Novelty Assessment Report

Paper: Adaptive gradient descent on Riemannian manifolds and its applications to Gaussian variational inference

PDF URL: https://openreview.net/pdf?id=2TTQjRkgFn

Venue: ICLR 2026 Conference Submission

Year: 2026

Report Generated: 2026-01-05

Abstract

We propose RAdaGD, a novel family of adaptive gradient descent methods on general Riemannian manifolds. RAdaGD adapts the step size parameter without line search, and includes instances that achieve a non-ergodic convergence guarantee, $f(x_k) - f(x_\star) \le \mathcal{O}(1/k)$, under local geodesic smoothness and generalized geodesic convexity. A core application of RAdaGD is Gaussian Variational Inference, where our method provides the first convergence guarantee in the absence of L-smoothness of the target log-density, under additional technical assumptions. We also investigate the empirical performance of RAdaGD in numerical simulations and demonstrate its competitiveness in comparison to existing algorithms.

Core Task Landscape

This paper addresses: **Adaptive Gradient Descent on Riemannian Manifolds**

A total of **50 papers** were analyzed and organized into a taxonomy with **17 categories**.

Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Core Adaptive Gradient Methods**
- **Second-Order and Trust-Region Methods**
- **Energy-Adaptive and Metric-Adaptive Methods**
- **Specialized Problem Formulations**
- **Acceleration and Momentum Techniques**
- **Theoretical Foundations and Geometric Frameworks**
- **Meta-Learning and Learned Optimization**
- **Domain Applications and Problem-Specific Adaptations**
- **Computational Tools and Software Frameworks**
- **Specialized Manifolds and Geometric Structures**

Complete Taxonomy Tree

- Adaptive Gradient Descent on Riemannian Manifolds Survey Taxonomy
- Core Adaptive Gradient Methods
  - Deterministic Adaptive Gradient Descent ★ (3 papers)
  - [0] Adaptive gradient descent on Riemannian manifolds and its applications to Gaussian variational inference (Anon et al., 2026) View paper
  - [1] Adaptive Gradient Descent on Riemannian Manifolds with Nonnegative Curvature (Malitsky, 2025) View paper
  - [49] Gradient method for optimization on Riemannian manifolds with lower bounded curvature (Orizon P. Ferreira, 2019) View paper
  - Stochastic Adaptive Gradient Methods (6 papers)
  - [2] Adaptive stochastic gradient algorithms on Riemannian manifolds (Hiroyuki Kasai, 2022) View paper
  - [6] Riemannian adaptive stochastic gradient algorithms on matrix manifolds (Hiroyuki Kasai, 2019) View paper
  - [8] Riemannian adaptive optimization methods (Gary BÃ©cigneul, 2018) View paper
  - [13] Adaptive Riemannian stochastic gradient descent and reparameterization for Gaussian mixture model fitting (Chunlin Ji, 2023) View paper
  - [25] A general framework of Riemannian adaptive optimization methods with a convergence analysis (Sakai Hiroyuki, 2024) View paper
  - [36] Riemannian Adaptive Optimization Algorithm and its Application to Natural Language Processing (Hiroyuki Sakai, 2021) View paper
  - Variance Reduction Methods (2 papers)
  - [15] Improved Variance Reduction Methods for Riemannian Non-Convex Optimization (Andi Han, 2021) View paper
  - [30] Riemannian SVRG: Fast Stochastic Optimization on Riemannian Manifolds (Zhang Hong-yi, 2022) View paper
- Second-Order and Trust-Region Methods
  - Adaptive Newton and Regularized Methods (3 papers)
  - [7] Adaptive Regularized Newton Method for Riemannian Optimization (Hu Jiang, 2022) View paper
  - [18] Adaptive Trust-Region Method on Riemannian Manifold (Shimin Zhao, 2023) View paper
  - [20] Adaptive quadratically regularized Newton method for Riemannian optimization (Jiang Hu, 2018) View paper
  - Conjugate Gradient Methods (2 papers)
  - [11] Two hybrid conjugate gradient based algorithms on Riemannian manifolds with adaptive restart strategy for nonconvex optimization problems (Meixuan Jiang, 2024) View paper

