Novelty Assessment Report

Paper: Policy Newton Algorithm in Reproducing Kernel Hilbert Space

PDF URL: https://openreview.net/pdf?id=PYnLd91wZY

Venue: ICLR 2026 Conference Submission

Year: 2026

Report Generated: 2025-12-29

Abstract

Reinforcement learning (RL) policies represented in Reproducing Kernel Hilbert Spaces (RKHS) offer powerful representational capabilities. While second-order optimization methods like Newton's method demonstrate faster convergence than first-order approaches, current RKHS-based policy optimization remains constrained to first-order techniques. This limitation stems primarily from the intractability of explicitly computing and inverting the infinite-dimensional Hessian operator in RKHS. We introduce Policy Newton in RKHS, the first second-order optimization framework specifically designed for RL policies represented in RKHS. Our approach circumvents direct computation of the inverse Hessian operator by optimizing a cubic regularized auxiliary objective function. Crucially, we leverage the Representer Theorem to transform this infinite-dimensional optimization into an equivalent, computationally tractable finite-dimensional problem whose dimensionality scales with the trajectory data volume. We establish theoretical guarantees proving convergence to a local optimum with a local quadratic convergence rate. Empirical evaluations on a toy financial asset allocation problem validate these theoretical properties, while experiments on standard RL benchmarks demonstrate that Policy Newton in RKHS achieves superior convergence speed and higher episodic rewards compared to established first-order RKHS approaches and parametric second-order methods. Our work bridges a critical gap between non-parametric policy representations and second-order optimization methods in reinforcement learning.

Core Task Landscape

This paper addresses: **Second-Order Policy Optimization in Reproducing Kernel Hilbert Space**

A total of **6 papers** were analyzed and organized into a taxonomy with **7 categories**.

Taxonomy Overview

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

- **Second-Order Optimization Methods in RKHS**
- **Online Kernel Learning with Adaptive Complexity Reduction**
- **First-Order and Evolutionary Policy Optimization in RKHS**
- **Integral Reinforcement Learning with Computational Considerations**

Complete Taxonomy Tree

- Second-Order Policy Optimization in Reproducing Kernel Hilbert Space Survey Taxonomy
- Second-Order Optimization Methods in RKHS
  - Policy Optimization with Second-Order RKHS Methods ★ (1 papers)
  - [0] Policy Newton Algorithm in Reproducing Kernel Hilbert Space (Anon et al., 2026) View paper
  - Variational Inference with Second-Order RKHS Methods (1 papers)
  - [3] A Stein variational Newton method (Detommaso, 2018) View paper
  - Positive Function Optimization with Pseudo-Mirror Descent (1 papers)
  - [4] Sparse representations of positive functions via first-and second-order pseudo-mirror descent (Abhishek Chakraborty, 2022) View paper
- Online Kernel Learning with Adaptive Complexity Reduction (1 papers)
  - [5] Efficient second-order online kernel learning with adaptive embedding (Daniele Calandriello, 2017) View paper
- First-Order and Evolutionary Policy Optimization in RKHS
  - Evolutionary Strategy-Based Policy Search (1 papers)
  - [1] A covariance matrix adaptation evolution strategy for direct policy search in reproducing kernel Hilbert space (Ngo Anh Vien, 2017) View paper
  - Proximal Policy Optimization with RKHS Metrics (1 papers)
  - [6] PPO-CIM: Proximal Policy Optimization with Correntropy Induced Metric (Y Guoa, n.d.) View paper
- Integral Reinforcement Learning with Computational Considerations (1 papers)
  - [2] Impact of Computation in Integral Reinforcement Learning for Continuous-Time Control (Cao, 2024) View paper

