Novelty Assessment Report

Paper: Improving Online-to-Nonconvex Conversion for Smooth Optimization via Double Optimism

PDF URL: https://openreview.net/pdf?id=ZAflv4dxQ9

Venue: ICLR 2026 Conference Submission

Year: 2026

Report Generated: 2025-12-27

Abstract

A recent breakthrough in nonconvex optimization is the online-to-nonconvex conversion framework of Cutkosky et al. (2023), which reformulates the task of finding an ε-first-order stationary point as an online learning problem. When both the gradient and the Hessian are Lipschitz continuous, instantiating this framework with two different online learners achieves a complexity of $\mathcal{O}(\varepsilon^{-1.75}\log(1/\varepsilon))$ in the deterministic case and a complexity of $\mathcal{O}(\varepsilon^{-3.5})$ in the stochastic case. However, this approach suffers from several limitations: (i) the deterministic method relies on a complex double-loop scheme that solves a fixed-point equation to construct hint vectors for an optimistic online learner, introducing an extra logarithmic factor; (ii) the stochastic method assumes a bounded second-order moment of the stochastic gradient, which is stronger than standard variance bounds; and (iii) different online learning algorithms are used in the two settings. In this paper, we address these issues by introducing an online optimistic gradient method based on a novel **doubly optimistic hint function**. Specifically, we use the gradient at an extrapolated point as the hint, motivated by two optimistic assumptions: that the difference between the hint and the target gradient remains near constant, and that consecutive update directions change slowly due to smoothness. Our method eliminates the need for a double loop and removes the logarithmic factor. Furthermore, by simply replacing full gradients with stochastic gradients and under the standard assumption that their variance is bounded by σ^2, we obtain a unified algorithm with complexity $\mathcal{O}(\varepsilon^{-1.75} + \sigma^2 \varepsilon^{-3.5})$, smoothly interpolating between the best-known deterministic rate and the optimal stochastic rate.

Disclaimer

This report is **AI-GENERATED** using Large Language Models and WisPaper (a scholar search engine). It analyzes academic papers' tasks and contributions against retrieved prior work. While this system identifies **POTENTIAL** overlaps and novel directions, **ITS COVERAGE IS NOT EXHAUSTIVE AND JUDGMENTS ARE APPROXIMATE**. These results are intended to assist human reviewers and **SHOULD NOT** be relied upon as a definitive verdict on novelty.

Note that some papers exist in multiple, slightly different versions (e.g., with different titles or URLs). The system may retrieve several versions of the same underlying work. The current automated pipeline does not reliably align or distinguish these cases, so human reviewers will need to disambiguate them manually.

If you have any questions, please contact: mingzhang23@m.fudan.edu.cn

Core Task Landscape

This paper addresses: **Finding First-Order Stationary Points in Smooth Nonconvex Optimization**

A total of **50 papers** were analyzed and organized into a taxonomy with **26 categories**.

Taxonomy Overview

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

- **Complexity Analysis and Lower Bounds**
- **Algorithm Design and Convergence**
- **Constrained Nonconvex Optimization**
- **Nonsmooth and Composite Optimization**
- **Higher-Order and Second-Order Methods**
- **Specialized Problem Structures**
- **Theoretical Foundations and Landscape Analysis**