- ◦ [14] Cayley-transform-based gradient and conjugate gradient algorithms on Grassmann manifolds (Xiaojing Zhu, 2021) View paper
- Energy-Adaptive and Metric-Adaptive Methods (4 papers)
  - [16] Energy-adaptive Riemannian optimization on the Stiefel manifold (Altmann, 2022) View paper
  - [17] Adaptive Riemannian Metrics on SPD Manifolds (Chen, 2023) View paper
  - [40] Adaptive Preconditioned Gradient Descent with Energy (Hailiang Liu, 2023) View paper
  - [41] Energy-Adaptive Riemannian Conjugate Gradient Method for Density Functional Theory (Peterseim, 2025) View paper
- Specialized Problem Formulations
  - Bilevel and Minimax Optimization (3 papers)
  - [3] An Adaptive Algorithm for Bilevel Optimization on Riemannian Manifolds (Shi Xu, 2025) View paper
  - [28] Gradient descent ascent for minimax problems on riemannian manifolds (Feihu Huang, 2023) View paper
  - [29] Riemannian Hamiltonian methods for min-max optimization on manifolds (Andi Han, 2022) View paper
  - Equilibrium Computation and Game-Theoretic Settings (1 papers)
  - [47] Last-Iterate Convergence of Adaptive Riemannian Gradient Descent for Equilibrium Computation (Cai Yang, 2023) View paper
  - Online and Dynamic Regret Optimization (2 papers)
  - [33] Online optimization on hadamard manifolds: Curvature independent regret bounds on horospherically convex objectives (Shahrampour, 2025) View paper
  - [45] Riemannian Online Optimistic Algorithms With Dynamic Regret (Xi Wang, 2025) View paper
- Acceleration and Momentum Techniques (2 papers)
  - [21] Momentum Improves Optimization on Riemannian Manifolds (Foivos Alimisis, 2021) View paper
  - [31] A variational formulation of accelerated optimization on Riemannian manifolds (Duruisseaux, 2022) View paper
- Theoretical Foundations and Geometric Frameworks (4 papers)
  - [12] Hide & Seek: Transformer Symmetries Obscure Sharpness & Riemannian Geometry Finds It (Dangel, 2025) View paper
  - [19] Trivializations for Gradient-Based Optimization on Manifolds (Mario Lezcano Casado, 2022) View paper
  - [24] FAdam: Adam is a natural gradient optimizer using diagonal empirical Fisher information (Hwang, 2024) View paper
  - [26] Mirror Descent and Exponentiated Gradient Algorithms Using Trace-Form Entropies (Andrzej Cichocki, 2025) View paper
- Meta-Learning and Learned Optimization (2 papers)
  - [5] Learning to optimize on Riemannian manifolds (Zhi Gao, 2022) View paper
  - [32] Large-scale Riemannian meta-optimization via subspace adaptation (Yu Peilin, 2025) View paper
- Domain Applications and Problem-Specific Adaptations
  - Quantum and Computational Physics (1 papers)
  - [4] Riemannian optimization methods for ground states of multicomponent Boseâ□□Einstein condensates (Robert Altmann, 2025) View paper
  - Signal Processing and Radar Systems (4 papers)
  - [9] Learning-enhanced Riemannian gradient descent method for transmit-receive joint design towards ISRJ suppression (Xiangfeng Qiu, 2025) View paper
  - [34] Adaptive filter with Riemannian manifold constraint (JosÃ© MejÃa, 2023) View paper
  - [35] MIMO Radar Waveform Design for Range-ISL Optimization via Iterative Deep Unfolding Network (Ziwei Zhao, 2024) View paper
  - [37] Riemannian Manifold Approach for RIS/IOS-Assisted Wireless Networks Design (Bin Li, 2025) View paper
  - Computer Vision and Image Processing (6 papers)
  - [10] Fireants: Adaptive riemannian optimization for multi-scale diffeomorphic matching (Jena, 2024) View paper
  - [38] NRDF: Neural Riemannian Distance Fields for Learning Articulated Pose Priors (Yannan He, 2024) View paper
  - [39] Manifold integrated gradients: Riemannian geometry for feature attribution (Eslam Zaher, 2024) View paper
  - [42] Adaptive group Riemannian manifold learning for hyperspectral image classification (Haoxiang Tao, 2022) View paper
  - [46] Locally adaptive sparse representation on Riemannian manifolds for robust classification (Ming Yin, 2018) View paper
  - [50] Adaptive Riemannian Optimization for Multi-Scale Diffeomorphic Matching (James Gee, 2025) View paper
- Computational Tools and Software Frameworks (3 papers)
  - [23] Geoopt: Riemannian Optimization in PyTorch (Kochurov, 2022) View paper
  - [44] A Diffusion-Map-Based Algorithm for Gradient Computation on Manifolds and Applications (Gomez, 2023) View paper
  - [48] Rieoptax: Riemannian Optimization in JAX (Utpala, 2022) View paper
- Specialized Manifolds and Geometric Structures (3 papers)
  - [22] Neural manifolds and gradient-based adaptation in neural-interface tasks (A Payeur, 2023) View paper
  - [27] Adaptive hamiltonian and riemann manifold monte carlo (Ziyu Wang, 2013) View paper
  - [43] Efficient Riemannian Optimization on the Stiefel Manifold via the Cayley Transform (Jun Li, 2022) View paper