Narrative

Core task: second-order policy optimization in reproducing kernel Hilbert space. The field structure reflects a spectrum of approaches for learning policies in RKHS, ranging from computationally intensive second-order methods that exploit curvature information to more scalable first-order and evolutionary strategies. The taxonomy organizes work into several main branches: one focuses on second-order optimization methods that leverage Hessian or Fisher information within the RKHS framework, another addresses online kernel learning with mechanisms to manage growing model complexity, a third encompasses first-order gradient-based and derivative-free evolutionary techniques, and a fourth examines integral reinforcement learning with an eye toward computational tractability. Representative works such as Stein Variational Newton[3] and Second-Order Online Kernel[5] illustrate how curvature-aware updates can be formulated in infinite-dimensional spaces, while methods like CMA-ES Direct Policy RKHS[1] demonstrate evolutionary alternatives that sidestep explicit gradient computation.

A particularly active line of inquiry concerns the trade-off between sample efficiency and computational overhead: second-order methods promise faster convergence by incorporating curvature, yet they often require expensive matrix operations or approximations to remain feasible in high-dimensional or online settings. Policy Newton RKHS[0] sits squarely within the branch of second-order RKHS methods, emphasizing Newton-type updates that exploit the geometry of the policy space. Its approach contrasts with lighter-weight schemes such as Sparse Pseudo-Mirror Descent[4], which sacrifices some curvature information to maintain sparsity and scalability, and with Computation Integral Reinforcement Learning[2], which prioritizes computational considerations in a related but distinct integral formulation. By adopting a full second-order perspective, Policy Newton RKHS[0] aligns closely with works that seek principled curvature exploitation, positioning itself as a rigorous yet computationally demanding option among the spectrum of RKHS-based policy optimization techniques.

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

These three subtopics represent distinct applications of second-order optimization methods in reproducing kernel Hilbert spaces (RKHS). The original leaf focuses on reinforcement learning policy optimization using Newton-type methods, while siblings address constrained function estimation (pseudo-mirror descent for nonnegative functions) and probabilistic inference (variational methods including Stein approaches). All leverage second-order functional derivatives but target different problem domains.

Similarities: - All three employ second-order optimization techniques in RKHS framework - Each involves functional optimization over infinite-dimensional spaces - All utilize kernel methods for representing and optimizing functions - Second-order information (Hessian or Fisher information analogs) is central to convergence properties

Differences: - Original leaf targets sequential decision-making (RL policies), while siblings focus on function estimation and distribution approximation - Pseudo-mirror descent enforces nonnegativity constraints (probability densities, trajectories), whereas policy optimization may not require such constraints - Variational inference methods approximate posterior distributions, while policy optimization maximizes expected returns - Policy optimization typically involves on-policy/off-policy data collection, while variational inference works with fixed datasets - Pseudo-mirror descent uses mirror descent geometry specific to positive functions, policy optimization uses policy gradient geometry

Suggested Search Directions: - Natural policy gradient methods that bridge policy optimization and information geometry - Actor-critic methods using kernel embeddings with second-order updates - Connections between policy optimization and variational inference (e.g., control as inference) - Second-order methods for constrained policy optimization with safety constraints

Sibling Subtopics

- **Positive Function Optimization with Pseudo-Mirror Descent** (leaves: 1, papers: 1)
- Scope: Second-order pseudo-mirror descent methods for nonnegative function estimation in RKHS, including maximum likelihood and trajectory optimization.
- Exclude: General policy optimization without nonnegativity constraints belongs under policy optimization categories.
- **Variational Inference with Second-Order RKHS Methods** (leaves: 1, papers: 1)
- Scope: Second-order functional optimization for variational inference and distribution approximation in RKHS, including Stein variational approaches.
- Exclude: Policy optimization and trajectory optimization belong under policy-focused or positive function categories.

Contributions Analysis

Overall novelty summary. The paper introduces a second-order optimization framework for reinforcement learning policies represented in reproducing kernel Hilbert spaces, specifically addressing the challenge of computing and inverting infinite-dimensional Hessian operators. According to the taxonomy, this work resides in the 'Policy Optimization with Second-Order RKHS Methods' leaf, which currently contains only this paper as its sole member. This positioning suggests the paper occupies a relatively sparse research direction within the broader field of RKHS-based policy optimization, where most existing work employs first-order or evolutionary approaches.

