \hat{\mathcal{T}}_{S_{\alpha_k \eta_k}} - \hat{\mathcal{T}}_{S_{\zeta_k}}^{-1}\| \|\hat{\mathcal{B}}_k \hat{\mathcal{T}}_{S_{\alpha_k \eta_k}}^{-1} \hat{\mathcal{T}}_{S_{\zeta_{k+1}}} \hat{H}_*^{-1}\| \\
&\quad + \|\hat{\mathcal{T}}_{S_{\alpha_k \eta_k}}^{-1} \hat{\mathcal{T}}_{S_{\zeta_{k+1}}} - \hat{\mathcal{T}}_{S_{\zeta_k}}\| \|\hat{H}_*^{-1} \hat{\mathcal{T}}_{S_{\zeta_k}}^{-1} \hat{\mathcal{B}}_k\|) \text{ (by [Hua13, Lemma 6.2.6])} \\
&\le b_0 (\|\hat{\mathcal{T}}_{S_{\zeta_{k+1}}}^{-1} \hat{\mathcal{T}}_{S_{\alpha_k \eta_k}} \hat{\mathcal{T}}_{S_{\zeta_k}} - I\| \|\hat{\mathcal{B}}_k\| \|\hat{H}_*^{-1}\| + \|\hat{\mathcal{T}}_{S_{\zeta_k}}^{-1} \hat{\mathcal{T}}_{S_{\alpha_k \eta_k}}^{-1} \hat{\mathcal{T}}_{S_{\zeta_{k+1}}} - I\| \|\hat{H}_*^{-1}\| \|\hat{\mathcal{B}}_k\|) \\
&\le b_1 \varsigma_k \|\hat{\bar{\mathcal{B}}}_k\|_F \le b_2 \varsigma_k \operatorname{tr}(\hat{\bar{\mathcal{B}}}_k), \text{ (by [Hua13, Lemma 6.2.6])}
\end{aligned} \tag{39}$$

where b_0, b_1 and b_2 are positive constants, $\varsigma_k = \|\hat{\mathcal{T}}_{S_{\zeta_k}}^{-1} \hat{\mathcal{T}}_{S_{\alpha_k \eta_k}}^{-1} \hat{\mathcal{T}}_{S_{\zeta_{k+1}}} - I\|$. It follows that

$$\begin{aligned}
\operatorname{tr}(\hat{\mathcal{C}}_{k+1}) &\le (1 + b_2 \varsigma_{k+1}) \operatorname{tr}(\hat{\bar{\mathcal{B}}}_{k+1}) = (1 + b_2 \varsigma_{k+1}) \left(\operatorname{tr}(\hat{\mathcal{C}}_k) + \frac{\|\bar{y}_k\|^2}{g(\bar{y}_k, \bar{s}_k)} - \frac{\bar{q}_k}{\cos^2 \bar{\theta}_k} \right) \\
&\le \operatorname{tr}(\hat{\mathcal{C}}_0) \prod_{i=1}^{k+1} (1 + b_2 \varsigma_i) + \sum_{i=0}^{k} \left(\left(\frac{\|\bar{y}_i\|^2}{g(\bar{y}_i, \bar{s}_i)} - \frac{\bar{q}_i}{\cos^2 \bar{\theta}_i} \right) \prod_{j=i+1}^{k+1} (1 + b_2 \varsigma_j) \right)
\end{aligned}$$

Note that $1 \leq \prod_{i=1}^{k+1}(1 + b_2\varsigma_i) \leq \prod_{i=1}^{\infty}(1 + b_2\varsigma_i) = \exp(\sum_{i=1}^{\infty} \log(1 + b_2\varsigma_i)) \leq \exp(\sum_{i=1}^{\infty} b_2\varsigma_i) = b_3 < \infty$ by (27) in Theorem 5.2, [HAG15, Lemma 3] and [Hua13, Lemma 6.2.5]. We have

$$\mathrm{tr}(\hat{\mathcal{C}}_{k+1}) \leq b_3 \,\mathrm{tr}(\hat{\mathcal{C}}_0) + b_3 \sum_{i=0}^{k} \left(\frac{\|\bar{y}_i\|^2}{g(\bar{y}_i, \bar{s}_i)} - \frac{\bar{q}_i}{\cos^2 \bar{\theta}_i} \right). \tag{40}$$

It follows from (38) and (40) that

$$\begin{aligned} \mathrm{tr}(\hat{\mathcal{C}}_{k+1})/b_3 - \ln(\det(\hat{\mathcal{C}}_{k+1})) \leq& \psi(\hat{\mathcal{C}}_0) + \sum_{i=0}^{k} \left(\frac{\|\bar{y}_i\|^2}{g(\bar{y}_i, \bar{s}_i)} - 1 - \ln \frac{g(\bar{y}_i, \bar{s}_i)}{\|\bar{s}_i\|^2} + \ln \cos^2 \bar{\theta}_i \right. \\ & \left. + 1 - \frac{\bar{q}_i}{\cos^2 \bar{\theta}_i} + \ln \frac{\bar{q}_i}{\cos^2 \bar{\theta}_i} \right). \end{aligned} \tag{41}$$

On one hand, it holds that $t/b_3 - \ln t = t/b_3 - \ln(t/b_3) - \ln b_3 \geq 1 - \ln b_3$. Therefore, $\mathrm{tr}(\hat{\mathcal{C}}_{k+1})/b_3 - \ln(\det(\hat{\mathcal{C}}_{k+1})) = \sum_{i=1}^{d}(\lambda_i/b_3 - \ln \lambda_i) \geq d - d \ln b_3$, where d is the dimension of the manifold $\mathcal{M}$ and λ_i is the i-th eigenvalue of $\hat{\mathcal{C}}_{k+1}$. On the other hand, using Lemma 5.2 yields that $\frac{\|\bar{y}_i\|^2}{g(\bar{y}_i,\bar{s}_i)} - 1 - \ln \frac{g(\bar{y}_i,\bar{s}_i)}{\|\bar{s}_i\|^2} \leq b_4 \epsilon_k^{\min(1,a_{11})}$ holds for all sufficient large i, where b_4 is a positive constant. Therefore, it follows from (41), (27) in Theorem 5.2 and [HAG15, Lemma 3] that