Narrative

Core task: adaptive gradient descent on Riemannian manifolds. This field extends classical optimization to curved spaces where constraints and geometry are intrinsic to the problem. The taxonomy reveals a rich structure spanning ten major branches. Core Adaptive Gradient Methods form the algorithmic backbone, encompassing deterministic and stochastic variants that generalize Adam-style updates to manifolds, as seen in Riemannian Adaptive Methods[8] and Adaptive Stochastic Riemannian[2]. Second-Order and Trust-Region Methods leverage curvature information for faster convergence, while Energy-Adaptive and Metric-Adaptive Methods tailor the Riemannian metric itself to problem structure. Specialized Problem Formulations address bilevel optimization, minimax games, and constrained settings, whereas Acceleration and Momentum Techniques bring Nesterov-style ideas to curved spaces. Theoretical Foundations provide convergence guarantees under various geometric assumptions, and Meta-Learning and Learned Optimization explore data-driven tuning of manifold algorithms. Domain Applications range from signal processing to neural network training, supported by Computational Tools like Geoopt[23] and Rieoptax[48], with Specialized Manifolds focusing on structures such as Stiefel, Grassmann, and SPD matrices.

Several active lines of work highlight key trade-offs. Deterministic methods prioritize clean convergence theory, often assuming bounded curvature or lower-bounded gradients, as in Gradient Lower Bounded[49] and Adaptive Gradient Nonnegative Curvature[1]. Stochastic and variance-reduced approaches balance sample efficiency with geometric complexity, while energy-adaptive schemes dynamically adjust metrics to problem landscapes. The original paper, Adaptive Gradient Riemannian[0], sits squarely within the deterministic adaptive gradient branch, emphasizing rigorous convergence analysis under geometric regularity conditions. Its focus contrasts with the stochastic emphasis of Adaptive Stochastic Riemannian[2] and the bilevel setting of Adaptive Bilevel Riemannian[3], yet shares common ground with Adaptive Gradient Nonnegative Curvature[1] in leveraging curvature bounds. Open questions persist around optimal stepsize schedules, the interplay between metric adaptation and convergence speed, and scalability to high-dimensional manifolds encountered in modern machine learning.

Related Works in Same Category

The following **2 sibling papers** share the same taxonomy leaf node with the original paper:

1. Adaptive Gradient Descent on Riemannian Manifolds with Nonnegative Curvature

Authors: Malitsky, Yura, Aban Ansari-Onnestam, Yura Malitsky | **Year/Venue**: 2025 | **URL**: View paper

Abstract

In this paper, we present an adaptive gradient descent method for geodesically convex optimization on a Riemannian manifold with nonnegative sectional curvature. The method automatically adapts to the local geometry of the function and does not use additional expensive computations other than calculation of the derivative of the Riemannian exponential. We prove the convergence of the method under the assumption of geodesic completeness. The performance of the method is illustrated by experiments...