The taxonomy reveals that neighboring research directions include variational inference with second-order RKHS methods and positive function optimization using pseudo-mirror descent, both of which leverage curvature information but target different problem classes. The paper's approach diverges from the more populated 'First-Order and Evolutionary Policy Optimization in RKHS' branch, which encompasses gradient-based proximal methods and covariance matrix adaptation strategies. The taxonomy structure indicates that while second-order methods exist for related tasks like Stein variational inference, direct application to policy optimization in RKHS remains underexplored, with the paper attempting to bridge this gap.

Among the three identified contributions, the literature search examined 24 candidates total. The core algorithmic contribution (cubic regularization with finite-dimensional reduction) was assessed against 10 candidates, with 1 appearing to provide overlapping prior work. The theoretical convergence guarantees were evaluated against 10 candidates, with 2 potentially offering similar results. Notably, the claim of being the 'first second-order optimization framework for RKHS policies' was examined against 4 candidates with no clear refutations found. These statistics reflect a limited search scope rather than exhaustive coverage, suggesting that while some technical components have precedent, the specific integration for policy optimization may represent a novel synthesis.

Based on the available signals from 24 examined candidates, the work appears to occupy a genuinely sparse area within the taxonomy, though the limited search scope prevents definitive conclusions about absolute novelty. The absence of sibling papers in its taxonomy leaf and the mixed refutation results across contributions suggest the paper combines known techniques in a potentially novel configuration, though comprehensive assessment would require broader literature coverage beyond the top-K semantic matches analyzed here.

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

Contribution 1: Policy Newton in RKHS algorithm with cubic regularization and finite-dimensional reduction

Description: The authors introduce the first second-order optimization method for RL policies in RKHS by deriving the Hessian operator as a second-order Fréchet derivative and using a cubic regularized auxiliary function to avoid computing the intractable inverse. They leverage the Representer Theorem to transform the infinite-dimensional optimization into a tractable finite-dimensional problem whose dimension scales with trajectory data volume.

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. Second-order optimization for non-convex machine learning: An empirical study

URL: View paper

Brief Assessment

Second-Order Non-Convex Empirical[8] focuses on general non-convex machine learning problems using trust region and adaptive regularization with cubics (ARC) methods for neural networks, not on reinforcement learning policies in RKHS. The candidate does not address policy optimization, RKHS representations, or the Representer Theorem-based reduction specific to RL trajectories.

2. Second-order optimization with lazy Hessians

URL: View paper

Brief Assessment

Lazy Hessians[13] addresses second-order optimization in finite-dimensional Euclidean spaces with lazy (stale) Hessian updates, not infinite-dimensional RKHS policy optimization. The candidate does not address RL policies, RKHS representations, or the Representer Theorem-based reduction from infinite to finite dimensions that is central to the original contribution.

3. Efficient Second-Order Methods for Non-Convex Optimization and Machine Learning

URL: View paper

Brief Assessment

Efficient Second-Order Non-Convex[16] focuses on general non-convex optimization and machine learning problems using classic Newton-type methods (trust region, cubic regularization, Newton-CG) and a novel AdaHessian optimizer. It does not address reinforcement learning policies in RKHS or derive Hessian operators as second-order Fréchet derivatives in infinite-dimensional function spaces.

4. A Cubic-regularized Policy Newton Algorithm for Reinforcement Learning

URL: View paper

Prior Art Analysis

Cubic Policy Newton[10] demonstrates that cubic regularization for policy Newton methods in reinforcement learning was previously established. The candidate paper explicitly proposes 'a cubic regularized policy newton (cr-pn) method' and derives the cubic-regularized auxiliary objective function formulation. Both papers address the same core challenge: avoiding explicit computation of the Hessian inverse by using cubic regularization. The candidate also establishes the finite-dimensional reduction through the representer theorem is not novel, as it presents a cubic-regularized optimization problem that is 'optimized numerically using the newton-cg algorithm' after forming gradient and hessian estimates from trajectories.