$$\begin{aligned} d - d\ln b_3 \leq& \psi(\hat{\mathcal{C}}_0) + b_4 \sum_{i=0}^{k} \epsilon_i + \sum_{i=0}^{k} \left(\ln \cos^2 \bar{\theta}_i + [1 - \frac{\bar{q}_i}{\cos^2 \bar{\theta}_i} + \ln \frac{\bar{q}_i}{\cos^2 \bar{\theta}_i}] \right) + b_5 \\ =& \psi(\hat{\mathcal{C}}_0) + b_4 b_6 + \sum_{i=0}^{k} \left(\ln \cos^2 \bar{\theta}_i + [1 - \frac{\bar{q}_i}{\cos^2 \bar{\theta}_i} + \ln \frac{\bar{q}_i}{\cos^2 \bar{\theta}_i}] \right) + b_5, \end{aligned}$$

where b_5, b_6 are some constants. Both $\ln \cos^2 \bar{\theta}_i$ and $1 - \frac{\bar{q}_i}{\cos^2 \bar{\theta}_i} + \ln \frac{\bar{q}_i}{\cos^2 \bar{\theta}_i}$ are non-positive and we have $\lim_{i \to \infty} \ln \cos^2 \bar{\theta}_i = 0$ and $\lim_{i\to\infty} 1 - \frac{\bar{q}_i}{\cos^2 \bar{\theta}_i} + \ln \frac{\bar{q}_i}{\cos^2 \bar{\theta}_i} = 0$. Note that the function $\ell(t) = 1 - t + \ln(t), t > 0$ has a unique maximizer at $t = 1$. It follows that $\lim_{i\to\infty} \cos \bar{\theta}_i = \lim_{i\to\infty} \bar{q}_i = 1$, which implies

$$\lim_{k\to\infty} \frac{\|(\mathcal{C}_k - \mathrm{id})\bar{s}\|^2}{\|\bar{s}_k\|^2} = \lim_{k\to\infty} \frac{\bar{q}_k^2}{\cos \bar{\theta}_k^2} - 2\bar{q}_k + 1 = 0.$$

Since

$$\frac{\|H_*^{\frac{1}{2}}\|^2 \|(\mathcal{C}_k - I)\bar{s}_k\|}{\|\bar{s}_k\|} = \frac{\|H_{k+1}^{\frac{1}{2}}\|^2 \|(H_{k+1}^{-\frac{1}{2}} \tilde{B}_k H_{k+1}^{-\frac{1}{2}} - I) H_{k+1}^{\frac{1}{2}} s_k\|}{\|H_{k+1}^{\frac{1}{2}} s_k\|} \geq \frac{\|(\tilde{B}_k - H_{k+1})s_k\|}{\|s_k\|},$$

we have

$$\lim_{k\to\infty} \frac{\|(\tilde{B}_k - H_{k+1})s_k\|}{\|s_k\|} = 0. \tag{42}$$

Let $\tilde{s}_k = \alpha_k \eta_k = \mathcal{T}_{S_{\alpha_k \eta_k}}^{-1} s_k$. It follows that

$$\begin{aligned} \frac{\|(\mathcal{B}_k - H_k)\eta_k\|}{\|\eta_k\|} &= \frac{\|(\mathcal{B}_k - H_k)\tilde{s}_k\|}{\|\tilde{s}_k\|} = \frac{\|(\mathcal{B}_k \mathcal{T}_{S_{\alpha_k \eta_k}}^{-1} - H_k \mathcal{T}_{S_{\alpha_k \eta_k}}^{-1})s_k\|}{\|s_k\|} \\ &= \frac{\|(\mathcal{T}_{S_{\alpha_k \eta_k}} \mathcal{B}_k \mathcal{T}_{S_{\alpha_k \eta_k}}^{-1} - \mathcal{T}_{S_{\alpha_k \eta_k}} H_k \mathcal{T}_{S_{\alpha_k \eta_k}}^{-1})s_k\|}{\|s_k\|} \\ &= \frac{\|(\tilde{\mathcal{B}}_k - H_{k+1} + H_{k+1} - \mathcal{T}_{S_{\alpha_k \eta_k}} H_k \mathcal{T}_{S_{\alpha_k \eta_k}}^{-1})s_k\|}{\|s_k\|} \\ &\leq \frac{\|(\tilde{\mathcal{B}}_k - H_{k+1})s_k\|}{\|s_k\|} + \frac{\|(H_{k+1} - \mathcal{T}_{S_{\alpha_k \eta_k}} H_k \mathcal{T}_{S_{\alpha_k \eta_k}}^{-1})s_k\|}{\|s_k\|} \\ &\to 0. \text{ (by (42) and [Hua13, Lemma 6.2.5])} \end{aligned}$$

What is more, from Assumption 5.5 and $\mathcal{B}_k \eta_k = -\operatorname{grad} f(x_k)$, it holds that

$$\lim_{k \to \infty} \frac{\|\operatorname{grad} f(x_k) + \operatorname{Hess} f(x_k)\eta_k\|}{\|\eta_k\|} = 0.$$

Define $\hat{f}_x = f \circ R_x$. Using the continuity of $\operatorname{Hess} f$ and $\operatorname{Hess} \hat{f}$ and $\operatorname{Hess} f(x^*) = \operatorname{Hess} \hat{f}_{x^*}(0_{x^*})$ yields the desired result (33). □

Corollary 5.2. *Suppose Assumptions 5.1 holds; Assumption 5.2 holds with $\tilde{\Omega}$ satisfying A.1–A.3; Assumptions 5.3 and 5.4 hold; and Assumption 5.5 holds with $a_{11} = 1$. Then there exists an index k_0 such that $\alpha_k = 1$ satisfies either* (2) *or* (3) *for $k \geq k_0$. Moreover, if $\alpha_k = 1$ is used for all $k \geq k_0$, then x_k converges to x^* superlinearly.*

Proof. By [RW12, Proposition 5], (33) in Theorem 5.3 implies that $\alpha_k = 1$ satisfies the Wolfe condition for all $k \geq k_0$. Therefore, $\alpha_k = 1$ also satisfies either (2) or (3) for $k \geq k_0$. Furthermore, using [RW12, Proposition 8] yields the superlinear convergence result. □

Note that Assumption 5.5 with $a_{11} = 1$ reduces to a Riemannian generalization of twice Lipschitz continuous differentiability. We note here without proof that, analogous to the Euclidean setting, Corollary 5.2 still holds as long as $a_{11} > 0$.