Relationship Analysis

Both papers belong to the Deterministic Adaptive Gradient Descent category, focusing on adaptive gradient methods for Riemannian optimization without line search under geodesic convexity assumptions. They overlap in addressing adaptive step-size selection on Riemannian manifolds with nonnegative curvature and share applications to Bures-Wasserstein geometry. The original paper (RAdaGD) differs by providing convergence guarantees under local geodesic smoothness rather than global L-smoothness, extends to negatively curved spaces with bounded sectional curvature, and specifically targets Gaussian variational inference applications with non-L-smooth potentials, whereas the candidate paper focuses on nonnegative curvature settings with geodesic completeness assumptions.

---

2. Gradient method for optimization on Riemannian manifolds with lower bounded curvature

Authors: Orizon P. Ferreira, O. P. Ferreira, L. F. Prudente, M. S. Louzeiro | **Year/Venue**: 2019 | **URL**: View paper

Abstract

The gradient method for minimize a differentiable convex function on Riemannian manifolds with lower bounded sectional curvature is analyzed in this paper. The analysis of the method is presented with three different finite procedures for determining the stepsize, namely, Lipschitz stepsize, adaptive stepsize and Armijo's stepsize. The first procedure requires that the objective function has Lipschitz continuous gradient, which is not necessary for the other approaches. Convergence of the whole ...

Relationship Analysis

Both papers belong to the deterministic adaptive gradient descent category, focusing on adaptive gradient methods for optimization on Riemannian manifolds with geodesic convexity assumptions. The original paper proposes RAdaGD, a line-search-free adaptive method that relaxes global L-smoothness to local geodesic smoothness and achieves O(1/k) convergence rates, with applications to Gaussian variational inference. The candidate paper analyzes the standard gradient method (not adaptive) on manifolds with lower bounded curvature using three fixed stepsize strategies (Lipschitz, adaptive, and Armijo), establishing convergence without level set boundedness assumptions and providing iteration-complexity bounds for Lipschitz gradient functions.

Contributions Analysis

Overall novelty summary. The paper proposes RAdaGD, a family of adaptive gradient descent methods for Riemannian manifolds that achieve non-ergodic O(1/k) convergence under local geodesic smoothness and generalized geodesic convexity. It resides in the Deterministic Adaptive Gradient Descent leaf, which contains only three papers total, including the original work. This leaf sits within the broader Core Adaptive Gradient Methods branch, indicating a relatively sparse research direction focused on deterministic settings with rigorous convergence guarantees, as opposed to the more crowded stochastic variants.

The taxonomy reveals that the paper's immediate neighbors include Gradient Lower Bounded and Adaptive Gradient Nonnegative Curvature, both emphasizing geometric regularity conditions for convergence. The sibling category Stochastic Adaptive Gradient Methods contains six papers addressing mini-batch and Adam-like algorithms, reflecting a more active research direction. Nearby branches such as Second-Order Methods and Energy-Adaptive Methods explore alternative geometric frameworks, while Specialized Problem Formulations address bilevel and minimax settings. The deterministic leaf's scope explicitly excludes stochastic variants and variance reduction, positioning this work within a narrower but theoretically focused niche.

Among twenty candidates examined, the Gaussian Variational Inference contribution shows one refutable candidate, suggesting prior work addresses convergence without L-smoothness under certain conditions. The RAdaGD algorithm itself examined four candidates with zero refutations, indicating potential novelty in its adaptive step-size mechanism. The local geodesic smoothness framework examined ten candidates without clear refutation, though the limited search scope means broader prior work may exist. These statistics reflect a targeted semantic search rather than exhaustive coverage, so the apparent novelty should be interpreted cautiously.

Based on the limited search of twenty candidates, the work appears to occupy a sparsely populated deterministic niche within a broader field that increasingly emphasizes stochastic methods. The taxonomy structure suggests the deterministic adaptive gradient direction receives less attention than stochastic counterparts, though the search scope does not capture the full landscape of Riemannian optimization literature.