Evidence

Evidence 1 - **Rationale**: Both papers propose using cubic regularization to avoid computing the Hessian inverse in policy Newton methods for RL. The candidate explicitly introduces this approach, demonstrating prior work exists on this specific technique. - **Original**: we introduce policy newton in rkhs, the first second-order optimization framework specifically designed for rl policies represented in rkhs. our approach circumvents direct computation of the inverse hessian operator by optimizing a cubic regularized auxiliary objective function within the rkhs - **Candidate**: we propose a cubic regularized policy newton (cr-pn) method that avoids saddle points and converges to an approximate second-order stationary point (sosp), where the approximation is quantified by a parameter ϵ > 0. such a point is referred to as ϵ-sosp. in this algorithm, we derive an estimate of t...

Evidence 2 - **Rationale**: The candidate demonstrates that forming finite-dimensional estimates from trajectory data for policy Newton methods was already established. Both papers reduce the problem to tractable computation using trajectory samples. - **Original**: crucially, we leverage the representer theorem to transform this infinite-dimensional optimization into an equivalent, computationally tractable finite-dimensional problem whose dimensionality scales with the trajectory data volume. - **Candidate**: from theorem 1, the policy gradient and hessian can be seen as expectations of g and h defined below. g(θ; τ) := ∇φ(θ; τ), h(θ; τ) := ∇φ(θ; τ)∇⊤ log p(τ; θ) + ∇2φ(θ; τ). (5) the above estimates are calculated by the information obtained from a given trajectoryτ and policy parameterθ. we simulate mul...

5. Cubic regularized subspace Newton for non-convex optimization

URL: View paper

Brief Assessment

Cubic Subspace Newton[9] addresses non-convex optimization in Euclidean space using coordinate subspaces, not policy optimization in RKHS for reinforcement learning. The technical domains are fundamentally different.

6. Second-Order Methods with Cubic Regularization Under Inexact Information

URL: View paper

Brief Assessment

Cubic Inexact Information[15] addresses cubic regularization for general convex optimization with inexact Hessians in Euclidean space, not RL policies in RKHS. The candidate does not work with policy optimization, trajectory data, or infinite-dimensional function spaces.

7. Adaptive Cubic Regularized Second-Order Latent Factor Analysis Model

URL: View paper

Brief Assessment

Adaptive Cubic Latent Factor[14] addresses latent factor analysis for recommender systems using cubic regularization on Gauss-Newton approximations, not reinforcement learning policy optimization in RKHS. The technical domains and problem formulations are fundamentally different.

8. Faster Riemannian Newton-type optimizationÂ by subsampling and cubic regularization

URL: View paper

Brief Assessment

Riemannian Cubic Subsampling[11] addresses constrained optimization on Riemannian manifolds (e.g., Grassmann, Stiefel), not policy optimization in reinforcement learning. The candidate's cubic regularization applies to general manifold optimization, whereas the original contribution specifically derives Hessian operators for RL policies in RKHS and uses the Representer Theorem for trajectory-based finite-dimensional reduction.

9. A Variance-Reduced Cubic-Regularized Newton for Policy Optimization

URL: View paper

Brief Assessment

Variance-Reduced Cubic Newton[7] focuses on parametric policy optimization with variance reduction techniques in Euclidean space, not on non-parametric RKHS policy representations or infinite-dimensional Hessian operators.

10. Rapid DP Convex Optimization via Curvature-Aware (Second-Order) Algorithms.

URL: View paper

Brief Assessment

Curvature-Aware DP Convex[12] focuses on differentially private convex optimization using cubic regularization, not reinforcement learning policy optimization in RKHS. The candidate addresses privacy-preserving optimization in a different problem domain.