It is shown in [Hua13, Theorem 5.2.4] that $\alpha_k = 1$ eventually satisfies the two frequently used line search conditions, i.e., the Wolfe conditions

$$h_k(\alpha_k) \leq h_k(0) + c_1 \alpha_k h_k'(0) \tag{43}$$

$$h_k'(\alpha_k) \geq c_2 h_k'(0), \tag{44}$$

where $0 < c_1 < 0.5 < c_2 < 1$ and the Armijo-Goldstein condition

$$h_k(\alpha_k) \leq h_k(0) + \sigma \alpha_k h_k'(0), \tag{45}$$

where α_k is the largest value in the set $\{t^{(i)} | t^{(i)} \in [\varrho_1 t^{(i-1)}, \varrho_2 t^{(i-1)}], t^{(0)} = 1\}$, $0 < \varrho_1 < \varrho_2 < 1$ and $0 < \sigma < 0.5$.[2] Therefore, if $\alpha_k = 1$ is attempted first using one of the line search conditions, then the superlinear convergence of Algorithm 1 is obtained.

If $h'(t)$ must be evaluated at $t \neq 0$ in line search conditions, such as the Wolfe condition, then the action of vector transport by differentiated retraction is necessary at least on a single direction. More specifically, the term $h'(t) = g_{R_{x_k}(t\eta_k)}(\operatorname{grad} f(R_{x_k}(t\eta_k)), \mathcal{T}_{R_{t\eta_k}} \eta_k)$ requires the action of vector transport by differentiated retraction, $\mathcal{T}_{R_\eta} \xi$, with η and ξ in the same direction. This has been discussed in [HGA15] and one approach to resort to as little information on differentiated retraction as possible is also proposed therein. With line search conditions (such as the Armijo-Goldstein condition) that do not require $h'(t)$ at $t \neq 0$, the use of differentiated retraction can be completely avoided since $\mathcal{T}_{R_{0\eta_k}} \eta_k = \eta_k$.

6 Limited-memory Riemannian BFGS Method

Analogous to the Euclidean BFGS method, the RBFGS method uses a dense matrix as the Hessian approximation, which is not efficient in the sense of computations and storage especially in the case of large-scale problems. In addition, the computations of $\tilde{\mathcal{B}}_k = \mathcal{T}_{S_{\alpha_k \eta_k}} \circ \mathcal{B}_k \circ \mathcal{T}_{S_{\alpha_k \eta_k}}^{-1}$ may need matrix multiplication, which may not be cheap. In [HGA15], a limited-memory RBFGS was proposed to overcome the difficulties. The RBFGS method in this paper, Algorithm 1, also allows the definition of a limited-memory RBFGS method, stated in Algorithm 2. Compared to the LRBFGS in [HGA15], Algorithm 2 avoids the dependence on differentiated retraction.

[2] If $\varrho_1 = \varrho_2$, then α_k is the largest value in the set $\{1, \varrho_1, \varrho_1^2, \ldots\}$. The difference of ϱ_1 and ϱ_2 allows the use of a more sophisticated approach to choose a point between $\varrho_1 t^{(i-1)}$ and $\varrho_2 t^{(i-1)}$, such as based on a polynomial interpolation (see [DS83, Section 6.3.2]).

Algorithm 2 LRBFGS

Input: Riemannian manifold $\mathcal{M}$ with Riemannian metric g; retraction R; isometric vector transport $\mathcal{T}_{\mathrm{S}}$, with R as associated retraction; continuously differentiable real-valued function f on $\mathcal{M}$, bounded below; initial iterate $x_0 \in \mathcal{M}$; convergence tolerance $\varepsilon > 0$; integer $m > 0$.

1: $k = 0$, $\gamma_0 = 1$, $l = 0$.
2: $\mathcal{H}_k^0 = \gamma_k \,\mathrm{id}$. Obtain $\eta_k \in \mathrm{T}_{x_k}\mathcal{M}$ by the following algorithm:
3: $\quad q \leftarrow \operatorname{grad} f(x_k)$
4: $\quad$ **for** $i = k-1, k-2, \ldots, k-l$ **do**
5: $\quad\quad \xi_i \leftarrow \rho_i g(s_i^{(k)}, q)$;
6: $\quad\quad q \leftarrow q - \xi_i y_i^{(k)}$;
7: $\quad$ **end for**
8: $\quad r \leftarrow \mathcal{H}_k^0 q$;
9: $\quad$ **for** $i = k-l, k-l+1, \ldots, k-1$ **do**
10: $\quad\quad \omega \leftarrow \rho_i g(y_i^{(k)}, r)$;
11: $\quad\quad r \leftarrow r + s_i^{(k)}(\xi_i - \omega)$;
12: $\quad$ **end for**
13: set $\eta_k = -r$;
14: find α_k that satisfies either (2) or (3);
15: Set $x_{k+1} = R_{x_k}(\alpha_k \eta_k)$. If $\|\operatorname{grad} f(x_{k+1})\| > \varepsilon$, then break.
16: Define $s_k^{(k+1)} = \mathcal{T}_{\mathrm{S}_{\alpha_k \eta_k}} \alpha_k \eta_k$ and $y_k = \beta_k^{-1} \operatorname{grad} f(x_{k+1}) - \mathcal{T}_{\mathrm{S}_{\alpha_k \eta_k}} \operatorname{grad} f(x_k)$;
17: **if** $\frac{y_k^\flat s_k}{\|s_k\|^2} \geq \vartheta(\|\operatorname{grad} f(x_k)\|)$ **then**
18: $\quad$ Define $\rho_k = 1/g(s_k^{(k+1)}, y_k^{(k+1)})$ and $\gamma_{k+1} = g(s_k^{(k+1)}, y_k^{(k+1)})/\|y_k^{(k+1)}\|^2$.
19: $\quad$ Add $s_k^{(k+1)}$, $y_k^{(k+1)}$ and ρ_k into storage and if $l \geq m$, then discard vector pair $\{s_{k-l}^{(k)}, y_{k-l}^{(k)}\}$ and scalar ρ_{k-l} from storage, else $l \leftarrow l+1$; Transport $s_{k-l+1}^{(k)}, s_{k-l+2}^{(k)}, \ldots, s_{k-1}^{(k)}$ and $y_{k-l+1}^{(k)}, y_{k-l+2}^{(k)}, \ldots, y_{k-1}^{(k)}$ from $\mathrm{T}_{x_k}\mathcal{M}$ to $\mathrm{T}_{x_{k+1}}\mathcal{M}$ by $\mathcal{T}_{\mathrm{S}}$, to get $s_{k-l+1}^{(k+1)}, s_{k-l+2}^{(k+1)}, \ldots, s_{k-1}^{(k+1)}$ and $y_{k-l+1}^{(k+1)}, y_{k-l+2}^{(k+1)}, \ldots, y_{k-1}^{(k+1)}$.
20: **else**
21: $\quad$ Set $\gamma_{k+1} \leftarrow \gamma_k$, $\{\rho_k, \ldots, \rho_{k-l+1}\} \leftarrow \{\rho_{k-1}, \ldots, \rho_{k-l}\}$, $\{s_k^{(k+1)}, \ldots, s_{k-l+1}^{(k+1)}\} \leftarrow \{\mathcal{T}_{\mathrm{S}_{\alpha_k \eta_k}} s_{k-1}^{(k)}, \ldots, \mathcal{T}_{\mathrm{S}_{\alpha_k \eta_k}} s_{k-l}^{(k)}\}$ and $\{y_k^{(k+1)}, \ldots, y_{k-l+1}^{(k+1)}\} \leftarrow \{\mathcal{T}_{\mathrm{S}_{\alpha_k \eta_k}} y_{k-1}^{(k)}, \ldots, \mathcal{T}_{\mathrm{S}_{\alpha_k \eta_k}} y_{k-l}^{(k)}\}$
22: **end if**
23: $k = k+1$, goto Step 2.