This paper presents **3 main contributions**, each analyzed against relevant prior work:

Contribution 1: RAdaGD: Adaptive gradient descent on Riemannian manifolds

Description: The authors introduce RAdaGD, a family of line-search-free adaptive gradient descent algorithms for Riemannian optimization. These methods automatically tune step sizes and achieve a non-ergodic convergence rate of O(1/k) under local geodesic smoothness and generalized geodesic convexity, which is claimed to be the first such rate for Riemannian adaptive methods.

This contribution was assessed against **4 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Adaptive Preconditioned Gradient Descent with Energy

URL: View paper

Brief Assessment

Adaptive Preconditioned Energy[40] focuses on preconditioned gradient descent with energy-based adaptive step sizes for constrained optimization, not specifically on Riemannian adaptive methods with non-ergodic convergence rates on manifolds.

---

2. A Riemannian Optimization Perspective of the Gauss-Newton Method for Feedforward Neural Networks

URL: View paper

Brief Assessment

Riemannian Gauss Newton[57] focuses on the Gauss-Newton method for neural networks, not adaptive gradient descent algorithms on general Riemannian manifolds with non-ergodic convergence rates.

3. Riemannian SVRG: Fast Stochastic Optimization on Riemannian Manifolds

URL: View paper

Brief Assessment

Riemannian SVRG[30] focuses on variance-reduced stochastic optimization for finite-sum problems on Riemannian manifolds, not adaptive gradient descent methods with automatic step-size tuning.

4. Generalized Steepest Descent Methods on Riemannian Manifolds and Hilbert Spaces: Convergence Analysis and Stochastic Extensions

URL: View paper

Brief Assessment

Generalized Steepest Descent[65] focuses on steepest descent methods with generalized smoothness conditions on Riemannian manifolds and Hilbert spaces, not specifically on adaptive gradient descent with non-ergodic convergence rates. The candidate addresses different algorithmic approaches and theoretical frameworks.

Contribution 2: First convergence guarantee for GVI without L-smoothness

Description: The authors apply RAdaGD to Gaussian Variational Inference and claim to provide the first algorithm with provable convergence guarantees when the target log-density is not globally L-smooth, requiring only a weaker growth condition and additional technical assumptions.

This contribution was assessed against **6 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Validated Variational Inference via Practical Posterior Error Bounds

URL: View paper

Brief Assessment

Validated Variational Inference[64] focuses on error bounds for variational inference using Wasserstein distances and divergences, not on convergence guarantees for Gaussian Variational Inference algorithms under relaxed smoothness conditions.

2. Forward-backward Gaussian variational inference via JKO in the Bures-Wasserstein Space

URL: View paper

Brief Assessment

Forward Backward Gaussian[60] focuses on composite optimization structure (smooth potential + non-smooth entropy) using JKO operators in Bures-Wasserstein space, requiring β-smooth potentials. The original paper addresses local geodesic smoothness without global L-smoothness assumptions through Riemannian adaptive methods.

3. The computational asymptotics of Gaussian variational inference and the Laplace approximation

URL: View paper

Brief Assessment

Computational Asymptotics Gaussian[62] focuses on asymptotic convexity properties and data-asymptotic convergence guarantees for Gaussian variational inference, not on providing convergence guarantees for optimization algorithms under non-smooth conditions during the iterative optimization process itself.

4. Towards Understanding the Dynamics of Gaussian-Stein Variational Gradient Descent

URL: View paper

Brief Assessment

Gaussian Stein Dynamics[61] focuses on Gaussian-SVGD dynamics with bilinear kernels for Gaussian targets, not on general GVI convergence without L-smoothness assumptions. The candidate's theoretical framework addresses different technical challenges (mean-field limits, finite-particle systems) rather than relaxing smoothness conditions for general target distributions.

5. Provable convergence guarantees for black-box variational inference

URL: View paper

Prior Art Analysis

Black Box Variational[59] demonstrates prior work that provides convergence guarantees for Gaussian Variational Inference without requiring global L-smoothness of the target log-density. The candidate paper establishes convergence guarantees under a growth condition (equation 16) rather than L-smoothness, and explicitly states this is achieved for black-box variational inference. The candidate's theoretical framework predates the original paper's claim of being 'the first' to provide such guarantees, as evidenced by the candidate's analysis of the negative ELBO decomposition and its treatment of the non-smooth neg-entropy component.