Contribution 2: Theoretical convergence guarantees with local quadratic convergence rate

Description: The authors provide theoretical analysis proving that Policy Newton in RKHS converges to a local optimum and achieves a local quadratic convergence rate, establishing formal guarantees for the second-order method in the RKHS setting.

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. Quasi-Newton policy gradient algorithms

URL: View paper

Prior Art Analysis

Quasi-Newton Policy Gradient[28] demonstrates that prior work established local quadratic convergence for policy optimization algorithms before the original paper. The candidate paper proves that their quasi-newton method achieves local quadratic convergence near the optimal policy (Theorem 5), with the convergence rate $\|\pi(k+1) - \pi\| \leq c\|\pi(k) - \pi\|^2$. This directly refutes the novelty claim that the original paper was first to provide theoretical guarantees of local quadratic convergence for policy optimization in RKHS, as the candidate achieved the same convergence property in a different but related policy optimization framework.

Evidence

Evidence 1 - **Rationale**: Both papers claim to establish local quadratic convergence for policy optimization methods. The candidate explicitly states they demonstrate local quadratic convergence theoretically, which directly overlaps with the original's claimed contribution. - **Original**: we establish theoretical guarantees proving convergence to a local optimum with a local quadratic convergence rate - **Candidate**: in this paper, we investigate the quasi-newton approach for solving (5). the main contributions of this paper are the following. •first, we present a unified quasi-newton type method for the policy optimization problem... •second, we analyze the convergence property of the proposed quasi-newton algor...

Evidence 2 - **Rationale**: Both theorems establish the same mathematical property: local quadratic convergence with the form $\|x(k+1) - x\| \leq c\|x(k) - x\|^2$. The candidate's Theorem 5 predates the original paper and proves the same convergence rate for policy optimization, demonstrating that this theoretical guarantee was already established in prior work. - **Original**: theorem 4.6 (local quadratic convergence (d)) consider the deterministic policy newton rkhs method (algorithm 1 with ∇h ˆj(hk) = ∇hj(hk) and ∇2 h ˆj(hk) = ∇2 hj(hk)). assuming the norm of the inverse operator ‖(∇2 hj(hk) + β 2 ‖Δhk‖i)-1‖ is bounded by some constant b, and we assume that the update s... - **Candidate**: theorem 5. let η= 1, f(π)sa = -(ra s -((i-γpa)vπ)s) and φkl(π) = ∑ sa πa s log πa s, φrkl(π) = 1 |a| ∑ sa log 1 πas , φh(π) = -2 ∑ sa √πas |a|, φα(π) = 4 |a|(1 -α2) ∑ sa (1 -(|a|πa s)(1+α)/2),(α< 1). denote the k-th policy obtained in algorithm 1 by π(k), then for η = 1 the update scheme in algorith...

2. Augmented Proximal Policy Optimization for Safe Reinforcement Learning

URL: View paper

Brief Assessment

Augmented Proximal PPO[27] focuses on safe RL with constraint satisfaction using augmented Lagrangian methods, not on second-order policy optimization in RKHS or quadratic convergence rates for policy Newton methods.

3. Approximate Newton policy gradient algorithms

URL: View paper

Prior Art Analysis

Approximate Newton Policy[24] demonstrates that prior work exists establishing theoretical convergence guarantees with local quadratic convergence rates for policy optimization algorithms. The candidate paper explicitly proves local quadratic convergence for their approximate newton method applied to policy gradient algorithms with entropy regularization, predating the original paper's submission. Both papers establish formal guarantees for second-order methods achieving quadratic convergence rates in policy optimization settings.

Evidence

Evidence 1 - **Rationale**: Both papers claim to establish theoretical guarantees for local quadratic convergence. The candidate paper explicitly states they demonstrate local quadratic convergence both theoretically and numerically for their approximate newton method. - **Original**: we establish theoretical guarantees proving convergence to a local optimum with a local quadratic convergence rate. - **Candidate**: we analyze the convergence property of the proposed approximate newton algorithms and demonstrate local quadratic convergence both theoretically and numerically.