7 Experiments

In this section, we compare the numerical performance of the Wolfe condition and the Armijo-Goldstein condition on Algorithms 1 and 2. Since the Wolfe condition requires the evaluation of $h'(t)$ at $t \neq 0$, the evaluation of the vector transport by differentiated retraction is needed. We use the locking condition proposed in [HGA15], which restricts the retraction R and the isometric vector transport $\mathcal{T}_{\mathrm{S}}$:

$$\mathcal{T}_{\mathrm{S}_\xi} \xi = \beta \mathcal{T}_{R_\xi} \xi, \quad \beta = \frac{\|\xi\|}{\|\mathcal{T}_{R_\xi} \xi\|}. \tag{46}$$

This condition requires less information about the differentiated retraction than the approach in [RW12]. However, it may need more computations in the vector transport compared to vector transports without any information from vector transport by differentiated retraction. As shown in [HGA15, Section 4], a few extra low-rank updates may be necessary in the vector transport. In contrast with the Wolfe condition, the Armijo-Goldstein condition does not require any information from the differentiated retraction. Therefore, it allows a cheaper isometric vector transport in general.

7.1 Problem, Retraction, Vector Transport and Step Size

The joint diagonalization (JD) problem on the Stiefel manifold [TCA09] is used to illustrate the performance of the proposed algorithms:

$$\min_{X \in \mathrm{St}(p,n)} f(X) = \min_{X \in \mathrm{St}(p,n)} -\sum_{i=1}^{N} \| \operatorname{diag}(X^T C_i X)\|_2^2,$$

where $\mathrm{St}(p,n) = \{X \in \mathbb{R}^{n\times p} | X^T X = I_p\}$, matrices $C_1, \ldots, C_N$ are given symmetric matrices, $\operatorname{diag}(M)$ denotes a vector formed by the diagonal entries of matrix M, and $\|\cdot\|_2$ denotes the 2-norm.

The algorithms proposed in this paper have been used to efficiently solve applications such as the geometric mean of symmetric positive definite (SPD) matrices [YHAG17] and the phase retrieval problem [?]. We are also working on other applications including the matrix singular value problems [SI13], dictionary learning on SPD matrices [CS17], computations in elastic shape analysis of curves [HGSA15, YHGA15], and the matrix completion problem [Van13, BA14, Mis14].

The Riemannian metric g on $\mathrm{St}(p,n)$ is the Euclidean metric $g(\eta_X, \xi_X) = \mathrm{tr}(\eta_X^T \xi_X)$. With this Riemannian metric g, the gradient is given in [TCA09, Section 2.3]. As discussed in [HAG15, Section 2.2] or more specifically in [HAG16b], a tangent vector $\eta_X \in \mathrm{T}_X \mathrm{St}(p,n)$ can be represented by a vector in the embedding space $\mathbb{R}^{n\times p}$ or a d-dimensional coefficient vector of a basis of $\mathrm{T}_X \mathrm{St}(p,n)$, where $d = np - p(p+1)/2$ is the dimension of $\mathrm{St}(p,n)$. 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. Hence, an orthonormal basis of $\mathrm{T}_X \mathrm{St}(p,n)$ is given by $\{\frac{1}{\sqrt{2}} X(e_i e_j^T - e_j e_i^T) : i = 1, \ldots, p, j = i+1, \ldots, p\} \cup \{X_\perp \tilde{e}_i e_j^T, i = 1, \ldots, n-p, j = 1, \ldots, p\}$, 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 $\mathbb{R}^{n-p}$. In practice, one does not have to form the matrix $X_\perp$. Only the actions of $X_\perp$ and $X_\perp^T$ are needed, i.e., $X_\perp^T \eta_X$ and $X_\perp K$ are needed given $\eta_X \in \mathrm{T}_X \mathcal{M}$ and $K \in \mathbb{R}^{(n-p)\times p}$ respectively. It has been shown in [HAG16b] that the time and space complexities of computing d-dimensional representation of a tangent vector are $O(np^2)$ and $O(np)$ respectively.

By varying the basis and fixing the coefficients, one can define vector transport by parallelization [HAG15, Section 2.3.1 and 5]. The implementation of this vector transport using a d-dimensional representation for tangent vectors is the identity [Hua13, Section 9.5] and is used in our experiment. For detailed discussions about the d-dimensional representation for tangent vectors and vector transport by parallelization on the Stiefel manifold, we refer to [HAG16b].

Three retractions are used. The first is the qf retraction [AMS08, (4.7)]

$$R_X(\eta_X) = \mathrm{qf}(X + \eta_X), \tag{47}$$

where qf denotes the Q factor of the QR decomposition with nonnegative elements on the diagonal of R. The pair of qf retraction and vector transport by parallelization does not satisfy the locking condition, in general, so we use the approach in [HGA15, Section 4.2] to modify the vector transport by parallelization in order to satisfy the locking condition. The implementation of the modified vector transport is a rank-two update matrix.