Complete Taxonomy Tree

- Finding First-Order Stationary Points in Smooth Nonconvex Optimization Survey Taxonomy
- Complexity Analysis and Lower Bounds
  - Lower Complexity Bounds for Stochastic Methods (2 papers)
  - [1] Lower bounds for non-convex stochastic optimization (Arjevani, 2023) View paper
  - [13] Lower Complexity Bounds of First-Order Methods for Affinely Constrained Composite Nonconvex Problems (Liu Wei, 2025) View paper
  - Deterministic Complexity Bounds (3 papers)
  - [12] Regional complexity analysis of algorithms for nonconvex smooth optimization (Curtis, 2021) View paper
  - [24] The computational complexity of finding stationary points in non-convex optimization (Alexandros Hollender, 2024) View paper
  - [34] On the oracle complexity of first-order and derivative-free algorithms for smooth nonconvex minimization (Gould, 2012) View paper
  - Refined Asymptotic Complexity Analysis (2 papers)
  - [39] Refining asymptotic complexity bounds for nonconvex optimization methods, including why steepest descent is o (eps^{-2}) rather than O (eps^{-2}) (S Gratton, 2025) View paper
  - [50] Refining asymptotic complexity bounds for nonconvex optimization methods, including why steepest descent is rather than (S Gratton, 2024) View paper
- Algorithm Design and Convergence
  - Adaptive Gradient Methods (4 papers)
  - [3] Adaptive methods for nonconvex optimization (Manzil Zaheer, 2018) View paper
  - [8] On the convergence of adaptive gradient methods for nonconvex optimization (Zhou, 2018) View paper
  - [18] UAdam: Unified Adam-Type Algorithmic Framework for Nonconvex Optimization (Yiming Jiang, 2023) View paper
  - [19] Complexity of Adagrad and other first-order methods for nonconvex optimization problems with bounds constraints (Gratton Serge, 2024) View paper
  - Fixed-Stepsize and Momentum Methods (2 papers)

