**NIPS2018:**

On Markov Chain Gradient Descent:

<https://arxiv.org/pdf/1809.04216.pdf>

The Convergence of Sparsified Gradient Methods:

<https://arxiv.org/pdf/1809.10505.pdf>

Learning Overparameterized Neural Networks via Stochastic Gradient Descent on Structured Data

<https://arxiv.org/pdf/1808.01204.pdf>

LAG: Lazily Aggregated Gradient for Communication-Efficient Distributed Learning

<https://arxiv.org/pdf/1805.09965.pdf>

Convergence of Cubic Regularization for Nonconvex Optimization under KL Property

<https://arxiv.org/pdf/1808.07382.pdf>

The Physical Systems Behind Optimization Algorithms

<https://arxiv.org/pdf/1612.02803.pdf>

Accelerated Stochastic Matrix Inversion: General Theory and Speeding up BFGS Rules for Faster Second-Order Optimization

<https://arxiv.org/pdf/1802.04079.pdf>

**ICML2018:**

Dissecting Adam: The Sign, Magnitude and Variance of Stochastic Gradients

<http://proceedings.mlr.press/v80/balles18a.html>

DRACO: Byzantine-resilient Distributed Training via Redundant Gradients

<http://proceedings.mlr.press/v80/chen18l.html>

ADMM and Accelerated ADMM as Continuous Dynamical Systems

<http://proceedings.mlr.press/v80/franca18a.html>

Fourier Policy Gradients

<http://proceedings.mlr.press/v80/fellows18a.html>

Gradient Primal-Dual Algorithm Converges to Second-Order Stationary Solution for Nonconvex Distributed Optimization Over Networks

<http://proceedings.mlr.press/v80/hong18a.html>

Gradient Descent for Sparse Rank-One Matrix Completion for Crowd-Sourced Aggregation of Sparsely Interacting Workers

<http://proceedings.mlr.press/v80/ma18b.html>

A Delay-tolerant Proximal-Gradient Algorithm for Distributed Learning

<http://proceedings.mlr.press/v80/mishchenko18a.html>

A Conditional Gradient Framework for Composite Convex Minimization with Applications to Semidefinite Programming

<http://proceedings.mlr.press/v80/yurtsever18a.html>

Stabilizing Gradients for Deep Neural Networks via Efficient SVD Parameterization

<http://proceedings.mlr.press/v80/zhang18g.html>

Policy Optimization as Wasserstein Gradient Flows

<http://proceedings.mlr.press/v80/zhang18a.html>

Accelerating Greedy Coordinate Descent Methods

<http://proceedings.mlr.press/v80/lu18b.html>

Characterizing Implicit Bias in Terms of Optimization Geometry

<http://proceedings.mlr.press/v80/gunasekar18a.html>

Convergence guarantees for a class of non-convex and non-smooth optimization problems

<http://proceedings.mlr.press/v80/khamaru18a.html>

Nonconvex Optimization for Regression with Fairness Constraints

<http://proceedings.mlr.press/v80/komiyama18a.html>

Distributed Asynchronous Optimization with Unbounded Delays: How Slow Can You Go?

<http://proceedings.mlr.press/v80/zhou18b.html>

**IEEE Transactions on Image Processing:**

Fast Image Recovery Using Variable Splitting and Constrained Optimization

<http://www.lx.it.pt/~mtf/Afonso_BioucasDias_Figueiredo_twocolumn_v7.pdf>

An Augmented Lagrangian Approach to the Constrained Optimization Formulation of Imaging Inverse Problems

<https://ieeexplore.ieee.org/document/5570998>

Fastℓ1-Minimization Algorithms for Robust Face Recognition

<https://ieeexplore.ieee.org/abstract/document/6514938>

An Augmented Lagrangian Method for Total Variation Video Restoration

<https://ieeexplore.ieee.org/abstract/document/5779734>

Robust Multichannel Blind Deconvolution via Fast Alternating Minimization

<https://ieeexplore.ieee.org/abstract/document/6074943>