The second retraction is given in [Zhu17, (14)]

$$R_X(\eta_X) = \left(I - \frac{1}{2}W_{\eta_X}\right)^{-1}\left(I + \frac{1}{2}W_{\eta_X}\right)X, \tag{48}$$

where $W_{\eta_X} = P_X \eta_X X^T - X\eta_X^T P_X$ and $P_X = I - \frac{1}{2}XX^T$. A Cayley-based isometric vector transport associated with this retraction is also given in [Zhu17, (22)]. This pair of retraction and vector transport does not satisfy the locking condition and the approach in [HGA15, Section 4.2] also can be used, which yields extra cost of rank-two update. It can be shown that the complexities of this retraction and the action of this isometric vector transport, i.e., given $\xi_X \in \mathrm{T}_X\,\mathcal{M}$, the evaluation of $\mathcal{T}_{\eta_X}\xi_X$, are both $O(np^2)$. Note that the isometric vector transport in [Zhu17, (22)] uses vectors in $\mathbb{R}^{n\times p}$ to represent tangent vectors rather than a d-dimensional representation.

The third retraction [HGA15, (7.3)] is given by the approach in [HGA15, Section 4.3]

$$\begin{pmatrix} R_X(\eta_X) & (R_X(\eta_X))_\perp \end{pmatrix} = \begin{pmatrix} X & X_\perp \end{pmatrix} \exp\begin{pmatrix} A & -K^T \\ K & 0_{(n-p)\times(n-p)} \end{pmatrix} \tag{49}$$

where $R_X(\eta_X)$ defines the retraction, $A = X^T\eta_X$, $K = X_\perp^T \eta_X$. As shown in [HGA15], the pair of retraction (49) and vector transport by parallelization, which uses $(R_X(\eta_X))_\perp$ in (49) to form the basis of $\mathrm{T}_{R_X(\eta_X)}\,\mathrm{St}(p,n)$, satisfies the locking condition with $\beta \equiv 1$. Therefore, the $X_\perp$ in (49) is computed and stored explicitly. The time and space complexities are $O(n^3)$ and $O(n^2)$ respectively.

The initial step size at each line search follows from [NW06, (3.60) and p. 60]. In our Riemannian experiments, this initial step size in the line search algorithm eventually is one. The two algorithms for finding a step size α_k satisfying the Armijo-Goldstein condition and the Wolfe condition are both based on quadratic or cubic polynomial interpolation. Specifically, they are [DS83, Algorithm A6.3.1 and Algorithm A6.3.1mod] respectively. We refer to [DS83, Section 6.3.2] for the technical details.

7.2 Environment and Parameters

All experiments are performed in C++ with compiler g++-4.7 on a 64 bit Ubuntu platform with 3.6 GHz CPU(Intel(R) Core(TM) i7-4790).[3]

The inverse Hessian approximation update formula

$$\mathcal{H}_{k+1} = \left(\mathrm{id} - \frac{s_k y_k^\flat}{g(y_k, s_k)}\right)\tilde{\mathcal{H}}_k\left(\mathrm{id} - \frac{y_k s_k^\flat}{g(y_k, s_k)}\right) + \frac{s_k s_k^\flat}{g(y_k, s_k)}, \quad \tilde{\mathcal{H}}_k = \mathcal{T}_{\mathrm{S}_{\alpha_k \eta_k}} \circ \mathcal{H}_k \circ \mathcal{T}_{\mathrm{S}_{\alpha_k \eta_k}}^{-1} \tag{50}$$

is used since computing $\eta_k = -\mathcal{H}_k \,\mathrm{grad}\, f(x_k)$ is cheaper than solving the linear system $\mathcal{B}_k \eta_k = -\,\mathrm{grad}\, f(x_k)$.

The function $\vartheta(t)$ in Algorithms 1 and 2 is chosen to be $10^{-4}t$. The stopping criterion is $gf_f/gf_0 < 10^{-6}$, where gf_f and gf_0 denote the final and initial norms of the gradients. The initial inverse Hessian approximation is chosen to be the identity. The value m in Algorithm 2 is set to be 4.

The C_i matrices are selected as $C_i = \mathrm{diag}(n, n-1, \ldots, 1) + 0.1(R_i + R_i^T)$, where the elements of $R_i \in \mathbb{R}^{n\times n}$ are independently drawn from the standard normal distribution. The initial iterate X_0 is given by applying Matlab's function *orth* to a matrix whose elements are drawn from the standard normal distribution using Matlab's *randn*.

[3]The code is available at `www.math.fsu.edu/~whuang2/papers/ARBMDRNOP.htm`.

7.3 Tests and Results

The values n and p are chosen to be 12 and 8 respectively. The parameters c_1 and σ are set to be a same value 10^{-4} since they play a same role the in line search conditions. Experimental results of averages of 1000 random runs for RBFGS and LRBFGS with various values of N, c_2, ϱ_1 and ϱ_2 are reported in Tables 1 and 2.

Let RBFGSW and RBFGSA denote Algorithm 1 with the Wolfe condition and the Armijo-Goldstein condition respectively and LRBFGSW and LRBFGSA denote corresponding limited-memory versions. Let RV1 and RV4 respectively denote the pair of retraction (47) and the vector transport by parallelization and the pair of retraction (48) and the Cayley-based vector transport, both of which do not satisfy the locking condition; RV2 and RV5 respectively denote RV1 using approach [HGA15, Section 4.2] and RV4 using approach [HGA15, Section 4.2], both of which satisfy the locking condition but the vector transports are not smooth; RV3 denote retraction (49) and the vector transport by parallelization, which satisfy the locking condition and have a smooth vector transport.

Note that there are not results for RBFGSW and LRBFGSW with RV1 and RV4 since the well-definedness of RBFGSW and LRBFGSW requires the locking condition. We observed in our experiments that the search direction may not be a descent direction in these cases. It follows that the algorithms do not converge.

We do not report the experimental results of RBFGS with RV4 and RV5 since the complexities of $\tilde{\mathcal{H}}_k = \mathcal{T}_{S_{\alpha_k \eta_k}} \circ \mathcal{H}_k \circ \mathcal{T}_{S_{\alpha_k \eta_k}}^{-1}$ in RBFGS's update is $O(n^2 p^3)$ for the Cayley-based vector transport. In contrast, the complexity of vector transport by parallelization is only $O(1)$ when the intrinsic representation is used for tangent vectors..