  - [4] Learning-Rate-Free Momentum SGD with Reshuffling Converges in Nonsmooth Nonconvex Optimization (Xiaoyin Hu, 2024) View paper
  - [6] Lookahead converges to stationary points of smooth non-convex functions (Jianyu Wang, 2020) View paper
  - Online-to-Nonconvex Conversion Frameworks ★ (2 papers)
  - [0] Improving Online-to-Nonconvex Conversion for Smooth Optimization via Double Optimism (Anon et al., 2026) View paper
  - [31] Optimal stochastic non-smooth non-convex optimization through online-to-non-convex conversion (Cutkosky, 2023) View paper
  - Variance-Reduced and Retrospective Methods (2 papers)
  - [21] A retrospective approximation approach for smooth stochastic optimization (Newton, 2025) View paper
  - [25] Variance-reduced first-order methods for deterministically constrained stochastic nonconvex optimization with strong convergence guarantees (Lu Zhao-song, 2024) View paper
  - Accelerated First-Order Methods (3 papers)
  - [15] An accelerated first-order method for non-convex optimization on manifolds (Criscitiello, 2023) View paper
  - [22] Faster first-order methods for stochastic non-convex optimization on Riemannian manifolds (Pan Zhou, 2019) View paper
  - [43] An accelerated first-order regularized momentum descent ascent algorithm for stochastic nonconvex-concave minimax problems (Huiling Zhang, 2023) View paper
  - Parameter-Free and Linesearch-Free Methods (2 papers)
  - [36] Complexity-Optimal and Parameter-Free First-Order Methods for Finding Stationary Points of Composite Optimization Problems (Kong, 2022) View paper
  - [38] Simple linesearch-free first-order methods for nonconvex optimization (Yagishita, 2025) View paper
- Constrained Nonconvex Optimization
  - Functional and Inequality Constraints (3 papers)
  - [5] Stochastic first-order methods for convex and nonconvex functional constrained optimization (Boob, 2023) View paper
  - [11] Convergence of first-order methods for constrained nonconvex optimization with dependent data (Alacaoglu, 2023) View paper
  - [48] Level constrained first order methods for function constrained optimization (Digvijay Boob, 2022) View paper
  - Linearly Constrained Optimization (2 papers)
  - [30] Finding second-order stationary points efficiently in smooth nonconvex linearly constrained optimization problems (Songtao Lu, 2020) View paper
  - [46] A global dual error bound and its application to the analysis of linearly constrained nonconvex optimization (Jiawei Zhang, 2022) View paper
  - Proximal and Augmented Lagrangian Methods (2 papers)
  - [10] An implementable proximal-type method for computing critical points to minimization problems with a nonsmooth and nonconvex constraint (de Oliveira, 2024) View paper
  - [44] A proximal-type method for nonsmooth and nonconvex constrained minimization problems (Gregorio M. Sempere, 2025) View paper
- Nonsmooth and Composite Optimization
  - Gradient-Free and Zeroth-Order Methods (1 papers)
  - [7] Gradient-Free Methods for Deterministic and Stochastic Nonsmooth Nonconvex Optimization (Lin, 2022) View paper
  - Smoothing and Proximal Approaches (3 papers)
  - [29] ItsDEAL: Inexact two-level smoothing descent algorithms for weakly convex optimization (Ahookhosh, 2025) View paper
  - [42] ItsOPT: An inexact two-level smoothing framework for nonconvex optimization via high-order Moreau envelope (Ahookhosh, 2024) View paper
  - [47] Computing proximal points of nonconvex functions (Warren Hare, 2009) View paper
- Higher-Order and Second-Order Methods
  - Second-Order Stationary Point Methods (3 papers)
  - [14] A Newton-CG based augmented Lagrangian method for finding a second-order stationary point of nonconvex equality constrained optimization with complexity â¦ (C He, 2023) View paper
  - [16] Yet another fast variant of Newton's method for nonconvex optimization (Serge Gratton, 2025) View paper
  - [40] Efficient approaches for escaping higher order saddle points in non-convex optimization (Anandkumar, 2016) View paper
  - Higher-Order Critical Point Methods (1 papers)
  - [20] An adaptive high order method for finding third-order critical points of nonconvex optimization (Zhu, 2022) View paper
  - Saddle Point Identification and Escape (2 papers)
  - [17] Identifying and attacking the saddle point problem in high-dimensional non-convex optimization (Yann N. Dauphin, 2014) View paper
  - [32] First-order methods almost always avoid strict saddle points (Jason D. Lee, 2019) View paper
- Specialized Problem Structures
  - Minimax and Saddle Point Problems (1 papers)
  - [33] Efficient Algorithms for Smooth Minimax Optimization (Kiran Koshy Thekumparampil, 2019) View paper
  - Bilevel Optimization (2 papers)
  - [35] On the Complexity of Finding Stationary Points in Nonconvex Simple Bilevel Optimization (Cao Jincheng, 2025) View paper
  - [49] Non-convex bilevel games with critical point selection maps (Arbel, 2022) View paper
  - Distributed and Decentralized Optimization (2 papers)
  - [23] Linear convergence of first-and zeroth-order primalâdual algorithms for distributed nonconvex optimization (Xinlei Yi, 2021) View paper
  - [45] The geometric effects of distributing constrained nonconvex optimization problems (Qiuwei Li, 2019) View paper
  - Block Coordinate Descent Methods (1 papers)
  - [41] Block Coordinate Descent Methods for Structured Nonconvex Optimization with Nonseparable Constraints: Optimality Conditions and Global Convergence (Yuan Zhijie, 2024) View paper
  - Non-Hermitian and Quaternion-Valued Problems (2 papers)
  - [27] Identifying non-Hermitian critical points with quantum metric (Ren Jun-Feng, 2024) View paper
  - [28] A Quaternion-Valued Neural Network Approach to Nonsmooth Nonconvex Constrained Optimization in Quaternion Domain (Jingxin Liu, 2024) View paper
- Theoretical Foundations and Landscape Analysis
  - Structural Conditions for Tractability (1 papers)
  - [2] Recent theoretical advances in non-convex optimization (Danilova, 2022) View paper

- ◦ Relaxed Convexity and PL Conditions (1 papers)
- ◦ [9] Linear convergence of first order methods for non-strongly convex optimization (Ion Necoara, 2019) View paper
- ◦ Stationarity Criteria and Optimality Conditions (1 papers)
- ◦ [26] Corrigendum: On the complexity of finding first-order critical points in constrained nonlinear optimization (C Cartis, 2017) View paper
- ◦ Application-Specific Landscape Properties (1 papers)
- ◦ [37] Transformers Learn Nonlinear Features In Context: Nonconvex Mean-field Dynamics on the Attention Landscape (Kim Juno, 2024) View paper

