

READING LIST

COSC 594/650 – Optimization for Machine Learning

Instructor: Dongsheng Ding Fall 2026

This reading list helps students who have not yet identified a project explore possible topics in optimization for machine learning. The papers and suggested questions cover both foundational and application-oriented directions. Students may use them as starting points or explore other relevant topics. A focused special case, one application on a modest dataset, or adaptation of an existing model can provide a useful scope. Check data, code, and computing needs early; reproducing large-scale pretraining is not expected. Theoretical projects do not require code. Follow the *Project Guideline* for requirements, deadlines, and grading.

This reading list was prepared with assistance from OpenAI Codex. The readings pair foundations with selected 2025–2026 advances. Brief annotations explain their relevance; reported benchmark improvements concern the cited paper’s setting, rather than a universal ranking of methods. Students need only pursue readings relevant to their own interests.

Gradient Methods: Speed, Noise, and Generalization

Possible directions. When does restarting acceleration help, and when does noise undermine it? How does gradient clipping trade bias against rare large errors? Can step size or early stopping improve held-out accuracy rather than just training loss? These questions can motivate special-case analysis or a study of regression or image classification with noisy features, labels, or uneven sampling.

- O’Donoghue and Candès. *Adaptive Restart for Accelerated Gradient Schemes*. FoCM, 2015. A foundational reference on adaptive restart and accelerated optimization.
- Gorbunov et al. *Stochastic Optimization with Heavy-Tailed Noise via Accelerated Gradient Clipping*. NeurIPS 2020. High-probability convergence under finite-variance gradient noise in smooth convex optimization.
- Hübler, Fatkhullin, and He. *From Gradient Clipping to Normalization for Heavy Tailed SGD*. AISTATS 2025. Recent nonconvex stochastic-optimization guarantees under weak noise-moment assumptions.
- Hardt, Recht, and Singer. *Train faster, generalize better: Stability of stochastic gradient descent*. ICML 2016. Connects optimization choices, sensitivity to training examples, and expected generalization.

Fairness and Constraints Learned from Data

Possible directions. Can a classifier satisfy a fairness or false-positive-rate constraint on unseen data? How do rare groups, validation size, and repeated model selection affect that conclusion? On a small census-classification dataset, how do fixed penalties and dual updates compare in accuracy, held-out violations, and sensitivity to minibatch size? Basic questions include feasibility, multiplier behavior, and confidence bounds for estimated constraints.

- Cotter et al. *Training Well-Generalizing Classifiers for Fairness Metrics and Other Data-Dependent Constraints*. ICML 2019. A foundation for separating optimization and constraint generalization.
- Kliachkin et al. *Benchmarking Stochastic Approximation Algorithms for Fairness-Constrained Training of Deep Neural Networks*. ICLR 2026. A recent Folktables benchmark of constrained solvers, with practical comparisons and gaps between neural-network experiments and available guarantees.

Min-Max Optimization and Adversarial Learning

Possible directions. Why can descent-ascent cycle even when a saddle point exists? When do extragradient, anchoring, or unequal primal and dual step sizes stabilize the final iterate? In a small generative model, do those improvements also help sample quality and robustness to noisy minibatches? Linear critics

or quadratic saddle problems offer accessible settings for analysis; a limited image-generation experiment offers an application direction.

- Nemirovski. *Prox-Method with Rate of Convergence $O(1/t)$ for Variational Inequalities with Lipschitz Continuous Monotone Operators and Smooth Convex-Concave Saddle Point Problems*. SIOPT, 2004. The classical mirror-prox foundation.
- Gidel et al. *A Variational Inequality Perspective on Generative Adversarial Networks*. ICLR 2019. Connects saddle-point algorithms to adversarial generative learning.
- Gorbunov et al. *Clipped Stochastic Methods for Variational Inequalities with Heavy-Tailed Noise*. NeurIPS 2022. High-probability analysis of noisy saddle-point methods under stated moment assumptions.
- Cai and Zheng. *Last-Iterate Convergence of Anchored Gradient Descent*. arXiv preprint, 2026. A recent last-iterate result for monotone problems; it does not establish convergence for arbitrary neural adversarial models.

Efficient Training and Adaptation of Neural Networks

Possible directions. Do gains from a pretraining optimizer transfer to fine-tuning a smaller pretrained text or vision model? How do update normalization, matrix shape, or adapter learning rates affect accuracy at a fixed compute budget? A focused application could examine domain adaptation or image classification; a basic investigation could isolate conditioning or low-rank factorization. Training steps, total computation, and wall-clock time need not rank methods identically.