We can see that increasing the range of $[\varrho_1, \varrho_2]$ does not influence the performance of RBFGSA and LRBFGSA significantly. However, as c_2 increases, RBFGSW and LRBFGSW tend to perform better in the sense that the numbers of function and gradient evaluations and computational time decrease in most cases. These phenomena lead us to recommend choosing a large $c_2 < 1$ for the Wolfe condition. There is no significant difference between the performance of RBFGSW with the largest c_2 and RBFGSA when the same pair of retraction and vector transport is used.

RBFGS with RV1 performs worse than RBFGS with RV2 in the sense of number of function and gradient evaluations. This implies that the locking condition, to some extent, reduces the number of function and gradient evaluations in RBFGS with either the Armijo-Goldstein condition or the Wolfe condition. Note that even though $h'(t)$ at $t \neq 0$ is not used in the Armijo-Goldstein line search condition, the locking condition can still reduce the number of function and gradient evaluations. However, due to the lower complexities of vector transport (compared to RV2), the relationship between the efficiencies of RBFGS with RV1 and RBFGS with RV2 depends on the costs on the function, gradient evaluation and the vector transport. Specifically, if the cost on vector transport is not negligible compared to the function and gradient evaluation, then the RBFGS with RV2 can be slower, such as results of $N = 32$ in Table 1. Otherwise, the RBFGS with RV1 can be slower, such as results of $N = 512$ in Table 1.

The pair of vector transport and retraction in RV3 satisfies the locking condition (46). Unsurprisingly, the numbers of function and gradient evaluations and vector transports of RBFGS with RV3 are smaller than RBFGS with RV1. Therefore, if the costs on the retraction and vector transport of RV1 and RV3 are not significantly different, which is true for $n = 12$, $p = 8$, and $N = 512$, then RBFGS with RV3 is faster than RBFGS with RV1 in the sense of computational time. Otherwise, RBFGS with RV3 can be unacceptably slow, as we see in Table 3 discussed later.

Regarding LRBFGS method, Table 2 shows that the numbers of function, gradient evaluations and vector transports are not influenced significantly by the choices of retraction and vector transports. Therefore, the lower the complexity of retraction and vector transport is, the faster the algorithm is in terms of computational time. As discussed in Section 7.1, RV1 has lower complexity than RV2 and RV3, RV4 has lower complexity than RV5. It is pointed out first that RV1 has lower cost than RV4. In LRBFGS method, multiple tangent vectors need be transported to a new tangent space, as shown in Step 21 of Algorithm 2. Suppose the number of tangent spaces is k. In RV1, the complexity of vector transport is independent of k, i.e., $O(np^2)$. However, in the Cayley-based vector transport of RV4, the complexity is $O(kn^2 p)$. It follows that LRBFGSA with RV1 is the faster algorithm among all of them.

Table 1: An average of 1000 random runs of RBFGS. $iter$, nf, ng, nV and t denote the number of iterations, the number of function evaluations, the number of gradient evaluations, the number of vector transports and the computational wall time in seconds. The subscript $-k$ indicates a scale of 10^{-k}.

	N		Armijo-Goldstien: $[\varrho_1, \varrho_2]$				Wolfe: c_2			
			$[\frac{1}{2}, \frac{1}{2}]$	$[\frac{1}{4}, \frac{3}{4}]$	$[\frac{1}{16}, \frac{15}{16}]$	$[\frac{1}{64}, \frac{63}{64}]$	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{15}{16}$	$\frac{63}{64}$
RV1	32	$iter$	143	140	140	140	\	\	\	\
		nf	152	148	148	148	\	\	\	\
		ng	144	141	141	141	\	\	\	\
		nV	286	280	280	280	\	\	\	\
		t	6.39_{-3}	6.36_{-3}	6.40_{-3}	6.46_{-3}	\	\	\	\
	512	$iter$	163	161	160	160	\	\	\	\
		nf	175	171	171	170	\	\	\	\
		ng	164	162	161	161	\	\	\	\
		nV	326	321	321	321	\	\	\	\
		t	5.16_{-2}	4.94_{-2}	4.88_{-2}	4.84_{-2}	\	\	\	\
RV2	32	$iter$	96	95	95	95	89	91	94	95
		nf	103	102	101	101	113	105	102	101
		ng	97	96	96	96	107	99	96	96
		nV	192	191	190	190	285	280	282	284
		t	7.55_{-3}	7.54_{-3}	7.47_{-3}	7.46_{-3}	7.96_{-3}	7.68_{-3}	7.70_{-3}	7.80_{-3}
	512	$iter$	108	108	107	107	95	97	101	104
		nf	118	116	115	115	126	117	114	114
		ng	109	109	108	108	118	110	107	107
		nV	217	215	215	215	308	303	308	314
		t	3.97_{-2}	3.84_{-2}	3.79_{-2}	3.83_{-2}	4.14_{-2}	3.85_{-2}	3.63_{-2}	3.55_{-2}
RV3	32	$iter$	121	116	116	116	109	110	114	115
		nf	132	126	126	126	141	129	126	126
		ng	122	117	117	117	131	119	117	117
		nV	241	232	231	231	349	338	343	347
		t	9.76_{-3}	9.36_{-3}	9.40_{-3}	9.42_{-3}	1.01_{-2}	9.71_{-3}	9.16_{-3}	9.45_{-3}
	512	$iter$	141	135	134	134	122	121	127	131
		nf	156	147	146	146	163	149	146	147
		ng	142	136	135	135	150	137	135	136
		nV	283	269	268	269	392	379	388	397
		t	4.89_{-2}	4.53_{-2}	4.54_{-2}	4.48_{-2}	5.00_{-2}	4.91_{-2}	4.70_{-2}	4.61_{-2}

In Table 3, we use LRBFGS to compare the performance of RV3 to RV1, RV2, RV4, and RV5 when $p \ll n$. Only the Armijo-Goldstein condition is used and N is chosen to be 4. Multiple n and p are tested to show the trend of computational time as n and p increase. As discussed in Section 7.1, RV3 is not competitive with RV1 and RV2 in the sense of computational time due to the significant cost on $X_\perp$ and the matrix exponential. Specifically, using RV3 can be slower by orders of magnitude than using the other four pairs. RV1 again is the fastest methods in the sense of computational time.