Narrative

Core task: finding first-order stationary points in smooth nonconvex optimization. The field has matured into several well-defined branches that address complementary aspects of this central challenge. Complexity Analysis and Lower Bounds (e.g., Nonconvex Lower Bounds[1], Oracle Complexity[34]) establishes fundamental limits on what any algorithm can achieve, while Algorithm Design and Convergence explores practical schemes—ranging from adaptive methods like Adaptive Nonconvex[3] and UAdam[18] to momentum-based approaches such as Learning Rate Free Momentum[4] and online-to-nonconvex conversion frameworks like Online to Nonconvex[31]. Constrained settings are handled by a dedicated branch (Constrained Nonconvex Optimization) that includes works on functional constraints (Functional Constrained[5]) and affine constraints (Affinely Constrained Bounds[13]), while Nonsmooth and Composite Optimization tackles problems where smoothness assumptions break down (Gradient Free Nonsmooth[7], Proximal Nonsmooth[44]). Higher-Order and Second-Order Methods pursue faster convergence by exploiting curvature information (Fast Newton Variant[16], Second Order Stationary[30]), and Specialized Problem Structures investigates domains like manifolds (Manifold Acceleration[15]) and bilevel optimization (Bilevel Stationary Points[35]).

Recent activity has concentrated on bridging online learning techniques with nonconvex settings and on refining complexity guarantees under weaker assumptions. Double Optimism[0] sits within the online-to-nonconvex conversion framework, closely aligned with Online to Nonconvex[31], which systematically translates regret bounds from online convex optimization into nonconvex convergence rates. This line of work contrasts with more classical adaptive gradient methods (Adaptive Nonconvex[3], Adaptive Gradient Convergence[8]) that tune step sizes based on observed gradient norms, and with momentum schemes (Learning Rate Free Momentum[4]) that accelerate convergence through inertial dynamics. By leveraging optimism—a technique originally designed for online regret minimization—Double Optimism[0] offers a distinct algorithmic pathway that exploits predictive gradient information, positioning it as a natural extension of conversion-based approaches while differing in emphasis from direct adaptive or momentum-driven designs.

Related Works in Same Category

The following **1 sibling papers** share the same taxonomy leaf node with the original paper:

1. Optimal stochastic non-smooth non-convex optimization through online-to-non-convex conversion

Authors: Cutkosky, Ashok, Mehta, Harsh, Orabona, et al. (6 authors total) | **Year/Venue**: 2023 | **URL**: View paper

Abstract

We present new algorithms for optimizing non-smooth, non-convex stochastic objectives based on a novel analysis technique. This improves the current best-known complexity for finding a $(Î´,Îµ)$-stationary point from $O(Îµ^{-4}Î´^{-1})$ stochastic gradient queries to $O(Îµ^{-3}Î´^{-1})$, which we also show to be optimal. Our primary technique is a reduction from non-smooth non-convex optimization to online learning, after which our results follow from standard regret bounds in online learning. F...

Relationship Analysis

Both papers belong to the Online-to-Nonconvex Conversion Frameworks category, reformulating nonconvex optimization as online learning problems with regret-based analysis. The original paper focuses on improving the O2NC framework for smooth optimization by introducing a doubly optimistic hint function based on extrapolated gradients, achieving O(ε^-1.75) deterministic and O(ε^-1.75 + σ^{2}ε^-3.5) stochastic complexities without logarithmic factors. The candidate paper addresses non-smooth nonconvex optimization, establishing optimal O(ε^-3δ^-1) complexity for finding (δ,ε)-stationary points and recovering smooth case results as special cases, representing a broader but less refined treatment of the smooth setting.

Contributions Analysis

Overall novelty summary. The paper introduces a doubly optimistic hint function within the online-to-nonconvex conversion framework, aiming to find first-order stationary points in smooth nonconvex optimization. It sits in the 'Online-to-Nonconvex Conversion Frameworks' leaf, which contains only two papers including the original work. This is a notably sparse research direction within a taxonomy of 50 papers across 26 leaf nodes, suggesting that the conversion-based approach remains relatively underexplored compared to more established branches like adaptive gradient methods or momentum-based schemes.

The taxonomy reveals that neighboring leaves include 'Adaptive Gradient Methods' (four papers on Adam-type algorithms), 'Fixed-Stepsize and Momentum Methods' (two papers on classical acceleration), and 'Variance-Reduced and Retrospective Methods' (two papers on stochastic gradient refinement). The online-to-nonconvex conversion approach diverges from these by reformulating optimization as an online learning problem with regret-based analysis, rather than directly designing step-size rules or momentum schedules. This structural separation indicates that the paper's core methodology—leveraging optimism from online learning—occupies a distinct conceptual niche within the broader algorithm design landscape.