Evidence 2 - **Rationale**: The candidate paper provides a dedicated section proving quadratic convergence with a formal theorem (Theorem 3.1), establishing that their method enjoys quadratic local convergence to the optimal policy. - **Original**: we establish theoretical guarantees for the proposed algorithm, proving convergence to a local optimum, and demonstrating a quadratic convergence rate. - **Candidate**: in this section, we study the quadratic convergence of the approximate newton method at the learning rate η = 1, which corresponds to the step size used in the newton method. our analysis is inspired by the results in [8, 37]. the following theorem states the second-order convergence when η = 1

4. Global convergence of policy gradient methods to (almost) locally optimal policies

URL: View paper

Brief Assessment

Global Policy Gradient Convergence[26] focuses on policy gradient methods for general MDPs and establishes convergence to (almost) locally optimal policies under strict saddle point assumptions. The original paper addresses second-order policy optimization in RKHS with cubic regularization and proves local quadratic convergence for the specific RKHS setting, which is a distinct technical contribution not covered by the candidate.

5. Data-enabled policy optimization for the linear quadratic regulator

URL: View paper

Brief Assessment

Data-Enabled LQR[22] focuses on linear quadratic regulator problems with projected gradient dominance and sublinear convergence, not the RKHS policy optimization setting with quadratic convergence analyzed in the original paper.

6. Geometry and convergence of natural policy gradient methods

URL: View paper

Brief Assessment

Natural Policy Gradient Geometry[21] focuses on natural policy gradient methods in MDPs with different geometries (Fisher, conditional entropy). The original paper addresses second-order optimization in RKHS for RL policies, a distinct technical framework.

---

7. On the convergence rates of policy gradient methods

URL: View paper

Brief Assessment

Policy Gradient Convergence Rates[20] focuses on policy gradient and policy mirror descent methods for tabular MDPs with direct policy parametrization, establishing convergence rates but not specifically addressing second-order methods in RKHS settings or cubic regularization frameworks as in the original paper.

---

8. Robust Policy Optimization in Continuous-time Mixed H2/Hâ Stochastic Control

URL: View paper

Brief Assessment

Robust Mixed Control[25] addresses continuous-time stochastic control with mixed H2/H∞ objectives, not RKHS-based policy optimization. The convergence analysis applies to different problem settings and mathematical frameworks.

---

9. Fast global convergence of natural policy gradient methods with entropy regularization

URL: View paper

Brief Assessment

Natural Policy Gradient Entropy[23] focuses on linear convergence rates for entropy-regularized NPG methods in tabular MDPs, not quadratic convergence for second-order RKHS methods. The candidate addresses first-order policy gradient with entropy regularization, while the original develops second-order Newton methods in RKHS.

---

10. Solving time-continuous stochastic optimal control problems: Algorithm design and convergence analysis of actor-critic flow

URL: View paper

Brief Assessment

Actor-Critic Flow[29] addresses time-continuous stochastic optimal control with linear convergence rates for the actor-critic flow, not the local quadratic convergence of second-order policy optimization in RKHS that the original paper establishes.

Contribution 3: First second-order optimization framework for RKHS policies in reinforcement learning

Description: The work bridges a critical gap by developing the first second-order optimization framework tailored for reinforcement learning policies represented in Reproducing Kernel Hilbert Spaces, addressing the limitation that previous RKHS policy optimization was constrained to first-order 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. Robust Policy Gradient Optimization through Parameter Perturbation in Reinforcement Learning

URL: View paper

Brief Assessment

Robust Parameter Perturbation[19] focuses on parameter perturbations during optimization for standard policy gradient methods, not on second-order optimization in RKHS. The candidate does not address RKHS policy representations or second-order methods in functional spaces.

