Novelty Assessment Report

Paper: Discounted Online Convex Optimization: Uniform Regret Across a Continuous Interval
PDF URL: https://openreview.net/pdf?id=65iFtHZ8Cu
Venue: ICLR 2026 Conference Submission
Year: 2026
Report Generated: 2026-01-07

Abstract

Reflecting the greater significance of recent history over the distant past in non-stationary environments, λ-discounted regret has been introduced in online convex optimization (OCO) to gracefully forget past data as new information arrives. When the discount factor λ is given, online gradient descent with an appropriate step size achieves an $O(1/\sqrt{1-\lambda})$ discounted regret. However, the value of λ is often not predetermined in real-world scenarios. This gives rise to a significant \emph{open question}: is it possible to develop a discounted algorithm that adapts to an unknown discount factor. In this paper, we affirmatively answer this question by providing a novel analysis to demonstrate that smoothed OGD (SOGD) achieves a uniform $O(\sqrt{\log T/1-\lambda})$ discounted regret, holding for all values of λ across a continuous interval simultaneously. The basic idea is to maintain multiple OGD instances to handle different discount factors, and aggregate their outputs sequentially by an online prediction algorithm named as Discounted-Normal-Predictor (DNP). Our analysis reveals that DNP can combine the decisions of two experts, even when they operate on discounted regret with different discount factors.

Core Task Landscape

This paper addresses: **Adapting Online Convex Optimization to Unknown Discount Factors**

A total of **4 papers** were analyzed and organized into a taxonomy with **4 categories**.

Taxonomy Overview

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

- **Adaptive Discounting Mechanisms for Non-Stationary Environments**
- **Discounted Regret Analysis for Specific Problem Classes**

Complete Taxonomy Tree

- Adapting Online Convex Optimization to Unknown Discount Factors Survey Taxonomy
- Adaptive Discounting Mechanisms for Non-Stationary Environments
  - Uniform Regret Guarantees Across Discount Factor Intervals ★ (1 papers)
  - [0] Discounted Online Convex Optimization: Uniform Regret Across a Continuous Interval (Anon et al., 2026) View paper
  - Instance-Specific Adaptive Discounting (1 papers)
  - [2] Discounted adaptive online learning: Towards better regularization (Zhang Zhiyu, 2024) View paper
- Discounted Regret Analysis for Specific Problem Classes
  - Online Linear Regression with Dynamic Regret (1 papers)
  - [1] Online linear regression in dynamic environments via discounting (Jacobsen, 2024) View paper
  - Application-Driven Discounting in Constrained Optimization (2 papers)
  - [3] Online optimization for the smart (micro) grid (Narayanaswamy, 2012) View paper
  - [4] Online Posted Pricing with Unknown Time-Discounted Valuations (G, 2021) View paper

Narrative

Core task: Adapting online convex optimization to unknown discount factors. The field addresses how online learning algorithms can handle discounted objectives when the discount factor is not known in advance. The taxonomy reveals two main branches: one focusing on adaptive discounting mechanisms for non-stationary environments, and another on discounted regret analysis for specific problem classes. The first branch develops methods that adjust to changing or unknown discount rates, enabling algorithms to remain robust across different temporal weighting schemes. The second branch examines how discounted regret bounds can be derived for particular settings such as linear regression or pricing problems, often leveraging problem structure to achieve tighter guarantees. Representative works like Discounted Linear Regression[1] and Discounted Adaptive Learning[2] illustrate how specialized problem classes benefit from tailored analysis, while applications such as Smart Grid Optimization[3] and Posted Pricing Valuations[4] demonstrate the practical relevance of discounting in dynamic decision-making.

A central theme across these branches is the trade-off between generality and problem-specific efficiency: adaptive mechanisms aim for uniform performance across a range of discount factors, whereas specialized analyses exploit structure to improve bounds. Discounted Convex Optimization[0] sits within the adaptive discounting branch, specifically targeting uniform regret guarantees across discount factor intervals. This positions it closely to works like Discounted Adaptive Learning[2], which also emphasizes robustness to unknown parameters, but Discounted Convex Optimization[0] extends the scope to general convex settings rather than focusing on a narrower problem class. Compared to Discounted Linear Regression[1], which tailors its approach to linear models, the original paper pursues broader applicability at the cost of potentially looser constants. The main open question remains how to balance adaptivity with computational efficiency and whether problem-agnostic methods can match the performance of specialized algorithms in practice.

Related Works in Same Category

No sibling papers were found in the same taxonomy leaf. A taxonomy-subtopic-level comparison will be produced instead.

Taxonomy-Level Summary

Both subtopics address the challenge of handling unknown or varying discount factors in online convex optimization, but from complementary perspectives. The original leaf focuses on providing worst-case guarantees that hold uniformly across entire intervals of discount factors, while its sibling adapts discount factors dynamically based on observed problem characteristics.

Similarities: - Both address uncertainty about the appropriate discount factor in online optimization settings - Both aim to provide regret guarantees in discounted online convex optimization frameworks - Both represent adaptive approaches that avoid requiring prior knowledge of the optimal discount factor

Differences: - Uniform Regret Guarantees provides simultaneous bounds for all discount factors in a range (worst-case across the spectrum), while Instance-Specific Adaptive Discounting learns and exploits problem-specific structure - The original leaf emphasizes robustness across continuous discount factor intervals, whereas the sibling focuses on instance-dependent optimization - Uniform approaches likely have higher regret constants but broader applicability, while adaptive methods may achieve tighter bounds for specific problem instances

Suggested Search Directions: - Hybrid methods that combine uniform guarantees with instance-specific adaptation - Theoretical comparisons of regret bounds between uniform and adaptive discount factor approaches - Applications where discount factor uncertainty is critical (e.g., reinforcement learning with unknown horizons)

Sibling Subtopics

- **Instance-Specific Adaptive Discounting** (leaves: 1, papers: 1)
- Scope: Algorithms that learn optimal discount factors on-the-fly for specific problem instances.
- Exclude: Methods providing uniform guarantees across all discount factors belong to Uniform Regret Guarantees.