Evidence

Evidence 1 - **Rationale**: This pair shows that the candidate provides convergence guarantees for variational inference where the full objective is not L-smooth (only log p is smooth, while the neg-entropy h is non-smooth), directly challenging the original's claim of being first. - **Original**: to the best of our knowledge, ours is the first method with provable guarantees in the non-l-smooth setting. - **Candidate**: suppose that log p(·, x) is m-smooth and concave (resp. μ-strongly concave). generate a sequence wt by using the prox-sgd algorithm (def. 6) applied to l and h (eq. 16), using genergy (5) as an estimator of ∇l. let the stepsizes γt be constant and equal to 1/(p aenergyt) (resp. be decaying as in th...

Evidence 2 - **Rationale**: This demonstrates that the candidate explicitly addresses the non-L-smooth setting by handling the composite objective where the neg-entropy is non-smooth, providing theoretical guarantees despite this challenge. - **Original**: our method provides the first convergence guarantee in the absence of l-smoothness of the target log-density, under additional technical assumptions. - **Candidate**: the first issue is due to the non-smoothness of neg-entropy function h. this means that under the benign assumption that the target log p is smooth, the full objective l + h cannot be smooth, since a nonsmooth function plus a smooth function is always nonsmooth. this renders stochastic optimization ...

Evidence 3 - **Rationale**: This shows the candidate identifies and solves the same fundamental challenge (non-L-smoothness) that the original claims to be first addressing, establishing prior work on this problem. - **Original**: a core application of radagd is gaussian variational inference, where our method provides the first convergence guarantee in the absence of l-smoothness of the target log-density, under additional technical assumptions. - **Candidate**: we identify two fundamental barriers preventing from analysing this vi problem as a standard stochastic optimization problem. first, the gradient noise depends on the parameters w in a nonstandard way (sec. 2.3). this adds great technical complexity and renders many traditional stochastic optimizati...

6. Decoupled variational Gaussian inference

URL: View paper

Brief Assessment

Decoupled Variational Gaussian[63] focuses on computational efficiency for variational Gaussian inference through reparameterization and convex optimization sequences, not on convergence guarantees under relaxed smoothness conditions for the target log-density.

Contribution 3: Local geodesic smoothness framework for broader function classes

Description: The authors establish that their convergence analysis relies on local geodesic smoothness rather than global L-smoothness, which broadens the class of applicable functions. They prove that every twice continuously differentiable function on a complete Riemannian manifold satisfies local geodesic smoothness.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Karush-Kuhn-Tucker optimality conditions for non-smooth geodesic quasi-convex optimization on Riemannian manifolds

URL: View paper

Brief Assessment

KKT Geodesic Optimization[53] focuses on quasi-convex optimization and quasi-subdifferentials, not on local geodesic smoothness assumptions or their relationship to twice continuously differentiable functions.

---

2. Nonconvex Matrix Factorization is Geodesically Convex: Global Landscape Analysis for Fixed-rank Matrix Optimization From a Riemannian Perspective

URL: View paper

Brief Assessment

The candidate paper (Geodesically Convex Factorization[58]) focuses on matrix factorization problems with fixed-rank PSD constraints under Riemannian quotient geometry. While it discusses geodesic smoothness properties, it does not address the broader framework of local geodesic smoothness for general twice continuously differentiable functions on complete Riemannian manifolds that the original paper establishes.

---

3. Riemannian adaptive optimization methods

URL: View paper

Brief Assessment

Riemannian Adaptive Methods[8] focuses on adaptive optimization algorithms (Adam, Adagrad, AMSGrad) on product manifolds with geodesic convexity assumptions, not on establishing local geodesic smoothness as a relaxation of global L-smoothness for broader function classes.

---

4. Zeroth-order Optimization on Riemannian Manifolds

URL: View paper

Brief Assessment

Zeroth Order Riemannian[54] focuses on zeroth-order optimization using gradient-Lipschitz and Hessian-Lipschitz assumptions, not on establishing local geodesic smoothness as a relaxation of global L-smoothness for first-order methods.