Among 20 candidates examined through semantic search and citation expansion, no papers were found that clearly refute the three main contributions. The unified algorithm achieving best-known rates in both deterministic and stochastic settings was assessed against 10 candidates, none of which provided overlapping prior work. Similarly, the adaptive step-size scheme for parameter-free optimization was compared to 10 candidates without finding refutable overlap. The doubly optimistic hint function itself was not matched against any candidates in the search. These statistics reflect a limited search scope rather than exhaustive coverage, but suggest that within the examined literature, the specific combination of techniques appears novel.

Given the sparse population of the online-to-nonconvex conversion leaf and the absence of refutable prior work among 20 examined candidates, the paper's contributions appear to advance a relatively young research direction. However, the limited search scale means that closely related work in adjacent areas—such as optimistic online learning or parameter-free adaptive methods—may not have been fully captured. The analysis provides evidence of novelty within the examined scope but does not rule out relevant prior work beyond the top-20 semantic matches.

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

Contribution 1: Doubly optimistic hint function for online-to-nonconvex conversion

Description: The authors propose a new hint function that evaluates the gradient at an extrapolated point, motivated by two optimistic assumptions: that the difference between hint and target gradient remains near constant, and that consecutive update directions change slowly due to smoothness. This eliminates the need for a double-loop fixed-point iteration used in prior work.

This contribution was assessed against **0 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

Contribution 2: Unified algorithm achieving best-known rates in deterministic and stochastic settings

Description: The authors develop a single algorithm that achieves complexity $O(\varepsilon-1.75)$ in the deterministic setting (matching the best-known rate without logarithmic factors) and $O(\sigma^2\varepsilon-3.5)$ in the stochastic setting (matching the optimal rate), smoothly interpolating between both regimes using only standard variance assumptions.

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. High-probability complexity bounds for stochastic non-convex minimax optimization

URL: View paper

Brief Assessment

High Probability Minimax[51] focuses on minimax optimization problems with high-probability complexity guarantees, not general nonconvex optimization. The candidate addresses a fundamentally different problem class (minimax/saddle-point problems) compared to the original paper's focus on unconstrained smooth nonconvex optimization.

2. Data-Driven Compositional Optimization in Misspecified Regimes

URL: View paper

Brief Assessment

Compositional Misspecified[52] addresses misspecified compositional optimization problems with risk and nonconvexity, not general nonconvex optimization. The paper focuses on learning unknown parameters θ* while solving compositional problems, which is fundamentally different from the original paper's focus on unified deterministic/stochastic rates for smooth nonconvex optimization.

3. A guide through the zoo of biased SGD

URL: View paper

Brief Assessment

Biased SGD Guide[57] focuses on biased gradient estimators in optimization, not on unified algorithms for deterministic/stochastic nonconvex optimization with optimal rates.

4. Online Optimization Perspective on First-Order and Zero-Order Decentralized Nonsmooth Nonconvex Stochastic Optimization

URL: View paper

Brief Assessment

Decentralized Nonsmooth[55] focuses on decentralized nonsmooth nonconvex optimization with different complexity bounds (O(δ^-1ε^-3)), while the original paper addresses centralized smooth optimization achieving O(ε^-1.75 + σ^2ε^-3.5). The settings and problem classes are fundamentally different.

5. A Novel Unified Parametric Assumption for Nonconvex Optimization

URL: View paper

Brief Assessment

Unified Parametric Assumption[56] focuses on a novel parametric assumption framework for nonconvex optimization with convergence guarantees, not on achieving specific complexity rates O(ε^-1.75) and O(σ^2ε^-3.5) that smoothly interpolate between deterministic and stochastic settings as claimed in the original paper.

6. Optimal complexity in non-convex decentralized learning over time-varying networks

URL: View paper

Brief Assessment

Time Varying Networks[58] focuses on decentralized optimization over time-varying communication networks, not centralized nonconvex optimization with unified deterministic/stochastic rates.