8 Conclusion

In this paper, we generalize the cautious BFGS update in [LF01b] and the weak line search condition in [BN89] to the Riemannian setting and define a framework for the RBFGS method by merging those two ideas that allows differentiated retraction to be avoided completed. It is proven that the RBFGS method, Algorithm 1, converges globally to a stationary point for a general nonconvex function, which has not been done for earlier Riemannian versions of the BFGS method. A local superlinear convergence analysis is also given. Experiments show that Algorithms 1 and 2 i) avoid the requirement of differentiated retraction and therefore are easy to use in practice; and ii) have competitive performance compared with the RBFGS and LRBFGS methods in [HGA15] which require information from differentiated retraction. Therefore, Algorithms 1 and 2 should be the first choice of Riemannian BFGS method due to their robustness and efficiency. They have been implemented in ROPTLIB [HAGH16] as the default RBFGS and LRBFGS methods.

Table 2: An average of 1000 random runs of LRBFGS. The subscript $-k$ indicates a scale of 10^{-k}.

	N		Armijo-Goldstien: $[\varrho_1, \varrho_2]$				Wolfe: c_2			
			$[\frac{1}{2}, \frac{1}{2}]$	$[\frac{1}{4}, \frac{3}{4}]$	$[\frac{1}{16}, \frac{15}{16}]$	$[\frac{1}{64}, \frac{63}{64}]$	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{15}{16}$	$\frac{63}{64}$
RV1	32	$iter$	142	142	142	142	\	\	\	\
		nf	145	145	145	145	\	\	\	\
		ng	143	143	143	143	\	\	\	\
		nV	1124	1122	1122	1122	\	\	\	\
		t	4.35_{-3}	4.28_{-3}	4.28_{-3}	4.22_{-3}	\	\	\	\
	512	$iter$	139	138	138	138	\	\	\	\
		nf	142	142	142	142	\	\	\	\
		ng	140	139	139	139	\	\	\	\
		nV	1097	1093	1093	1093	\	\	\	\
		t	3.94_{-2}	3.87_{-2}	3.88_{-2}	3.87_{-2}	\	\	\	\
RV2	32	$iter$	141	140	140	140	136	139	140	140
		nf	144	144	144	144	170	146	144	144
		ng	142	141	141	141	168	144	142	141
		nV	1112	1111	1111	1111	1242	1244	1249	1251
		t	1.23_{-2}	1.22_{-2}	1.22_{-2}	1.22_{-2}	1.28_{-2}	1.19_{-2}	1.19_{-2}	1.19_{-2}
	512	$iter$	137	136	136	136	133	135	136	136
		nf	140	139	139	139	167	141	139	139
		ng	138	137	137	137	165	139	137	137
		nV	1082	1077	1077	1077	1218	1202	1210	1212
		t	5.17_{-2}	4.69_{-2}	4.68_{-2}	4.64_{-2}	5.44_{-2}	4.63_{-2}	4.60_{-2}	4.65_{-2}
RV3	32	$iter$	142	141	141	141	136	139	140	141
		nf	145	144	144	144	171	146	144	145
		ng	143	142	142	142	169	144	142	142
		nV	1122	1114	1114	1114	1247	1242	1248	1254
		t	8.68_{-3}	8.60_{-3}	8.60_{-3}	8.61_{-3}	1.02_{-2}	8.88_{-3}	8.62_{-3}	8.62_{-3}
	512	$iter$	138	136	136	136	133	136	136	136
		nf	142	140	140	140	168	143	140	140
		ng	139	137	137	137	165	140	138	138
		nV	1090	1080	1080	1080	1220	1214	1214	1216
		t	4.44_{-2}	4.34_{-2}	4.34_{-2}	4.30_{-2}	5.30_{-2}	4.49_{-2}	4.33_{-2}	4.35_{-2}
RV4	32	$iter$	141	141	141	141	\	\	\	\
		nf	145	145	145	145	\	\	\	\
		ng	142	142	142	142	\	\	\	\
		nV	1115	1115	1115	1115	\	\	\	\
		t	7.21_{-3}	6.97_{-3}	7.04_{-3}	7.05_{-3}	\	\	\	\
	512	$iter$	138	137	137	137	\	\	\	\
		nf	142	140	140	140	\	\	\	\
		ng	139	138	138	138	\	\	\	\
		nV	1093	1080	1081	1081	\	\	\	\
		t	4.27_{-2}	4.25_{-2}	4.23_{-2}	4.23_{-2}	\	\	\	\
RV5	32	$iter$	141	141	141	141	137	139	140	140
		nf	145	144	144	144	172	147	145	144
		ng	142	142	142	142	170	144	142	142
		nV	1117	1113	1112	1112	1250	1244	1249	1251
		t	1.68_{-2}	1.65_{-2}	1.69_{-2}	1.68_{-2}	1.74_{-2}	1.67_{-2}	1.65_{-2}	1.66_{-2}
	512	$iter$	137	136	136	136	133	135	136	136
		nf	140	140	140	140	167	143	140	140
		ng	138	137	137	137	165	140	138	137
		nV	1082	1077	1077	1077	1214	1208	1212	1212
		t	5.16_{-2}	5.15_{-2}	5.16_{-2}	5.17_{-2}	5.88_{-2}	5.10_{-2}	5.06_{-2}	5.06_{-2}

Table 3: An average of 10 random runs of LRBFGS with Armijo-Goldstein condition, $\varrho_1 = 1/64$, $\varrho_2 = 63/64$, and $N = 4$. The subscript $-k$ indicates a scale of 10^{-k}. NA denotes "not run" due to its cost.