---

5. Stochastic Zeroth-Order Riemannian Derivative Estimation and Optimization

URL: View paper

Brief Assessment

Stochastic Zeroth Order[56] focuses on zeroth-order optimization with gradient estimation techniques, not on establishing local geodesic smoothness frameworks or proving that C^2 functions satisfy such properties on Riemannian manifolds.

---

6. A Riemannian Optimization Perspective of the Gauss-Newton Method for Feedforward Neural Networks

URL: View paper

Brief Assessment

Riemannian Gauss Newton[57] establishes geodesic polyak-lojasiewicz and lipschitz-smoothness conditions for neural network training, not a general local geodesic smoothness framework for twice continuously differentiable functions on complete Riemannian manifolds.

---

7. An Introduction to Optimization on Smooth Manifolds

URL: View paper

Brief Assessment

Optimization Smooth Manifolds[52] is an introductory textbook on Riemannian optimization that covers standard geometric concepts. It does not address local geodesic smoothness assumptions or prove that C^2 functions satisfy such properties, which are the novel technical contributions of the original paper.

---

8. Differentially private Riemannian optimization

URL: View paper

Brief Assessment

Differentially Private Riemannian[55] focuses on differential privacy for Riemannian optimization, not on establishing local geodesic smoothness as a relaxation of global L-smoothness for convergence analysis. The candidate's local smoothness appears in a different context (privacy-preserving optimization) rather than as a core analytical framework for adaptive gradient methods.

---

9. Riemannian sam: Sharpness-aware minimization on Riemannian manifolds

URL: View paper

Brief Assessment

Riemannian SAM[51] focuses on sharpness-aware minimization with retraction smoothness assumptions, not on establishing local geodesic smoothness as a general framework for convergence analysis.

---

10. Riemannian SVRG: Fast Stochastic Optimization on Riemannian Manifolds

URL: View paper

Brief Assessment

Riemannian SVRG[30] assumes geodesically L-smooth functions globally, not local geodesic smoothness. The paper states 'each fi : m→ r is geodesically l-smooth' without discussing local smoothness relaxations.

Appendix: Text Similarity Detection

No high-similarity text segments were detected across any compared papers.