7. Cyclic Block Coordinate Descent With Variance Reduction for Composite Nonconvex Optimization

URL: View paper

Brief Assessment

Cyclic Variance Reduction[59] addresses block coordinate descent methods for composite nonconvex optimization, not the online-to-nonconvex conversion framework. The technical approaches and problem settings differ fundamentally.

8. Variance-reduced first-order methods for deterministically constrained stochastic nonconvex optimization with strong convergence guarantees

URL: View paper

Brief Assessment

Variance Reduced Constrained[25] addresses deterministically constrained stochastic optimization with constraint satisfaction guarantees, while the original paper focuses on unconstrained smooth optimization. These are fundamentally different problem settings with different algorithmic requirements and complexity measures.

9. A unified and refined convergence analysis for non-convex decentralized learning

URL: View paper

Brief Assessment

Decentralized Refined[54] focuses on decentralized multi-agent optimization with network topology considerations, not on unified online-to-nonconvex conversion frameworks for single-agent smooth optimization as in the original paper.

10. Trust region methods for nonconvex stochastic optimization beyond lipschitz smoothness

URL: View paper

Brief Assessment

Trust Region Nonsmooth[53] addresses nonconvex optimization under (L0,L1)-smoothness (generalized smoothness beyond Lipschitz), not the standard Lipschitz smooth setting studied in the original paper. The candidate focuses on trust region methods for a different smoothness regime, while the original develops online-to-nonconvex conversion with doubly optimistic hints for standard smooth functions.

Contribution 3: Adaptive step size scheme for parameter-free optimization

Description: The authors introduce an adaptive step size scheme that automatically adjusts based on past gradient information, eliminating the need for manual tuning while maintaining the same complexity guarantees and offering improved dependence on local rather than global Lipschitz constants.

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. Towards Stability of Parameter-free Optimization

URL: View paper

Brief Assessment

Parameter Free Stability[63] focuses on parameter-free optimization with adaptive step sizes for smooth nonconvex problems using online-to-nonconvex conversion. The candidate (AdamG) addresses parameter-free adaptive gradient methods for general optimization without the specific online-to-nonconvex framework or the doubly optimistic hint construction that characterizes the original paper's contribution.

2. Problem-Parameter-Free Federated Learning

URL: View paper

Brief Assessment

Parameter Free Federated[65] focuses on federated learning with adaptive stepsizes for distributed optimization across clients, not on general nonconvex optimization with online-to-nonconvex conversion frameworks. The candidate addresses client heterogeneity in FL settings rather than single-agent optimization with gradient/Hessian Lipschitz continuity.

3. Analysis of Schedule-Free Nonconvex Optimization

URL: View paper

Brief Assessment

Schedule Free[62] focuses on horizon-free step-size schedules for nonconvex optimization through interpolation between averaging and momentum, whereas the original paper develops adaptive step sizes based on past gradient information within the online-to-nonconvex conversion framework. These are fundamentally different algorithmic approaches addressing different aspects of parameter-free optimization.

4. Storm+: Fully adaptive sgd with recursive momentum for nonconvex optimization

URL: View paper

Brief Assessment

Storm Plus Recursive[61] focuses on parameter-free stochastic nonconvex optimization with recursive momentum, while the original paper's adaptive step size scheme is designed specifically for the online-to-nonconvex conversion framework with doubly optimistic hints. The technical approaches differ fundamentally in their algorithmic frameworks and hint construction mechanisms.

5. Agda+: Proximal alternating gradient descent ascent method with a nonmonotone adaptive step-size search for nonconvex minimax problems

URL: View paper

Brief Assessment

Agda Plus[66] focuses on minimax problems with adaptive step-size search for primal-dual methods, not general nonconvex minimization. The original paper addresses smooth nonconvex optimization via online-to-nonconvex conversion, which is a fundamentally different problem class and algorithmic framework.

6. STORM+: Fully adaptive SGD with momentum for nonconvex optimization

URL: View paper

Brief Assessment

Storm Plus Momentum[64] focuses on parameter-free stochastic nonconvex optimization with adaptive learning rate and momentum parameters, but does not address the online-to-nonconvex conversion framework or the doubly optimistic hint construction that are central to the original paper's contribution.