Contributions Analysis

Overall novelty summary. The paper addresses adapting online convex optimization to unknown discount factors by proposing a smoothed OGD algorithm that achieves uniform discounted regret bounds across a continuous interval of discount factors. According to the taxonomy, this work sits in the 'Uniform Regret Guarantees Across Discount Factor Intervals' leaf under 'Adaptive Discounting Mechanisms for Non-Stationary Environments'. Notably, this leaf contains only the original paper itself with no sibling papers, suggesting this specific research direction—achieving simultaneous guarantees for all discount factors in a range—is relatively unexplored within the examined literature.

The taxonomy reveals that the broader field divides into adaptive discounting mechanisms and problem-specific discounted regret analysis. The original paper's leaf sits alongside 'Instance-Specific Adaptive Discounting', which focuses on learning optimal discount factors for particular instances rather than providing uniform guarantees. The sibling branch 'Discounted Regret Analysis for Specific Problem Classes' contains work on linear regression and application-driven optimization, indicating that most prior research either targets specific problem structures or learns instance-specific parameters rather than pursuing the uniform guarantee approach taken here.

Among the three identified contributions, the literature search examined only five candidate papers total. The core contribution of uniform discounted regret bounds was examined against one candidate with no refutation found. The novel DNP analysis for combining experts was examined against four candidates, again with no refutations. The smoothed OGD framework itself was not examined against any candidates. Given this limited search scope of five papers, the analysis suggests these contributions appear distinct from the examined prior work, though the small candidate pool means substantial related work may exist beyond the top-K semantic matches retrieved.

Based on the limited search of five candidates and the sparse taxonomy leaf containing only the original paper, the work appears to occupy a relatively unexplored niche within discounted online learning. However, the small search scope and absence of sibling papers in the taxonomy leaf may reflect limitations in the literature retrieval process rather than definitive evidence of novelty. A more comprehensive search across online learning and non-stationary optimization literature would be needed to fully assess the contribution's originality.

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

Contribution 1: Uniform discounted regret bound across continuous discount factor interval

Description: The authors develop an algorithm that achieves $O(\sqrt{\log T}/(1-\lambda))$ discounted regret uniformly for all discount factors λ in a continuous interval $[1-1/\tau, 1-1/T]$, without requiring prior knowledge of the specific discount factor value.

This contribution was assessed against **1 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Minimax-regret climate policy with deep uncertainty in climate modeling and intergenerational discounting

URL: View paper

Brief Assessment

Minimax Climate Policy[5] addresses climate policy optimization under uncertainty about discount rates in economic modeling, not online convex optimization with uniform regret bounds across discount factor intervals.

Contribution 2: Novel analysis of DNP for combining experts with different discount factors

Description: The authors provide a novel theoretical analysis showing that Discounted-Normal-Predictor (DNP) can successfully aggregate decisions from experts operating under different discount factors, overcoming a key technical challenge in the meta-expert framework.

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. Dopamine neurons encode a multidimensional probabilistic map of future reward

URL: View paper

Brief Assessment

Dopamine Probabilistic Map[6] focuses on neuroscience and temporal difference learning in dopamine neurons, not on online convex optimization or meta-expert frameworks for combining algorithmic experts with different discount factors.

---

2. Competitive online algorithms for probabilistic prediction

URL: View paper

Brief Assessment

Competitive Probabilistic Prediction[9] focuses on competitive online algorithms for probabilistic prediction and expert aggregation in general online learning settings, not specifically on discounted regret or combining experts with different discount factors in the discounted online convex optimization framework.

3. Online Portfolio Hedging with the Weak Aggregating Algorithm

URL: View paper

Brief Assessment

Portfolio Hedging Algorithm[8] applies discounting in the context of portfolio hedging with the weak aggregating algorithm, not for combining experts operating under different discount factors in online convex optimization. The technical focus and application domain differ substantially from the original paper's contribution.

4. Temporal-difference reinforcement learning with distributed representations

URL: View paper

Brief Assessment

Distributed TD Learning[7] focuses on temporal-difference reinforcement learning with distributed representations of belief and discounting factors in a neuroscience context, not on the meta-expert framework for online convex optimization that combines experts with different discount factors as analyzed in the original paper.

Contribution 3: Smoothed OGD algorithm with sequential aggregation framework

Description: The authors propose a method that constructs multiple OGD instances with different step sizes corresponding to discretized discount factors, then uses DNP with conservative updating to sequentially aggregate their decisions, enabling adaptation to unknown discount factors.

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.

Appendix: Text Similarity Detection

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

References

- [0] Discounted Online Convex Optimization: Uniform Regret Across a Continuous Interval View paper
- [1] Online linear regression in dynamic environments via discounting View paper
- [2] Discounted adaptive online learning: Towards better regularization View paper
- [3] Online optimization for the smart (micro) grid View paper
- [4] Online Posted Pricing with Unknown Time-Discounted Valuations View paper
- [5] Minimax-regret climate policy with deep uncertainty in climate modeling and intergenerational discounting View paper
- [6] Dopamine neurons encode a multidimensional probabilistic map of future reward View paper
- [7] Temporal-difference reinforcement learning with distributed representations View paper
- [8] Online Portfolio Hedging with the Weak Aggregating Algorithm View paper
- [9] Competitive online algorithms for probabilistic prediction View paper