References

- [0] Adaptive gradient descent on Riemannian manifolds and its applications to Gaussian variational inference View paper
- [1] Adaptive Gradient Descent on Riemannian Manifolds with Nonnegative Curvature View paper
- [2] Adaptive stochastic gradient algorithms on Riemannian manifolds View paper
- [3] An Adaptive Algorithm for Bilevel Optimization on Riemannian Manifolds View paper
- [4] Riemannian optimization methods for ground states of multicomponent Boseâ□□Einstein condensates View paper
- [5] Learning to optimize on Riemannian manifolds View paper
- [6] Riemannian adaptive stochastic gradient algorithms on matrix manifolds View paper
- [7] Adaptive Regularized Newton Method for Riemannian Optimization View paper
- [8] Riemannian adaptive optimization methods View paper
- [9] Learning-enhanced Riemannian gradient descent method for transmit-receive joint design towards ISRJ suppression View paper
- [10] Fireants: Adaptive riemannian optimization for multi-scale diffeomorphic matching View paper
- [11] Two hybrid conjugate gradient based algorithms on Riemannian manifolds with adaptive restart strategy for nonconvex optimization problems View paper
- [12] Hide & Seek: Transformer Symmetries Obscure Sharpness & Riemannian Geometry Finds It View paper
- [13] Adaptive Riemannian stochastic gradient descent and reparameterization for Gaussian mixture model fitting View paper
- [14] Cayley-transform-based gradient and conjugate gradient algorithms on Grassmann manifolds View paper
- [15] Improved Variance Reduction Methods for Riemannian Non-Convex Optimization View paper
- [16] Energy-adaptive Riemannian optimization on the Stiefel manifold View paper
- [17] Adaptive Riemannian Metrics on SPD Manifolds View paper
- [18] Adaptive Trust-Region Method on Riemannian Manifold View paper
- [19] Trivializations for Gradient-Based Optimization on Manifolds View paper
- [20] Adaptive quadratically regularized Newton method for Riemannian optimization View paper
- [21] Momentum Improves Optimization on Riemannian Manifolds View paper
- [22] Neural manifolds and gradient-based adaptation in neural-interface tasks View paper
- [23] Geoopt: Riemannian Optimization in PyTorch View paper
- [24] FAdam: Adam is a natural gradient optimizer using diagonal empirical Fisher information View paper
- [25] A general framework of Riemannian adaptive optimization methods with a convergence analysis View paper
- [26] Mirror Descent and Exponentiated Gradient Algorithms Using Trace-Form Entropies View paper
- [27] Adaptive hamiltonian and riemann manifold monte carlo View paper
- [28] Gradient descent ascent for minimax problems on riemannian manifolds View paper
- [29] Riemannian Hamiltonian methods for min-max optimization on manifolds View paper
- [30] Riemannian SVRG: Fast Stochastic Optimization on Riemannian Manifolds View paper
- [31] A variational formulation of accelerated optimization on Riemannian manifolds View paper
- [32] Large-scale Riemannian meta-optimization via subspace adaptation View paper
- [33] Online optimization on hadamard manifolds: Curvature independent regret bounds on horospherically convex objectives View paper
- [34] Adaptive filter with Riemannian manifold constraint View paper
- [35] MIMO Radar Waveform Design for Range-ISL Optimization via Iterative Deep Unfolding Network View paper
- [36] Riemannian Adaptive Optimization Algorithm and its Application to Natural Language Processing View paper
- [37] Riemannian Manifold Approach for RIS/IOS-Assisted Wireless Networks Design View paper
- [38] NRDF: Neural Riemannian Distance Fields for Learning Articulated Pose Priors View paper
- [39] Manifold integrated gradients: Riemannian geometry for feature attribution View paper
- [40] Adaptive Preconditioned Gradient Descent with Energy View paper
- [41] Energy-Adaptive Riemannian Conjugate Gradient Method for Density Functional Theory View paper
- [42] Adaptive group Riemannian manifold learning for hyperspectral image classification View paper
- [43] Efficient Riemannian Optimization on the Stiefel Manifold via the Cayley Transform View paper
- [44] A Diffusion-Map-Based Algorithm for Gradient Computation on Manifolds and Applications View paper
- [45] Riemannian Online Optimistic Algorithms With Dynamic Regret View paper
- [46] Locally adaptive sparse representation on Riemannian manifolds for robust classification View paper
- [47] Last-Iterate Convergence of Adaptive Riemannian Gradient Descent for Equilibrium Computation View paper
- [48] Rieoptax: Riemannian Optimization in JAX View paper
- [49] Gradient method for optimization on Riemannian manifolds with lower bounded curvature View paper
- [50] Adaptive Riemannian Optimization for Multi-Scale Diffeomorphic Matching View paper
- [51] Riemannian sam: Sharpness-aware minimization on Riemannian manifolds View paper
- [52] An Introduction to Optimization on Smooth Manifolds View paper
- [53] Karush-Kuhn-Tucker optimality conditions for non-smooth geodesic quasi-convex optimization on Riemannian manifolds View paper
- [54] Zeroth-order Optimization on Riemannian Manifolds View paper
- [55] Differentially private Riemannian optimization View paper
- [56] Stochastic Zeroth-Order Riemannian Derivative Estimation and Optimization View paper
- [57] A Riemannian Optimization Perspective of the Gauss-Newton Method for Feedforward Neural Networks View paper
- [58] Nonconvex Matrix Factorization is Geodesically Convex: Global Landscape Analysis for Fixed-rank Matrix Optimization From a Riemannian Perspective View paper
- [59] Provable convergence guarantees for black-box variational inference View paper
- [60] Forward-backward Gaussian variational inference via JKO in the Bures-Wasserstein Space View paper

- [61] Towards Understanding the Dynamics of Gaussian-Stein Variational Gradient Descent View paper
- [62] The computational asymptotics of Gaussian variational inference and the Laplace approximation View paper
- [63] Decoupled variational Gaussian inference View paper
- [64] Validated Variational Inference via Practical Posterior Error Bounds View paper
- [65] Generalized Steepest Descent Methods on Riemannian Manifolds and Hilbert Spaces: Convergence Analysis and Stochastic Extensions View paper