7. Comparison of optimization techniques based on gradient descent algorithm: A review

URL: View paper

Brief Assessment

Gradient Descent Review[67] discusses adaptive learning rates in general optimization contexts, but does not address the specific parameter-free nonconvex optimization framework with complexity guarantees O(ε^-1.75 + σ^2ε^-3.5) that depends on local rather than global Lipschitz constants as proposed in the original paper.

8. Adaptive methods for nonconvex optimization

URL: View paper

Brief Assessment

Adaptive Nonconvex[3] focuses on adaptive learning rates for nonconvex optimization in deep learning contexts (Adam, Yogi variants), not on parameter-free optimization frameworks that eliminate manual tuning of global constants like Lipschitz parameters. The original paper's contribution addresses automatic adjustment based on local Lipschitz constants within an online-to-nonconvex conversion framework, which is a different technical approach.

9. Nest your adaptive algorithm for parameter-agnostic nonconvex minimax optimization

URL: View paper

Brief Assessment

Nested Adaptive[60] focuses on minimax optimization with nested inner/outer loops for primal-dual variables, not general nonconvex optimization. The adaptive mechanisms serve different algorithmic purposes in fundamentally different problem settings.

10. Simple linesearch-free first-order methods for nonconvex optimization

URL: View paper

Brief Assessment

Linesearch Free[38] focuses on auto-conditioned stepsize strategies for nonconvex optimization using local curvature estimates, while the original paper addresses online-to-nonconvex conversion with doubly optimistic hints for smooth optimization. The technical approaches and problem formulations differ fundamentally.