---

2. Inverse KKT: Learning cost functions of manipulation tasks from demonstrations

URL: View paper

Brief Assessment

Inverse KKT[17] focuses on inverse reinforcement learning (learning cost functions from demonstrations) using second-order optimization in RKHS, not on direct policy optimization in RL. The candidate's RKHS functional is for cost function recovery, not policy representation.

---

3. Impact of Computation in Integral Reinforcement Learning for Continuous-Time Control

URL: View paper

Brief Assessment

Computation Integral Reinforcement Learning[2] focuses on continuous-time control and integral reinforcement learning with quadrature methods for policy evaluation, not on developing second-order optimization frameworks for RKHS-represented policies in general reinforcement learning settings.

---

4. Deep Policy Gradient Methods, RKHS and Convergence Guarantees of Neural Network Parameterized Policies

URL: View paper

Brief Assessment

Deep Policy Gradient RKHS[18] is a 2016 proposal document that mentions investigating second-order methods (approximate Hessians) for RKHS policies as future work, but does not present a complete framework, derivation, or implementation. The original paper provides the first actual second-order optimization framework with cubic regularization, finite-dimensional reduction via representer theorem, and convergence guarantees.

Appendix: Text Similarity Detection

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

References

- [0] Policy Newton Algorithm in Reproducing Kernel Hilbert Space View paper

- [1] A covariance matrix adaptation evolution strategy for direct policy search in reproducing kernel Hilbert space View paper
- [2] Impact of Computation in Integral Reinforcement Learning for Continuous-Time Control View paper
- [3] A Stein variational Newton method View paper
- [4] Sparse representations of positive functions via first-and second-order pseudo-mirror descent View paper
- [5] Efficient second-order online kernel learning with adaptive embedding View paper
- [6] PPO-CIM: Proximal Policy Optimization with Correntropy Induced Metric View paper
- [7] A Variance-Reduced Cubic-Regularized Newton for Policy Optimization View paper
- [8] Second-order optimization for non-convex machine learning: An empirical study View paper
- [9] Cubic regularized subspace Newton for non-convex optimization View paper
- [10] A Cubic-regularized Policy Newton Algorithm for Reinforcement Learning View paper
- [11] Faster Riemannian Newton-type optimizationÂ by subsampling and cubic regularization View paper
- [12] Rapid DP Convex Optimization via Curvature-Aware (Second-Order) Algorithms. View paper
- [13] Second-order optimization with lazy Hessians View paper
- [14] Adaptive Cubic Regularized Second-Order Latent Factor Analysis Model View paper
- [15] Second-Order Methods with Cubic Regularization Under Inexact Information View paper
- [16] Efficient Second-Order Methods for Non-Convex Optimization and Machine Learning View paper
- [17] Inverse KKT: Learning cost functions of manipulation tasks from demonstrations View paper
- [18] Deep Policy Gradient Methods, RKHS and Convergence Guarantees of Neural Network Parameterized Policies View paper
- [19] Robust Policy Gradient Optimization through Parameter Perturbation in Reinforcement Learning View paper
- [20] On the convergence rates of policy gradient methods View paper
- [21] Geometry and convergence of natural policy gradient methods View paper
- [22] Data-enabled policy optimization for the linear quadratic regulator View paper
- [23] Fast global convergence of natural policy gradient methods with entropy regularization View paper
- [24] Approximate Newton policy gradient algorithms View paper
- [25] Robust Policy Optimization in Continuous-time Mixed H2/Hâ Stochastic Control View paper
- [26] Global convergence of policy gradient methods to (almost) locally optimal policies View paper
- [27] Augmented Proximal Policy Optimization for Safe Reinforcement Learning View paper
- [28] Quasi-Newton policy gradient algorithms View paper
- [29] Solving time-continuous stochastic optimal control problems: Algorithm design and convergence analysis of actor-critic flow View paper