- Hayou, Ghosh, and Yu. *LoRA+: Efficient Low Rank Adaptation of Large Models*. ICML 2024. Analysis and experiments on unequal learning rates for the two adapter factors.
- Liu et al. *Muon is Scalable for LLM Training*. Technical report, 2025. Reports matched pretraining loss at about 52% of AdamW’s FLOPs in fitted scaling experiments; fine-tuning gains are less consistent.
- Pethick et al. *Training Deep Learning Models with Norm-Constrained LMOs*. ICML 2025. Scion connects norm constraints and gradient geometry to efficient vision and language-model training.
- Li et al. *NorMuon: Making Muon more efficient and scalable*. ICML 2026. Combines orthogonalized updates with neuron-wise adaptation; reports 21.74% fewer training steps than AdamW in its 1.1B-parameter pretraining comparison.

Preference Learning and Language-Model Safety

Possible directions. Does improving a preference-training loss improve response quality on new prompts? How do reference-policy regularization and imperfect reward scores affect that relationship? For a small pretrained assistant, can dual updates better control a helpfulness–harm trade-off than a fixed penalty? A response-selection study with a frozen model is a lighter alternative. Basic questions include duality, approximation error, and uncertainty in estimated constraints.

- Rafailov et al. *Direct Preference Optimization: Your Language Model is Secretly a Reward Model*. NeurIPS 2023. An influential foundation for direct optimization from preference data.
- Huang et al. *One-Shot Safety Alignment for Large Language Models via Optimal Dualization*. NeurIPS 2024. Connects constrained alignment to an idealized distributional dual problem and practical alignment.
- Zhang et al. *Alignment of Large Language Models with Constrained Learning*. NeurIPS 2025. Recent iterative dual-based alignment with empirical helpfulness–safety trade-offs; restricted models and imperfect inner optimization matter.

Diffusion Models: Training, Control, and Unlearning

Possible directions. How do diffusion-time sampling and stopping time affect gradient variance and unseen-data performance? Can reward-based fine-tuning improve prompt following, object counts, or text rendering without reducing diversity? Can a model remove one concept while preserving unrelated capabilities, and when do dual updates become unstable? A small image model or a narrow adaptation of an existing checkpoint can keep an application study manageable.

- Nichol and Dhariwal. *Improved Denoising Diffusion Probabilistic Models*. ICML 2021. A training foundation; Section 3.3 studies diffusion-time importance sampling.
- Han, Razaviyayn, and Xu. *Neural Network-Based Score Estimation in Diffusion Models: Optimization and Generalization*. ICLR 2024. Early-stopping theory in a specialized wide-network setting.
- Zheng et al. *DiffusionNFT: Online Diffusion Reinforcement with Forward Process*. ICLR 2026. Recent diffusion post-training: reports GenEval 0.98 in roughly 1,000 steps and up to 25-fold wall-clock speedup over FlowGRPO in tested reward-training comparisons.
- Khalafi, Ribeiro, and Ding. *Unlearning in Diffusion Models: A Unified Framework with KL Divergence and Likelihood Constraints*. ICML 2026. Recent constrained unlearning and retention–removal trade-offs; some neural likelihood-constrained runs are unstable.

Constrained Reinforcement Learning and Robust Control

Possible directions. How does tightening a cost budget affect reward and final-policy feasibility? In a small navigation or control simulator, does a safety margin reduce observed violations, and what reward is lost? Does a nominally feasible policy remain feasible after transition perturbations? Tabular models permit basic convergence or sensitivity analysis; one lightweight control task permits an empirical study of constraints and model mismatch.

- Ding, Zhang, Duan, Başar, and Jovanović. *Convergence and Sample Complexity of Natural Policy Gradient Primal-Dual Methods for Constrained MDPs*. JMLR, 2025. A theoretical reference for policy optimization, feasibility, and sampling; the exact tabular case is an accessible starting point.
- Ding, Wei, Zhang, and Ribeiro. *Last-Iterate Convergent Policy Gradient Primal-Dual Methods for Constrained MDPs*. NeurIPS 2023. Regularized and optimistic approaches to final-policy convergence.
- Milosevic, Müller, and Scherf. *Embedding Safety into RL: A New Take on Trust Region Methods*. ICML 2025. Constraint-aware geometry; reports fewer violations with competitive returns in tested tasks, without a pathwise safety guarantee for the sampled implementation.
- Ganguly et al. *Efficient Policy Optimization in Robust Constrained MDPs with Iteration Complexity Guarantees*. NeurIPS 2025. Recent complexity results with transition uncertainty, under specific feasibility and structural assumptions.

AI Use Statement

OpenAI Codex assisted with literature discovery, scope refinement, drafting, and L^AT_EX formatting. References were checked against primary records. The questions are optional starting points for independent work, not verified claims of novelty.