		$n:50$ $p:2$	$n:100$ $p:2$	$n:100$ $p:4$	$n:200$ $p:4$	$n:200$ $p:8$	$n:400$ $p:8$	$n:400$ $p:16$	$n:1600$ $p:32$
RV1	$iter$	137	203	729	954	1602	2461	3644	6993
	nf	142	209	737	966	1616	2484	3672	7045
	ng	138	204	730	955	1603	2462	3645	6994
	nV	1087	1615	5816	7623	12807	19677	29137	55924
	t	2.65_{-3}	5.66_{-3}	2.86_{-2}	6.20_{-2}	1.66_{-1}	8.69_{-1}	2.07	6.99_{1}
RV2	$iter$	124	205	777	1187	1785	2713	3233	7307
	nf	127	210	785	1198	1800	2736	3258	7357
	ng	125	206	778	1188	1786	2714	3234	7308
	nV	979	1628	6206	9485	14270	21690	25849	58435
	t	9.13_{-3}	1.96_{-2}	8.06_{-2}	1.70_{-1}	3.63_{-1}	1.75	3.25	8.54_{1}
RV3	$iter$	140	220	768	1128	1809	2581	3409	NA
	nf	145	225	779	1139	1825	2603	3440	NA
	ng	141	221	769	1129	1810	2582	3410	NA
	nV	1110	1750	6134	9015	14458	20632	27254	NA
	t	1.44_{-1}	6.76_{-1}	2.43	1.99_{1}	3.29_{1}	1.71_{2}	2.32_{2}	NA
RV4	$iter$	133	205	783	1122	1605	2641	3421	7557
	nf	138	210	793	1133	1618	2663	3450	7614
	ng	134	206	784	1123	1606	2642	3422	7558
	nV	1051	1628	6254	8966	12820	21114	27351	60437
	t	3.24_{-3}	6.63_{-3}	3.98_{-2}	1.20_{-1}	2.69_{-1}	1.34	2.47	8.13_{1}
RV5	$iter$	133	211	749	1096	1842	2895	3320	7897
	nf	137	217	758	1105	1858	2916	3347	7957
	ng	134	212	750	1097	1843	2896	3321	7898
	nV	1054	1679	5981	8759	14723	23145	26544	63161
	t	1.13_{-2}	1.89_{-2}	9.34_{-2}	2.29_{-1}	6.07_{-1}	2.58	4.62	1.07_{2}

References

[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.

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

[BA14] N. Boumal and P.-A. Absil. Low-rank matrix completion via trust-regions on the Grassmann manifold. *Linear Algebra and its Applications*, 475:200–239, 2014.

[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.

[BN89] R. H. Byrd and J. Nocedal. A tool for the analysis of quasi-Newton methods with application to unconstrained minimization. *SIAM Journal on Numerical Analysis*, 26(3):727–739, 1989.

[Boo86] W. M. Boothby. *An introduction to differentiable manifolds and Riemannian geometry*. Academic Press, second edition, 1986.

[CS17] A. Cherian and S. Sra. Riemannian dictionary learning and sparse coding for positive definite matrices. *IEEE Transactions on Neural Networks and Learning Systems*, 28(12):2859–2871, Dec 2017.

[Dai13] Y.-H. Dai. A perfect example for the BFGS method. *Mathematical Programming*, 138(1-2):501–530, March 2013. doi:10.1007/s10107-012-0522-2.

[DS83] J. E. Dennis and R. B. Schnabel. *Numerical methods for unconstrained optimization and nonlinear equations*. Springer, New Jersey, 1983.

[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-Moré 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. doi:10.1007/978-1-4471-2437-5_14.

[HAG15] W. Huang, P.-A. Absil, and K. A. Gallivan. A Riemannian symmetric rank-one trust-region method. *Mathematical Programming*, 150(2):179–216, February 2015.

[HAG16a] W. Huang, P.-A. Absil, and K. A. Gallivan. A Riemannian BFGS Method for Nonconvex Optimization Problems. *Lecture Notes in Computational Science and Engineering*, 112:627–634, 2016.

[HAG16b] W. Huang, P.-A. Absil, and K. A. Gallivan. Intrinsic representation of tangent vectors and vector transport on matrix manifolds. *Numerische Mathematik*, 136(2):523–543, 2016.

[HAGH16] Wen Huang, P.-A. Absil, K. A. Gallivan, and Paul Hand. ROPTLIB: an object-oriented C++ library for optimization on Riemannian manifolds. Technical Report FSU16-14, Florida State University, 2016.

[HGA15] W. Huang, K. A. Gallivan, and P.-A. Absil. A Broyden Class of Quasi-Newton Methods for Riemannian Optimization. *SIAM Journal on Optimization*, 25(3):1660–1685, 2015.

[HGSA15] W. Huang, K. A. Gallivan, Anuj Srivastava, and P.-A. Absil. Riemannian optimization for registration of curves in elastic shape analysis. *Journal of Mathematical Imaging and Vision*, 54(3):320–343, 2015. DOI:10.1007/s10851-015-0606-8.

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

[LF01a] D.-H. Li and M. Fukushima. A modified BFGS method and its global convergence in nonconvex minimization. *Journal of Computational and Applied Mathematics*, 129:15–35, 2001.

[LF01b] D.-H. Li and M. Fukushima. On the global convergence of the BFGS method for nonconvex unconstrained optimization problems. *SIAM Journal on Optimization*, 11(4):1054–1064, January 2001. doi:10.1137/S1052623499354242.

[Mis14] B. Mishra. *A Riemannian approach to large-scale constrained least-squares with symmetries*. PhD thesis, University of Liege, 2014.

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

[QGA10] C. Qi, K. A. Gallivan, and P.-A. Absil. Riemannian BFGS algorithm with applications. *Recent Advances in Optimization and its Applications in Engineering*, pages 183–192, 2010.

[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. doi:10.1137/11082885X.

[SI13] H. Sato and T. Iwai. A Riemannian optimization approach to the matrix singular value decomposition. *SIAM Journal on Optimization*, 23(1):188–212, 2013.

[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.

[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.

[Van13] B. Vandereycken. Low-rank matrix completion by Riemannian optimization—extended version. *SIAM Journal on Optimization*, 23(2):1214–1236, 2013.

[WLQ06] Z. Wei, G. Li, and L. Qi. New quasi-Newton methods for unconstrained optimization problems. *Applied Mathematics and Computation*, 175(2):1156–1188, April 2006. doi:10.1016/j.amc.2005.08.027.

[WYYL04] Z. Wei, G. Yu, G. Yuan, and Z. Lian. The superlinear convergence of a modified BFGS-type method for unconstrained optimization. *Computational optimization and applications*, pages 315–332, 2004.

[YHAG17] Xinru Yuan, Wen Huang, P.-A. Absil, and K. A. Gallivan. A Riemannian quasi-newton method for computing the Karcher mean of symmetric positive definite matrices. Technical Report FSU17-02, Florida State University, 2017. www.math.fsu.edu/ whuang2/papers/RMKMSPDM.htm.

[YHGA15] Y. You, W. Huang, K. A. Gallivan, and P. A. Absil. A Riemannian approach for computing geodesics in elastic shape analysis. In *2015 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pages 727–731, Dec 2015.

[Zhu17] X. Zhu. A Riemannian conjugate gradient method for optimization on the Stiefel manifold. *Computational Optimization and Applications*, 67(1):73–110, 2017.