Appendix: Text Similarity Detection

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

References

- [0] Improving Online-to-Nonconvex Conversion for Smooth Optimization via Double Optimism View paper
- [1] Lower bounds for non-convex stochastic optimization View paper
- [2] Recent theoretical advances in non-convex optimization View paper
- [3] Adaptive methods for nonconvex optimization View paper
- [4] Learning-Rate-Free Momentum SGD with Reshuffling Converges in Nonsmooth Nonconvex Optimization View paper
- [5] Stochastic first-order methods for convex and nonconvex functional constrained optimization View paper
- [6] Lookahead converges to stationary points of smooth non-convex functions View paper
- [7] Gradient-Free Methods for Deterministic and Stochastic Nonsmooth Nonconvex Optimization View paper
- [8] On the convergence of adaptive gradient methods for nonconvex optimization View paper
- [9] Linear convergence of first order methods for non-strongly convex optimization View paper
- [10] An implementable proximal-type method for computing critical points to minimization problems with a nonsmooth and nonconvex constraint View paper
- [11] Convergence of first-order methods for constrained nonconvex optimization with dependent data View paper
- [12] Regional complexity analysis of algorithms for nonconvex smooth optimization View paper
- [13] Lower Complexity Bounds of First-Order Methods for Affinely Constrained Composite Nonconvex Problems View paper
- [14] A Newton-CG based augmented Lagrangian method for finding a second-order stationary point of nonconvex equality constrained optimization with complexity â¦ View paper
- [15] An accelerated first-order method for non-convex optimization on manifolds View paper
- [16] Yet another fast variant of Newton's method for nonconvex optimization View paper
- [17] Identifying and attacking the saddle point problem in high-dimensional non-convex optimization View paper
- [18] UAdam: Unified Adam-Type Algorithmic Framework for Nonconvex Optimization View paper
- [19] Complexity of Adagrad and other first-order methods for nonconvex optimization problems with bounds constraints View paper
- [20] An adaptive high order method for finding third-order critical points of nonconvex optimization View paper
- [21] A retrospective approximation approach for smooth stochastic optimization View paper
- [22] Faster first-order methods for stochastic non-convex optimization on Riemannian manifolds View paper
- [23] Linear convergence of first-and zeroth-order primalâdual algorithms for distributed nonconvex optimization View paper
- [24] The computational complexity of finding stationary points in non-convex optimization View paper
- [25] Variance-reduced first-order methods for deterministically constrained stochastic nonconvex optimization with strong convergence guarantees View paper
- [26] Corrigendum: On the complexity of finding first-order critical points in constrained nonlinear optimization View paper
- [27] Identifying non-Hermitian critical points with quantum metric View paper
- [28] A Quaternion-Valued Neural Network Approach to Nonsmooth Nonconvex Constrained Optimization in Quaternion Domain View paper
- [29] ItsDEAL: Inexact two-level smoothing descent algorithms for weakly convex optimization View paper
- [30] Finding second-order stationary points efficiently in smooth nonconvex linearly constrained optimization problems View paper
- [31] Optimal stochastic non-smooth non-convex optimization through online-to-non-convex conversion View paper
- [32] First-order methods almost always avoid strict saddle points View paper
- [33] Efficient Algorithms for Smooth Minimax Optimization View paper
- [34] On the oracle complexity of first-order and derivative-free algorithms for smooth nonconvex minimization View paper
- [35] On the Complexity of Finding Stationary Points in Nonconvex Simple Bilevel Optimization View paper
- [36] Complexity-Optimal and Parameter-Free First-Order Methods for Finding Stationary Points of Composite Optimization Problems View paper
- [37] Transformers Learn Nonlinear Features In Context: Nonconvex Mean-field Dynamics on the Attention Landscape View paper
- [38] Simple linesearch-free first-order methods for nonconvex optimization View paper
- [39] Refining asymptotic complexity bounds for nonconvex optimization methods, including why steepest descent is o (eps^{-2}) rather than O (eps^{-2}) View paper
- [40] Efficient approaches for escaping higher order saddle points in non-convex optimization View paper
- [41] Block Coordinate Descent Methods for Structured Nonconvex Optimization with Nonseparable Constraints: Optimality Conditions and Global Convergence View paper
- [42] ItsOPT: An inexact two-level smoothing framework for nonconvex optimization via high-order Moreau envelope View paper
- [43] An accelerated first-order regularized momentum descent ascent algorithm for stochastic nonconvex-concave minimax problems View paper
- [44] A proximal-type method for nonsmooth and nonconvex constrained minimization problems View paper
- [45] The geometric effects of distributing constrained nonconvex optimization problems View paper
- [46] A global dual error bound and its application to the analysis of linearly constrained nonconvex optimization View paper
- [47] Computing proximal points of nonconvex functions View paper
- [48] Level constrained first order methods for function constrained optimization View paper
- [49] Non-convex bilevel games with critical point selection maps View paper
- [50] Refining asymptotic complexity bounds for nonconvex optimization methods, including why steepest descent is rather than View paper
- [51] High-probability complexity bounds for stochastic non-convex minimax optimization View paper
- [52] Data-Driven Compositional Optimization in Misspecified Regimes View paper
- [53] Trust region methods for nonconvex stochastic optimization beyond lipschitz smoothness View paper
- [54] A unified and refined convergence analysis for non-convex decentralized learning View paper

- [55] Online Optimization Perspective on First-Order and Zero-Order Decentralized Nonsmooth Nonconvex Stochastic Optimization View paper
- [56] A Novel Unified Parametric Assumption for Nonconvex Optimization View paper
- [57] A guide through the zoo of biased SGD View paper
- [58] Optimal complexity in non-convex decentralized learning over time-varying networks View paper
- [59] Cyclic Block Coordinate Descent With Variance Reduction for Composite Nonconvex Optimization View paper
- [60] Nest your adaptive algorithm for parameter-agnostic nonconvex minimax optimization View paper
- [61] Storm+: Fully adaptive sgd with recursive momentum for nonconvex optimization View paper
- [62] Analysis of Schedule-Free Nonconvex Optimization View paper
- [63] Towards Stability of Parameter-free Optimization View paper
- [64] STORM+: Fully adaptive SGD with momentum for nonconvex optimization View paper
- [65] Problem-Parameter-Free Federated Learning View paper
- [66] Agda+: Proximal alternating gradient descent ascent method with a nonmonotone adaptive step-size search for nonconvex minimax problems View paper
- [67] Comparison of optimization techniques based on gradient descent algorithm: A review View paper